7450 Ethernet Service Switch
7750 Service Router
7950 Extensible Routing System
Virtualized Service Router

Release 23.3.R1

Layer 3 Services Guide: IES and VPRN Services

3HE 19222 AAAA TQZZA 01
Edition 01
March 2023

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

Table of contents

1 Getting started

1.1 About this guide

This guide describes Layer 3 service functionality provided by SR-series routers and presents examples to configure and implement various protocols and services.

This guide is organized into functional chapters and provides concepts and descriptions of the implementation flow, as well as Command Line Interface (CLI) syntax and command usage.

Note: Unless otherwise indicated, this guide uses classic CLI command syntax and configuration examples.

The topics and commands described in this document apply to the:

- 7450 ESS
- 7750 SR
- 7950 XRS
- Virtualized Service Router

For a list of unsupported features by platform and chassis, see the *SR OS R23.x.Rx Software Release Notes*, part number 3HE 19269 000 *x* TQZZA.

Command outputs shown in this guide are examples only; actual displays may differ depending on supported functionality and user configuration.

Note:
The SR OS CLI trees and command descriptions can be found in the following guides:

- *7450 ESS, 7750 SR, 7950 XRS, and VSR Classic CLI Command Reference Guide*
- *7450 ESS, 7750 SR, 7950 XRS, and VSR Clear, Monitor, Show, and Tools Command Reference Guide* (for both MD-CLI and Classic CLI)
- *7450 ESS, 7750 SR, 7950 XRS, and VSR MD-CLI Command Reference Guide*

Note:
This guide generically covers Release 23.*x*.R*x* content and may contain some content that will be released in later maintenance loads. Please see the *SR OS R23.x.Rx Software Release Notes*, part number 3HE 19269 000 *x* TQZZA, or for information about features supported in each load of the Release 23.*x*.R*x* software.

1.2 Layer 3 services configuration process

Table 1: Configuration process lists tasks related to the configuration and implementation of Layer 3 Services.

This guide is presented in an overall logical configuration flow. Each section describes a software area and provides CLI syntax and command usage to configure parameters for a functional area.

Table 1: Configuration process

Area	Task	Section
Internet Enhanced Service (IES)	Configure an IES service	Configuring an IES service with CLI
	Configure IES components	Configuring IES components
	Service management	Service management tasks
Virtual Private Routed Network (VPRN) Service	Configure a VPRN service	Configuring a VPRN service with CLI
	Configure VPRN components	Configuring VPRN components
	Service management	Service management tasks

1.3 Conventions

This section describes the general conventions used in this guide.

1.3.1 Precautionary and information messages

The following information symbols are used in the documentation.

DANGER: Danger warns that the described activity or situation may result in serious personal injury or death. An electric shock hazard could exist. Before you begin work on this equipment, be aware of hazards involving electrical circuitry, be familiar with networking environments, and implement accident prevention procedures.

WARNING: Warning indicates that the described activity or situation may, or will, cause equipment damage, serious performance problems, or loss of data.

Caution: Caution indicates that the described activity or situation may reduce your component or system performance.

Note: Note provides additional operational information.

Tip: Tip provides suggestions for use or best practices.

1.3.2 Options or substeps in procedures and sequential workflows

Options in a procedure or a sequential workflow are indicated by a bulleted list. In the following example, at step 1, the user must perform the described action. At step 2, the user must perform one of the listed options to complete the step.

Example: Options in a procedure

1. User must perform this step.
2. This step offers three options. User must perform one option to complete this step.
   - This is one option.
   - This is another option.
   - This is yet another option.

Substeps in a procedure or a sequential workflow are indicated by letters. In the following example, at step 1, the user must perform the described action. At step 2, the user must perform two substeps (a. and b.) to complete the step.

Example: Substeps in a procedure

1. User must perform this step.
2. User must perform all substeps to complete this action.
   a. This is one substep.
   b. This is another substep.

2 IES

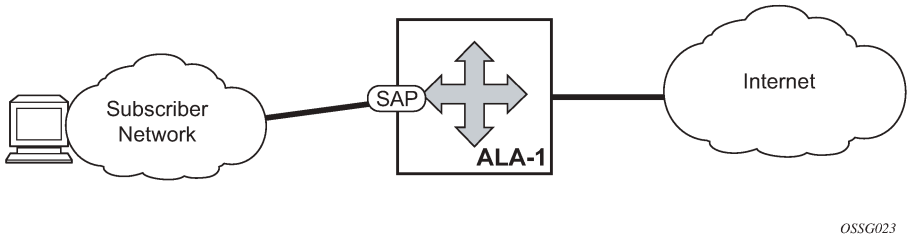
2.1 IES service overview

Internet Enhanced Service (IES) is a routed connectivity service where the subscriber communicates with an IP router interface to send and receive Internet traffic. An IES has one or more logical IP routing interfaces each with a SAP which acts as the access point to the subscriber's network. IES allows customer-facing IP interfaces to participate in the same routing instance used for service network core routing connectivity. IES services require that the IP addressing scheme used by the subscriber be unique between other provider addressing schemes and potentially the entire Internet.

While IES is part of the routing domain, the usable IP address space may be limited. This allows a portion of the service provider address space to be reserved for service IP provisioning, and be administered by a separate but subordinate address authority.

IP interfaces defined within the context of an IES service must have a SAP associated as the uplink access point to the subscriber network. Multiple IES services are created to segregate subscriber-owned IP interfaces.

Figure 1: Internet enhanced service



The IES service provides Internet connectivity. Other features include:

- multiple IES services are created to separate customer-owned IP interfaces
- more than one IES service can be created for a single customer ID
- more than one IP interface can be created within a single IES service ID

  All IP interfaces created within an IES service ID belong to the same customer.

These features apply to the 7750 SR and 7450 ESS.

See the *7450 ESS, 7750 SR, and VSR Triple Play Service Delivery Architecture Guide* for information about how subscriber group-interfaces function in the Routed Central Office model.

2.2 IES features

This section describes the 7450 ESS and 7750 SR service features and any special capabilities or considerations as they relate to IES services.

2.2.1 IP interfaces

IES customer IP interfaces can be configured with most of the same options found on the core IP interfaces. The advanced configuration options supported are:

- QoS Policy Propagation Using BGP (QPPB)
- VRRP - for IES services with more than one IP interface
- Cflowd
- Secondary IP addresses
- ICMP Options

Configuration options found on core IP interfaces not supported on IES IP interfaces are:

- MPLS forwarding
- NTP broadcast receipt

2.2.1.1 QoS Policy Propagation Using BGP

This section discusses QPPB as it applies to VPRN, IES, and router interfaces. See the IES section and the "IP Router Configuration" section in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

QoS policy propagation using BGP (QPPB) is a feature that allows a route to be installed in the routing table with a forwarding-class and priority so that packets matching the route can receive the associated QoS. The forwarding-class and priority associated with a BGP route are set using BGP import route policies. In the industry, this feature is called QPPB, and even though the feature name refers to BGP specifically. On SR OS, QPPB is supported for BGP (IPv4, IPv6, VPN-IPv4, VPN-IPv6), RIP and static routes.

While SAP ingress and network QoS policies can achieve the same end result as QPPB, the effort involved in creating the QoS policies, keeping them up-to-date, and applying them across many nodes is much greater than with QPPB. This is because of assigning a packet, arriving on a particular IP interface, to a specific forwarding-class and priority/profile, based on the source IP address or destination IP address of the packet. In a typical application of QPPB, a BGP route is advertised with a BGP community attribute that conveys a particular QoS. Routers that receive the advertisement accept the route into their routing table and set the forwarding-class and priority of the route from the community attribute.

2.2.1.1.1 QPPB applications

There are two typical applications of QPPB:

- coordination of QoS policies between different administrative domains
- traffic differentiation within a single domain, based on route characteristics

2.2.1.1.2 Inter-AS coordination of QoS policies

The operator of an administrative domain A can use QPPB to signal to a peer administrative domain B that traffic sent to specific prefixes advertised by domain A should receive a particular QoS treatment in domain B. More specifically, an ASBR of domain A can advertise a prefix XYZ to domain B and include a BGP

community attribute with the route. The community value implies a particular QoS treatment, as agreed by the two domains (in their peering agreement or service level agreement, for example). When the ASBR and other routers in domain B accept and install the route for XYZ into their routing table, they apply a QoS policy on selected interfaces that classifies traffic toward network XYZ into the QoS class implied by the BGP community value.

QPPB may also be used to request that traffic sourced from specific networks receive appropriate QoS handling in downstream nodes that may span different administrative domains. This can be achieved by advertising the source prefix with a BGP community, as discussed above. However, in this case other approaches are equally valid, such as marking the DSCP or other CoS fields based on source IP address so that downstream domains can take action based on a common understanding of the QoS treatment implied by different DSCP values.

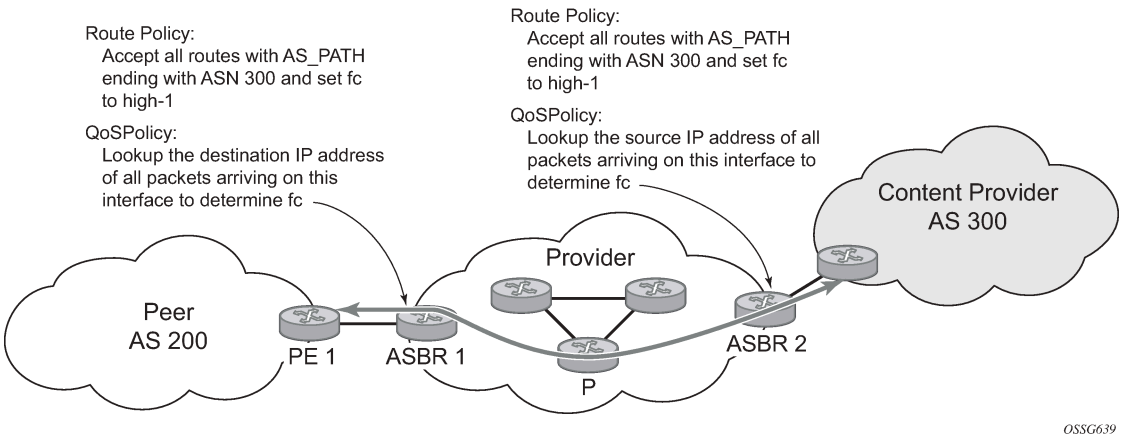
In the above examples, coordination of QoS policies using QPPB could be between a business customer and its IP VPN service provider, or between one service provider and another.

2.2.1.1.3 Traffic differentiation based on route characteristics

There may be times when a network operator wants to provide differentiated service to specific traffic flows within its network, and these traffic flows can be identified with known routes. For example, the operator of an ISP network may want to give priority to traffic originating in a particular ASN (the ASN of a content provider offering over-the-top services to the ISP's customers), following a specific AS_PATH, or destined for a particular next-hop (remaining on-net vs. off-net).

Figure 2: Use of QPPB to differentiate traffic in an ISP network shows an example of an ISP that has an agreement with the content provider managing AS300 to provide traffic sourced and terminating within AS300 with differentiated service appropriate to the content being transported. In this example, we presume that ASBR1 and ASBR2 mark the DSCP of packets terminating and sourced, respectively, in AS300 so that other nodes within the ISP's network do not need to rely on QPPB to determine the correct forwarding-class to use for the traffic. The DSCP or other CoS markings could be left unchanged in the ISP's network and QPPB used on every node.

Figure 2: Use of QPPB to differentiate traffic in an ISP network



2.2.1.2 QPPB

There are two main aspects of the QPPB feature:

- the ability to associate a forwarding-class and priority with specific routes in the routing table.
- the ability to classify an IP packet arriving on a particular IP interface to the forwarding-class and priority associated with the route that best matches the packet.

2.2.1.2.1 Associating an FC and priority with a route

This feature uses a command in the route-policy hierarchy to set the forwarding class and optionally the priority associated with routes accepted by a route-policy entry. The command has the following structure:

fc fc-name [priority {low | high}]

The use of this command is illustrated by the following example:

```
config>router>policy-options
begin
community gold members 300:100
policy-statement qppb_policy
entry 10
from
protocol bgp
community gold
exit
action accept
fc h1 priority high
exit
exit
exit
commit
```

The **fc** command is supported with all existing from and to match conditions in a route policy entry and with any action other than reject, it is supported with next-entry, next-policy and accept actions. If a next-entry or next-policy action results in multiple matching entries, then the last entry with a QPPB action determines the forwarding class and priority.

A route policy that includes the **fc** command in one or more entries can be used in any import or export policy but the **fc** command has no effect except in the following types of policies:

- VRF import policies:
  - config>service>vprn>vrf-import
- BGP import policies:
  - config>router>bgp>import
  - config>router>bgp>group>import
  - config>router>bgp>group>neighbor>import
  - config>service>vprn>bgp>import
  - config>service>vprn>bgp>group>import
  - config>service>vprn>bgp>group>neighbor>import
- RIP import policies:
  - config>router>rip>import
  - config>router>rip>group>import
  - config>router>rip>group>neighbor>import

- config>service>vprn>rip>import
- config>service>vprn>rip>group>import
- config>service>vprn>rip>group>neighbor>import

As evident from above, QPPB route policies support routes learned from RIP and BGP neighbors of a VPRN as well as for routes learned from RIP and BGP neighbors of the base/global routing instance.

QPPB is supported for BGP routes belonging to any of the address families listed below:

- IPv4 (AFI=1, SAFI=1)
- IPv6 (AFI=2, SAFI=1)
- VPN-IPv4 (AFI=1, SAFI=128)
- VPN-IPv6 (AFI=2, SAFI=128)

A VPN-IP route may match both a VRF import policy entry and a BGP import policy entry (if vpn-apply-import is configured in the base router BGP instance). In this case the VRF import policy is applied first and then the BGP import policy, so the QPPB QoS is based on the BGP import policy entry.

This feature also introduces the ability to associate a forwarding-class and optionally priority with IPv4 and IPv6 static routes. This is achieved by specifying the forwarding-class within the static-route-entry next-hop or indirect context.

Priority is optional when specifying the forwarding class of a static route, but when configured it can only be deleted and returned to unspecified by deleting the entire static route.

2.2.1.2.2 Displaying QoS information associated with routes

The following commands are enhanced to show the forwarding-class and priority associated with the displayed routes:

- **show router route-table**
- **show router fib**
- **show router bgp routes**
- **show router rip database**
- **show router static-route**

This feature uses a **qos** keyword to the **show>router>route-table** command. When this option is specified the output includes an additional line per route entry that displays the forwarding class and priority of the route. If a route has no fc and priority information then the third line is blank. The following CLI shows an example:

show router route-table [**family**] [*ip-prefix*[/*prefix-length*]] [**longer** | **exact**] [**protocol** *protocol-name*] **qos**

An example output of this command is shown below:

```
A:Dut-A# show router route-table 10.1.5.0/24 qos
===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix                                 Type    Proto    Age         Pref
       Next Hop[Interface Name]                                  Metric
       QoS
-------------------------------------------------------------------------------
10.1.5.0/24                                 Remote  BGP      15h32m52s   0
```

```
        PE1_to_PE2                                                      0
        h1, high
-------------------------------------------------------------------------------
No. of Routes: 1
===============================================================================
A:Dut-A#
```

2.2.1.2.3 Enabling QPPB on an IP interface

To enable QoS classification of ingress IP packets on an interface based on the QoS information associated with the routes that best match the packets the **qos-route-lookup** command is necessary in the configuration of the IP interface. The **qos-route-lookup** command has parameters to indicate whether the QoS result is based on lookup of the source or destination IP address in every packet. There are separate **qos-route-lookup** commands for the IPv4 and IPv6 packets on an interface, which allows QPPB to enabled for IPv4 only, IPv6 only, or both IPv4 and IPv6. The current QPPB based on a source IP address is not supported for IPv6 packets nor is it supported for ingress subscriber management traffic on a group interface.

The **qos-route-lookup** command is supported on the following types of IP interfaces:

- base router network interfaces (config>router>interface)
- VPRN SAP and spoke SDP interfaces (config>service>vprn>interface)
- VPRN group-interfaces (config>service>vprn>sub-if>grp-if)
- IES SAP and spoke SDP interfaces (config>service>ies>interface)
- IES group-interfaces (config>service>ies>sub-if>grp-if)

When the **qos-route-lookup** command with the destination parameter is applied to an IP interface and the destination address of an incoming IP packet matches a route with QoS information the packet is classified to the fc and priority associated with that route, overriding the fc and priority/profile determined from the sap-ingress or network qos policy associated with the IP interface. If the destination address of the incoming packet matches a route with no QoS information the fc and priority of the packet remain as determined by the sap-ingress or network qos policy.

Similarly, when the **qos-route-lookup** command with the source parameter is applied to an IP interface and the source address of an incoming IP packet matches a route with QoS information the packet is classified to the fc and priority associated with that route, overriding the fc and priority/profile determined from the sap-ingress or network qos policy associated with the IP interface. If the source address of the incoming packet matches a route with no QoS information the fc and priority of the packet remain as determined by the sap-ingress or network qos policy.

Currently, QPPB is not supported for ingress MPLS traffic on network interfaces or on CsC PE'-CE' interfaces (config>service>vprn>nw-if).

Note: QPPB based on a source IP address is not supported for ingress subscriber management traffic on a group interface.

2.2.1.2.4 QPPB when next hops are resolved by QPPB routes

In some circumstances (IP VPN inter-AS model C, Carrier Supporting Carrier, indirect static routes, and so on) an IPv4 or IPv6 packet may arrive on a QPPB-enabled interface and match a route A1 whose next-hop N1 is resolved by a route A2 with next-hop N2 and perhaps N2 is resolved by a route A3 with next-hop N3, and so on. The QPPB result is based only on the forwarding-class and priority of route A1. If A1 does not

have a forwarding-class and priority association then the QoS classification is not based on QPPB, even if routes A2, A3, and so on. have forwarding-class and priority associations.

2.2.1.2.5 QPPB and multiple paths to a destination

When ECMP is enabled some routes may have multiple equal-cost next-hops in the forwarding table. When an IP packet matches such a route the next-hop selection is typically based on a hash algorithm that tries to load balance traffic across all the next-hops while keeping all packets of a specific flow on the same path. The QPPB configuration model described in Associating an FC and priority with a route allows different QoS information to be associated with the different ECMP next-hops of a route. The forwarding-class and priority of a packet matching an ECMP route is based on the particular next-hop used to forward the packet.

When BGP fast reroute [1] is enabled some BGP routes may have a backup next-hop in the forwarding table in addition to the one or more primary next-hops representing the equal-cost best paths allowed by the ECMP/multipath configuration. When an IP packet matches such a route a reachable primary next-hop is selected (based on the hash result) but if all the primary next-hops are unreachable then the backup next-hop is used. The QPPB configuration model described in Associating an FC and priority with a route allows the forwarding-class and priority associated with the backup path to be different from the QoS characteristics of the equal-cost best paths. The forwarding class and priority of a packet forwarded on the backup path is based on the **fc** and priority of the backup route.

2.2.1.2.6 QPPB and policy-based routing

When an IPv4 or IPv6 packet with destination address X arrives on an interface with both QPPB and policy-based-routing enabled:

- There is no QPPB classification if the IP filter action redirects the packet to a directly connected interface, even if X is matched by a route with a forwarding-class and priority
- QPPB classification is based on the forwarding-class and priority of the route matching IP address Y if the IP filter action redirects the packet to the indirect next-hop IP address Y, even if X is matched by a route with a forwarding-class and priority

2.2.1.3 QPPB and GRT lookup

Source-address based QPPB is not supported on any SAP or spoke SDP interface of a VPRN configured with the **grt-lookup** command.

2.2.1.3.1 QPPB interaction with SAP ingress QoS policy

When QPPB is enabled on a SAP IP interface the forwarding class of a packet may change from **fc1**, the original **fc** determined by the SAP ingress QoS policy to fc2, the new fc determined by QPPB. In the ingress datapath SAP ingress QoS policies are applied in the first P chip and route lookup/QPPB occurs in the second P chip. This has the implications listed below:

- Ingress remarking (based on profile state) is always based on the original fc (fc1) and sub-class (if defined)
- The profile state of a SAP ingress packet that matches a QPPB route depends on the configuration of **fc2** only. If the de-1-out-profile flag is enabled in **fc2** and **fc2** is not mapped to a priority mode queue,

then the packet is marked out of profile if its DE bit = 1. If the profile state of **fc2** is explicitly configured (in or out) and **fc2** is not mapped to a priority mode queue then the packet is assigned this profile state. In both cases, there is no consideration of whether **fc1** was mapped to a priority mode queue.

- The priority of a SAP ingress packet that matches a QPPB route depends on several factors. If the de-1-out-profile flag is enabled in **fc2** and the DE bit is set in the packet then priority is low regardless of the QPPB priority or **fc2** mapping to profile mode queue, priority mode queue or policer. If **fc2** is associated with a profile mode queue then the packet priority is based on the explicitly configured profile state of **fc2** (in profile = high, out profile = low, undefined = high), regardless of the QPPB priority or **fc1** configuration. If **fc2** is associated with a priority mode queue or policer then the packet priority is based on QPPB (unless DE=1), but if no priority information is associated with the route then the packet priority is based on the configuration of **fc1** (if **fc1** mapped to a priority mode queue then it is based on DSCP/IP prec/802.1p and if **fc1** mapped to a profile mode queue then it is based on the profile state of **fc1**).

Table 2: QPPB interactions with SAP ingress QoS summarizes the interactions.

Table 2: QPPB interactions with SAP ingress QoS

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
Profile mode queue	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class
Priority mode queue	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Policer	Policer	From new base FC unless overridden by DE=1	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Priority mode queue	Policer	From new base FC unless overridden by DE=1	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
Policer	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Profile mode queue	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then follows original FC's profile mode rules.	From new base FC	From original FC and sub-class
Priority mode queue	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class
Profile mode queue	Policer	From new base FC unless overridden by DE=1	If DE=1 override, then low otherwise from QPPB. If no DEI or QPPB overrides, then follows original FC's profile mode rules.	From new base FC	From original FC and sub-class
Policer	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class

2.2.1.4 Object grouping and state monitoring

This feature introduces a generic operational group object which associates different service endpoints (pseudowires and SAPs) located in the same or in different service instances. The operational group status is derived from the status of the individual components using specific rules specific to the application using the concept. A number of other service entities, the monitoring objects, can be configured to monitor the operational group status and to perform specific actions as a result of status transitions. For example, if the operational group goes down, the monitoring objects are brought down.

2.2.1.4.1 IES IP interface applicability

This concept is used by an IPv4 IES interface to affect the operational state of the IP interface monitoring the operational group. Individual SAP and spoke SDPs are supported as monitoring objects.

The following rules apply:

- An object can only belong to one group at a time.
- An object that is part of a group cannot monitor the status of a group.
- An object that monitors the status of a group cannot be part of a group.
- An operational group may contain any combination of member types, SAP or Spoke-SDPs.
- An operational group may contain members from different VPLS service instances.
- Objects from different services may monitor the oper-group.

There are two steps involved in enabling the functionality:

1. Identify a set of objects whose forwarding state should be considered as a whole group then group them under an operational group using the **oper-group** command.
2. Associate the IP interface to the oper-group using the **monitor-group** command.

The status of the operational group (oper-group) is dictated by the status of one or more members according to the following rules:

- The oper-group goes down if all the objects in the oper-group go down. The oper-group comes up if at least one component is up.
- An object in the group is considered down if it is not forwarding traffic in at least one direction. That could be because the operational state is down or the direction is blocked through some validation mechanism.
- If a group is configured but no members are specified yet, then its status is considered up.
- As soon as the first object is configured the status of the operational group is dictated by the status of the provisioned members.

The following configuration shows the oper-group g1, the VPLS SAP that is mapped to it and the IP interfaces in IES service 2001 monitoring the oper-group g1. This example uses an R-VPLS context. The VPLS instance includes the **allow-ip-int-bind** and the **name v1**. The IES interface links to the VPLS using the **vpls v1** option. All commands are under the configuration service hierarchy.

To further describe the configuration. Oper-group g1 has a single SAP (1/1/1:2001) mapped to it and the IP interfaces in the IES service 2001 derive its state from the state of oper-group g1.

```
oper-group g1 create

vpls 1 name "v1" customer 1 create
            allow-ip-int-bind
            stp
                shutdown
            exit
            sap 1/1/1:2001 create
                oper-group g1
                eth-cfm
                   mep domain 1 association 1 direction down
 ccm-enable
       no shutdown
                exit
            exit
            sap 1/1/2:2001 create
```

```
            exit
            sap 1/1/3:2001 create
            exit
no shutdown


ies 2001 customer 1 create
            interface "i2001"  create
                address 192.168.1.1/24
                monitor-oper-group "g1"
                vpls "v1"
            exit
          no shutdown
          exit
```

2.2.2 Subscriber interfaces

Subscriber interfaces are composed of a combination of two key technologies, subscriber interfaces and group interfaces. While the subscriber interface defines the subscriber subnets, the group interfaces are responsible for aggregating the SAPs. Subscriber interfaces apply to the 7450 ESS and 7750 SR only.

- **subscriber interface**

  This is an interface that allows the sharing of a subnet among one or many group interfaces in the routed CO model.

- **group interface**

  This interface aggregates multiple SAPs on the same port.

- **redundant interfaces**

  This is a special spoke-terminated Layer 3 interface. It is used in a Layer 3 routed CO dual-homing configuration to shunt downstream (network to subscriber) to the active node for a specific subscriber. Redundant interfaces apply to the 7750 SR only.

2.2.2.1 IPv6 ESM

All IPv6 Enhanced Subscriber Management (ESM) services require either Routed CO (IES), or Routed CO for VPRN as a supporting service construct. Because of the complexities of the IPv6 link-model, there is currently no support for IPv6 ESM in a VPLS. There is also currently no support for IPv6 in combination with Basic Subscriber Management (BSM). This feature applies to the 7450 ESS and 7750 SR only.

2.2.2.2 RADIUS accounting

In the 7750 SR OS, the accounting paradigm is based on sla-profile instances, yet this is at odds with traditional RADIUS authentication and accounting which is host-centric. In previous OS releases, it was possible to have many hosts sharing a common sla-profile instance, and therefore accounting and QoS parameters. Complications would arise with RADIUS accounting because Accounting-Start and Accounting-Stop are a function of sla-profile instance and not the hosts – this meant that some host-specific parameters (like Framed-Ip-Address) would not be consistently included in RADIUS accounting.

Dual-stack subscribers are now two different hosts sharing a single sla-profile instance. A new RADIUS accounting mode has been introduced to support multiple-host environments.

A new command, **host-accounting**, is introduced under **accounting-policy**, which allows configurable behavior.

No host-accounting:

When **no host-accounting** is configured, accounting behavior is as follows:

- A RADIUS accounting start message is sent when the SLA-profile instance is created. It contains accounting (octets/packets) and the Framed-Ip-Address of the host which caused the sla-profile instance to be created.
- Additional hosts may bind to the sla-profile instance at any time, but no additional accounting messages are sent during these events.
- If the original host disconnects, then future accounting messages use an IP address of one of the remaining hosts.
- When the final host associated with an sla-profile instance disconnects, an accounting stop message is sent.

Host-accounting enabled:

When **host-accounting** is configured, additional RADIUS accounting messages are created for host activity in addition to messages for common queue accounting. The behavior is as follows:

- A RADIUS accounting start message is sent each time a host is authenticated. It contains the Framed-Ip-Address among other things. It does not contain any octet or packet counts.
- A RADIUS accounting start message is sent each time a sla-profile instance is created.
- Whenever a host disconnects a RADIUS, accounting stop message is sent for that host.
- If all host associated with an sla-profile instance disconnect, a RADIUS accounting stop message is sent for that instance.

This behavior means specific AVP may be in either host, sla-profile instance, or both accounting records. See Table 3: RADIUS accounting table .

Note:
Interim-Acct records are not sent for hosts, only the start- and stop-accounting messages.

Table 3: RADIUS accounting table

RADIUS accounting AVP	Host accounting	SLA-profile accounting
User-Name	Yes	—
NAS-Identifier	Yes	Yes
NAS-IP-Address	Yes	Yes
Nas-Port-Id	Yes	—
Nas-Port	Yes	—
Nas-Port-Type	Yes	—
Service-Type	Yes	—
Framed-Protocol	Yes	—

RADIUS accounting AVP	Host accounting	SLA-profile accounting
Framed-Ip-Address	Yes	—
Framed-Ip-Netmask	Yes	—
Framed-Route	Yes	—
Class	Yes	—
Session-Timeout	Yes	Yes
Circuit-Id VSA	Yes	—
Called-Station-Id	Yes	—
Calling-Station-Id	Yes	—
MAC-Addr VSA	Yes	—
Remote-Id VSA	Yes	—
Acct-Input-Octets	—	Yes
Acct-Output-Octets	—	Yes
Acct-Input-Gigawords	—	Yes
Acct-Output-Gigawords	—	Yes
Acct-Session-Id	Yes	Yes
Acct-Session-Time	Yes	Yes
Acct-Input-Packets	—	Yes
Acct-Output-Packets	—	Yes
Agent-Circuit-Id	Yes	—
Agent-Remote-Id	Yes	—
Actual-Data-Rate-Upstream	Yes	—
Actual-Data-Rate-Downstream	Yes	—
Access-Loop-Encapsulation	Yes	—
Alc-Accounting	—	Yes
Alc-Subscriber-Id	Yes	Yes
Alc-Subscriber-Profile-String	Yes	Yes
Alc-Sla-Profile-String	Yes	Yes

2.2.2.3 RADIUS connectivity

SR OS supports IPv4/IPv6 RADIUS connectivity for dot1x host authentication. IPv4 connectivity is supported through legacy configuration using the **config>system>security>radius** command. IPv4/IPv6 connectivity is supported through common RADIUS backend using the **config>router>radius-server** command.

2.2.2.3.1 RADIUS connectivity on VPRN or GRT

The RADIUS connectivity used for dot1x authentication is reachable through the following services:

- GRT
- VPRN
- R-VPLS (VPRN, IES)

For R-VPLS, the port terminating the dot1x is a VPLS/EVPN SAP, but the RADIUS server is reachable through the same R-VPLS L3 Service, that is the VPRN or IES that the R-VPLS is connected to.

2.2.2.3.2 RADIUS support for IPv6 server connectivity

The current RADIUS profile used for dot1x authentication under **config>system>security>dot1x radius-policy** does not support IPv6 connectivity. When the system-wide RADIUS server is configured under the **config>router>radius-server** or **config>service>vprn>radius-server** contexts, IPv6 connectivity for dot1x host authentication is supported.

2.2.2.3.3 Common backend RADIUS server support

The dot1x RADIUS servers use the common backend RADIUS server. The common backend RADIUS server can be configured using the **config>router>radius-server** or **config>service>vprn>radius-server** contexts, and supports both IPv4 and IPv6 RADIUS connectivity.

These RADIUS servers can be placed under a RADIUS policy that is configured under **config>aaa>radius-server-policy** context. Multiple RADIUS servers can be added to a RADIUS server policy. See the **radius-server-policy** command description for more information.

The RADIUS policy can be assigned to a dot1x through **radius-server-policy** under dot1x configuration.

The following example shows the command usage for a base router configuration:

```
*A:swsim100>config>router>radius-server# server 1
  - no server <server-name>
  - server <server-name> [address <ip-address>] [secret <key >] [hash|hash2|
    custom] [create]
 <server-name>        : [32 chars max]
 <ip-address>         : ipv4-address   - a.b.c.d
                        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
 <key >               : secret-key - [64 chars max]
                        hash-key
                        hash2-key
```

```
                        custom-key
 <hash|hash2|custom>  : keywords - specify hashing scheme
 <create>             : keyword
```

The following example shows the command usage for a service VPRN configuration:

```
*A:swsim100>config>service>vprn>radius-server# server 1
  - no server <server-name>
  - server <server-name> [address <ip-address>] [secret <key >] [hash|hash2|
    custom] [create]
 <server-name>        : [32 chars max]
 <ip-address>         : ipv4-address   - a.b.c.d
                        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
 <key >               : secret-key - [64 chars max]
                        hash-key
                        hash2-key
                        custom-key
 <hash|hash2|custom>  : keywords - specify hashing scheme
 <create>             : keyword
```

After a server is configured, a RADIUS policy can be created from the configured servers. For server redundancy, there can be multiple servers for a policy.

The following is an example configuration output:

```
*A:swsim100>config>aaa# info
----------------------------------------------
        radius-server-policy "test" create
            servers
                router 60
                server 1 name "test"
                server 2 name "test2"
            exit
        exit
```

These RADIUS server policies can be configured against a port dot1x configuration using the **dot1x radius-server-policy** command.

```
*A:swsim100> configure port 1/1/c1/1 ethernet dot1x radius-server-policy "test"
```

2.2.2.3.4 Authentication and accounting RADIUS server

The user can configure separate authentication and accounting RADIUS servers for dot1x. Two separate RADIUS policies must be set and assigned correctly to the dot1x RADIUS server policy for authentication and accounting.

2.2.2.3.5 Accounting considerations for per-host configurations

When the dot1x is configured as per-host, accounting is disabled and a zero (0) value is provided for the following attributes:

- Acct-Input-Octets (RFC 2866)

- Acct-Output-Octets (RFC 2866)
- Acct-Input-Packets (RFC 2866)
- Acct-Output-Packets (RFC 2866)
- Acct-Input-Gigawords (RFC 2869)
- Acct-Output-Gigawords (RFC 2869)

2.2.3 SAPs

2.2.3.1 Encapsulations

The following SAP encapsulations are supported on the IES services:

- Ethernet null
- Ethernet dot1q

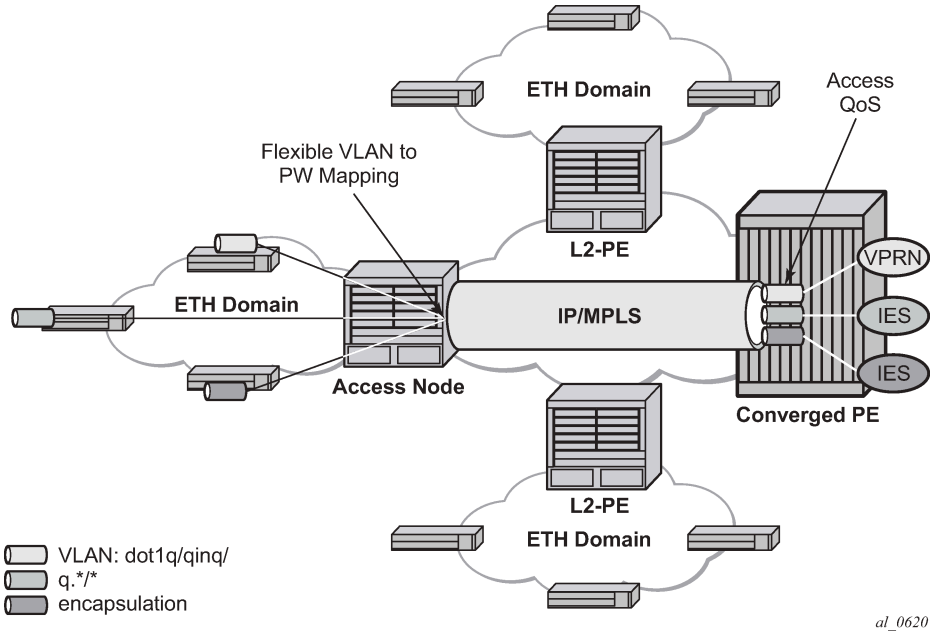
2.2.3.2 Pseudowire SAPs

This feature allows customers of an IES, VPRN, or Epipe VLL service and connected to an Ethernet SAP on an Access PE to be backhauled through an Ethernet aggregation network using MPLS pseudowires terminating directly on a converged PE hosting the IES, VPRN, or Epipe VLL service. If Enhanced Subscriber Management over PW is also used, then the converged PE may also act as a BNG. This service is different from VLL Spoke-SDP termination on an IES or VPRN because access QoS policies can be applied directly at a centralized PE hosting the IES or VPRN instance. This feature uses the same concepts of pseudowire ports and pseudowire SAPs that are used for ESM over MPLS pseudowires, described in the 7450 ESS, 7750 SR, and VSR Triple Play Service Delivery Architecture Guide.

The MPLS pseudowire originates from the first hop aggregation PE (referred to as access PE) upstream of the Access-Node (or directly from a multi-service AN), and terminates on the converged PE. Multiple customers from a specific access-port on the Access-PE can be backhauled over a single MPLS pseudowire toward the converged PE. This capability allows the network to scale and does not require an MPLS pseudowire per customer between the Access-PE and the converged PE. The access-port on the Access-PE can be dot1q, q-in-q or NULL encapsulated. The converged PE terminates the MPLS pseudowire, decapsulates the received frames, and provides access QoS functions including HQoS, without requiring an internal or external loopback. Each MPLS pseudowire is represented on the BNG as a "PW-port" for which SAPs are created. These SAPs are termed "PW SAPs", and must be statically configured on IES or VPRN interfaces or under the VLL service (unlike the ESM case where a capture SAP can be configured). The underlying Ethernet port must be in hybrid mode. Pseudowire SAPs are supported on Ethernet MDAs.

Figure 3: Network architecture using pseudowire SAPs illustrates the architecture of an aggregation network that uses pseudowire SAPs.

Figure 3: Network architecture using pseudowire SAPs



2.2.3.3 Encapsulation

The packet is encapsulated on an Ethernet pseudowire, which is associated with a pseudowire port on the converged PE, and a spoke SDP on the access PE. The SDP could use an LDP LSP, RSVP LSP, segment routed tunnel, BGP RFC 8277 tunnel, or LDP over RSVP tunnel. Hash labels are not supported. The SDP may be bound to a port or a LAG, although shaping Vports for pseudowire ports on LAGs in distributed mode is not supported. If an SDP is rerouted, then the corresponding pseudowire ports are brought operationally down. Pseudowire ports are associated with an SDP by configuration.

2.2.3.4 Pseudowire SAP configuration

Prerequisites

The following prerequisites are required at the access PE:

1. Configure an Epipe VLL service.
2. Configure a NULL, 1q or q-in-q SAP on the Epipe service.

The steps in this procedure are used to configure a pseudowire SAP on the IES or VPRN service at the Layer 3 PE.

The PW SAP may be mated to an Ethernet SAP or an Ethernet Spoke-SDP in the Epipe VLL service. A PW SAP may also form a part of an Epipe service that contains a PBB tunnel, as shown in the following example.

```
configure service epipe 300 name "access-vc-100" customer 1 create
    pbb tunnel 30123 backbone-dest-mac 00-00-5e-00-53-01 isid 100
    sap lag-3:25.100 create
       no shutdown
```

```
        exit
    sap pw-8:10.20 create
       no shutdown
       exit
```

Procedure

Step 1. Define a pseudowire port.

```
pw-port 1 create
exit
pw-port 2 create
 exit
```

Step 2. Bind a physical port or LAG, in hybrid mode, with the pseudowire port.

```
service customer 1 create
   multi-service-site “abc” create
     assignment port pw-1
     egress
      policer-control-policy “abc”
      exit
     exit description “Default customer”
    exit
  sdp 1 mpls create
   far-end 10.1.1.2
   ldp
   path-mtu 1514
   keep-alive
   shutdown
   exit
   binding
    port lag-1
    pw-port 1 vc-id 1 create
     no shutdown
     exit
    pw-port 2 vc-id 2 create
     no shutdown
     exit
  exit
no shutdown
exit
```

Step 3. Depending on whether the PW SAP is on an IES/VPRN or an Epipe VLL, perform one of the following steps:

- For a PW SAP on an IES/VPRN, configure a SAP on the IES or VPRN interface, with a SAP ID that uses the form **pw-***id*.

```
ies 1 customer 1 create
 interface “ies if” create
 address 192.168.1.1/24
 mac 00:00:00:00:00:ff
  static-arp 192.168.1.2 00:00:00:00:00:aa
  sap pw-1:1 create
   exit
  exit
  no shutdown
   exit
```

- For a PW SAP on an Epipe VLL, configure a SAP on the service, with a SAP ID that uses the form **pw-***id*.

```
epipe 1 customer 1 create
   sap pw-1:1 create
    exit
   exit
   no shutdown
    exit
```

2.2.3.5 QoS for pseudowire ports and pseudowire SAPs

Pseudowire SAPs support the QoS models allowed for regular VLL, IES or VPRN SAPs. These include:

- Per-service HQoS

  Note: This allows shaping of the total traffic per access node (and total traffic per class per AN), assuming one pseudowire per AN from the A-PE.

- SAP QoS support as available on the IOM3-XP, including
  - H-QoS (service scheduler child to port scheduler parent)
    - SAP queues attached to H-QoS scheduler by 'parent' statement
    - Scheduler attached to Port Scheduler by 'port-parent' statement
  - Direct service queue to port scheduler mapping
    - Aggregate-rate-limit

  Support for the redirection of SAP egress queues to an access queue group instance. It is possible to redirect SAP queues of a pseudowire SAP using the SAP based redirection for the IOM with Ethernet MDA, and policy based redirection for the IOM with Ethernet MDA, as applicable.
- Policing and H-POL

2.2.3.6 Shaping and bandwidth control

Pseudowire SAPs can be shaped on egress by a Vport on a physical port. The pseudowire SAP egress cannot explicitly declare which Vport to use, but they inherit the Vport used by the PW port egress shaping.

The intermediate destination identifier, used for ESM on MPLS pseudowires, is not applicable to VLL, IES and VPRN pseudowire SAPs.

If a pseudowire port is configured on a LAG, then Vport shaping is only supported if the LAG is in link mode.

Per-access node shaping is configured as follows:

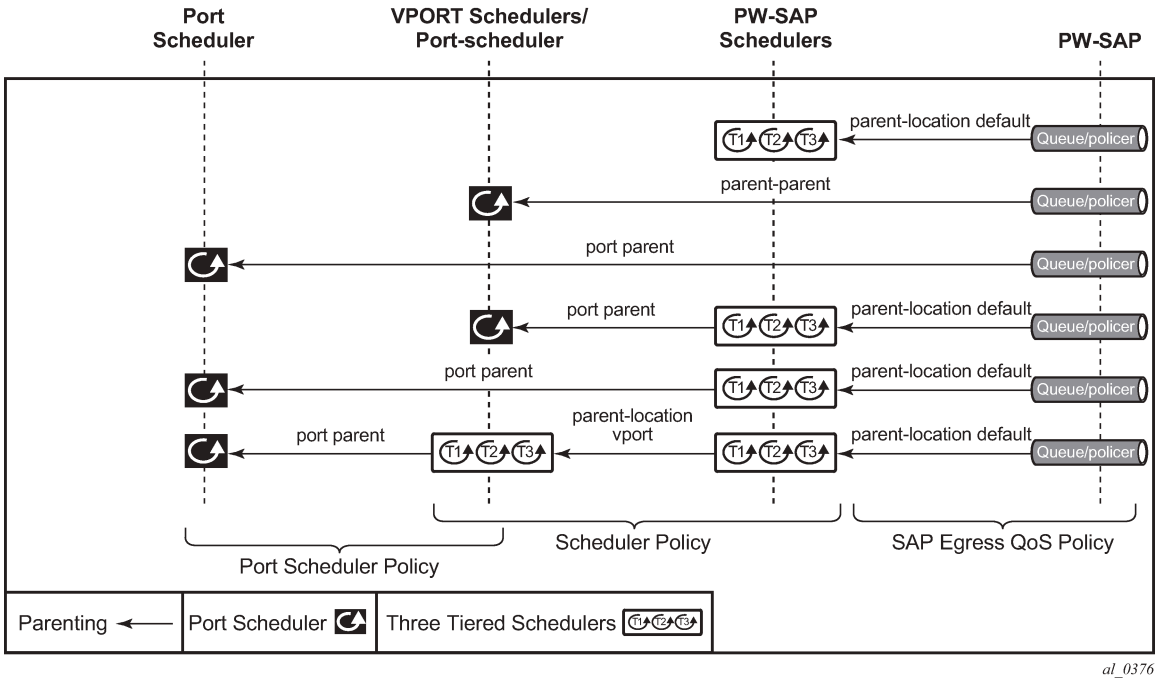
1. Configure a Vport per AN under the port (or LAG) to which the SDP corresponding to the pseudowire SAP is bound. The Vport would be configured with **aggregate rate-limit**. (**config>port>ethernet>access>egress>vport vport-name create**).
2. Explicitly assign (by static configuration) a pseudowire port to a Vport. For limiting the total traffic to an AN, all pseudowire ports for an AN-port would refer to the same Vport.

For bandwidth control per pseudowire, the following configuration steps are used:

1. Create multiple Vports under the port to which SDP is bound. Each Vport can be configured with **agg-rate rate**, a scheduler or port-scheduler.
2. Assign each pseudowire to an AN to a unique Vport shaper (regular IOM/MDA).

To make use of the **agg-rate rate** or **port-scheduler** under a Vport, PW SAP queues and schedulers must be configured with the **port-parent** command. To make use of a scheduler under a Vport, PW SAP schedulers must be configured with a **parent** command and the **parent-location vport** under the tier 1 of the scheduler policy. The egress hierarchical parenting relationship options are shown in Figure 4: PW SAP egress scheduling hierarchy options. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR Quality of Service Guide Quality of Service guide* for more information.

Figure 4: PW SAP egress scheduling hierarchy options



2.2.3.7 Lag considerations

PW ports may be bound to Vport schedulers bound to a LAG. However, if the LAG is configured in distributed mode, then bandwidth is shared according to the active LAG members across a single IOM. If the LAG spans multiple IOMs, then it effectively operates in link mode across the IOMs. That is, the full LAG bandwidth is allocated to the LAG members on each IOM. Therefore, the use of a Vport on a distributed mode LAG with a port scheduler on the port or Vport and PW SAPs is explicitly not supported and is not a recommended configuration. It is recommended that port-fair mode is used instead.

2.2.3.8 Last mile packet size adjustment

In the application where pseudowire SAPs are used to apply access QoS for services aggregated from an Ethernet access network, MPLS labels may not be present on the last-mile and link from an access node. In these cases, policers, queues and H-QoS schedulers should account for packets without MPLS overhead, modeled as "encaps-offset". Vport and port schedulers behave as per the table below. In the

data-path, the actual pseudowire encap overhead (taking into account the MPLS labels) added to the packet is tracked, and may be applied to the scheduler calculations via the configured packet-byte-offset.

The rate limit configured for the pseudowire SAP accounts for subscriber or service frame wire rate: without MPLS overhead and including the last mile overhead (unless a packet-byte-offset is configured).

Table 4: Packet sizes used for pseudowire SAPs summarizes the default packet sizes used at each of the schedulers on the IOM/Ethernet MDA, assuming a 1000byte customer packet.

Table 4: Packet sizes used for pseudowire SAPs

Type	Size
exp-secondary-shaper	20B preamble + 26 MPLS + 1000B pkt
port-scheduler rate	20B preamble + 1000B pkt
regular queue/policer rate	1000B pkt
vport agg-limit-rate	20B preamble + 1000B pkt
vport port-scheduler rate	20B preamble + 1000B pkt
vport scheduler rate	1000B pkt
vport scheduler to port-scheduler rates	20B preamble + 1000B pkt

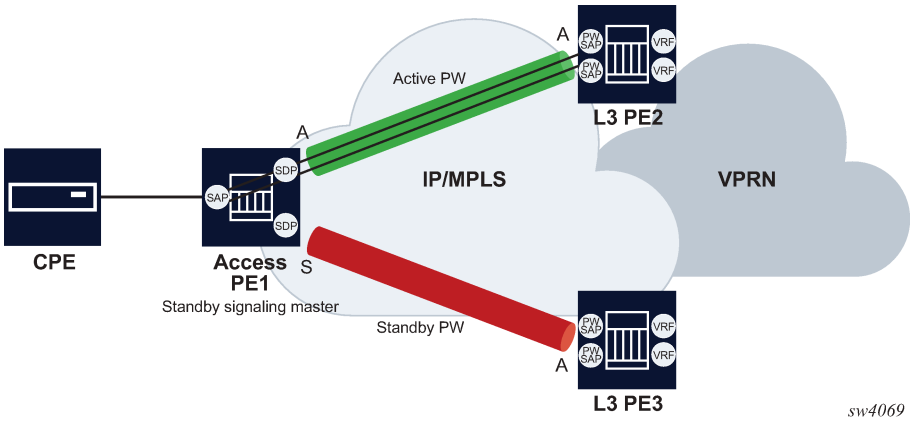
2.2.3.9 Redundancy with pseudowire SAPs

This section describes a mechanism in which one end on a pseudowire (the "master") dictates the active PW selection, which is followed by the other end of the PW (the 'slave'). This mechanism and associated terminology is specified in RFC6870.

Within a chassis, IOM and port based redundancy is based on active/backup LAG. The topology for the base MPLS LSP used by the SDP could be constrained such that it could get re-routed in the aggregation network, but would always appear on the LAG ports on the Layer 3 PE. In the case that the tunnel is re-routed to a different port, the MPLS pseudowire SAPs would be brought down.

To provide Layer 3 PE redundancy, dual homing of the access PE into separate Layer 3 PEs using active/standby pseudowire status is supported. This is shown in Figure 5: Dual homing into multiple Layer 3 PEs.

Figure 5: Dual homing into multiple Layer 3 PEs



Dual homing operates in a similar manner to Spoke-SDP termination on IES/VPRN. Figure 5: Dual homing into multiple Layer 3 PEs displays the access PE is dual-homed to the Layer 3 PEs using two spoke-SDPs. The endpoint in the access PE is configured to be the master from a pseudowire redundancy perspective using the **standby-signaling-master** command. The access PE picks one of the Spoke-SDPs to make active, and one to make standby, based on the local configuration of primary or spoke SDP precedence.

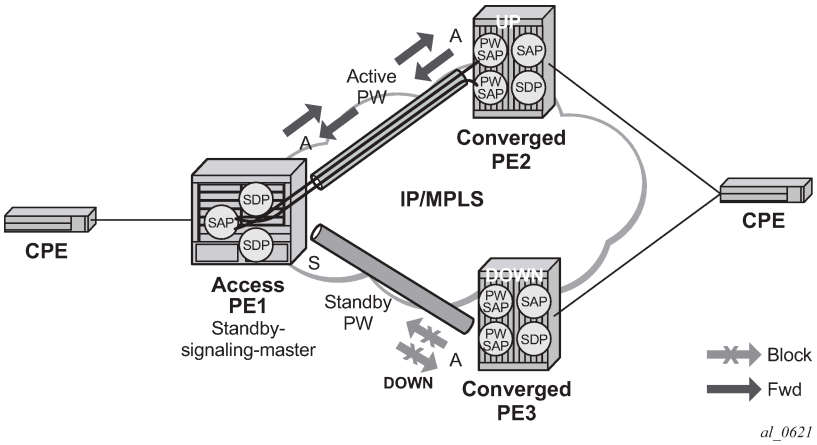
The pseudowire port at the Layer 3 PE behaves as a slave from the perspective of pseudowire status signaling. That is, if its peer signals "PW FWD standby (0x20)" status bit for the specific Spoke-SDP and the local configuration does not allow this bit to be ignored, the PE takes the pseudowire port to a local operationally down state. This is consistent with the Spoke-SDP behavior for the case of Spoke-SDP termination on IES/VPRN.

As a consequence, all of the pseudowire SAPs bound to the pseudowire port are taken down, which causes the corresponding IES or VPRN interface to go to a local operationally down state and therefore stops forwarding packets toward this pseudowire port.

Conversely, the formerly standby pseudowire is made active and then the corresponding pseudowire port on the second Layer 3 PE is taken locally operationally up. Therefore, all of the pseudowire SAPs bound to the pseudowire port are brought up, which causes the corresponding IES or VPRN interface to go to a local operationally up state allowing forwarding of packets toward this pseudowire port.

For VLLs, a PW port always behaves as a slave from the perspective of PW redundancy. This is because the PW port is taken locally operationally down if any non-zero PW status (including a PW Preferential Forwarding status of **standby**) is received. Master-slave PW redundancy mechanisms for dual homing of the access PE into separate converged PEs using active or standby PW status are supported as shown in Figure 6: Master-slave PW redundancy . However, this is only applicable to VLL services consisting of only SAPs, PW-SAPs, or spoke-SDPs. Dual-homing redundancy, taking into account the status of the PW SAP, is not supported where a PBB tunnel between the two converged PEs exists in the Epipe VLL service.

Figure 6: Master-slave PW redundancy



As in the existing implementation, **standby-signaling-master** is configured on the Spoke-SDP at the access PE. However explicit configuration of **standby-signaling-slave** on the PW port is not required, as this is the default behavior.

The forwarding behavior is the same as when **standby-signaling-slave** is configured for Epipe Spoke-SDPs. That is, when enabled, if a PW Forwarding Standby (0x20) LDP status message is received for the P11111111W, then the transmit direction is blocked for the PW port. All PW SAPs bound to the corresponding PW port are treated from a SAP OAM perspective in the same manner as a fault on the service, such as an SDP-binding down or remote SAP down.

PW redundancy with multiple active/standby PW ports or PW SAPs bound to the same Ethernet SAP in the converged PE is not supported. The independent mode of operation for PW redundancy is not also supported for a PW port.

2.2.3.10 Operational group support for PW ports

A PW port state may be linked to the state of an oper-group, such that if the oper-group goes down, the SDP binding for the PW port also goes operationally down, and therefore the corresponding PW status bit signaled (0x00000001 - Pseudowire Not Forwarding). If a status of 0x00000001 is signaled for a currently active PW, and active/standby dual homing is in use then the access PE fails over to the standby PW to the standby converged PE.

This is achieved by linking an SDP binding to an operational group for PW SAPs belonging to any supported service types (including those with group interfaces) bound to that PW port, such as IES, VPRN, or Epipe VLL. The association to an operational group is configured under the PW port config at the SDP binding level, as follows:

```
config
   service
      sdp
         binding
            [no] pw-port <pw-port-id> [vc-id <vc-id>] [create]
               monitor-oper-group <group-name>
```

The **monitor-oper-group** command specifies the operational group to be monitored by the PW-Port under which it is configured. The oper-group name must be already configured under the **config>service** context before its name is referenced in this command.

The following illustrates how a PW port can track the status of VPRN uplinks using monitor-oper-group.

Uplinks in a VPRN may be monitored using a BFD session on the network facing IP interfaces in a VPRN or on the network IP interfaces supporting the uplinks.

Oper-groups monitor the state of these BFD sessions inside the VPRN as follows:

```
config>service>
    oper-group "test-oper-grp" create
       bfd-enable interface "vprn-if" dest-ip 10.0.0.20 service 105
```

Alternatively, the state of network interfaces can be monitored as follows:

```
config>service>
   oper-group "test-oper-grp" create
      bfd-enable interface "network-if" dest-ip 10.0.1.20
```

The PW port is then configured with monitor-oper-group as follows:

```
config>service>sdp>binding
            pw-port 100 vc-id 25
            monitor-oper-group "test-oper-group"
```

2.2.4 Routing protocols

The IES IP interfaces are restricted as to the routing protocols that can be defined on the interface based on the fact that the customer has a different routing domain for this service. The IES IP interfaces support the following routing protocols:

- RIP
- OSPF
- IS-IS
- BGP
- IGMP
- PIM

Note: The SAP for the IES IP interface is created at the IES service level, but the routing protocols for the IES IP interface are configured at the routing protocol level for the main router instance.

2.2.4.1 CPE connectivity check

Static routes are used within many IES services. Unlike dynamic routing protocols, there is no way to change the state of routes based on availability information for the associated CPE. CPE connectivity check adds flexibility so that unavailable destinations are removed from the service provider’s routing tables dynamically and minimize wasted bandwidth.

The availability of the far-end static route is monitored through periodic polling. The polling period is configured. If the poll fails a specified number of sequential polls, the static route is marked as inactive.

An ICMP ping mechanism is used to test the connectivity.

If the connectivity check fails and the static route is deactivated, the router continues to send polls and re-activate any routes that are restored.

2.2.5 QoS policies

When applied to IES services, service ingress QoS policies only create the unicast queues defined in the policy. The multipoint queues are not created on the service. With IES services, service egress QoS policies function as with other services where the class-based queues are created as defined in the policy. Both Layer 2 or Layer 3 criteria can be used in the QoS policies for traffic classification in an IES.

2.2.6 Filter policies

Only IP filter policies can be applied to IES services.

2.2.7 MPLS entropy label and hash label

The router supports both the MPLS entropy label (RFC 6790) and the Flow Aware Transport label, known as the hash label (RFC 6391). LSR nodes in a network to load-balance labeled packets in a more granular way than by hashing on the standard label stack. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide* for more information.

The entropy label is supported for Epipe and Ipipe spoke-SDP termination on IES interfaces. To configure insertion of the entropy label on a spoke-SDP that terminates on the service, use the **entropy-label** command in the **spoke-sdp** context of a specific interface.

The hash label is also supported for Epipe and Ipipe spoke-SDP termination on IES services. Configure it using the **hash-label** command in the **spoke-sdp** context for an IES interface context.

Either the hash label or the entropy label can be configured on one object, but not both.

2.2.8 Spoke SDPs

Distributed services use service distribution points (SDPs) to direct traffic to another router through service tunnels. SDPs are created on each participating router and then bound to a specific service. SDP can be created as either GRE or MPLS. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR Services Overview Guide* for information about configuring SDPs.

This feature provides the ability to cross-connect traffic entering on a spoke SDP, used for Layer 2 services (VLLs or VPLS), on to an IES or VPRN service. From a logical point of view, the spoke SDP entering on a network port is cross-connected to the Layer 3 service as if it entered by a service SAP. The main exception to this is traffic entering the Layer 3 service by a spoke SDP is handled with network QoS policies and not access QoS policies.

This feature applies to the 7450 ESS and 7750 SR only.

Figure 7: SDP-ID and VC label service identifiers depicts traffic terminating on a specific IES or VPRN service that is identified by the SDP-ID and VC label present in the service packet.

Figure 7: SDP-ID and VC label service identifiers

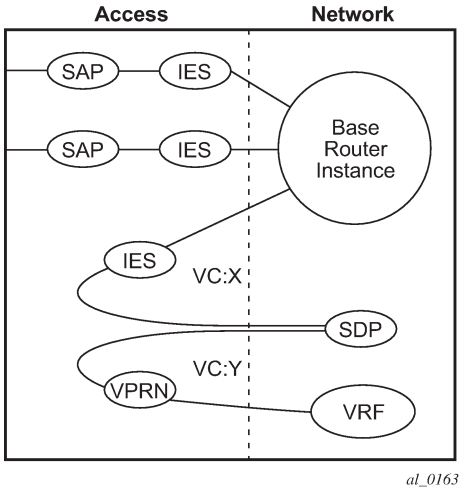


Figure 8: IES spoke-SDP termination

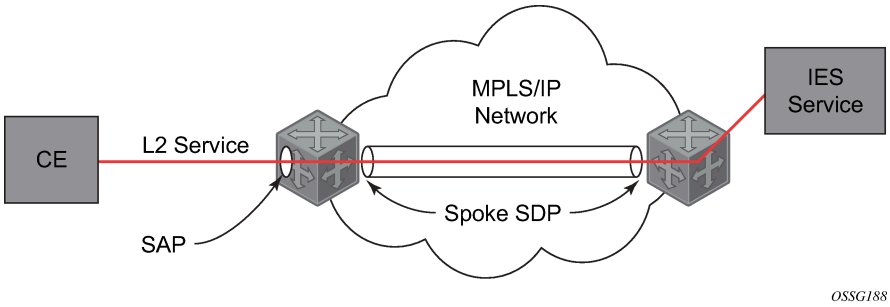


Figure 8: IES spoke-SDP termination depicts a spoke-SDP terminating directly into a Layer 3 service interface (IES or VPRN) at one end, and a Layer 2 service (Epipe, Ipipe, or VPLS) at the other. There is no special configuration required on the Layer 2 service.

If the terminating Layer 2 service is an Ipipe, then on the IES/VPRN interface end, the spoke-SDP must be created with the vc-type Ipipe option. Spoke-SDPs created with vc-type ether (the default) are compatible with Epipe and VPLS services, as well as with other IES/VPRN interfaces.

If the MPLS network uses LDP signaling, then in order for a spoke-SDP to function, the LDP binding MTUs at each end must match. For a Layer 2 service, the MTU of the local binding is 14 octets less than the configured service-mtu (such as, binding MTU = service-mtu - 14). For an IES or VPRN interface, the binding MTU is equal to either the configured ip-mtu of the interface, or the SDP's path-mtu minus 14, whichever is lower. The local and remote MTUs of all bindings can be found using the CLI command **show router ldp bindings**.

All routing protocols that are supported by IES/VPRN are supported for spoke-SDP termination.

See VCCV BFD support for VLL, Spoke SDP Termination on IES and VPRN, and VPLS Services in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Layer 2 Services and EVPN Guide* for information about using VCCV BFD in spoke-SDP termination.

Note: Spoke-SDP termination of Ipipe VLLs on IES is not supported in System Profile B. To determine if Ipipes are currently bound to an IES interface, use the **show router ldp bindings services** command before configuring profile B.

2.2.9 SRRP

Subscriber Router Redundancy Protocol (SRRP) is used on the 7750 SR and 7450 ESS and is closely tied to the multi-chassis synchronization (MCS) protocol used to synchronize information between redundant nodes. An MCS peer must be configured and operational when subscriber hosts have a redundant connection to two nodes. Subscriber hosts are identified by the ingress SAP, the host's IP and MAC addresses. After a host is identified on one node, the MCS peering is used to inform the other node that the host exists and conveys the dynamic DHCP lease state information of the host. MCS creates a common association between the virtual ports (SAPs) shared by a subscriber. This association is configured at the MCS peering level by defining a tag for a port and range of SAPs. The same tag is defined on the other nodes peering context for another port (does not need to be the same port-ID) with the same SAP range. In this manner, a subscriber host and Dot1Q tag sent across the peering with the appropriate tag is mapped to the redundant SAP on the other node.

SRRP can only be configured on group interfaces. When SRRP is active on a group IP interface, the SRRP instance attempts to communicate through in-band (over the group IP interfaces SAPs) and out-of-band (over the group IP interfaces redundant IP interface) messages to a remote router. If the remote router is also running SRRP with the same SRRP instance ID, one router enters a master state while the other router enters a backup state. Because both routers are sharing a common SRRP gateway MAC address that is used for the SRRP gateway IP addresses and for proxy ARP functions, either node may act as the default gateway for the attached subscriber hosts.

For correct operation, each subscriber subnet associated with the SRRP instance must have a gw-address defined. The SRRP instance cannot be activated (no shutdown) unless each subscriber subnet associated with the group IP interface has an SRRP gateway IP address. Once the SRRP instance is activated, new subscriber subnets cannot be added without a corresponding SRRP gateway IP address. Table 5: SRRP state effect on subscriber hosts associated with group IP interface describes how the SRRP instance state is used to manage access to subscriber hosts associated with the group IP interface.

SRRP instances are created in the disabled state (shutdown). To activate SRRP the **no shutdown** command in the SRRP context must be executed.

Before activating an SRRP instance on a group IP interface, the following actions are required:

1. Add a SRRP gateway IP addresses to all subscriber subnets associated with the group IP interface, including subnets on subscriber IP interfaces associated as retail routing contexts (at least one subnet must be on the subscriber IP interface containing the group IP interface and its SRRP instance).
2. Create a redundant IP interface and associate it with the SRRP instances group IP interface for shunting traffic to the remote router when master.
3. Specify the group IP interface SAP used for SRRP advertisement and Information messaging.

Before activating an SRRP instance on a group IP interface, the following actions should be considered:

1. Associate the SRRP instance to a Multi-Chassis Synchronization (MCS) peering terminating on the neighboring router (the MCS peering should exist as the peering is required for redundant subscriber host management).
2. Define a description string for the SRRP instance.
3. Specify the SRRP gateway MAC address used by the SRRP instance (must be the same on both the local and remote SRRP instance participating in the same SRRP context).

4. Change the base priority for the SRRP instance.
5. Specify one or more VRRP policies to dynamically manage the SRRP instance base priority.
6. Specify a new keep alive interval for the SRRP instance.

Table 5: SRRP state effect on subscriber hosts associated with group IP interface lists the SRRP's state effect on subscriber hosts associated with group IP interfaces.

Table 5: SRRP state effect on subscriber hosts associated with group IP interface

SRRP state	ARP	Local proxy ARP enabled	Remote proxy ARP enabled	Subscriber host routing
Disabled	- Responds to ARP for all owned subscriber subnet IP addresses. - Will not respond to ARP for subscriber subnet SRRP gateway IP addresses. - All ARP responses contain the native MAC of the group IP interface (not the SRRP gateway MAC).	- Responds to ARP for all subscriber hosts on the subscriber subnet.	- Responds to ARP for all reachable remote IP hosts.	- All routing out the group IP interface uses the native group IP interface MAC address. - The group IP interface redundant IP interface is not used. - Does not accept packets destined for the SRRP gateway MAC received on the group IP interface.
Becoming Master (To enter becoming master state, a master must currently exist)	- Responds to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC). - Responds to ARP for subscriber subnet SRRP gateway IP addresses. (hardware address = SRRP gateway IP address, source MAC = group IP interface native MAC).	- Responds to ARP for all subscriber hosts on the subscriber subnet (hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC).	- Responds to ARP for all reachable remote IP hosts (hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC).	- All routing out the group IP interface uses the native group IP interface MAC address. - Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface. - Does not accept packets destined for the SRRP gateway MAC received on the group IP interface.
Master	- Responds to ARP for all owned subscriber subnet	- Responds to ARP for all subscriber hosts on the	- Responds to ARP for all reachable remote IP hosts	- All routing out the group IP interface uses the SRRP

SRRP state	ARP	Local proxy ARP enabled	Remote proxy ARP enabled	Subscriber host routing
	IP addresses (hardware address and source MAC = group IP interface native MAC). - Responds to ARP for subscriber subnet SRRP gateway IP addresses (hardware address = SRRP gateway IP address, source MAC = group IP interface native MAC).	subscriber subnet (hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC).	(hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC).	gateway MAC address. - Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface. - Accepts packets destined for the SRRP gateway MAC received on the group IP interface.
Becoming Backup (redundant IP interface operational)	- Responds to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC). - Does not respond to ARP for subscriber subnet SRRP gateway IP addresses.	- Does not respond to ARP for any subscriber hosts on the subscriber subnet.	- Does not respond to ARP for any remote IP hosts.	- Does not route out the group IP interface for subscriber hosts associated with the subscriber subnet. - Accepts packets destined for the SRRP gateway MAC received on the group IP interface. - Subscriber hosts mapped to the group IP interface are remapped to the redundant IP interface.
Becoming Backup (redundant IP interface not available)	- Responds to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC). - Does not respond to ARP for subscriber subnet	- Does not respond to ARP for any subscriber hosts on the subscriber subnet.	- Does not respond to ARP for any remote IP hosts.	- Will route out the group IP interface for subscriber hosts associated with the subscriber subnet using the group IP interface native MAC address. - Subscriber hosts mapped to the redundant IP interface are

SRRP state	ARP	Local proxy ARP enabled	Remote proxy ARP enabled	Subscriber host routing
	SRRP gateway IP addresses.			remapped to the group IP interface. - Accepts packets destined for the SRRP gateway MAC received on the group IP interface.
Backup (redundant IP interface operational)	- Responds to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC). - Will not respond to ARP for subscriber subnet SRRP gateway IP addresses.	- Does not respond to ARP for any subscriber hosts on the subscriber subnet.	- Does not respond to ARP for any remote IP hosts.	- Does not route out the group IP interface for subscriber hosts associated with the subscriber subnet. - Subscriber hosts mapped to the group IP interface are remapped to the redundant IP interface. - Does not accept packets destined for the SRRP gateway MAC received on the group IP interface.
Backup (redundant IP interface not available)	- Responds to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC). - Does not respond to ARP for subscriber subnet SRRP gateway IP addresses.	- Does not respond to ARP for any subscriber hosts on the subscriber subnet.	- Does not respond to ARP for any remote IP hosts.	- Routes out the group IP interface for subscriber hosts associated with the subscriber subnet using the group IP interface native MAC address. - Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface. - Does not accept packets destined for the SRRP gateway MAC received

SRRP state	ARP	Local proxy ARP enabled	Remote proxy ARP enabled	Subscriber host routing
				on the group IP interface.

2.2.9.1 SRRP messaging

SRRP uses the same messaging format as VRRP with slight modifications. The source IP address is derived from the system IP address assigned to the local router. The destination IP address and IP protocol are the same as VRRP (224.0.0.18 and 112, respectively).

The message type field is set to 1 (advertisement) and the protocol version is set to 8 to differentiate SRRP message processing from VRRP message processing.

The vr-id field supports an SRRP instance ID of 32 bits.

Because of the large number of subnets backed up by SRRP, only one message every minute carries the gateway IP addresses associated with the SRRP instance. These gateway addresses are stored by the local SRRP instance and are compared with the gateway addresses associated with the local subscriber IP interface.

Unlike VRRP, only two nodes may participate in an SRRP instance due the explicit association between the SRRP instance group IP interface, the associated redundant IP interface and the multi-chassis synchronization (MCS) peering. Because only two nodes are participating, the VRRP skew timer is not used when waiting to enter the master state. Also, SRRP always preempts when the local priority is better than the current master and the backup SRRP instance always inherits the master's advertisement interval from the SRRP advertisement messaging.

SRRP advertisement messages carry a *becoming-master* indicator flag. The *becoming-master* flag is set by a node that is attempting to usurp the master state from an existing SRRP master router. When receiving an SRRP advertisement message with a better priority and with the *becoming-master* flag set, the local master initiates its *becoming-backup* state, stops routing with the SRRP gateway MAC and sends an SRRP advertisement message with a priority set to zero. The new master continues to send SRRP advertisement messages with the *becoming-master* flag set until it either receives a return priority zero SRRP advertisement message from the previous master or its *becoming-master* state timer expires. The new backup node continues to send zero priority SRRP advertisement messages every time it receives an SRRP advertisement message with the *becoming-master* flag set. After the new master either receives the old masters priority zero SRRP advertisement message or the *become-master* state timer expires, it enters the *master* state. The *become-master* state timer is set to 10 seconds upon entering the *become-master* state.

The SRRP advertisement message is always evaluated to see if it has a higher priority than the SRRP advertisement that would be sent by the local node. If the advertised priority is equal to the current local priority, the source IP address of the received SRRP advertisement is used as a tie breaker. The node with the lowest IP address is considered to have the highest priority. SRRP does not preempt when priorities are equal. Preemption occurs only when priorities are specified. The lower IP address is only used as a tie-breaker when there is no master in the network. In other words, when both routers are changing from the "init" state, the lower IP is used to choose the master. If a master already exists, despite having the lower IP address, the system does not preempt the current master.

The SRRP instance maintains the source IP address of the current master. If an advertisement is received with the current masters source IP address and the local priority is higher priority than the masters advertised priority, the local node immediately enters the *becoming-master* state unless the advertised

priority is zero. If the advertised priority is zero, the local node bypasses the *becoming-master* state and immediately enters the *master* state. Priority zero is a special case and is sent when an SRRP instance is relinquishing the master state.

2.2.9.2 SRRP and multi-chassis synchronization

To take full advantage of SRRP resiliency and diagnostic capabilities, the SRRP instance should be tied to a MCS peering that terminates on the redundant node. The SRRP instance is tied to the peering using the **srrp srrp-id** command within the appropriate MCS peering configuration. Once the peering is associated with the SRRP instance, MCS synchronizes the local information about the SRRP instance with the neighbor router. MCS automatically derives the MCS key for the SRRP instance based on the SRRP instance ID. For example, an SRRP instance ID of 1 would appear in the MCS peering database with a MCS-key srrp-0000000001.

The SRRP instance information stored and sent to the neighbor router consists of the following:

- SRRP instance MCS key
- containing service type and ID
- containing subscriber IP interface name
- subscriber subnet information
- containing group IP interface information
- SRRP group IP interface redundant IP interface name, IP address and mask
- SRRP advertisement message SAP
- local system IP address (SRRP advertisement message source IP address)
- group IP interface MAC address
- SRRP gateway MAC address
- SRRP instance administration state (up / down)
- SRRP instance operational state (disabled / becoming-backup / backup / becoming-master / master)
- current SRRP priority
- remote redundant IP interface availability (available / unavailable)
- local receive SRRP advertisement SAP availability (available / unavailable)

2.2.9.3 SRRP instance

The SRRP instance uses the received information to verify provisioning and obtain operational status of the SRRP instance on the neighboring router.

2.2.9.3.1 SRRP instance MCS key

The SRRP instance MCS key ties the received MCS information to the local SRRP instance with the same MCS key. If the received key does not match an existing SRRP instance, the MCS information associated with the key is ignored. After an SRRP instance is created and mapped to an MCS peering, the SRRP instance evaluates received information with the same MCS key to verify it corresponds to the same peering. If the received MCS key is on a different peering than the local MCS key an SRRP peering

mismatch event is generated detailing the SRRP instance ID, the IP address of the peering the MCS key is received on and the IP address to which the local MCS key is mapped. If the peering association mismatch is corrected, an SRRP peering mismatch clear event is generated.

2.2.9.3.2 Containing service type and ID

The Containing Service Type is the service type (IES or VPRN) that contains the local SRRP instance. The Containing Service ID is the service ID of that service. This information is supplied for troubleshooting purposes only and is not required to be the same on both nodes.

2.2.9.3.3 Containing subscriber IP interface name

The containing subscriber IP interface name is the subscriber IP interface name that contains the SRRP instance and its group IP interface. This information is supplied for troubleshooting purposes only and is not required to be the same on both nodes.

2.2.9.3.4 Subscriber subnet information

The subscriber subnet information includes all subscriber subnets backed up by the SRRP instance. The information for each subnet includes the Owned IP address, the mask and the gateway IP address. If the received subscriber subnet information does not match the local subscriber subnet information, an SRRP Subscriber Subnet Mismatch event is generated describing the SRRP instance ID and the local and remote node IP addresses. When the subscriber subnet information matches, an SRRP Subscriber Subnet Mismatch Clear event is generated.

2.2.9.3.5 Containing group IP interface information

The containing group IP interface information is the information about the group IP interface that contains the SRRP instance. The information includes the name of the group IP interface, the list of all SAPs created on the group IP interface, the administrative and operational state of each SAP and the MCS key and the peering destination IP address associated with each SAP. To obtain the MCS information, the SRRP instance queries MCS to determine the peering association of the SRRP instance and then queries MCS for each SAP on the group IP interface. If the local SRRP instance is associated with a different MCS peering than any of the SAPs or if one or more SAPs are not tied to an MCS peering, an SRRP group interface SAP peering mismatch event is generated detailing the SRRP instance ID, and the group IP interface name.

When receiving the remote containing group IP interface information, the local node compares the received SAP information with the local group IP interface SAP information. If a local SAP is not included in the SAP information or a remote SAP is not included in the local group IP interface, an SRRP Remote SAP mismatch event is generated detailing the SRRP instance ID and the local and remote group IP interface names. If a received SAP's MCS key does not match a local SAP's MCS Key, an SRRP SAP MCS key mismatch event is generated detailing the SRRP instance ID, the local and remote group IP interface names, the SAP-ID and the local and remote MCS keys.

2.2.9.3.6 Remote redundant IP interface mismatch

If the group IP remote redundant IP interface address space does not exist, is not within the local routing context for the SRRP instances group IP interface or is not on a redundant IP interface, the local node sends redundant IP interface unavailable to prevent the remote neighbor from using its redundant IP interface. An SRRP redundant IP interface mismatch event is generated for the SRRP instance detailing the SRRP instance, the local and remote system IP addresses, the local and remote group IP interface names and the local and remote redundant IP interface names and IP addresses and masks. The local redundant IP interface may still be used if the remote node is not sending redundant IP interface unavailable.

2.2.9.3.7 Remote sending redundant IP interface unavailable

If the remote node is sending redundant IP interface unavailable, the local node treats the local redundant IP interface associated with the SRRP instances group IP interface as down. A Local Redundant IP Interface Unavailable event is generated detailing the SRRP instance ID, the local and remote system IP addresses, the local group IP interface name, the local redundant IP interface name and the redundant IP interface IP address and mask.

2.2.9.3.8 Remote SRRP advertisement SAP non-existent

If the remote node's SRRP advertisement SAP does not exist on the local SRRP instances group IP interface, the local node sends local receive SRRP advertisement SAP unavailable to the remote node. An SRRP receive advertisement SAP non-existent event is generated detailing the SRRP instance ID, the local and remote system IP addresses, the local group IP interface name and the received remote SRRP advertisement SAP. Because SRRP advertisement messages cannot be received, the local node immediately becomes master if it has the lower system IP address.

2.2.9.3.9 Remote sending local receive SRRP advertisement SAP unavailable

If the local node is receiving local receive SRRP advertisement SAP unavailable from the remote node, an SRRP Remote Receive advertisement SAP Unavailable event is generated detailing the SRRP instance ID, the local and remote system IP addresses, the remote group IP interface name and the local SRRP advertisement SAP. Because the remote node cannot receive SRRP advertisement messages, the local node immediately becomes master if it has the lower system IP address.

2.2.9.3.10 Local and remote dual master detected

If the local SRRP state is master and the remote SRRP state is master, an SRRP dual master event is generated detailing the SRRP instance ID and the local, remote system IP addresses and the local and remote group IP interface names and port numbers.

2.2.9.4 Subscriber subnet owned IP address connectivity

In order for the network to reliably reach the owned IP addresses on a subscriber subnet, it is not necessary for the owning node to advertise the IP addresses as /32 host routes into the core. Network reachability to the subscriber subnet is advertised into the IGP core by both of the dual homing nodes.

The shortest path to the subscriber may not always traverse the active path for a subscriber. In this case, the path traverses the non-active/primary node for the subscriber and the traffic is redirected through the redundant interface to the other node through the redundant interface to the active path. This ensures that all downstream traffic to a subscriber always flows through one node.

2.2.9.5 Subscriber subnet SRRP gateway IP address connectivity

The SRRP gateway IP addresses on the subscriber subnets cannot be advertised as /32 host routes because they may be active (master) on multiple group IP interfaces on multiple SRRP routers. Without a /32 host route path, the network forwards any packet destined for an SRRP gateway IP address to the closest router advertising the subscriber subnet. While a case may be made that only a node that is currently forwarding for the gateway IP address in a master state should respond to ping or other diagnostic messages, the distribution of the subnet and the case of multiple masters make any resulting response or non-response inconclusive at best. To provide some ability to ping the SRRP gateway address from the network side reliably, any node receiving the ICMP ping request responds if the gateway IP address is defined on its subscriber subnet.

2.2.9.6 Receive SRRP advertisement SAP and anti-spoof

The group IP interface SAPs are designed to support subscriber hosts and perform an ingress anti-spoof function that ensures that any IP packet received on the group IP interface is coming in the correct SAP with the correct MAC address. If the IP and MAC are not registered as valid subscriber hosts on the SAP, the packet is silently discarded. Because the SRRP advertisement source IP addresses are not subscriber hosts, an anti-spoof entry does not exist and SRRP advertisement messages would normally be silently discarded. To avoid this issue, when a group IP interface SAP is configured to send and receive SRRP advertisement messages, anti-spoof processing on the SAP is disabled. This precludes subscriber host management on the SRRP messaging SAP.

2.2.9.7 BFD with SRRP/VRRP

BFD with SRRP is supported. This allows the use of longer timers inside SRRP resulting in more SRRP instances while still retaining fast failure detection with BFD.

2.3 Configuring an IES service with CLI

This section provides information to configure IES services using the command line interface.

2.3.1 Basic configuration

The most basic IES service configuration has the following entities:

- customer ID (see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Services Overview Guide* for more information)
- an interface to create and maintain IP routing interfaces within IES service ID
- a SAP on the interface specifying the access port and encapsulation values

The following is an example configuration of an IES service on ALA-48:

```
*A:ALA-48>config>service# info
----------------------------------------------
ies 1000 customer 50 vpn 1000 create
            description "to internet"
            interface "to-web" create
                address 10.1.1.1/24
                sap 1/1/5:0.* create
                exit
            exit
            no shutdown
----------------------------------------------
*A:ALA-48>config>service#
```

2.3.2 Common configuration tasks

About this task

This section provides a brief overview of the tasks that must be performed to configure IES services and provides the CLI commands.

Procedure

Step 1. Associate an IES service with a customer ID.

Step 2. Associate customer ID with the service.

Step 3. Assign an IP address.

Step 4. Create a subscriber interface (for 7450 ESS and 7750 SR only and optional).

Step 5. Create an interface.

Step 6. Define SAP parameters on the interface

- Select nodes and ports.
- Optionally, select the following policies:
  - QoS policies other than the default (configured in the config>qos context); for 7450 ESS and 7750 SR only
  - filter policies (configured in the config>filter context)
  - accounting policy (configured in the config>log context); for 7450 ESS and 7750 SR only

Step 7. Enable service.

2.3.3 Configuring IES components

Use the CLI syntax to configure IES components.

2.3.3.1 Configuring an IES service

Use the following CLI syntax to create an IES service:

The following example displays a basic IES service configuration for the 7450 ESS and 7750 SR:

```
A:ALA-48>config>service#
------------------------------------------
...
ies 1001 customer 1730 vpn 1001 create
            description "to-internet"
            no shutdown
exit
------------------------------------------
A:ALA-48>config>service#
```

The following example displays a basic IES service configuration for the 7950 XRS:

```
A:ALA-48>config>service#
------------------------------------------
...
ies 1001 customer 1730 create
            description "to-internet"
            no shutdown
exit
------------------------------------------
A:ALA-48>config>service#
```

2.3.3.2 Configuring IES subscriber interface parameters

Note: This section applies only to the 7750 SR only.

Subscriber interfaces operate only with basic (or enhanced) subscriber management. At the very least, a host, either statically configured or dynamically learned by DHCP must be present in order for the interface to be useful.

The following example displays a subscriber interface configuration for the 7750 SR:

```
A:ALA-48>config>service>ies>sub-if# info
----------------------------------------------
                address 192.168.140.1/24
                group-interface "abc-if" create
                    sap 1/1/19:0 create
                        ingress
                            qos 2
                            filter ip 10
                        exit
                        static-host ip 192.168.145.100 mac 00:01:00:00:00:01 create
                        exit
                    exit
                exit
----------------------------------------------
A:ALA-48>config>service>ies>sub-if#
```

2.3.3.3 Configuring IES interface parameters

The following example displays an IES configuration with interface parameters for the 7450 ESS and 7750 SR:

```
A:ALA-48>config>service>ies>if# info
----------------------------------------------
                address 10.1.1.1/24
                sap 1/1/10:0.* create
                    ingress
                        qos 100
                    exit
                    egress
                        scheduler-policy "SLA1"
                    exit
                exit
                vrrp 1 owner
 authentication-key "3WErEDozxyQ" hash
                exit
----------------------------------------------
A:ALA-48>config>service>ies>if#
```

The following example displays an IES configuration with interface parameters for the 7950 XRS:

```
A:ALA-48>config>service>ies>if# info
----------------------------------------------
                address 10.1.1.1/24
                sap 1/1/10:0.* create
 exit
----------------------------------------------
A:ALA-48>config>service>ies>if#
```

2.3.3.4 Configuring spoke-SDP parameters

The following example displays a spoke SDP configuration for the 7450 ESS and 7750 SR:

```
A:ALA-48>config>service>ies# info
----------------------------------------------
            description "to internet"
            interface "spokeSDP-test" create
                spoke-sdp 2:100 create
                    egress
                        filter ip 10
                    exit
                exit
            exit
            no shutdown
----------------------------------------------
A:ALA-48>config>service>ies#
```

2.3.3.5 Configuring SAP parameters

A SAP is a combination of a port and encapsulation parameters which identifies the service access point on the interface and within the router. Each SAP must be unique within a router.

When configuring IES SAP parameters on a 7450 ESS or 7750 SR, a default QoS policy is applied to each ingress and egress SAP. Additional QoS policies and scheduler policies must be configured in the **config>qos** context. Filter policies are configured in the **config>filter** context and must be explicitly applied to a SAP. There are no default filter policies.

This example displays an IES SAP configuration for the 7450 ESS and 7750 SR:

```
*A:ALA-A>config>service>ies>if# info
----------------------------------------------
                address 10.10.36.2/24
                sap 5/1/3.1:0 create
                    ingress
                        qos 101
                    exit
                    egress
                        scheduler-policy "alpha"
                        qos 1010
                    exit
                exit
----------------------------------------------
*A:ALA-A>config>service>ies>if#
```

This example displays an IES SAP configuration for the 7950 XRS:

```
*A:ALA-A>config>service>ies>if# info
----------------------------------------------
                address 10.10.36.2/24
 exit
----------------------------------------------
*A:ALA-A>config>service>ies>if#
```

2.3.3.6 Configuring VRRP

Configuring VRRP parameters on an IES interface is optional and applies only to the 7450 ESS and 7750 SR. VRRP can be configured in either an owner or non-owner mode. The owner is the VRRP router whose virtual router IP address is the same as the real interface IP address. This is the router that responds to packets addressed to one of the IP addresses for ICMP pings, TCP connections, and so on. All other virtual router instances participating in this message domain must have the same VRID configured and cannot be configured as owner.

For further information about VRRP CLI syntax and command descriptions, see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

The following example displays the IES configuration:

```
*A:ALA-A>config>service>ies>if# info
----------------------------------------------
                address 10.10.36.2/24
                vrrp 2 owner
                    backup 10.10.36.2
 authentication-key "3WErEDozxyQ" hash
                exit
----------------------------------------------
*A:ALA-A>config>service#
```

2.3.3.7 Configuring IPsec parameters

The following output displays an IES service with IPsec parameters configured for the 7750 SR:

```
*A:ALA-49>config# info
----------------------------------------------
...
    service
        ies 100 customer 1 create
            interface "ipsec-public" create
                address 10.10.10.1/24
                sap ipsec-1.public:1 create
                exit
            exit
            no shutdown
        exit
    exit
...
----------------------------------------------
*A:ALA-49>config#
```

2.3.3.8 IGMP host tracking

The following output displays an IES service with IGMP host tracking parameters configured.

```
*A:ALA-49>config>service# info
----------------------------------------------
...
         ies 25 customer 1 create
             interface "ip_if_4" create
                 loopback
                 hold-time down ip 1200
                 address 10.64.64.64/24
                 sap lag-64:64 create
                     no shutdown
                 exit
                 allow-directed-broadcasts
                 host-connectivity-verify
                 ip-mtu 9000
                 local-dhcp-server "server 1"
                 local-proxy-arp
                 proxy-arp-policy treetrace-1
                 remote-proxy-arp
                 secondary 10.3.4.5 255.255.255.0
                 secondary 10.3.4.5/24
                 tos-marking-state trusted
                 tos-marking-state untrusted
                 urpf-check
                 exit
             exit
             igmp-host-tracking
                 expiry-time 65535
                 no shutdown
             exit
...
----------------------------------------------
*A:ALA-49>config>service#
```

2.4 Service management tasks

This section describes the IES service management tasks.

2.4.1 Modifying IES service parameters

Existing IES service parameters in the CLI or NMS can be modified, added, removed, enabled or disabled. The changes are applied immediately to all services when the charges are applied.

To display a list of customer IDs, use the **show service customer** command. Enter the parameters (such as description, SAP information and SDP information) and then enter the new information.

The following example displays the modified service for 7450 ESS and 7750 SR:

```
*A:ALA-A>config>service>ies# info
----------------------------------------------
ies 1000 customer 50 vpn 1000 create
           description "This is a new description"
           interface "to-web" create
               address 10.1.1.1/24
               mac 00:dc:98:1d:00:00
               allow-directed-broadcast
               sap 22/1/50:0 create
               exit
           exit
           no shutdown
exit
----------------------------------------------
*A:ALA-A>config>service#
```

The following example displays the modified service for the 7950 XRS:

```
*A:ALA-A>config>service>ies# info
----------------------------------------------
ies 1000 customer 50 vpn 1000 create
           description "This is a new description"
           interface "to-web" create
               address 10.1.1.1/24
               mac 00:dc:98:1d:00:00
               allow-directed-broadcast
 exit
           exit
           no shutdown
exit
----------------------------------------------
*A:ALA-A>config>service#
```

2.4.2 Deleting a spoke-SDP

To delete the spoke SDP from the service interface must be shut down. This cleans up the associated VC labels.

Use the following CLI syntax to delete a spoke SDP from an interface for the 7450 ESS and 7750 SR:

```
config>service# ies service-id [customer customer-id] [vpn vpn-id]
    – interface ip-int-name
```

```
        – [no] spoke-sdp sdp-id:vc-id
            – shutdown
```

The following example displays the spoke SDP configuration for the 7450 ESS and 7750 SR:

```
A:ALA-48>config>service>ies# info
----------------------------------------------
            description "to internet"
            interface "spokeSDP-test" create
            exit
            no shutdown
----------------------------------------------
A:ALA-48>config>service>ies#
```

2.4.3 Deleting an IES service

An IES service cannot be deleted until SAPs and interfaces are shut down and deleted and the service is shutdown on the service level.

Use the following CLI syntax to delete an IES service:

```
config>service#
    – [no] ies service-id
    – shutdown
        – [no] interface ip-int-name
        – shutdown
            – [no] sap sap-id
            – shutdown
```

2.4.4 Disabling an IES service

An IES service can be shut down without deleting the service parameters.

```
config>service> ies service-id
    – shutdown
```

2.4.5 Re-enabling an IES service

To re-enable an IES service that was shut down.

```
config>service> ies service-id
    – [no] shutdown
```

```
config>service# ies 2000
    – config>service>ies# no shutdown
    – config>service>ies# exit
```

3 Virtual Private Routed Network service

3.1 VPRN service overview

RFC 2547b is an extension to the original RFC 2547, *BGP/MPLS VPNs*, which details a method of distributing routing information using BGP and MPLS forwarding data to provide a Layer 3 Virtual Private Network (VPN) service to end customers.

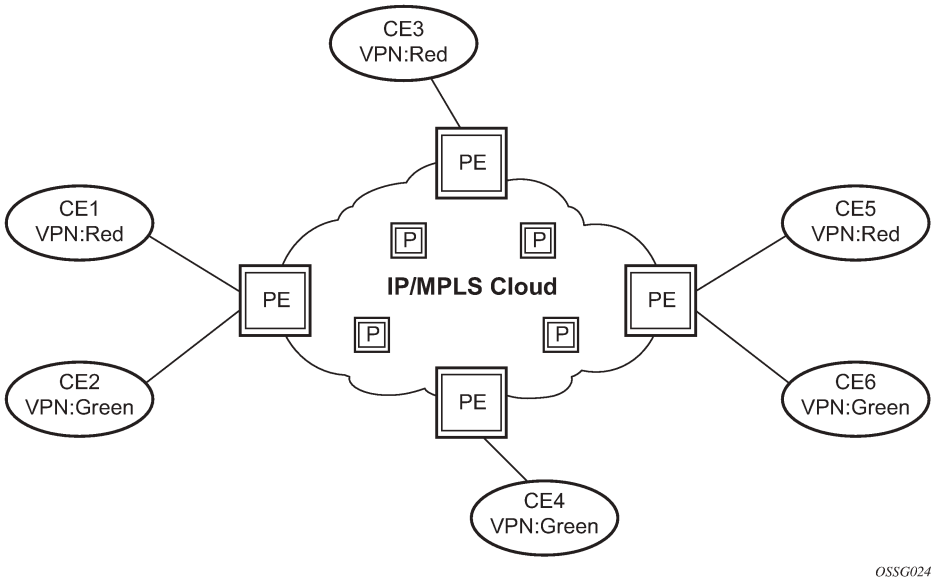
Each Virtual Private Routed Network (VPRN) consists of a set of customer sites connected to one or more PE routers. Each associated PE router maintains a separate IP forwarding table for each VPRN. Additionally, the PE routers exchange the routing information configured or learned from all customer sites via MP-BGP peering. Each route exchanged via the MP-BGP protocol includes a Route Distinguisher (RD), which identifies the VPRN association and handles the possibility of IP address overlap.

The service provider uses BGP to exchange the routes of a particular VPN among the PE routers that are attached to that VPN. This is done in a way which ensures that routes from different VPNs remain distinct and separate, even if two VPNs have an overlapping address space. The PE routers peer with locally connected CE routers and exchange routes with other PE routers to provide end-to-end connectivity between CEs belonging to a specific VPN. Because the CE routers do not peer with each other there is no overlay visible to the CEs.

When BGP distributes a VPN route it also distributes an MPLS label for that route. On an SR series router, the method of allocating a label to a VPN route depends on the VPRN label mode and the configuration of the VRF export policy. SR series routers support three label allocation methods: label per VRF, label per next hop, and label per prefix.

Before a customer data packet travels across the service provider's backbone, it is encapsulated with the MPLS label that corresponds, in the customer's VPN, to the route which best matches the packet's destination address. The MPLS packet is further encapsulated with one or additional MPLS labels or GRE tunnel header so that it gets tunneled across the backbone to the correct PE router. Each route exchanged by the MP-BGP protocol includes a route distinguisher (RD), which identifies the VPRN association. Thus the backbone core routers do not need to know the VPN routes. Figure 9: Virtual Private Routed Network displays a VPRN network diagram example.

Figure 9: Virtual Private Routed Network



3.1.1 Routing prerequisites

RFC 4364 requires the following features:

- multi-protocol extensions to BGP
- extended BGP community support
- BGP capability negotiation

Tunneling protocol options are as follows:

- Label Distribution Protocol (LDP)
- MPLS RSVP-TE tunnels
- Generic Router Encapsulation (GRE) tunnels
- BGP route tunnel (RFC 8277)

3.1.2 Core MP-BGP support

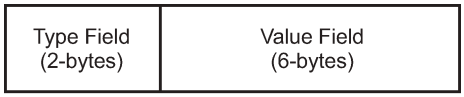
BGP is used with BGP extensions mentioned in Routing prerequisites to distribute VPRN routing information across the service provider's network.

BGP was initially designed to distribute IPv4 routing information. Therefore, multi-protocol extensions and the use of a VPN-IP address were created to extend BGP's ability to carry overlapping routing information. A VPN-IPv4 address is a 12-byte value consisting of the 8-byte route distinguisher (RD) and the 4-byte IPv4 IP address prefix. A VPN-IPv6 address is a 24-byte value consisting of the 8-byte RD and 16-byte IPv6 address prefix. Service providers typically assign one or a small number of RDs per VPN service network-wide.

3.1.3 Route distinguishers

The route distinguisher (RD) is an 8-byte value consisting of two major fields, the **Type** field and **Value** field. The **Type** field determines how the **Value** field should be interpreted. The 7750 SR and 7950 XRS implementation supports the three (3) **Type** values as defined in the standard.

Figure 10: Route distinguisher



L3 guide 10

The three Type values are:

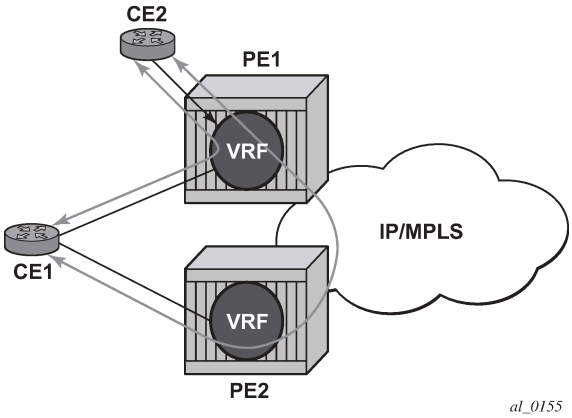
- Type 0: Value Field — Administrator subfield (2 bytes)
  - Assigned number subfield (4 bytes)
  - The administrator field must contain an AS number (using private AS numbers is discouraged). The Assigned field contains a number assigned by the service provider.
- Type 1: Value Field — Administrator subfield (4 bytes)
  - Assigned number subfield (2 bytes)
  - The administrator field must contain an IP address (using private IP address space is discouraged). The Assigned field contains a number assigned by the service provider.
- Type 2: Value Field — Administrator subfield (4 bytes)
  - Assigned number subfield (2 bytes)
  - The administrator field must contain a 4-byte AS number (using private AS numbers is discouraged). The Assigned field contains a number assigned by the service provider.

3.1.3.1 eiBGP load balancing

eiBGP load balancing allows a route to have multiple next hops of different types, using both IPv4 next hops and MPLS LSPs simultaneously.

Figure 11: Basic eiBGP topology displays a basic topology that could use eiBGP load balancing. In this topology CE1 is dual homed and therefore reachable by two separate PE routers. CE 2 (a site in the same VPRN) is also attached to PE1. With eiBGP load balancing, PE1 uses its own local IPv4 nexthop as well as the route advertised by MP-BGP, by PE2.

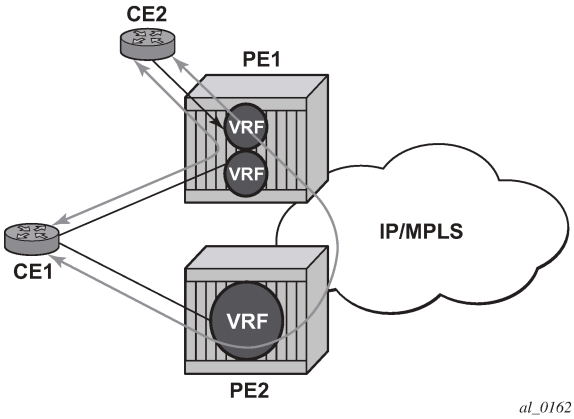
Figure 11: Basic eiBGP topology



Another example displayed in Figure 12: Extranet load balancing shows an extra net VPRN (VRF). The traffic ingressing the PE that should be load balanced is part of a second VPRN and the route over which the load balancing is to occur is part of a separate VPRN instance and are leaked into the second VPRN by route policies.

Here, both routes can have a source protocol of VPN-IPv4 but one still has an IPv4 nexthop and the other can have a VPN-IPv4 nexthop pointing out a network interface. Traffic is still load balanced (if eiBGP is enabled) as if only a single VRF was involved.

Figure 12: Extranet load balancing



Traffic is load balanced across both the IPv4 and VPN-IPv4 next hops. This helps to use all available bandwidth to reach a dual-homed VPRN.

3.1.4 Route reflector

The use of Route Reflectors is supported in the service provider core. Multiple sets of route reflectors can be used for different types of BGP routes, including IPv4 and VPN-IPv4 as well as multicast and IPv6 (multicast and IPv6 apply to the 7750 SR only).

3.1.5 CE to PE route exchange

Routing information between the Customer Edge (CE) and Provider Edge (PE) can be exchanged by the following methods:

- Static Routes
- EBGP
- RIP
- OSPF
- OSPF3

Each protocol provides controls to limit the number of routes learned from each CE router.

3.1.5.1 Route redistribution

Routing information learned from the CE-to-PE routing protocols and configured static routes should be injected in the associated local VPN routing/forwarding (VRF). In the case of dynamic routing protocols, there may be protocol specific route policies that modify or reject specific routes before they are injected into the local VRF.

Route redistribution from the local VRF to CE-to-PE routing protocols is to be controlled via the route policies in each routing protocol instance, in the same manner that is used by the base router instance.

The advertisement or redistribution of routing information from the local VRF to or from the MP-BGP instance is specified per VRF and is controlled by VRF route target associations or by VRF route policies.

VPN-IP routes imported into a VPRN, have the protocol **type bgp-vpn** to denote that it is an VPRN route. This can be used within the route policy match criteria.

3.1.5.2 CPE connectivity check

Static routes are used within many IES services and VPRN services. Unlike dynamic routing protocols, there is no way to change the state of routes based on availability information for the associated CPE. CPE connectivity check adds flexibility so that unavailable destinations are removed from the VPRN routing tables dynamically and minimize wasted bandwidth. Figure 13: Directly connected IP target shows a setup with a directly connected IP target and Figure 14: Multiple hops to IP target shows a setup with multiple hops to an IP target.

Figure 13: Directly connected IP target

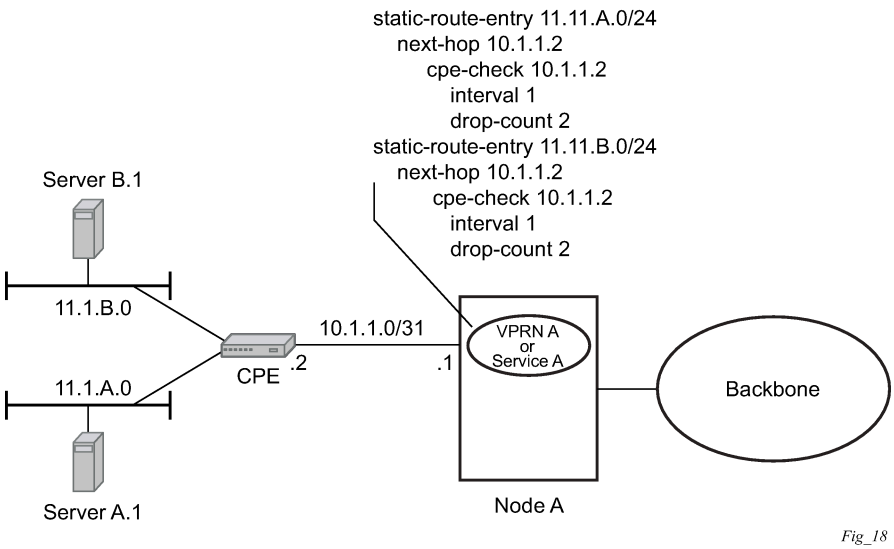
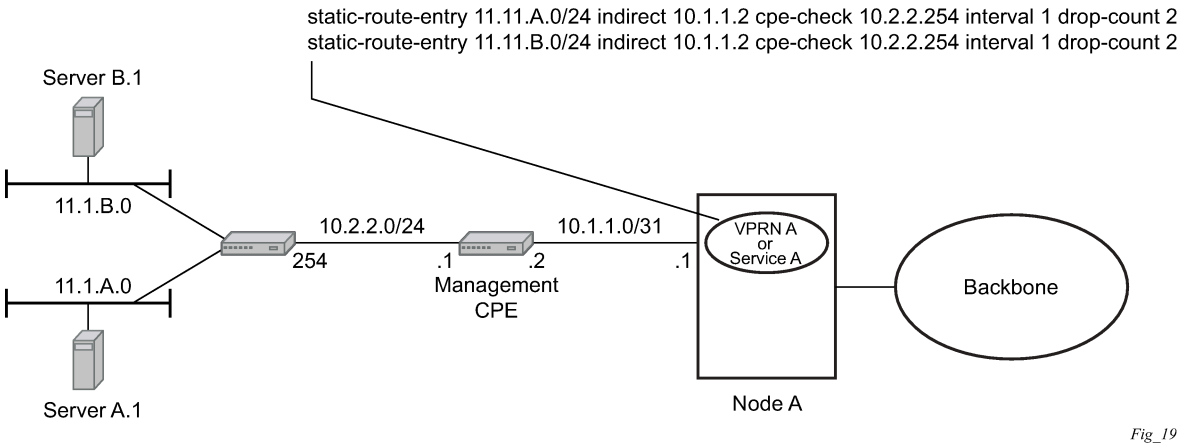


Figure 14: Multiple hops to IP target



The availability of the far-end static route is monitored through periodic polling. The polling period is configured. If the poll fails a specified number of sequential polls, the static route is marked as inactive.

Either ICMP ping or unicast ARP mechanism can be used to test the connectivity. ICMP ping is preferred.

If the connectivity check fails and the static route is deactivated, the router continues to send polls and re-activate any routes that are restored.

3.1.6 RT Constraint

3.1.6.1 Constrained VPN Route Distribution based on route targets

Constrained Route Distribution (or RT Constraint) is a mechanism that allows a router to advertise Route Target membership information to its BGP peers to indicate interest in receiving only VPN routes tagged with specific Route Target extended communities. Upon receiving this information, peers restrict the advertised VPN routes to only those requested, minimizing control plane load in terms of protocol traffic and possibly also RIB memory.

The Route Target membership information is carried using MP-BGP, using an AFI value of 1 and SAFI value of 132. In order for two routers to exchange RT membership NLRI they must advertise the corresponding AFI/SAFI to each other during capability negotiation. The use of MP-BGP means RT membership NLRI are propagated, loop-free, within an AS and between ASes using well-known BGP route selection and advertisement rules.

ORF can also be used for RT-based route filtering, but ORF messages have a limited scope of distribution (to direct peers) and therefore do not automatically create pruned inter-cluster and inter-AS route distribution trees.

3.1.6.2 Configuring the route target address family

RT Constraint is supported only by the base router BGP instance. When the **family** command at the BGP router group or neighbor CLI context includes the **route-target** keyword, the RT Constraint capability is negotiated with the associated set of EBGP and IBGP peers.

ORF and RT Constraint are mutually exclusive on a particular BGP session. The CLI does not attempt to block this configuration, but if both capabilities are enabled on a session, the ORF capability is not included in the OPEN message sent to the peer.

3.1.6.3 Originating RT constraint routes

When the base router has one or more RTC peers (BGP peers with which the RT Constraint capability has been successfully negotiated), one RTC route is created for each RT extended community imported into a locally-configured L2 VPN or L3 VPN service. These imported route targets are configured in the following contexts:

- **config**>**service**>**vprn**
- **config**>**service**>**vprn**>**mvpn**

By default, these RTC routes are automatically advertised to all RTC peers, without the need for an export policy to explicitly "accept" them. Each RTC route has a prefix, a prefix length and path attributes. The prefix value is the concatenation of the origin AS (a 4-byte value representing the 2- or 4-octet AS of the originating router, as configured using the **config>router>autonomous-system** command) and 0 or 16-64 bits of a route target extended community encoded in one of the following formats: 2-octet AS specific extended community, IPv4 address specific extended community, or 4-octet AS specific extended community.

A router may be configured to send the default RTC route to any RTC peer. This is done using the new **default-route-target** group/neighbor CLI command. The default RTC route is a special type of RTC route that has zero prefix length. Sending the default RTC route to a peer conveys a request to receive all VPN routes (regardless of route target extended community) from that peer. The default RTC route is typically advertised by a route reflector to its clients. The advertisement of the default RTC route to a peer does not suppress other more specific RTC routes from being sent to that peer.

3.1.6.4 Receiving and re-advertising RT Constraint routes

All received RTC routes that are deemed valid are stored in the RIB-IN. An RTC route is considered invalid and treated as withdrawn, if any of the following applies:

- The prefix length is 1-31.
- The prefix length is 33-47.
- The prefix length is 48-96 and the 16 most-significant bits are not 0x0002, 0x0102 or 0x0202.

If multiple RTC routes are received for the same prefix value then standard BGP best path selection procedures are used to determine the best of these routes.

The best RTC route per prefix is re-advertised to RTC peers based on the following rules:

- The best path for a default RTC route (prefix-length 0, origin AS only with prefix-length 32, or origin AS plus 16 bits of an RT type with prefix-length 48) is never propagated to another peer.
- A PE with only IBGP RTC peers that is neither a route reflector nor an ASBR does not re-advertise the best RTC route to any RTC peer because of standard IBGP split horizon rules.
- A route reflector that receives its best RTC route for a prefix from a client peer re-advertises that route (subject to export policies) to all of its client and non-client IBGP peers (including the originator), per standard RR operation. When the route is re-advertised to client peers, the RR (i) sets the ORIGINATOR_ID to its own router ID and (ii) modifies the NEXT_HOP to be its local address for the sessions (for example, system IP).
- A route reflector that receives its best RTC route for a prefix from a non-client peer re-advertises that route (subject to export policies) to all of its client peers, per standard RR operation. If the RR has a non-best path for the prefix from any of its clients, it advertises the best of the client-advertised paths to all non-client peers.
- An ASBR that is neither a PE nor a route reflector that receives its best RTC route for a prefix from an IBGP peer re-advertises that route (subject to export policies) to its EBGP peers. It modifies the NEXT_HOP and AS_PATH of the re-advertised route per standard BGP rules. No aggregation of RTC routes is supported.
- An ASBR that is neither a PE nor a route reflector that receives its best RTC route for a prefix from an External Border Gateway Protocol (EBGP) peer re-advertises that route (subject to export policies) to its EBGP and IBGP peers. When re-advertised routes are sent to EBGP peers, the ASBR modifies the NEXT_HOP and AS_PATH per standard BGP rules. No aggregation of RTC routes is supported.

Note:
These advertisement rules do not handle hierarchical RR topologies properly. This is a limitation of the current RT constraint standard.

3.1.6.5 Using RT Constraint routes

In general (ignoring IBGP-to-IBGP rules, Add-Path, Best-external, and so on), the best VPN route for every prefix/NLRI in the RIB is sent to every peer supporting the VPN address family, but export policies may be used to prevent some prefix/NLRI from being advertised to specific peers. These export policies may be configured statically or created dynamically based on use of ORF or RT constraint with a peer. ORF and RT Constraint are mutually exclusive on a session.

When RT Constraint is configured on a session that also supports VPN address families using route targets (that is: vpn-ipv4, vpn-ipv6, l2-vpn, mvpn-ipv4, mvpn-ipv6, mcast-vpn-ipv4 or evpn), the advertisement of the VPN routes is affected as follows:

1. When the session comes up, the advertisement of the VPN routes is delayed for a short while to allow RTC routes to be received from the peer.
2. After the initial delay, the received RTC routes are analyzed and acted upon. If S1 is the set of routes previously advertised to the peer and S2 is the set of routes that should be advertised based on the most recent received RTC routes, the following applies:
   - The set of routes in S1 but not in S2 should be withdrawn immediately (subject to MRAI).
   - The set of routes in S2 but not in S1 should be advertised immediately (subject to MRAI).

If a default RTC route is received from a peer P1, the VPN routes that are advertised to P1 is the set that meets all of the following requirements:

- The set is eligible for advertisement to P1 per BGP route advertisement rules.
- The set has not been rejected by manually configured export policies.
- The set has not been advertised to the peer.

Note: This applies whether P1 advertised the best route for the default RTC prefix or not.

In this context, a default RTC route is any of the following:

- a route with NLRI length = zero
- a route with NLRI value = origin AS and NLRI length = 32
- a route with NLRI value = {origin AS+0x0002 | origin AS+0x0102 | origin AS+0x0202} and NLRI length = 48
  - If an RTC route for prefix A (origin-AS = A1, RT = A2/n, n > 48) is received from an IBGP peer I1 in autonomous system A1, the VPN routes that are advertised to I1 is the set that meets all of the following requirements:
    - The set is eligible for advertisement to I1 per BGP route advertisement rules.
    - The set has not been rejected by manually configured export policies.
    - The set carries at least one route target extended community with value A2 in the n most significant bits.
    - The set has not been advertised to the peer.

    Note: This applies whether I1 advertised the best route for A or not.

  - If the best RTC route for a prefix A (origin-AS = A1, RT = A2/n, n > 48) is received from an IBGP peer I1 in autonomous system B, the VPN routes that are advertised to I1 is the set that meets all of the following requirements:
    - The set is eligible for advertisement to I1 per BGP route advertisement rules.
    - The set has not been rejected by manually configured export policies.
    - The set carries at least one route target extended community with value A2 in the n most significant bits.

- The set has not been advertised to the peer.

Note: This applies only if I1 advertised the best route for A.

– If the best RTC route for a prefix A (origin-AS = A1, RT = A2/n, n > 48) is received from an EBGP peer E1, the VPN routes that are advertised to E1 is the set that meets all of the following requirements:

- The set is eligible for advertisement to E1 per BGP route advertisement rules.
- The set has not been rejected by manually configured export policies.
- The set carries at least one route target extended community with value A2 in the n most significant bits.
- The set has not been advertised to the peer.

Note: This applies only if E1 advertised the best route for A.

3.1.7 BGP fast reroute in a VPRN

BGP fast reroute is a feature that brings together indirection techniques in the forwarding plane and pre-computation of BGP backup paths in the control plane to support fast reroute of BGP traffic around unreachable/failed next-hops. In a VPRN context BGP fast reroute is supported using unlabeled IPv4, unlabeled IPv6, VPN-IPv4, and VPN-IPv6 VPN routes. The supported VPRN scenarios are described in Table 6: BGP fast reroute scenarios (VPRN context).

BGP fast reroute information specific to the base router BGP context is described in the BGP Fast Reroute section of the *7450 ESS, 7750 SR, 7950 XRS, and VSR Unicast Routing Protocols Guide*.

Table 6: BGP fast reroute scenarios (VPRN context)

Ingress packet	Primary route	Backup route	Prefix independent convergence
IPv4 (ingress PE)	IPv4 route with next-hop A resolved by an IPv4 route	IPv4 route with next-hop B resolved by an IPv4 route	Yes
IPv4 (ingress PE)	VPN-IPv4 route with next-hop A resolved by a GRE, LDP, RSVP or BGP tunnel	VPN-IPv4 route with next-hop A resolved by a GRE, LDP, RSVP or BGP tunnel	Yes
MPLS (egress PE)	IPv4 route with next-hop A resolved by an IPv4 route	IPv4 route with next-hop B resolved by an IPv4 route	Yes
MPLS (egress PE)	IPv4 route with next-hop A resolved by an IPv4 rout	VPN-IPv4 route* with next-hop B resolved by a GRE, LDP, RSVP or BGP tunnel	Yes
IPv6 (ingress PE)	IPv6 route with next-hop A resolved by an IPv6 route	IPv6 route with next-hop B resolved by an IPv6 route	Yes

Ingress packet	Primary route	Backup route	Prefix independent convergence
IPv6 (ingress PE)	VPN-IPv6 route with next-hop A resolved by a GRE, LDP, RSVP or BGP tunnel	VPN-IPv6 route with next-hop B resolved by a GRE, LDP, RSVP or BGP tunnel	Yes
MPLS (egress)	IPv6 route with next-hop A resolved by an IPv6 route	IPv6 route with next-hop B resolved by an IPv6 route	Yes
MPLS (egress)	IPv6 route with next-hop A resolved by an IPv6 route	Yes	VPRN label mode must be VRF. VPRN must export its VPN-IP routes with RD ≠ y. For the best performance the backup next-hop must advertise the same VPRN label value with all routes (per VRF label).

3.1.7.1 BGP fast reroute in a VPRN configuration

In a VPRN context, BGP fast reroute is optional and must be enabled. Fast reroute can be applied to all IPv4 prefixes, all IPv6 prefixes, all IPv4 and IPv6 prefixes, or to a specific set of IPv4 and IPv6 prefixes.

If all IP prefixes require backup path protection, use a combination of the BGP instance-level **backup-path** and VPRN-level **enable-bgp-vpn-backup** commands. The VPRN BGP **backup-path** command enables BGP fast reroute for all IPv4 prefixes or all IPv6 prefixes, or both, that have a best path through a VPRN BGP peer. The VPRN-level **enable-bgp-vpn-backup** command enables BGP fast reroute for all IPv4 prefixes or all IPv6 prefixes, or both, that have a best path through a remote PE peer.

If only some IP prefixes require backup path protection, use route policies to apply the **install-backup-path** action to the best paths of the IP prefixes requiring protection. See the "BGP Fast Reroute" section of the *7450 ESS, 7750 SR, 7950 XRS, and VSR Unicast Routing Protocols Guide* for more information.

3.1.8 BGP best-external in a VPRN context

If two or more PE routers connect to a multi-homed site and learn routes for a common set of IP prefixes from that site, then the failure of one of the PE routers or a PE-CE link can be handled by rerouting the traffic over the alternate paths. The traffic failover time in this situation can be reduced if all the PE routers have advance knowledge of the potential backup paths and do not have to wait for BGP route advertisements or withdrawals, or both, to reprogram their forwarding tables. This can be challenging with normal BGP procedures because a PE router is not allowed to advertise, to other PE routers, a BGP route that it has learned from a connected CE device if that route is not its active route for the destination in the route table. If the multi-homing scenario calls for all traffic destined for an IP prefix to be carried over a preferred primary path (passing through PE1-CE1 for example), then all other PE routers (PE2, PE3, and so on) have that VPN route as their active route for the destination, and they are not able to advertise their own routes for the same IP prefix.

The SR OS supports a VPRN feature, configured using the **export-inactive-bgp** command, that resolves the issue described above. When a VPRN is configured with this command, it is allowed to advertise (as a VPN-IP route toward other PEs) its best CE-BGP route for an IP prefix, even when that CE-BGP route

is inactive in the route table because of the presence of a more-preferred VPN-IP route from another PE. In order for the CE-BGP route to be advertised, the CE-BGP route must be accepted by the VRF export policy. When a VPN-IP route is advertised because of the **export-inactive-bgp** command, the label carried in the route is a per-next-hop label corresponding to the next-hop IP address of the CE-BGP route, or a per-prefix label; this helps avoid packet looping issues because of unsynchronized IP FIBs.

When a PE router that advertised a backup path for an IP prefix receives a withdrawal for the VPN-IP route that it was using as the primary/active route, its backup path may be promoted to the primary path; that is, the CE-BGP route may become the active route for the destination. In this case, the PE router is required to re-advertise the VPN-IP route with a per-VRF label if that is the default allocation policy and there is no label-per-prefix policy override. It takes some time for the new VPN-IP route to reach all the ingress routers and for them to update their forwarding tables. In the meantime, traffic continues to be received with the old per-next-hop label. The egress PE drops this in-flight traffic unless **label retention** is configured using the **bgp-labels-hold-timer** command in the **config>router>mpls-labels** context. This command configures a delay (in seconds) between the withdrawal of a VPN-IP route with a per-next-hop label and the deletion of the corresponding label forwarding entry in the IOM. The value of **bgp-labels-hold-timer** should be large enough to account for the propagation delay of the route withdrawal to all the ingress routers.

3.2 VPRN features

This section describes various VPRN features and any special capabilities or considerations as they relate to VPRN services.

3.2.1 IP interfaces

VPRN customer IP interfaces can be configured with most of the same options found on the core IP interfaces. The advanced configuration options supported are as follows:

- VRRP
- Cflowd
- secondary IP addresses
- ICMP options

Note: NTP broadcast receipt is a configuration option found on core IP interfaces that is not supported on VPRN IP interfaces.

3.2.1.1 QoS Policy Propagation Using BGP

This section discusses QPPB as it applies to VPRN, IES, and router interfaces. See the QoS Policy Propagation Using BGP (QPPB) section and the IP Router Configuration section in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

The QoS Policy Propagation Using BGP (QPPB) feature applies only to the 7450 ESS and 7750 SR.

QoS policy propagation using BGP (QPPB) is a feature that allows a route to be installed in the routing table with a forwarding-class and priority so that packets matching the route can receive the associated QoS. The forwarding-class and priority associated with a BGP route are set using BGP import route policies. In the industry, this feature is called QPPB, and even though the feature name refers to BGP

specifically. On SR OS, QPPB is supported for BGP (IPv4, IPv6, VPN-IPv4, VPN-IPv6), RIP and static routes.

While SAP ingress and network QoS policies can achieve the same end result as QPPB, the effort involved in creating the QoS policies, keeping them up-to-date, and applying them across many nodes is much greater than with QPPB. This is because of assigning a packet, arriving on a particular IP interface, to a specific forwarding-class and priority/profile, based on the source IP address or destination IP address of the packet. In a typical application of QPPB, a BGP route is advertised with a BGP community attribute that conveys a particular QoS. Routers that receive the advertisement accept the route into their routing table and set the forwarding-class and priority of the route from the community attribute.

3.2.1.2 QPPB applications

The typical applications of QPPB are as follows:

- coordination of QoS policies between different administrative domains
- traffic differentiation within a single domain, based on route characteristics

3.2.1.3 Inter-AS coordination of QoS policies

The operator of an administrative domain A can use QPPB to signal to a peer administrative domain B that traffic sent to specific prefixes advertised by domain A should receive a particular QoS treatment in domain B. More specifically, an ASBR of domain A can advertise a prefix XYZ to domain B and include a BGP community attribute with the route. The community value implies a particular QoS treatment, as agreed by the two domains (in their peering agreement or service level agreement, for example). When the ASBR and other routers in domain B accept and install the route for XYZ into their routing table, they apply a QoS policy on selected interfaces that classifies traffic toward network XYZ into the QoS class implied by the BGP community value.

QPPB may also be used to request that traffic sourced from specific networks receive appropriate QoS handling in downstream nodes that may span different administrative domains. This can be achieved by advertising the source prefix with a BGP community, as discussed above. However, in this case other approaches are equally valid, such as marking the DSCP or other CoS fields based on source IP address so that downstream domains can take action based on a common understanding of the QoS treatment implied by different DSCP values.

In the above examples, coordination of QoS policies using QPPB could be between a business customer and its IP VPN service provider, or between one service provider and another.

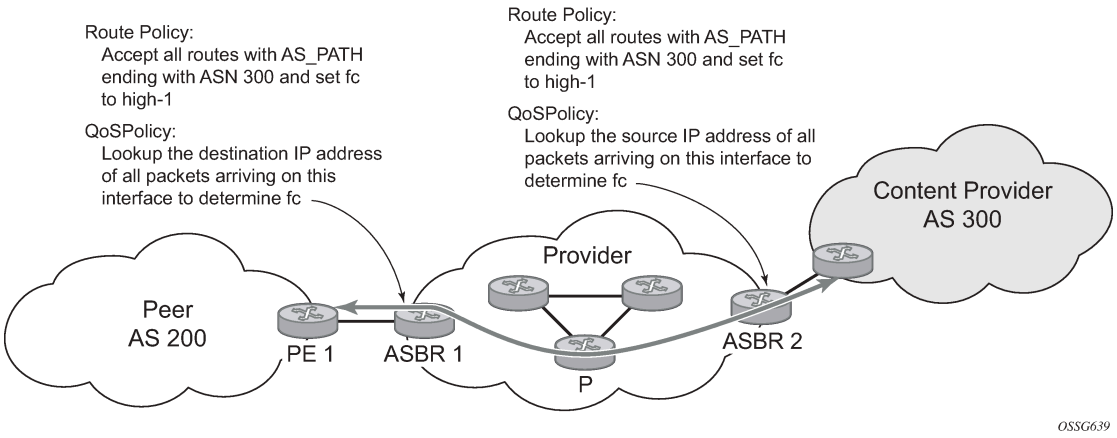
3.2.1.4 Traffic differentiation based on route characteristics

There may be times when a network operator wants to provide differentiated service to certain traffic flows within its network, and these traffic flows can be identified with known routes. For example, the operator of an ISP network may want to give priority to traffic originating in a particular ASN (the ASN of a content provider offering over-the-top services to the ISP's customers), following a specific AS_PATH, or destined for a particular next-hop (remaining on-net vs. off-net).

Figure 15: Use of QPPB to differentiate traffic in an ISP network shows an example of an ISP that has an agreement with the content provider managing AS300 to provide traffic sourced and terminating within AS300 with differentiated service appropriate to the content being transported. In this example we presume that ASBR1 and ASBR2 mark the DSCP of packets terminating and sourced, respectively, in AS300 so

that other nodes within the ISP's network do not need to rely on QPPB to determine the correct forwarding-class to use for the traffic. The DSCP or other CoS markings could be left unchanged in the ISP's network and QPPB used on every node.

Figure 15: Use of QPPB to differentiate traffic in an ISP network



3.2.1.5 QPPB

There are two main aspects of the QPPB feature on the 7450 ESS and 7750 SR:

- the ability to associate a forwarding-class and priority with specific routes in the routing table
- the ability to classify an IP packet arriving on a particular IP interface to the forwarding-class and priority associated with the route that best matches the packet

3.2.1.6 Associating an FC and priority with a route

This feature uses the **fc** command in the route-policy hierarchy to set the forwarding class and optionally the priority associated with routes accepted by a route-policy entry. The command has the following structure:

fc *fc-name* [**priority** {**low** | **high**}]

The use of this command is illustrated by the following example:

```
config>router>policy-options
begin
community gold members 300:100
policy-statement qppb_policy
entry 10
from
protocol bgp
community gold
exit
action accept
fc h1 priority high
exit
exit
exit
commit
```

The **fc** command is supported with all existing from and to match conditions in a route policy entry and with any action other than reject, it is supported with next-entry, next-policy and accept actions. If a next-entry or next-policy action results in multiple matching entries then the last entry with a QPPB action determines the forwarding class and priority.

A route policy that includes the **fc** command in one or more entries can be used in any import or export policy, but the **fc** command has no effect except in the following types of policies:

- VRF import policies
  - config>service>vprn>vrf-import
- BGP import policies
  - config>router>bgp>import
  - config>router>bgp>group>import
  - config>router>bgp>group>neighbor>import
  - config>service>vprn>bgp>import
  - config>service>vprn>bgp>group>import
  - config>service>vprn>bgp>group>neighbor>import
- RIP import policies
  - config>router>rip>import
  - config>router>rip>group>import
  - config>router>rip>group>neighbor>import
  - config>service>vprn>rip>import
  - config>service>vprn>rip>group>import
  - config>service>vprn>rip>group>neighbor>import

As evident from above, QPPB route policies support routes learned from RIP and BGP neighbors of a VPRN as well as for routes learned from RIP and BGP neighbors of the base/global routing instance.

QPPB is supported for BGP routes belonging to any of the address families listed below:

- IPv4 (AFI=1, SAFI=1)
- IPv6 (AFI=2, SAFI=1)
- VPN-IPv4 (AFI=1, SAFI=128)
- VPN-IPv6 (AFI=2, SAFI=128)

A VPN-IP route may match both a VRF import policy entry and a BGP import policy entry (if vpn-apply-import is configured in the base router BGP instance). In this case the VRF import policy is applied first and then the BGP import policy, so the QPPB QoS is based on the BGP import policy entry.

This feature also introduces the ability to associate a forwarding-class and optionally priority with IPv4 and IPv6 static routes. This is achieved by specifying the forwarding-class within the static-route-entry next-hop or indirect context.

Priority is optional when specifying the forwarding class of a static route, but once configured it can only be deleted and returned to unspecified by deleting the entire static route.

3.2.1.7 Displaying QoS information associated with routes

The following commands are enhanced to show the forwarding-class and priority associated with the displayed routes:

- **show router route-table**
- **show router fib**
- **show router bgp routes**
- **show router rip database**
- **show router static-route**

This feature uses a **qos** keyword to the **show>router>route-table** command. When this option is specified the output includes an additional line per route entry that displays the forwarding class and priority of the route. If a route has no fc and priority information, then the third line is blank. The following CLI shows an example:

show router route-table [**family**] [*ip-prefix*[/*prefix-length*]] [**longer** | **exact**] [**protocol** *protocol-name*] **qos**

An example output of this command is shown below:

```
A:Dut-A# show router route-table 10.1.5.0/24 qos
===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix                                Type    Proto    Age         Pref
       Next Hop[Interface Name]                                 Metric
       QoS
-------------------------------------------------------------------------------
10.1.5.0/24                                Remote  BGP      15h32m52s   0
       PE1_to_PE2                                               0
       h1, high
-------------------------------------------------------------------------------
No. of Routes: 1
===============================================================================
A:Dut-A#
```

3.2.1.8 Enabling QPPB on an IP interface

To enable QoS classification of ingress IP packets on an interface based on the QoS information associated with the routes that best match the packets the **qos-route-lookup** command is necessary in the configuration of the IP interface. The **qos-route-lookup** command has parameters to indicate whether the QoS result is based on lookup of the source or destination IP address in every packet. There are separate qos-route-lookup commands for the IPv4 and IPv6 packets on an interface, which allows QPPB to enabled for IPv4 only, IPv6 only, or both IPv4 and IPv6. Current QPPB based on a source IP address is not supported for IPv6 packets nor is it supported for ingress subscriber management traffic on a group interface.

The qos-route-lookup command is supported on the following types of IP interfaces:

- base router network interfaces (config>router>interface)
- VPRN SAP and spoke SDP interfaces (config>service>vprn>interface)
- VPRN group-interfaces (config>service>vprn>sub-if>grp-if)
- IES SAP and spoke SDP interfaces (config>service>ies>interface)

- IES group-interfaces (config>service>ies>sub-if>grp-if)

When the qos-route-lookup command with the destination parameter is applied to an IP interface and the destination address of an incoming IP packet matches a route with QoS information the packet is classified to the fc and priority associated with that route, overriding the fc and priority/profile determined from the SAP-Ingress or network qos policy associated with the IP interface. If the destination address of the incoming packet matches a route with no QoS information the fc and priority of the packet remain as determined by the SAP-Ingress or network qos policy.

Similarly, when the qos-route-lookup command with the source parameter is applied to an IP interface and the source address of an incoming IP packet matches a route with QoS information the packet is classified to the fc and priority associated with that route, overriding the fc and priority/profile determined from the SAP-Ingress or network qos policy associated with the IP interface. If the source address of the incoming packet matches a route with no QoS information the fc and priority of the packet remain as determined by the SAP-Ingress or network qos policy.

Currently, QPPB is not supported for ingress MPLS traffic on network interfaces or on CsC PE'-CE' interfaces (config>service>vprn>nw-if).

3.2.1.9 QPPB when next hops are resolved by QPPB routes

In some circumstances (IP VPN inter-AS model C, Carrier Supporting Carrier, indirect static routes, and so on) an IPv4 or IPv6 packet may arrive on a QPPB-enabled interface and match a route A1 whose next-hop N1 is resolved by a route A2 with next-hop N2 and perhaps N2 is resolved by a route A3 with next-hop N3, and so on. The QPPB result is based only on the forwarding-class and priority of route A1. If A1 does not have a forwarding-class and priority association then the QoS classification is not based on QPPB, even if routes A2, A3, and so on have forwarding-class and priority associations.

3.2.1.10 QPPB and multiple paths to a destination

When ECMP is enabled some routes may have multiple equal-cost next-hops in the forwarding table. When an IP packet matches such a route the next-hop selection is typically based on a hash algorithm that tries to load balance traffic across all the next-hops while keeping all packets of a specific flow on the same path. The QPPB configuration model described in Associating an FC and priority with a route allows different QoS information to be associated with the different ECMP next hops of a route. The forwarding-class and priority of a packet matching an ECMP route is based on the particular next-hop used to forward the packet.

When BGP FRR is enabled some BGP routes may have a backup next-hop in the forwarding table in addition to the one or more primary next-hops representing the equal-cost best paths allowed by the ECMP/multipath configuration. When an IP packet matches such a route a reachable primary next-hop is selected (based on the hash result) but if all the primary next-hops are unreachable then the backup next-hop is used. The QPPB configuration model described in Associating an FC and priority with a route allows the forwarding-class and priority associated with the backup path to be different from the QoS characteristics of the equal-cost best paths. The forwarding class and priority of a packet forwarded on the backup path is based on the **fc** and priority of the backup route.

3.2.1.11 QPPB and policy-based routing

When an IPv4 or IPv6 packet with destination address X arrives on an interface with both QPPB and policy-based-routing enabled:

- There is no QPPB classification if the IP filter action redirects the packet to a directly connected interface, even if X is matched by a route with a forwarding-class and priority
- QPPB classification is based on the forwarding-class and priority of the route matching IP address Y if the IP filter action redirects the packet to the indirect next-hop IP address Y, even if X is matched by a route with a forwarding-class and priority.

3.2.1.12 QPPB and GRT lookup

Source-address based QPPB is not supported on any SAP or spoke SDP interface of a VPRN configured with the **grt-lookup** command.

3.2.1.13 QPPB interaction with SAP ingress QoS policy

When QPPB is enabled on a SAP IP interface the forwarding class of a packet may change from **fc1**, the original **fc** determined by the SAP ingress QoS policy to fc2, the new fc determined by QPPB. In the ingress datapath SAP ingress QoS policies are applied in the first P chip and route lookup/QPPB occurs in the second P chip. This has the implications listed below:

- Ingress remarking (based on profile state) is always based on the original fc (fc1) and sub-class (if defined).
- The profile state of a SAP ingress packet that matches a QPPB route depends on the configuration of **fc2** only. If the de-1-out-profile flag is enabled in **fc2** and **fc2** is not mapped to a priority mode queue then the packet is marked out of profile if its DE bit = 1. If the profile state of **fc2** is explicitly configured (in or out) and **fc2** is not mapped to a priority mode queue, then the packet is assigned this profile state. In both cases there is no consideration of whether **fc1** was mapped to a priority mode queue.
- The priority of a SAP ingress packet that matches a QPPB route depends on several factors. If the de-1-out-profile flag is enabled in **fc2** and the DE bit is set in the packet then priority is low regardless of the QPPB priority or **fc2** mapping to profile mode queue, priority mode queue or policer. If **fc2** is associated with a profile mode queue then the packet priority is based on the explicitly configured profile state of **fc2** (in profile = high, out profile = low, undefined = high), regardless of the QPPB priority or **fc1** configuration. If **fc2** is associated with a priority mode queue or policer then the packet priority is based on QPPB (unless DE=1), but if no priority information is associated with the route then the packet priority is based on the configuration of **fc1** (if **fc1** mapped to a priority mode queue then it is based on DSCP/IP prec/802.1p and if **fc1** mapped to a profile mode queue then it is based on the profile state of **fc1**).

Table 7: QPPB interactions with SAP ingress QoS summarizes these interactions.

Table 7: QPPB interactions with SAP ingress QoS

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
Profile mode queue	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
Priority mode queue	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Policer	Policer	From new base FC unless overridden by DE=1	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Priority mode queue	Policer	From new base FC unless overridden by DE=1	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Policer	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then from original dot1p/ exp/DSCP mapping or policy default.	From new base FC	From original FC and sub-class
Profile mode queue	Priority mode queue	Ignored	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then follows original FC's profile mode rules.	From new base FC	From original FC and sub-class
Priority mode queue	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class
Profile mode queue	Policer	From new base FC unless	If DE=1 override then low otherwise from QPPB. If no DEI or QPPB overrides then	From new base FC	From original FC and sub-class

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
		overridden by DE=1	follows original FC's profile mode rules.		
Policer	Profile mode queue	From new base FC unless overridden by DE=1	From QPPB, unless packet is marked in or out of profile in which case follows profile. Default is high priority.	From new base FC	From original FC and sub-class

3.2.1.14 Object grouping and state monitoring

This feature introduces a generic operational group object which associates different service endpoints (pseudowires and SAPs) located in the same or in different service instances. The operational group status is derived from the status of the individual components using specific rules specific to the application using the concept. A number of other service entities, the monitoring objects, can be configured to monitor the operational group status and to perform specific actions as a result of status transitions. For example, if the operational group goes down, the monitoring objects are brought down.

3.2.1.15 VPRN IP interface applicability

Prerequisites

This concept is used by an IPv4 VPRN interface to affect the operational state of the IP interface monitoring the operational group. Individual SAP and spoke SDPs are supported as monitoring objects.

The following rules apply:

- An object can only belong to one group at a time.
- An object that is part of a group cannot monitor the status of a group.
- An object that monitors the status of a group cannot be part of a group.
- An operational group may contain any combination of member types: SAP or Spoke-SDPs.
- An operational group may contain members from different VPLS service instances.
- Objects from different services may monitor the oper-group.

Procedure

Step 1. Identify a set of objects whose forwarding state should be considered as a whole group then group them under an operational group using the **oper-group** command.

Step 2. Associate the IP interface to the oper-group using the **monitor-group** command.

What to do next

The status of the operational group (oper-group) is dictated by the status of one or more members according to the following rules:

- The oper-group goes down if all the objects in the oper-group go down. The oper-group comes up if at least one of the components is up.

- An object in the group is considered down if it is not forwarding traffic in at least one direction. That could be because the operational state is down or the direction is blocked through some validation mechanism.
- If a group is configured but no members are specified yet then its status is considered up.
- As soon as the first object is configured the status of the operational group is dictated by the status of the provisioned members.

The simple configuration below shows the oper-group g1, the VPLS SAP that is mapped to it and the IP interfaces in VPRN service 2001 monitoring the oper-group g1. This is example uses an R-VPLS context. The VPLS instance includes the **allow-ip-int-bind** and the **name v1**. The VPRN interface links to the VPLS using the **vpls v1** option. All commands are under the configuration service hierarchy.

To further describe the configuration. Oper-group g1 has a single SAP (1/1/1:2001) mapped to it and the IP interfaces in the VPRN service 2001 derive its state from the state of oper-group g1.

```
oper-group g1 create

vpls 1 name "v1" customer 1 create
            allow-ip-int-bind
            stp
                shutdown
            exit
            sap 1/1/1:2001 create
                oper-group g1
                eth-cfm
                   mep domain 1 association 1 direction down
 ccm-enable
       no shutdown
                exit
            exit
            sap 1/1/2:2001 create
            exit
            sap 1/1/3:2001 create
            exit
no shutdown


vprn 2001 customer 1 create
            interface "i2001"  create
                address 10.1.1.1/24
                monitor-oper-group "g1"
                vpls "v1"
            exit
          no shutdown
          exit
```

3.2.2 Subscriber interfaces

Subscriber interfaces are composed of a combination of two key technologies, subscriber interfaces and group interfaces. While the subscriber interface defines the subscriber subnets, the group interfaces are responsible for aggregating the SAPs.

Subscriber interfaces apply only to the 7450 ESS and 7750 SR.

- **subscriber interface**

  This is an interface that allows the sharing of a subnet among one or many group interfaces in the routed CO model.

- **group interface**

  This interface aggregates multiple SAPs on the same port.

- **redundant interfaces**

  This is a special spoke-terminated Layer 3 interface. It is used in a Layer 3 routed CO dual-homing configuration to shunt downstream (network to subscriber) to the active node for a specific subscriber.

3.2.3 SAPs

3.2.3.1 Encapsulations

The following SAP encapsulations are supported on the 7750 SR and 7950 XRS VPRN service:

- Ethernet null
- Ethernet dot1q
- QinQ
- LAG
- Tunnel (IPsec or GRE)

3.2.3.2 Pseudowire SAPs

Pseudowire SAPs are supported on VPRN interfaces for the 7750 SR in the same way as on IES interfaces.

3.2.4 QoS policies

When applied to a VPRN SAP, service ingress QoS policies only create the unicast queues defined in the policy if PIM is not configured on the associated IP interface; if PIM is configured, the multipoint queues are applied as well.

With VPRN services, service egress QoS policies function as with other services where the class-based queues are created as defined in the policy.

Both Layer 2 and Layer 3 criteria can be used in the QoS policies for traffic classification in an VPRN.

3.2.5 Filter policies

Ingress and egress IPv4 and IPv6 filter policies can be applied to VPRN SAPs.

3.2.6 DSCP marking

Specific DSCP, forwarding class, and Dot1P parameters can be specified to be used by every protocol packet generated by the VPRN. This enables prioritization or de-prioritization of every protocol (as required). The markings effect a change in behavior on ingress when queuing. For example, if OSPF is not

enabled, then traffic can be de-prioritized to best effort (be) DSCP. This change de-prioritizes OSPF traffic to the CPU complex.

DSCP marking for internally generated control and management traffic by marking the DSCP value should be used for the specific application. This can be configured per routing instance. For example, OSPF packets can carry a different DSCP marking for the base instance and then for a VPRN service. ISIS and ARP traffic is not an IP-generated traffic type and is not configurable. See Table 8: DSCP/FC marking .

When an application is configured to use a specified DSCP value then the MPLS EXP, Dot1P bits are marked in accordance with the network or access egress policy as it applies to the logical interface the packet is egressing.

The DSCP value can be set per application. This setting is forwarded to the egress line card. The egress line card does not alter the coded DSCP value and marks the LSP-EXP and IEEE 802.1p (Dot1P) bits according to the appropriate network or access QoS policy.

Table 8: DSCP/FC marking

Protocol	IPv4	IPv6	DSCP marking	Dot1P marking	Default FC
ARP	—	—	—	Yes	NC
BGP	Yes	Yes	Yes	Yes	NC
BFD	Yes	—	Yes	Yes	NC
RIP	Yes	Yes	Yes	Yes	NC
PIM (SSM)	Yes	Yes	Yes	Yes	NC
OSPF	Yes	Yes	Yes	Yes	NC
SMTP	Yes	—	—	—	AF
IGMP/MLD	Yes	Yes	Yes	Yes	AF
Telnet	Yes	Yes	Yes	Yes	AF
TFTP	Yes	—	Yes	Yes	AF
FTP	Yes	—	—	—	AF
SSH (SCP)	Yes	Yes	Yes	Yes	AF
SNMP (get, set, and so on)	Yes	Yes	Yes	Yes	AF
SNMP trap/log	Yes	Yes	Yes	Yes	AF
syslog	Yes	Yes	Yes	Yes	AF
OAM ping	Yes	Yes	Yes	Yes	AF
ICMP ping	Yes	Yes	Yes	Yes	AF
Traceroute	Yes	Yes	Yes	Yes	AF

Protocol	IPv4	IPv6	DSCP marking	Dot1P marking	Default FC
TACPLUS	Yes	Yes	Yes	Yes	AF
DNS	Yes	Yes	Yes	Yes	AF
SNTP/NTP	Yes	—	—	—	AF
RADIUS	Yes	—	—	—	AF
Cflowd	Yes	—	—	—	AF
DHCP 7450 ESS and 7750 SR only	Yes	Yes	Yes	Yes	AF
Bootp	Yes	—	—	—	AF
IPv6 Neighbor Discovery	Yes	—	—	—	NC

3.2.6.1 Default DSCP mapping table

```
DSCP NameDSCP ValueDSCP ValueDSCP ValueLabel
Decimal Hexadecimal Binary
============================================================
Default 00x00 0b000000be
nc1 48 0x30 0b110000h1
nc2 56 0x38 0b111000nc
ef 46 0x2e 0b101110ef
af11100x0a0b001010assured
af12120x0c0b001100assured
af13140x0e0b001110assured
af21 18 0x12 0b010010l1
af22 20 0x14 0b010100l1
af23220x160b010110l1
af31 26 0x1a 0b011010l1
af32 28 0x1c 0b011100l1
af33 30 0x1d 0b011110l1
af41 34 0x22 0b100010h2
af42 36 0x24 0b100100h2
af43 38 0x26 0b100110h2

default*0
```

*The default forwarding class mapping is used for all DSCP names or values for which there is no explicit forwarding class mapping.

3.2.7 Configuration of TTL propagation for VPRN routes

This feature allows the separate configuration of TTL propagation for in transit and CPM generated IP packets, at the ingress LER within a VPRN service context. The following commands are supported:

- config router ttl-propagate vprn-local [none | vc-only | all]
- config router ttl-propagate vprn-transit [none | vc-only | all]

You can enable TTL propagation behavior separately as follows:

- for locally generated packets by CPM (vprn-local)
- for user and control packets in transit at the node (vprn-transit)

The following parameters can be specified:

- The **all** parameter enables TTL propagation from the IP header into all labels in the stack, for VPN-IPv4 and VPN-IPv6 packets forwarded in the context of all VPRN services in the system.
- The **vc-only** parameter reverts to the default behavior by which the IP TTL is propagated into the VC label but not to the transport labels in the stack. You can explicitly set the default behavior by configuring the vc-only value.
- The **none** parameter disables the propagation of the IP TTL to all labels in the stack, including the VC label. This is needed for a transparent operation of UDP traceroute in VPRN inter-AS Option B such that the ingress and egress ASBR nodes are not traced.

This command does not use a no version.

The user can override the global configuration within each VPRN instance using the following commands:

- config service vprn ttl-propagate local [inherit | none | vc-only | all]
- config service vprn ttl-propagate transit [inherit | none | vc-only | all]

The default behavior for a VPRN instance is to inherit the global configuration for the same command. You can explicitly set the default behavior by configuring the inherit value.

This command does not have a no version.

The commands do not apply when the VPRN packet is forwarded over GRE transport tunnel.

If a packet is received in a VPRN context and a lookup is done in the Global Routing Table (GRT), (when leaking to GRT is enabled for example), the behavior of the TTL propagation is governed by the LSP shortcut configuration as follows:

- when the matching route is an RSVP LSP shortcut: **configure router mpls shortcut-transit-ttl-propagate**
- when the matching route is an LDP LSP shortcut: **configure router ldp shortcut-transit-ttl-propagate**

When the matching route is a RFC 8277 label route or a 6PE route, It is governed by the BGP label route configuration

When a packet is received on one VPRN instance and is redirected using Policy Based Routing (PBR) to be forwarded in another VPRN instance, the TTL propagation is governed by the configuration of the outgoing VPRN instance.

Packets that are forwarded in different contexts can use different TTL propagation over the same BGP tunnel, depending on the TTL configuration of each context. An example of this may be VPRN using a BGP tunnel and an IPv4 packet forwarded over a BGP label route of the same prefix as the tunnel.

3.2.8 CE to PE routing protocols

The 7750 SR and 7950 XRS VPRN supports the following PE to CE routing protocols:

- BGP
- Static
- RIP
- OSPF

3.2.8.1 PE to PE tunneling mechanisms

The 7750 SR and 7950 XRS support multiple mechanisms to provide transport tunnels for the forwarding of traffic between PE routers within the 2547bis network.

The 7750 SR and 7950 XRS VPRN implementation supports the use of:

- RSVP-TE protocol to create tunnel LSPs between PE routers
- LDP protocol to create tunnel LSP's between PE routers
- GRE tunnels between PE routers

These transport tunnel mechanisms provide the flexibility of using dynamically created LSPs where the service tunnels are automatically bound (the autobind feature) and the ability to provide specific VPN services with their own transport tunnels by explicitly binding SDPs if needed. When the autobind is used, all services traverse the same LSPs and do not allow alternate tunneling mechanisms (like GRE) or the ability to craft sets of LSPs with bandwidth reservations for specific customers as is available with explicit SDPs for the service.

3.2.8.2 Per VRF route limiting

The 7750 SR and 7950 XRS allow setting the maximum number of routes that can be accepted in the VRF for a VPRN service. There are options to specify a percentage threshold at which to generate an event that the VRF table is near full and an option to disable additional route learning when full or only generate an event.

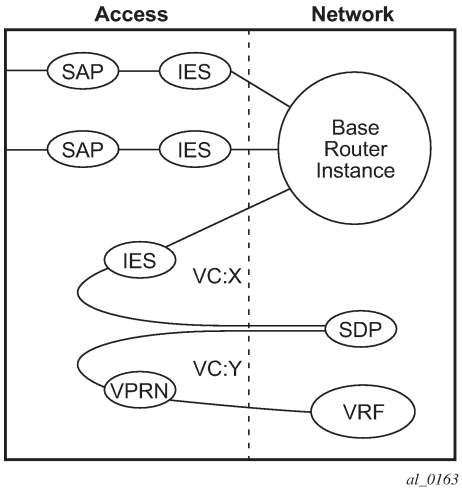
3.2.9 Spoke SDPs

Distributed services use service distribution points (SDPs) to direct traffic to another router via service tunnels. SDPs are created on each participating router and then bound to a specific service. SDP can be created as either GRE or MPLS. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR Services Overview Guide* for information about configuring SDPs.

This feature provides the ability to cross-connect traffic entering on a spoke SDP, used for Layer 2 services (VLLs or VPLS), on to an IES or VPRN service. From a logical point of view, the spoke SDP entering on a network port is cross-connected to the Layer 3 service as if it entered by a service SAP. The main exception to this is traffic entering the Layer 3 service by a spoke SDP is handled with network QoS policies and not access QoS policies.

Figure 16: SDP-ID and VC label service identifiers depicts traffic terminating on a specific IES or VPRN service that is identified by the **sdp-id** and VC label present in the service packet.

Figure 16: SDP-ID and VC label service identifiers



See "VCCV BFD support for VLL, Spoke SDP Termination on IES and VPRN, and VPLS Services" in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Layer 2 Services and EVPN Guide* for information about using VCCV BFD in spoke-SDP termination.

Note: Spoke-SDP termination of Ipipe VLLs on VPRN is not supported in System Profile B. To determine if Ipipes are currently bound to an VPRN interface, use the **show router ldp bindings services** command before configuring profile B.

3.2.9.1 T-LDP status signaling for spoke-SDPs terminating on IES/VPRN

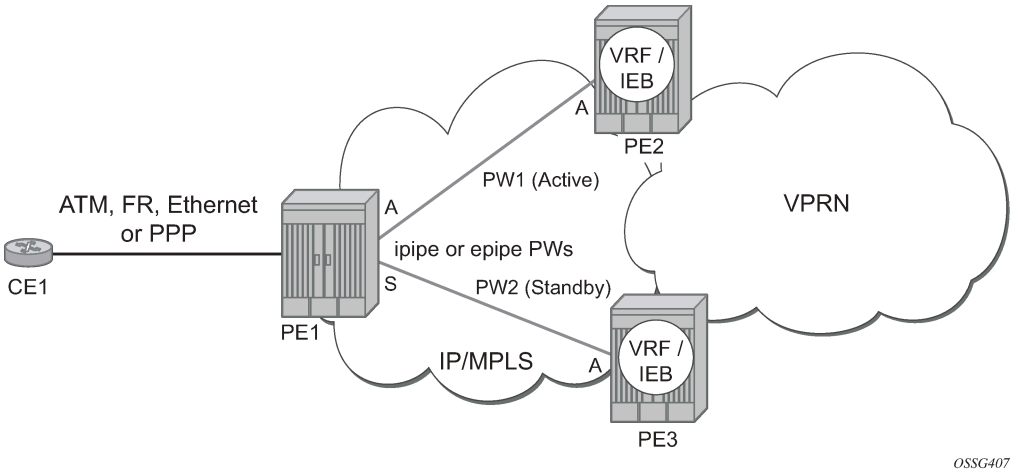
T-LDP status signaling and PW active/standby signaling capabilities are supported on Ipipe and Epipe spoke SDPs.

Spoke SDP termination on an IES or VPRN provides the ability to cross-connect traffic entering on a spoke SDP, used for Layer 2 services (VLLs or VPLS), on to an IES or VPRN service. From a logical point of view the spoke SDP entering on a network port is cross-connected to the Layer 3 service as if it had entered using a service SAP. The main exception to this is traffic entering the Layer 3 service using a spoke SDP is handled with network QoS policies instead of access QoS policies.

When a SAP down or SDP binding down status message is received by the PE in which the Ipipe or Ethernet Spoke-SDP is terminated on an IES or VPRN interface, the interface is brought down and all associated routes are withdrawn in a similar way when the Spoke-SDP goes down locally. The same actions are taken when the standby T-LDP status message is received by the IES/VPRN PE.

This feature can be used to provide redundant connectivity to a VPRN or IES from a PE providing a VLL service, as shown in Figure 17: Active/standby VRF using resilient Layer 2 circuits.

Figure 17: Active/standby VRF using resilient Layer 2 circuits



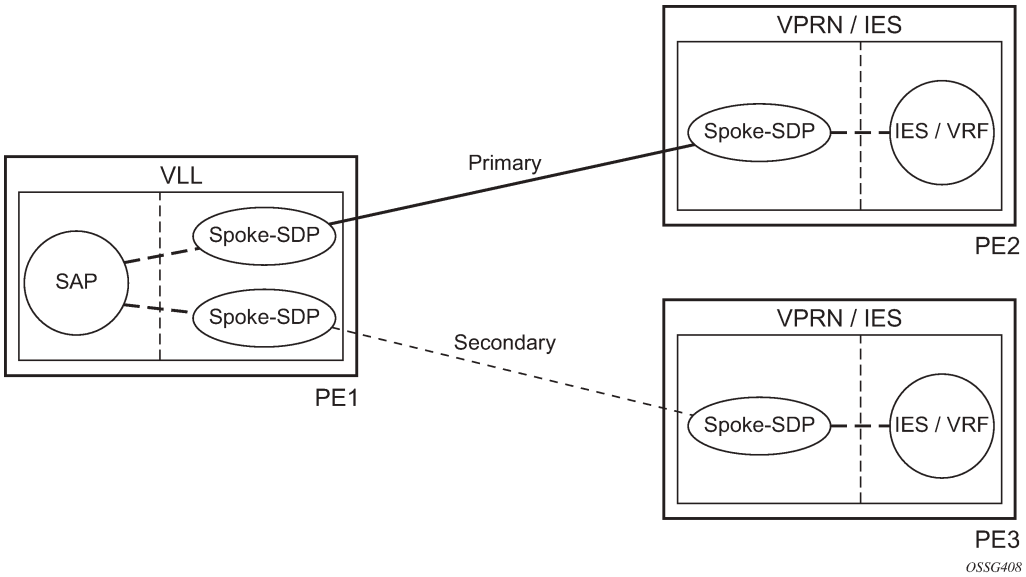
3.2.9.2 Spoke SDP redundancy into IES/VPRN

This feature can be used to provide redundant connectivity to a VPRN or IES from a PE providing a VLL service, as shown in Figure 17: Active/standby VRF using resilient Layer 2 circuits, using either Epipe or Ipipe spoke-SDPs. This feature is supported on the 7450 ESS and 7750 SR only.

In Figure 17: Active/standby VRF using resilient Layer 2 circuits, PE1 terminates two spoke SDPs that are bound to one SAP connected to CE1. PE1 chooses to forward traffic on one of the spoke SDPs (the active spoke-SDP), while blocking traffic on the other spoke SDP (the standby spoke SDP) in the transmit direction. PE2 and PE3 take any spoke SDPs for which PW forwarding standby has been signaled by PE1 to an operationally down state.

The 7450 ESS, 7750 SR, and 7950 XRS routers are expected to fulfill both functions (VLL and VPRN/IES PE), while the 7705 SAR must be able to fulfill the VLL PE function. Figure 18: Spoke SDP redundancy model illustrates the model for spoke SDP redundancy into a VPRN or IES.

Figure 18: Spoke SDP redundancy model



3.2.9.3 Weighted ECMP for spoke-SDPs terminating on IES/VPRN and R-VPLS interfaces

ECMP and weighted ECMP into RSVP-TE and SR-TE LSPs is supported for Ipipe and Epipe spoke SDPs terminating on IP interfaces in an IES or VPRN, or for spoke SDP termination on a routed VPLS. It is also supported for SDPs using LDP over RSVP tunnels. Weighted ECMP is configured under the SDP used by the service, as follows:

```
config
   service
      sdp <sdp-id>
         [no] weighted-ecmp
```

Default: no weighted-ecmp

When a service uses a provisioned SDP on which weighted ECMP is configured, a path is selected based on the configured hash. Paths are then load balanced across LSPs used by an SDP according to normalized LSP load balancing weights. If one or more LSPs in the ECMP set to a specific next hop has no **load-balancing-weight** value configured, then regular ECMP spraying is used.

3.2.10 IP-VPNs

3.2.10.1 Using OSPF in IP-VPNs

Using OSPF as a CE to PE routing protocol allows OSPF that is currently running as the IGP routing protocol to migrate to an IP-VPN backbone without changing the IGP routing protocol, introducing BGP as the CE-PE or relying on static routes for the distribution of routes into the service providers IP-VPN. The following features are supported:

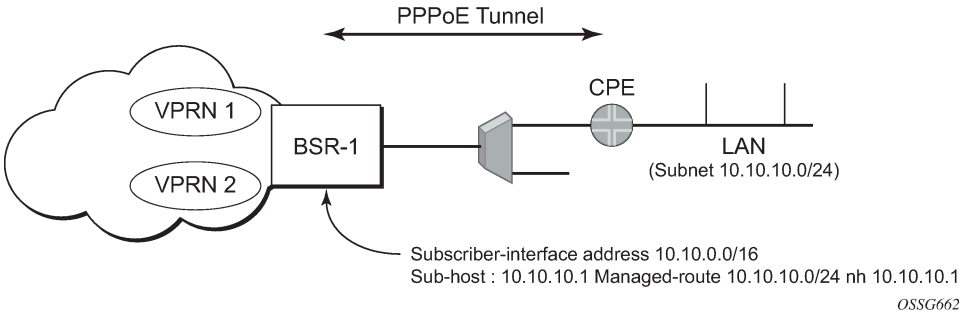
- Advertisement/redistribution of BGP-VPN routes as summary (type 3) LSAs flooded to CE neighbors of the VPRN OSPF instance. This occurs if the OSPF route type (in the OSPF route type BGP extended community attribute carried with the VPN route) is not external (or NSSA) and the locally configured domain-id matches the domain-id carried in the OSPF domain ID BGP extended community attribute carried with the VPN route.
- OSPF sham links; a sham link is a logical PE-to-PE unnumbered point-to-point interface that essentially rides over the PE-to-PE transport tunnel. A sham link can be associated with any area and can therefore appear as an intra-area link to CE routers attached to different PEs in the VPN.

3.2.11 IPCP subnet negotiation

This feature enables negotiation between Broadband Network Gateway (BNG) and customer premises equipment (CPE) so that CPE is allocated to both ip-address and associated subnet.

Some CPEs use the network up-link in PPPoE mode and perform dhcp-server function for all ports on the LAN side. Instead of wasting 1 subnet for p2p uplink, CPEs use allocated subnet for the LAN portion as shown in Figure 19: CPEs network up-link mode.

Figure 19: CPEs network up-link mode



From a BNG perspective, the specific PPPoE host is allocated a subnet (instead of /32) by RADIUS, external dhcp-server, or local-user-db. And locally, the host is associated with managed-route. This managed-route is subset of the subscriber-interface subnet (on a 7450 ESS or 7750 SR), and also, subscriber-host ip-address is from managed-route range. The negotiation between BNG and CPE allows CPE to be allocated both ip-address and associated subnet.

3.2.12 Cflowd for IP-VPNs

The cflowd feature allows service providers to collect IP flow data within the context of a VPRN. This data can used to monitor types and general proportion of traffic traversing an VPRN context. This data can also be shared with the VPN customer to see the types of traffic traversing the VPN and use it for traffic engineering.

This feature should not be used for billing purposes. Existing queue counters are designed for this purpose and provide very accurate per bit accounting records.

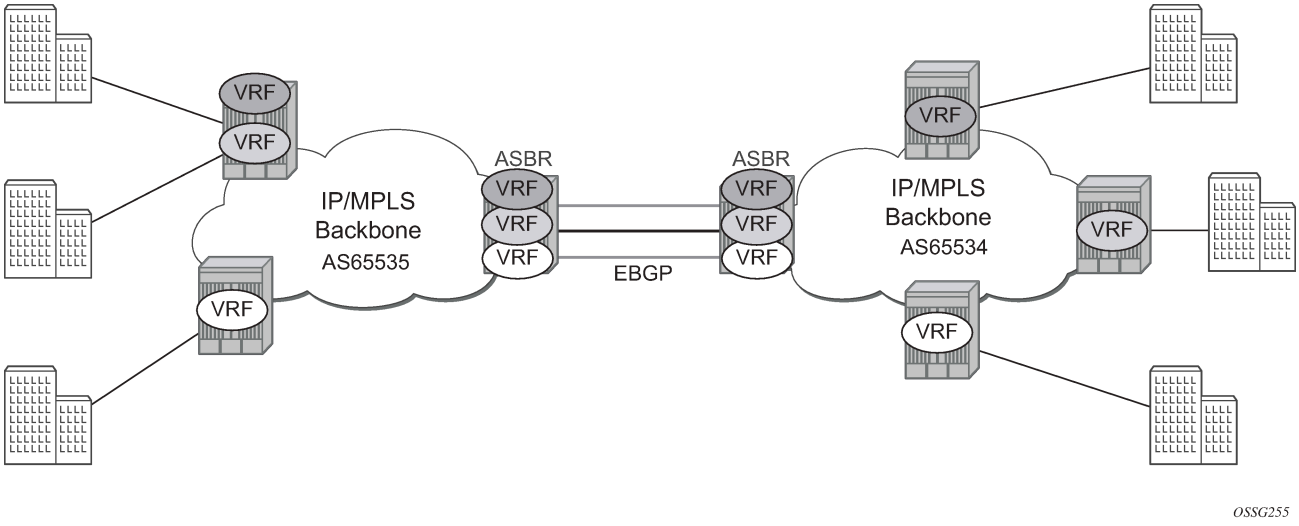
3.2.13 Inter-AS VPRNs

Inter-AS IP-VPN services have been driven by the popularity of IP services and service provider expansion beyond the borders of a single Autonomous System (AS) or the requirement for IP VPN services to cross the AS boundaries of multiple providers. Three options for supporting inter-AS IP-VPNs are described in RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*.

This feature applies to the 7450 ESS and 7750 SR only.

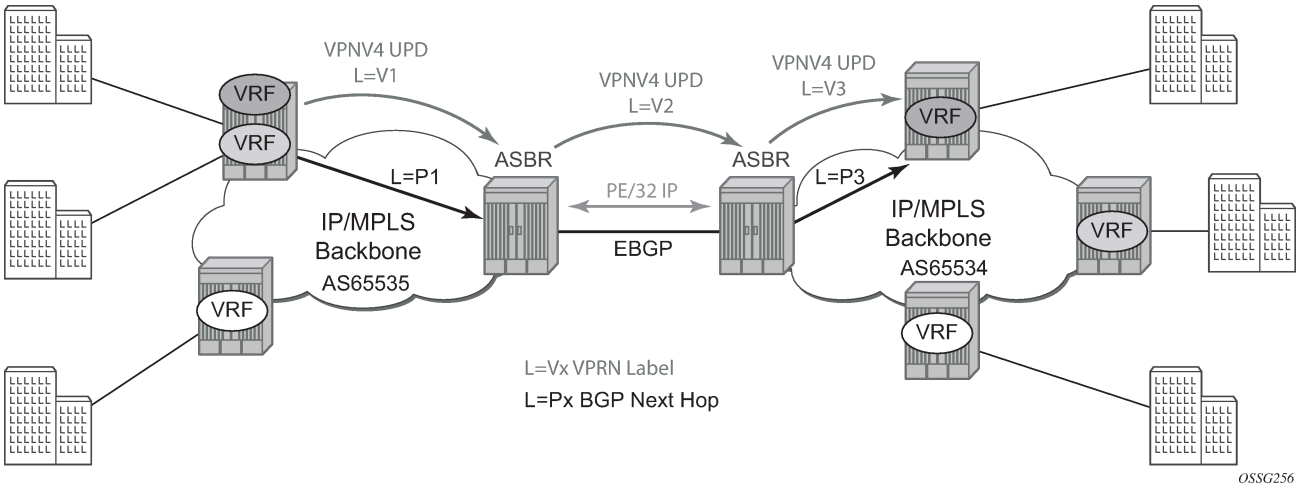
The first option, referred to as Option-A (Figure 20: Inter-AS Option-A: VRF-to-VRF model), is considered inherent in any implementation. This method uses a back-to-back connection between separate VPRN instances in each AS. As a result, each VPRN instance views the inter-AS connection as an external interface to a remote VPRN customer site. The back-to-back VRF connections between the ASBR nodes require individual sub-interfaces, one per VRF.

Figure 20: Inter-AS Option-A: VRF-to-VRF model



The second option, referred to as Option-B (Figure 21: Inter-AS Option-B), relies heavily on the AS Boundary Routers (ASBRs) as the interface between the autonomous systems. This approach enhances the scalability of the EBGP VRF-to-VRF solution by eliminating the need for per-VPRN configuration on the ASBRs. However it requires that the ASBRs provide a control plan and forwarding plane connection between the autonomous systems. The ASBRs are connected to the PE nodes in its local autonomous system using Interior Border Gateway Protocol (IBGP) either directly or through route reflectors. This means the ASBRs receive all the VPRN information and forward these VPRN updates, VPN-IPV4, to all its EBGP peers, ASBRs, using itself as the next-hop. It also changes the label associated with the route. This means the ASBRs must maintain an associate mapping of labels received and labels issued for those routes. The peer ASBRs in turn forward those updates to all local IBGP peers.

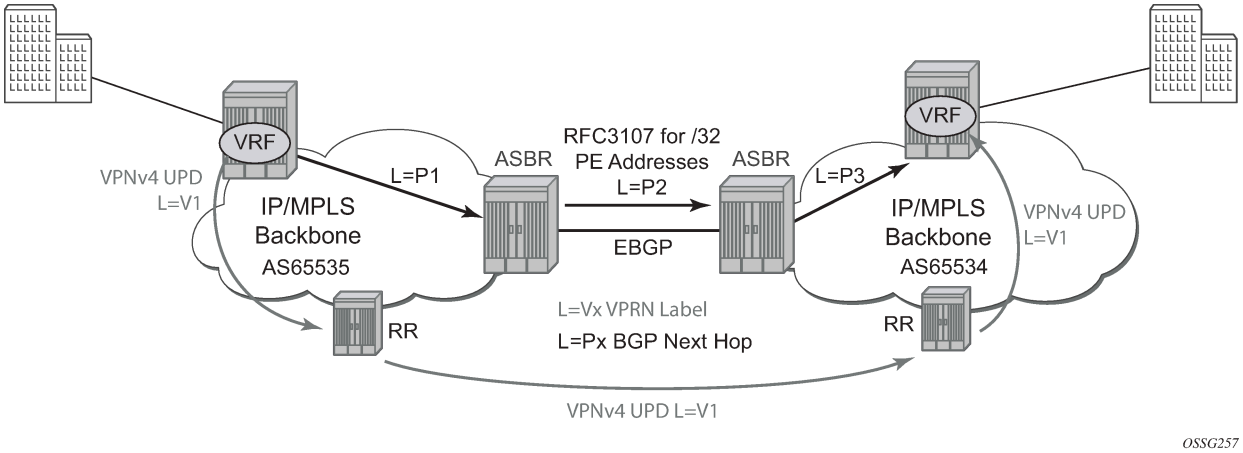
Figure 21: Inter-AS Option-B



This form of inter-AS VPRNs performs all necessary mapping functions and the PE routers do not need to perform any additional functions than in a non-Inter-AS VPRN.

On the 7750 SR, this form of inter-AS VPRNs does not require instances of the VPRN to be created on the ASBR, as in option-A, as a result there is less management overhead. This is also the most common form of Inter-AS VPRNs used between different service providers as all routes advertised between autonomous systems can be controlled by route policies on the ASBRs by the **config>router>bgp>transport-tunnel** CLI command. The third option, referred to as Option-C (Figure 22: Option C example), allows for a higher scale of VPRNs across AS boundaries but also expands the trust model between ASNs. As a result this model is typically used within a single company that may have multiple ASNs for various reasons. This model differs from Option-B, in that in Option-B all direct knowledge of the remote AS is contained and limited to the ASBR. As a result, in option-B the ASBR performs all necessary mapping functions and the PE routers do not need to perform any additional functions than in a non-Inter-AS VPRN.

Figure 22: Option C example



With Option-C, knowledge from the remote AS is distributed throughout the local AS. This distribution allows for higher scalability but also requires all PEs and ASBRs involved in the Inter-AS VPRNs to participate in the exchange of inter-AS routing information.

In Option-C, the ASBRs distribute reachability information for remote PE's system IP addresses only. This is done between the ASBRs by exchanging MP-EBGP labeled routes, using RFC 8277, *Using BGP to Bind MPLS Labels to Address Prefixes*. Either RSVP-TE or LDP LSP can be selected to resolve next-hop for multi-hop EBGP peering by the **config>router>bgp>transport-tunnel** CLI command.

Distribution of VPRN routing information is handled by either direct MP-BGP peering between PEs in the different ASNs or more likely by one or more route reflectors in ASN.

3.2.14 VPRN label security at inter-AS boundary

This feature allows the user to enforce security at an inter-AS boundary and to configure a router, acting in a PE role or in an ASBR role, or both, to accept packets of VPRN prefixes only from direct EBGP neighbors to which it advertised a VPRN label.

3.2.14.1 Feature configuration

To use this feature, network IP interfaces that can have the feature enabled must first be identified. Participating interfaces are identified as having the untrusted state. The router supports a maximum of 15 network interfaces that can participate in this feature.

The following command is used to enable or disable this feature.

config>**router**>**interface**>**untrusted** [**default-forwarding** {**forward** | **drop**}]

Normally, the user applies the **untrusted** command to an inter-AS interface and PIP keeps track of the untrusted status of each interface. In the data path, an inter-AS interface that is flagged by PIP causes the default forwarding to be set to the value of the **default-forwarding** keyword (**forward** or **drop**).

For backward compatibility, **default-forwarding** on the interface is set to the **forward** value. This means that labeled packets are checked in the normal way against the table of programmed ILMs to decide if it should be dropped or forwarded in a GRT, a VRF, or a Layer 2 service context.

If the user sets the **default-forwarding** argument to the **drop** value, all labeled packets received on that interface are dropped. For details, see Data path forwarding behavior.

This feature sets the default behavior for an untrusted interface in the data path and for all ILMs. To allow the data path to provide an exception to the normal way of forwarding handling away from the default for VPRN ILMs, BGP must flag those ILMs to the data path.

The user enables exceptional ILM forwarding behavior, on a per-VPN-family basis, by using the following command:

configure>**router**>**bgp**>**neighbor-trust** [**vpn-ipv4**] [**vpn-ipv6**]

At a high level, BGP tracks each direct EBGP neighbor over an untrusted interface and to which it sent a VPRN prefix label. For each of those VPRN prefixes, BGP programs a bit map in the ILM that indicates, on a per-untrusted interface basis, whether the matching packets must be forwarded or dropped. For details, see CPM behavior.

3.2.14.2 CPM behavior

This feature affects PIP behavior for management of network IP interfaces and in BGP for the resolution of BGP VPN-IPv4 and VPN-IPv6 prefixes.

The following are characteristics of CPM behavior related to PIP and the VPRN label security at inter-AS boundary feature:

- PIP manages the status of an untrusted interface based on the user configuration on the interface, as described in Feature configuration. It programs the interface record in the data path using a 4-bit untrusted interface identification number. A trusted interface has no untrusted record.
- BGP determines the status of trusted or untrusted of an EBGP neighbor by checking the untrusted record provided by PIP for the index of the interface used by the EBGP session to the neighbor.
- BGP only tracks the status of trusted or untrusted for directly connected EBGP neighbors. The neighbor address and the local address must be on the same local subnet.
- BGP includes the neighbor status of trusted or untrusted in the tribe criteria. For example, if a group consists of two untrusted EBGP neighbors and one trusted EBGP neighbor and all three neighbors have the same local-AS, neighbor-AS, and export policy, then the result is two different tribes.

  As a result, if the interface status changes from trusted to untrusted or untrusted to trusted, the EBGP neighbors on that interface bounce.
- When the feature is enabled for a specified VPN family and BGP advertises a label for one or more resolved VPN prefixes to a group of trusted and untrusted EBGP neighbors, it creates a 16-bit map in the ILM record in which it sets the bit position corresponding to the identification number of each untrusted interface used by a EBGP session to a neighbor to which it sent the label.

  A bit in the ILM record bit-map is referred to as the untrusted interface forwarding bit. The bit position corresponding to the identification number of any other untrusted interface is left clear.

  For details on the data path of the ILM bit-map record, see Data path forwarding behavior.
- Because the same label value is advertised for prefixes in the same VRF (label per-VRF mode) and for prefixes with the same next hop (label per-next-hop mode), BGP programs the forwarding bit position in the ILM bit map for both VPN IPv4 and VPN IPv6 prefixes sharing the same label, as long as the feature is enabled for at least one of the two VPN families.
- BGP tracks, on a per-untrusted interface basis, the number of RIB-Out entries to EBGP neighbors that reference a specific VPN label. When that reference transitions from zero to a positive value or from a positive value to zero, the label for the ILM of the VPN prefix is re-downloaded to the IOM with the forwarding bit position in the ILM bit map record updated accordingly (set or unset, respectively).

This feature supports label per-VRF and label per-next-hop modes for the PE role. The feature supports label per-next-hop mode for the ASBR role.

The feature is not supported with label per-prefix mode in a PE role and is not supported in a Carrier Supporting Carrier (CSC) PE role.

3.2.14.3 Data path forwarding behavior

ILM forwarding on a trusted interface behaves as in prior releases and is not changed. The ILM forwarding bit map is ignored and packets are forwarded normally.

ILM forwarding on an untrusted interface follows these rules:

- Only the top-most label in the label stack in a received packet is checked against the next set of rules. The top label can correspond to any one of the following applications:
  - a transport label with a pop or swap operation of static, RSVP-TE, SR-TE, LDP, SR-ISIS, SR-OSPF, or BGP-LU

  - a BGP VPRN inter-AS option B label with a swap operation when the router acts in the ASBR role for VPN routes
  - a service delimiting label for a local VRF when the router acts as a PE in a VPRN service
- The data path checks the bit position in the bit map in the ILM record, when present, that corresponds to the untrusted interface identification number in the interface record and then makes a forwarding decision to drop or forward.

  A decision to forward means that a labeled packet proceeds to the regular ILM processing and its label stack is checked against the table of programmed ILMs to decide if the packet should be:

  - dropped
  - forwarded to CPM
  - forwarded as an MPLS packet
  - forwarded as an IP packet in a GRT or a VRF context
  - forwarded as a packet in a Layer 2 service context
- The following are the processing rules of the ILM:
  - interface **default-forwarding=forward** and ILM bit-map not present ⇒ forward packet
  - interface **default-forwarding=forward** and interface forwarding bit position in the ILM bit-map 1 ⇒ forward packet
  - interface **default-forwarding=forward** and interface forwarding bit position in the ILM bit-map zero ⇒ drop packet
  - interface **default-forwarding=drop** and ILM bit-map not present ⇒ drop packet
  - interface **default-forwarding=drop** and interface forwarding bit position in the ILM bit-map zero ⇒ drop packet
  - interface **default-forwarding=drop** and interface forwarding bit position in the ILM bit-map 1 ⇒ forward packet
- When the EBGP neighbor is not directly connected, BGP does not track that neighbor (see CPM behavior). In this case, the VPRN packet is received with a transport label or without a transport label if implicit-null is enabled in LDP or RSVP-TE for the transport label. Either way, the forwarding decision for the packet is solely dictated by the configuration of the **default-forwarding** value on the incoming interface.
- If the direct EBGP neighbor sends a VPRN packet using the MPLS-over-GRE encapsulation, the data path does not check the interface forwarding bit position in the ILM bit map. In this case, the forwarding decision of the packet is solely dictated by the configuration of the **default-forwarding** value on the incoming interface.

  SR OS EBGP neighbors never use the MPLS-over-GRE encapsulation over an inter-AS link, but third party implementations may do this.

3.2.15 CSC

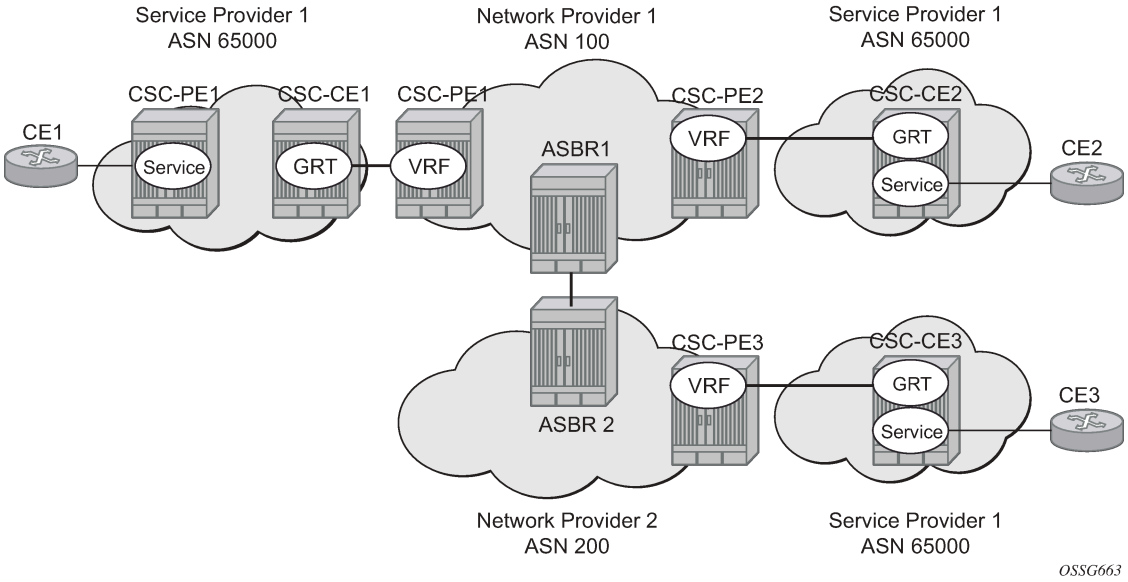
Carrier Supporting Carrier (CSC) is a solution for the 7750 SR and 7950 XRS that allows one service provider (the Customer Carrier) to use the IP VPN service of another service provider (the Super Carrier) for some or all of its backbone transport. RFC 4364 defines a Carrier Supporting Carrier solution for BGP/

MPLS IP VPNs that uses MPLS on the interconnections between the two service providers to provide a scalable and secure solution.

CSC support in SR OS allows a 7750 SR or 7950 XRS to be deployed as any of the following devices shown in Figure 23: Carrier Supporting Carrier reference diagram:

- PE1 (service provider PE)
- CSC-CE1, CSC-CE2 and CSC-CE3 (CE device from the point of view of the backbone service provider)
- CSC-PE1, CSC-PE2 and CSC-PE3 (PE device of the backbone service provider)
- ASBR1 and ASBR2 (ASBR of the backbone service provider)

Figure 23: Carrier Supporting Carrier reference diagram



3.2.15.1 Terminology

CE	customer premises equipment dedicated to one particular business/enterprise
PE	service provider router that connects to a CE to provide a business VPN service to the associated business/enterprise
CSC-CE	an ASBR/peering router that is connected to the CSC-PE of another service provider for purposes of using the associated CsC IP VPN service for backbone transport
CSC-PE	a PE router belonging to the backbone service provider that supports one or more CSC IP VPN services

3.2.15.2 CSC connectivity models

A PE router deployed by a customer service provider to provide Internet access, IP VPNs, or L2 VPNs may connect directly to a CSC-PE device, or it may back haul traffic within its local "site" to the CSC-CE that

provides this direct connection. Here, "site" means a set of routers owned and managed by the customer service provider that can exchange traffic through means other than the CSC service. The function of the CSC service is to provide IP/MPLS reachability between isolated sites.

The CSC-CE is a "CE" from the perspective of the backbone service provider. There may be multiple CSC-CEs at a specific customer service provider site and each one may connect to multiple CSC-PE devices for resiliency/multi-homing purposes.

The CSC-PE is owned and managed by the backbone service provider and provides CSC IP VPN service to connected CSC-CE devices. In many cases, the CSC-PE also supports other services, including regular business IP VPN services. A single CSC-PE may support multiple CSC IP VPN services. Each customer service provider is allocated its own VRF within the CSC-PE; VRFs maintain routing and forwarding separation and allow the use of overlapping IP addresses by different customer service providers.

A backbone service provider may not have the geographic span to connect, with reasonable cost, to every site of a customer service provider. In this case, multiple backbone service providers may coordinate to provide an inter-AS CSC service. Different inter-AS connectivity options are possible, depending on the trust relationships between the different backbone service providers.

The CSC Connectivity Models apply to the 7750 SR and 7950 XRS only.

3.2.15.3 CSC-PE configuration and operation

This section applies to CSC-PE1, CSC-PE2 and CSC-PE3 in Figure 23: Carrier Supporting Carrier reference diagram.

This section applies only to the 7750 SR.

3.2.15.4 CSC interface

From the point of view of the CSC-PE, the IP/MPLS interface between the CSC-PE and a CSC-CE has these characteristics:

1. The CSC interface is associated with one (and only one) VPRN service. Routes with the CSC interface as next-hop are installed only in the routing table of the associated VPRN.
2. The CSC interface supports EBGP or IBGP for exchanging labeled IPv4 routes (RFC 8277). The BGP session may be established between the interface addresses of the two routers or else between a loopback address of the CSC-PE VRF and a loopback address of the CSC-CE. In the latter case, the BGP next-hop is resolved by either a static or OSPFv2 route.
3. An MPLS packet received on a CSC interface is dropped if the top-most label was not advertised over a BGP (RFC 8277) session associated with one of the VPRN's CSC interfaces.
4. The CSC interface supports ingress QoS classification based on 802.1p or MPLS EXP. It is possible to configure a default FC and default profile for the CSC interface.
5. The CSC interface supports QoS (re)marking for egress traffic. Policies to remark 802.1p or MPLS EXP based on forwarding-class and profile are configurable per CSC interface.
6. By associating a port-based egress queue group instance with a CSC interface, the egress traffic can be scheduled/shaped with per-interface, per-forwarding-class granularity.
7. By associating a forwarding-plane based ingress queue group instance with a CSC interface, the ingress traffic can be policed to per-interface, per-forwarding-class granularity.

8. Ingress and egress statistics and accounting are available per CSC interface. The exact set of collected statistics depends on whether a queue-group is associated with the CSC interface, the traffic direction (ingress vs. egress), and the stats mode of the queue-group policers.

An Ethernet port or LAG with a CSC interface can be configured in hybrid mode or network mode. The port or LAG supports null, dot1q or qinq encapsulation. To create a CSC interface on a port or LAG in null mode, the following commands are used:

config>service>vprn>nw-if>port port-id config>service>vprn>nw-if>lag lag-id

To create a CSC interface on a port or LAG in dot1q mode, the following commands are used:

config>service>vprn>nw-if>port port-id:qtag1 config>service>vprn>nw-if>lag port-id:qtag1

To create a CSC interface on a port or LAG in qinq mode, the following commands are used:

config>service>vprn>nw-if>port port-id:qtag1.qtag2 config>service>vprn>nw-if>port port-id:qtag1.* config>service>vprn>nw-if>lag port-id:qtag1.qtag2 config>service>vprn>nw-if>lag port-id:qtag1.*

A CSC interface supports the same capabilities (and supports the same commands) as a base router network interface, except it does not support:

- IPv6
- LDP
- RSVP
- Proxy ARP (local/remote)
- Network domain configuration
- DHCP
- Ethernet CFM
- Unnumbered interfaces

3.2.15.5 QoS

3.2.15.5.1 Egress

Egress traffic on a CSC interface can be shaped and scheduled by associating a port-based egress queue-group instance with the CSC interface. The steps for doing this are summarized below:

Procedure

Step 1. Create an egress queue-group-template.

Step 2. Define one or more queues in the egress queue-group. For each one specify scheduling parameters such as CIR, PIR, CBS and MBS and, if using H-QoS, the parent scheduler.

Step 3. Apply an instance of the egress queue-group template to the network egress context of the Ethernet port with the CSC interface. When doing so, and if applicable, associate an accounting policy or a scheduler policy, or both, with this instance.

Step 4. Create a network QoS policy.

Step 5. In the egress part of the network QoS policy define EXP remarking rules, if necessary.

Step 6. In the egress part of the network QoS policy map a forwarding-class to a queue-id using the port-redirect-group command.

For example:
config>qos>network>egress>fc$ port-redirect-group queue 5

Step 7. Apply the network QoS policy created in step 4 to the CSC interface and specify the name of the egress queue-group created in step 1 and the specific instance defined in step 3.

3.2.15.5.2 Ingress

Ingress traffic on a CSC interface can be policed by associating a forwarding-plane based ingress queue-group instance with the CSC interface. The steps for doing this are summarized below:

Procedure

Step 1. Create an ingress queue-group-template.

Step 2. Define one or more policers in the ingress queue-group. For each one specify parameters such as CIR, PIR, CBS and MBS and, if using H-Pol, the parent arbiter.

Step 3. Apply an instance of the ingress queue-group template to the network ingress context of the forwarding plane with the CSC interface.

When doing so, and if applicable, associate an accounting policy or a policer-control-policy, or both, with this instance.

Step 4. Create a network QoS policy.

Step 5. In the ingress part of the network QoS policy define EXP classification rules, if necessary.

Step 6. In the ingress part of the network QoS policy map a forwarding-class to a policer-id using the fp-redirect-group policer command.

For example:
config>qos>network>ingress>fc$ fp-redirect-group policer 5

Step 7. Apply the network QoS policy created in step 4 to the CSC interface and specify the name of the ingress queue-group created in step 1 and the specific instance defined in step 3.

3.2.15.6 MPLS

BGP-8277 is used as the label distribution protocol on the CSC interface. When BGP in a CSC VPRN needs to distribute a label corresponding to a received VPN-IPv4 route, it takes the label from the global label space. The allocated label is not re-used for any other FEC regardless of the routing instance (base router or VPRN). If a label L is advertised to the BGP peers of CSC VPRN A then a received packet with label L as the top most label is only valid if received on an interface of VPRN A, otherwise the packet is discarded.

To use BGP-8277 as the label distribution protocol on the CSC interface, add the **family label-ipv4** command to the family configuration at the instance, group, or neighbor level. This causes the capability to send and receive labeled-IPv4 routes {AFI=1, SAFI=4} to be negotiated with the CSC-CE peers.

3.2.15.7 CSC VPRN service configuration

To configure a VPRN to support CSC service, the **carrier-carrier-vpn** command must be enabled. The command fails if the VPRN service has any existing SAP or spoke SDP interfaces. A CSC interface can

be added to a VPRN (using the **network-interface** command) only if the **carrier-carrier-vpn** command is enabled.

A VPRN service with the **carrier-carrier-vpn** command may be provisioned to use **auto-bind-tunnel**, configured spoke SDPs, or some combination. All SDP types are supported except for:

- GRE SDPs
- LDP over RSVP-TE tunnel SDPs

Other aspects of VPRN configuration are the same in a CSC model as in a non-CSC model.

3.2.16 Node management using VPRN

There are two basic approaches that can be used to manage a node using a VPRN. In both cases, management traffic is received and sent in the VPRN router instance:

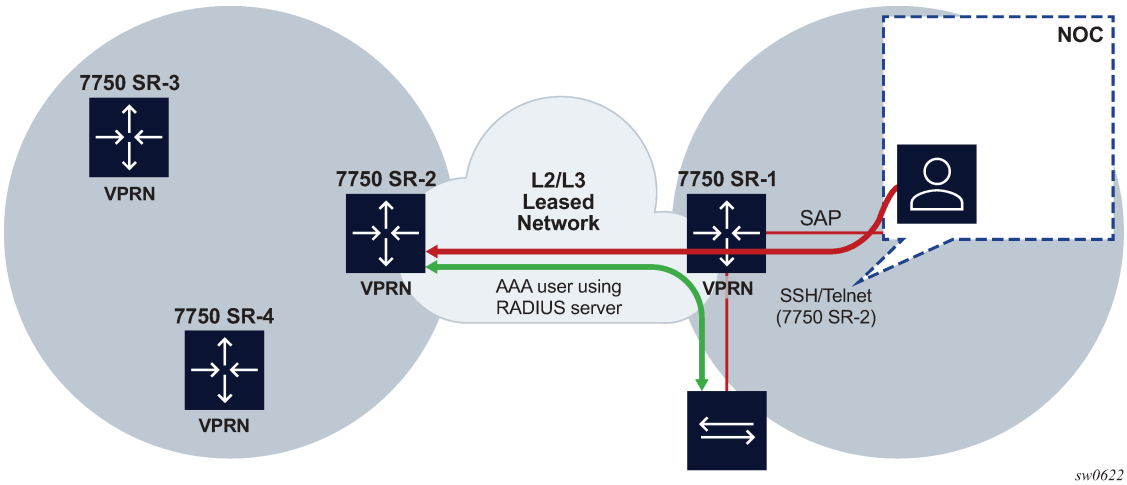
- Management traffic can target the IP address of a VPRN interface
- Management traffic can target the system address in the base router instance (using grt-leaking)

In the first approach, node management can be enabled using the local interface of any VPRN service. A management VPRN is separated from other traffic using an MPLS transport tunnel. This provides IP domain separation and security for the management domain. The management domain supports IPv4 and IPv6 address families and the AAA server is connected to the same VPRN for authentication, authorization, and accounting. The SR OS allows management using a VPRN as long the management packet is destined for a local interface of the VPRN, in addition it allows configuration of the AAA servers within a VPRN; see Figure 24: VRF network example.

In the second approach, node management is achieved using GRT leaking. In this case the management traffic uses an IP address in the Base routing context. See Management via VPRN using GRT leaking for details on this method.

The remainder of this section describes node management using the local VPRN interfaces (non grt-leaking).

Figure 24: VRF network example



3.2.16.1 VPRN management

VPRN management can be enabled by configuring the appropriate management protocol within the VPRN from the **vprn**>**management** context. The following protocols can be enabled in this context:

- FTP
- gRPC
- NETCONF
- SSH
- Telnet
- Telnetv6

SNMP access is not controlled in the **vprn**>**management** context. See section SNMP management.

NTP and PTP access are not controlled in the **vprn**>**management** context. See NTP within a VPRN service and PTP within a VPRN service.

By default, all protocols are disabled. When one of these protocols is enabled, that VRF becomes a management VRF.

Note: An SSH server can only be enabled or disabled in VPRN. All other SSH configurations such as key re-exchange, SSH client/server HMAC/cipher list, and so on, are configured under the system and used globally.

All other gRPC configurations remain global under the **config**>**system**>**grpc** context.

All other NETCONF configurations remain global under the **config**>**system**>**management-interface**>**netconf** context.

The TLS profiles needed for gRPC can be configured under the **config**>**system**>**security**>**tls** context.

3.2.16.2 AAA management

To configure the authentication, authorization, and accounting order for the system, including VPRNs, use the following CLI contexts:

- Classic CLI
  config>**system**>**security**>**password**>**authentication-order**
- MD-CLI
  configure system security user-params authentication-order order

To configure the system local user profile configuration for local user authentication and authorization, including VPRNs, use the following CLI contexts:

- Classic CLI
  config>**system**>**security**>**profile**
- MD-CLI
  configure system security aaa local-profiles profile

To configure AAA servers, use the following CLI contexts:

- Classic CLI
  config>**system**>**security** for system AAA servers

  config>**service**>**vprn**>**aaa**>**remote-servers** for AAA remote servers under the VPRN

- MD-CLI
  configure system security aaa for system AAA servers

  configure service vprn aaa remote-servers for AAA remote servers under the VPRN

When AAA servers are configured using the preceding commands, they are used as follows:

- If servers are configured under the VPRN AAA, only the VPRN AAA servers are used.

  For example, the **authentication-order** command lists the order as local, TACACS+, and RADIUS, while the VPRN only has a RADIUS server configured, and under the system AAA servers both TACACS+ and RADIUS are configured. In this case, if a management session connects to the VPRN and the destination IP matches a local interface in the VPRN, the SR OS tries the local AAA first, and then RADIUS as configured in the VPRN. The SR OS does not try the system AAA servers because there is a AAA server configured in the VPRN.
- If servers are configured under VPRN AAA and the VPRN AAA parameters are configured for in-band, out-of-band, or VPRN, the servers can be used for the VPRN and the system.
- If no AAA servers are configured under VPRN AAA, the system AAA servers are used.

3.2.16.3 SNMP management

The SR OS SNMP agent can be reached via a VPRN interface address when **configure**>**service**>**vprn**>**snmp**>**access** is enabled.

Using an SNMP community defined inside the VPRN context (**configure**>**service**>**vprn**>**snmp**>**community**) or a user associated with an SNMPv3 USM access group defined in the system context (**configure**>**system**>**snmp**>**access**) allows access to a subset of the full SNMP data model. This subset can be seen in the output of **show system security view "vprn-view"**.

Using an SNMP community defined in the system context (**configure**>**system**>**security**>**snmp**>**community**) allows access to the full SNMP data model (unless otherwise restricted used SNMP views).

Alternatively, grt leaking and a Base routing IP address can be used (along with an SNMP community defined at the system context) to get access to the entire SNMP data model (see the **allow-local-management** command).

A network manager using SNMP, cannot discover or fully manage an SR OS router using an SNMP community defined inside the VPRN context. Full SNMP access requires using one of the approaches described above.

SNMP communities configured under a VPRN are associated with the SNMP context "vprn". For example, walking the ifTable (IF-MIB) using the community configured for VPRN 5 returns counters and status for interfaces in VPRN 5 only.

Note: To access Base router ifTable entries in a VPRN, use the community string that is defined in the system context (**config**>**system**>**security**>**snmp**>**community**).
To access VPRN ifTable entries, use the community string that is defined inside that VPRN context (**config**>**service**>**vprn**>**snmp**>**community**).

3.2.16.4 Events and notifications

Syslog, SNMP traps and Netconf notifications are generated via the Event Logging System.

VPRN syslog destinations are defined under the **config**>**service**>**vprn**>**log**>**syslog** context.

SNMP trap destination in a VPRN are defined using the **config**>**service**>**vprn**>**log**>**snmp-trap-group** context.

Events for the whole system can be directed to a destination within the management VPRN by using the **config**>**log**>**services-all-events** command. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR System Management Guide* for details of this command.

3.2.16.5 DNS resolution

DNS default domain name and DNS servers for domain name resolution can be defined within a VPRN in the **config**>**service**>**vprn**>**dns** context.

3.2.17 Traffic leaking to GRT

Traffic leaking to Global Route Table (GRT) for the 7750 SR and 7950 XRS allows service providers to offer VPRN and Internet services to their customers over a single VRF interface. This supports IPv4 and, for the 7750 SR, requires the customer VPRN interfaces to terminate on a minimum of IOM3-XP and IMM hardware.

Packets entering a local VRF interface can have route processing results derived from the VPRN forwarding table or the GRT. The leaking and preferred lookup results are configured on a per VPRN basis. Configuration options can be general (for example, any lookup miss in the VPRN forwarding table can be resolved in the GRT), or specific (for example, specific routes should only be looked up in the GRT and ignored in the VPRN). To provide operational simplicity and improve streamlining, the CLI configuration is contained within the context of the VPRN service.

This feature is enabled within the VPRN service context under **config>service>vprn>grt-lookup**. This is an administrative context and provides the container under which the user can enter all specific commands, except policy definition. Policy definitions remain unchanged but are referenced from this context.

The **enable-grt** command establishes the basic functionality. When it is configured, any lookup miss in the VRF table is resolved in the GRT, if available. By itself, this only provides part of the solution. Packet forwarding within GRT must route packets back to the correct node and to the specific VPRN from which the destination exists. Destination prefixes must be leaked from the VPRN to the GRT through the use of policy. Policies are created under the **config>router>policy-options** hierarchy. By default, the number of prefixes leaked from the VPRN to the GRT is limited to five. The **export-limit** command under the **grt-lookup** hierarchy allows a service provider to override the default, or remove the limit.

When a VPRN forwarding table consists of a default route or an aggregate route, the customer may require the service provider to poke holes in those, or provide more specific route resolution in the GRT. In this case, the service provider may configure a static-route-entry and specify the GRT as the nexthop type.

The lookup result prefers any successful lookup in the GRT that is equal to or more specific than the static route, bypassing any successful lookup in the local VPRN.

This feature and Unicast Reverse Path Forwarding (uRPF) are mutually exclusive. When a VPRN service is configured with either of these functions, the other cannot be enabled. Also, prefixes leaked from any VPRN should never conflict with prefixes leaked from any other VPRN or existing prefixes in the GRT. Prefixes should be globally unique within the service provider network and if these are propagated outside a single provider network, they must be from the public IP space and globally unique. Network Address Translation (NAT) is not supported as part of this feature. The following type of routes are not leaked from VPRN into the Global Routing Table (GRT):

- Aggregate routes
- Blackhole routes
- BGP VPN extranet routes

3.2.17.1 Management via VPRN using GRT leaking

In addition to node management using the IP addresses of VPRN interfaces, see Node management using VPRN , management via a VPRN can also be achieved using IP addresses in the Base routing instance and GRT leaking.

When a management packet arrives on a VPRN, there is a lookup for the destination IP address of the packet. If the destination IP is resolved using VPRN and the corresponding protocol is enabled under VPRN management, then the packet is extracted to CPM.

If the destination IP address is not a VRF IP and GRT leaking is enabled, a second lookup is done in the GRT FIB. If the IP belongs to a local interface in GRT and **allow-local-management** is enabled under **enable-grt**, then the packet is extracted using GRT leaking to the CPM.

3.2.18 Traffic leaking from VPRN to GRT for IPv6

This feature allows IPv6 destination lookups in two distinct routing tables and applies only to the 7750 SR and 7950 XRS. IPv6 packets within a Virtual Private Routed Network (VPRN) service is able to perform a lookup for IPv6 destination against the Global Route Table (GRT) as well as within the local VPRN.

Currently, VPRN to VPRN routing exchange is accomplished through the use of import and export policies based on Route Targets (RTs), the creation of extranets. This new feature allows the use of a single VPRN interface for both corporate VPRN routing and other services (for example, Internet) that are reachable outside the local routing context. This feature takes advantage of the dual lookup capabilities in two separate routing tables in parallel.

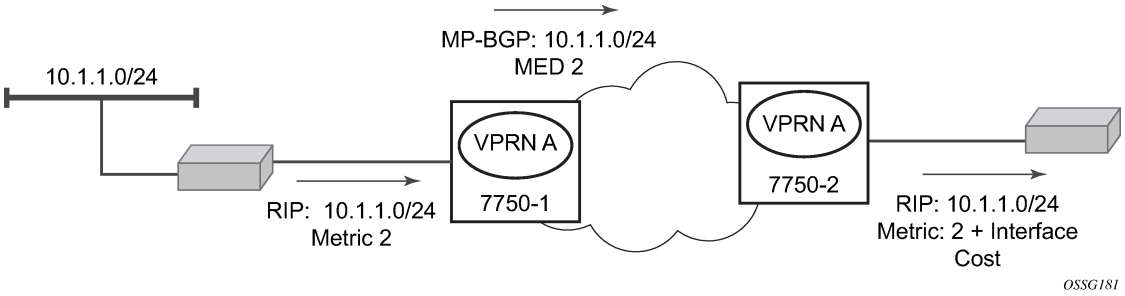
This feature enables IPv6 capability in addition to the existing IPv4 VPRN-to-GRT Route Leaking feature.

3.2.19 RIP metric propagation in VPRNs

When RIP is used as the PE-CE protocol for VPRNs (IP-VPNs), the RIP metric is only used by the local node running RIP with the Customer Equipment (CE). The metric is not used to or encoded into and MP-BGP path attributes exchanged between PE routers.

The RIP metric can also be used to be exchanged between PE routers so if a customer network is dual homed to separate PEs the RIP metric learned from the CE router can be used to choose the best route to the destination subnet. By using the learned RIP metric to set the BGP MED attribute, remote PEs can choose the lowest MED and in turn the PE with the lowest advertised RIP metric as the preferred egress point for the VPRN. Figure 25: RIP metric propagation in VPRNs shows RIP metric propagation in VPRNs.

Figure 25: RIP metric propagation in VPRNs



3.2.20 NTP within a VPRN service

Communication to external NTP clocks through VPRNs is supported in two ways: communication with external servers and peers, and communication with external clients.

Communication with external servers and peers are controlled using the same commands as used for access via base routing (see Network Time Protocol (NTP) in the *7450 ESS, 7750 SR, 7950 XRS, and VSR Basic System Configuration Guide*). Communication with external clients is controlled via the VPRN routing configuration. The support for the external clients can be as unicast or broadcast service. In addition, authentication keys for external clients are configurable on a per-VPRN basis.

Only a single instance of NTP remains in the node that is time sourced to as many as five NTP servers attached to the base or management network.

The NTP show command displays NTP servers and all known clients. Because NTP is UDP-based only, no state is maintained. As a result, the **show** command output only displays when the last message from the client was received.

3.2.21 PTP within a VPRN service

The PTP within a VPRN service provides access to the PTP clock within the 7750 SR through one or more VPRN services. Only one VPRN or the base routing instance may have configured peers, but all may have discovered peers. If needed, a limit on the maximum number of dynamic peers allowed may be configured on a per routing instance basis.

For more detail on PTP see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Basic System Configuration Guide*.

3.2.22 VPN route label allocation

The method used for allocating a label value to an originated VPN-IP route (exported from a VPRN) depends on the configuration of the VPRN service and its VRF export policies. SR OS supports three 3 label modes:

- label per VRF
- label per next hop (LPN)
- label per prefix (LPP)

Label per VRF is the label allocation default. It is used when the label mode is configured as VRF (or not configured) and the VRF export policies do not apply an advertise-label per-prefix action. All routes exported from the VPRN with the per-VRF label have the same label value. When the PE receives a terminating MPLS packet with a per-VRF label, the label value selects the VRF context in which to perform a forwarding table lookup and this lookup determines the outgoing interface (or set of interfaces if ECMP applies).

Label per next hop is used when the exported route is not a local or aggregate route, the label mode is configured as next-hop, and the VRF export policies do not apply an advertise-label per-prefix override. It is also used when an inactive (backup path) BGP route is exported by effect of the export-inactive-bgp command if there is no advertise-label per-prefix override. All LPN-exported routes with the same primary next hop have the same per-next-hop label value. When the PE receives a terminating MPLS packet with a per-next-hop label, the label lookup selects the outgoing interface for forwarding, without any FIB lookup that may cause problems with overlapping prefixes. LPN does not support ECMP, BGP fast reroute, QPPB, or policy accounting features that may otherwise apply.

Label per-prefix is used when a qualifying IP route is exported by matching a VRF export policy action with advertise-label per-prefix. Any IPv4 or IPv6 route that is not a local route, aggregate route, BGP-VPN route, or GRT lookup static route qualifies. With LPP, every prefix is associated with its own unique label value that does not change while the route is present in the route table. When the PE receives a terminating MPLS packet with a per-prefix label value, the packet is forwarded as if the FIB lookup found only the matching prefix route and not any of the more specific prefix routes that would normally be selected. LPP supports ECMP, QPPB, and policy accounting as part of the egress forwarding decision. It does not support BGP fast reroute or BGP sticky ECMP.

The following points summarize the logic that determines the label allocation method for an exported route:

- If the IP route is LOCAL, AGGREGATE, or BGP-VPN always use the VRF label.
- If the IP route is accepted by a VRF export policy with the advertise-label per-prefix action, use LPP.
- If the IP (BGP) route is exported by the export-inactive-bgp command (VPRN best external), use LPN.
- If the IP route is exported by a VPRN configured for label-mode next-hop, use LPN.
- Else, use the per-VRF label.

3.2.22.1 Configuring the service label mode

To change the service label mode of the VPRN for the 7750 SR, the **config>service>vprn>label-mode** {**vrf** | **next-hop**} command is used:

The default mode (if the command is not present in the VPRN configuration) is vrf meaning distribution of one service label for all routes of the VPRN. When a VPRN X is configured with the label-mode next-hop command the service label that it distributes with an IPv4 or IPv6 route that it exports depends on the type of route as summarized in Table 9: Service labels distributed in service label per next hop mode.

Table 9: Service labels distributed in service label per next hop mode

Route type	Distributed service label
remote route with IP A (associated with a SAP) as resolved next-hop	platform-wide unique label allocated to next-hop A
remote route with IP B (associated with a spoke SDP) as resolved next-hop	platform-wide unique label allocated to next-hop B

Route type	Distributed service label
local route	platform-wide unique label allocated to VPRN X
aggregate route	platform-wide unique label allocated to VPRN X
ECMP route	platform-wide unique label allocated to next-hop A (the lowest next-hop address in the ECMP set)
BGP route with a backup next-hop (BGP FRR)	platform-wide unique label allocated to next-hop A (the lowest next-hop address of the primary next-hops)

A change to the label-mode of a VPRN requires the VPRN to first be shutdown.

3.2.22.2 Restrictions and usage notes

The service label per next-hop mode has the following restrictions (applies only to the 7750 SR):

- **ECMP**

  The VPRN label mode should be set to VRF if distribution of traffic across the multiple PE-CE next-hop interfaces of an ECMP route is needed.

- **hub and spoke VPN**

  The VPRN label mode should not be set to next-hop if the operator does not want the hub-connected CE to be involved in the forwarding of spoke-to-spoke traffic.

- **BGP next-hop indirection**

  BGP next-hop indirection has no benefit in service label per next-hop mode. When the resolved next-hop interface of a BGP next-hop changes all of the affected BGP routes must be re-advertised to VPRN peers with the new service label corresponding to the new resolved next-hop.

- **BGP anycast**

  When a PE failure results in redirection of MPLS packets to the other PE in a dual-homed pair, the service label mode is forced to VRF, for example, FIB lookup determines the next-hop even if the label mode of the VPRN is configured as next-hop.

- **U-turn routing**

  U-turn routing is effectively disabled by service-label per next-hop.

- **Carrier Supporting Carrier**

  The label-mode configuration of a VPRN with CSC interfaces is ignored for BGP-8277 routes learned from connected CSC-CE devices.

3.2.23 VPRN Support for BGP FlowSpec

When a VPRN BGP instance receives an IPv4 or IPv6 flow route, and that route is valid/best, the system attempts to construct an IPv4 or IPv6 filter entry from the NLRI contents and the actions encoded in the

UPDATE message. If the attempt is successful, the filter entry is added to the system-created "fSpec-*n*" IPv4 or IPv6 embedded filter, where *n* is the service-id of the VPRN. These embedded filters may be inserted into configured IPv4 and IPv6 filter policies that are applied to ingress traffic on a selected set of the VPRN's IP interfaces. These interfaces can include SAP and spoke SDP interfaces, but not CsC network interfaces.

When FlowSpec rules are embedded into a user-defined filter policy, the insertion point of the rules is configurable through the **offset** parameter of the **embed-filter** command. The sum of the **ip-filter-max-size** and **offset** must not exceed the maximum filter **entry-id** range.

3.2.24 MPLS entropy label and hash label

The router supports both the MPLS entropy label (RFC 6790) and the Flow Aware Transport label, known as the hash label (RFC 6391). LSR nodes in a network can load-balance labeled packets in a more granular way than by hashing on the standard label stack. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide* for more information.

The entropy label is supported for VPRN services, as well as Epipe and Ipipe spoke-SDP termination on VPRN interfaces. To configure insertion of the entropy label, use the **entropy-label** command in the **vprn** context or **spoke-sdp** context of an interface.

The hash label is also supported for Epipe and Ipipe spoke-SDP termination on VPRN and VPRN services bound to any MPLS-type encapsulated SDP, as well as to a VPRN service using **auto-bind-tunnel** with the **resolution-filter** configured as any MPLS tunnel type. Configure it using the **hash-label** command in the **vprn**, **vprn**>**spoke-sdp**, and **vprn**>**if**>**spoke-sdp** contexts.

Either the hash label or the entropy label can be configured on one object, but not both.

3.2.25 LSP tagging for BGP next hops or prefixes and BGP-LU

It is possible to constrain the tunnels used by the system for resolution of BGP next-hops or prefixes and BGP labeled unicast routes using LSP administrative tags. See "LSP Tagging and Auto-Bind Using Tag Information" section of the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide* for more information.

3.2.26 Route leaking from the global route table to VPRN instances

The Global Route Table (GRT) to VPRN route leaking feature allows actual routes from the global route table to be exported into specific VPRN instances allowing those routes to be used for forwarding as well as being re-advertised within the VPRN context.

There are two stages for route leaking. The first stage requires the configuration of a set of **leak-export** route policies that identifies which GRT routes are subject to being exported into VPRN services. The **leak-export** command in the **configure>router** context is used to configure between one and four route policies. The GRT routes must match the policy entries configured with **action accept**. In addition, the **leak-export-limit** command is used to specify the maximum number of GRT routes that can be included in the GRT leak pool.

The second stage requires the configuration of an **import-grt** policy that specifies which routes within the GRT leak pool that are leaked into the associated VPRN instances route table. The **import-grt** command in the **config>service>vprn>grt-lookup** context is used and accepts one route policy. In order for the GRT route to be leaked into the local VPRN, the route must match a policy entry with **action accept**.

If a GRT route passes both stages, it is added into the VPRN route table which it to be used for IP forwarding as well as re-advertisement within other routing protocols within the VPRN context.

Both IPv4 and IPv6 routes can be leaked using this process from the GRT into one or more VPRN instances. The GRT route types that can be leaked using this process are:

- RIP, OSPF, and IS-IS routes
- Direct routes
- Static routes

3.2.27 Class-based forwarding of VPN-v4/v6 prefixes over RSVP-TE or SR-TE LSPs

This feature enables class-based forwarding (CBF) with ECMP of BGP VPN-v4/v6 prefixes that are resolved using RSVP-TE or SR-TE as **auto-bind-tunnel**.

3.2.27.1 Feature configuration

To configure this feature:

- Enable resolution to RSVP-TE or SR-TE tunnels in the **auto-bind-tunnel** context.
- Enable ECMP in the **auto-bind-tunnel** context.
- Enable class-forwarding in the **vprn** context.
- Define at least one class forwarding policy in the **mpls** context, the FC to sets associations and the LSP to (policy, set) associations.

The SR OS CBF implementation supports spraying of packets over a maximum of six forwarding sets of ECMP LSPs only when the system profile is **profile-b** that is supported on an FP4 or later-based CPM. In any other case, the maximum number of forwarding sets of ECMP LSPs is four.

```
config
      router
            [no] mpls
                  class-forwarding-policy policy-name create
                  fc be forwarding-set set-id <1..6>
                  fc l2 forwarding-set set-id <1..6>
                  fc af forwarding-set set-id <1..6>
                  fc l1 forwarding-set set-id <1..6>
                  fc h2 forwarding-set set-id <1..6>
                  fc ef forwarding-set set-id <1..6>
                  fc h1 forwarding-set set-id <1..6>
                  fc nc forwarding-set set-id <1..6>
                  [no] default-set set-id <1..6>
```

All FCs are mapped to set 1 as soon as the policy is created. The user can make changes to the mapping of FCs as required. An FC that is not added to the class-forwarding policy, is always mapped to set 1. At most, an FC can be mapped to a single forwarding set. One or more FCs can be mapped to the same set. The user can indicate the initial default set by including the *default-set* option.

The default forwarding set is used to forward packets of any FC in cases where all LSPs of the forwarding set the FC maps to become operationally DOWN. The router uses the user-configured default set as the initial default set. Otherwise, the router selects the lowest numbered set as the default forwarding set in a class-forwarding policy. When the last LSP in a default forwarding set goes into an operationally DOWN state, the router designates the next lowest-numbered set as the new default forwarding set.

A mapping to a class-forwarding policy and set is added to the existing CBF configuration of an RSVP-TE or SR-TE LSP or to an LSP template. The following commands perform this function:

config>router>mpls>lsp>class-forwarding forward-set policy *policy-name* **set** *set-id*

config>router>mpls>lsp-template>class-forwarding forwarding-set policy *policy-name* **set** *set-id*

An MPLS LSP only maps to a single class-forwarding policy and forwarding set. Multiple LSPs can map to the same policy and set. If they form an ECMP set, from the IGP shortcut perspective, packets of the FCs mapped to this set are sprayed over these LSPs based on a modulo operation of the output of the hash routine on the headers of the packet and the number of LSPs in the set.

3.2.27.2 Feature behavior

When a VPN-v4/v6 prefix is resolved, the default behavior of the data path is to spray the packets over the entire ECMP set using a modulo operation of the number of resolved next hops in the ECMP set and the output of the hash on the packet header fields. With class-based forwarding enabled, the FC of the packet, is used to look up the forwarding set ID. Then, a modulo operation is performed on the tunnel next hops of this set ID only, to spray packets of this FC. The data path concurrently implements ECMP within the tunnels of each set ID.

The CBF information of the LSPs forming the ECMP set is checked for consistency before programming. If more than a single class-forwarding policy exists, the set is considered inconsistent from a CBF perspective and no CBF information is programmed in the data-path and regular ECMP occurs.

Also, regardless of the CBF consistency check, the system programs the data-path with the full ECMP set.

The following describes the fall-back behavior in data path of the CBF feature.

An FC, for which all LSPs in the forwarding set are operationally DOWN, has its packets forwarded over the default forwarding set. The default forwarding set is either the initial default forwarding set configured by the user or the lowest numbered set in the class-forwarding policy that has one or more LSPs in the operationally UP state. If the initial or subsequently elected default forwarding set has all its LSPs operationally DOWN, the next lower numbered forwarding set, which has at least one LSP in the operationally UP state, is elected as the default forwarding set.

If all LSPs of all forwarding sets become operationally DOWN, the router resumes regular ECMP spraying on the remaining LSPs in the full ECMP set.

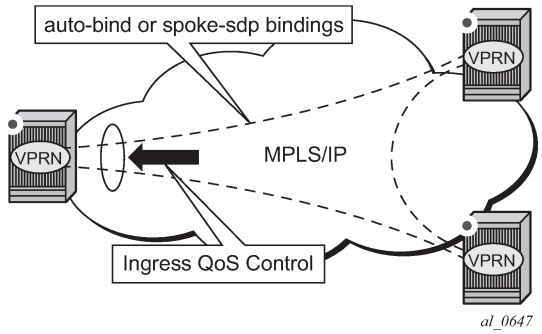
Whenever the first LSP in a forwarding set becomes operationally UP, the router triggers the re-election of the default set and selects this set as the new default set, if it is the initial default set, otherwise, it selects the lowest numbered set.

SR OS implements a hierarchical ECMP architecture for BGP prefixes. The first level is the ECMP at the VPRN Level between different BGP next hop, and the second level is ECMP at the auto-bind-tunnel level, having the same next hop. This CBF feature is applied at the auto-bind-tunnel level. Weighted ECMP and the CBF feature are mutually exclusive on a per-BGP next-hop basis. When both are configured, Weighted ECMP takes the preference. CPM-originated packets on the router, including control plane and OAM packets, are forwarded over a single LSP from the set of LSPs that the packet's FC is mapped to, as per the CBF configuration.

3.3 QoS on ingress bindings

Traffic is tunneled between VPRN service instances on different PEs over service tunnels bound to MPLS LSPs or GRE tunnels. The binding of the service tunnels to the underlying transport is achieved either automatically (using the **auto-bind-tunnel** command) or statically (using the **spoke-sdp** command; not that under the VPRN IP interface). QoS control can be applied to the service tunnels for traffic ingressing into a VPRN service; see Figure 26: Ingress QoS control on VPRN bindings.

Figure 26: Ingress QoS control on VPRN bindings



An ingress queue group must be configured and applied to the ingress network FP where the traffic is received for the VPRN. All traffic received on that FP for any binding in the VPRN (either automatically or statically configured) which is redirected to a policer in the FP queue group (using **fp-redirect-group** in the network QoS policy) is controlled by that policer. As a result, the traffic from all such bindings is treated as a single entity (per forwarding class) with regard to ingress QoS control. Any **fp-redirect-group mcast-policer, broadcast-policer** or **unknown-policer** commands in the network QoS policy are ignored for this traffic (IP multicast traffic would use the ingress network queues or queue group related to the network interface).

Ingress classification is based on the configuration of the ingress section of the specified network QoS policy, noting that the dot1p and exp classification is based on the outer Ethernet header and MPLS label whereas the DSCP applies to the outer IP header if the tunnel encapsulation is GRE, or the DSCP in the first IP header in the payload if **ler-use-dscp** is enabled in the ingress section of the referenced network QoS policy.

Ingress bandwidth control does not take into account the outer Ethernet header, the MPLS labels/control word or GRE headers, or the FCS of the incoming frame.

The following command configures the association of the network QoS policy and the FP queue group and instance within the network ingress of a VPRN:

```
configure
    vprn
        network
            ingress
                qos <network-policy-id> fp-redirect-group <queue-group-name>
                                        instance <instance-id>
```

When this command is configured, it overrides the QoS applied to the related network interfaces for unicast traffic arriving on bindings in that VPRN. The IP and IPv6 criteria statements are not supported in the applied network QoS policy.

This is supported for all available transport tunnel types and is independent of the label mode (**vrf** or **next-hop**) used within the VPRN. It is also supported for Carrier-Supporting-Carrier VPRNs.

The ingress network interfaces on which the traffic is received must be on FP2- and higher-based hardware.

3.4 Multicast in IP-VPN applications

This section and its subsections focuses on Multicast in IP VPN functionality. See the *7450 ESS, 7750 SR, 7950 XRS, and VSR Multicast Routing Protocols Guide* for information about multicast protocols.

Applications for this feature include enterprise customer implementing a VPRN solution for their WAN networking needs, customer applications including stock-ticker information, financial institutions for stock and other types of trading data and video delivery systems.

Implementation of multicast in IP VPNs entails the support and separation of the providers core multicast domain from the various customer multicast domains and the various customer multicast domains from each other.

Figure 27: Multicast in IP-VPN applications

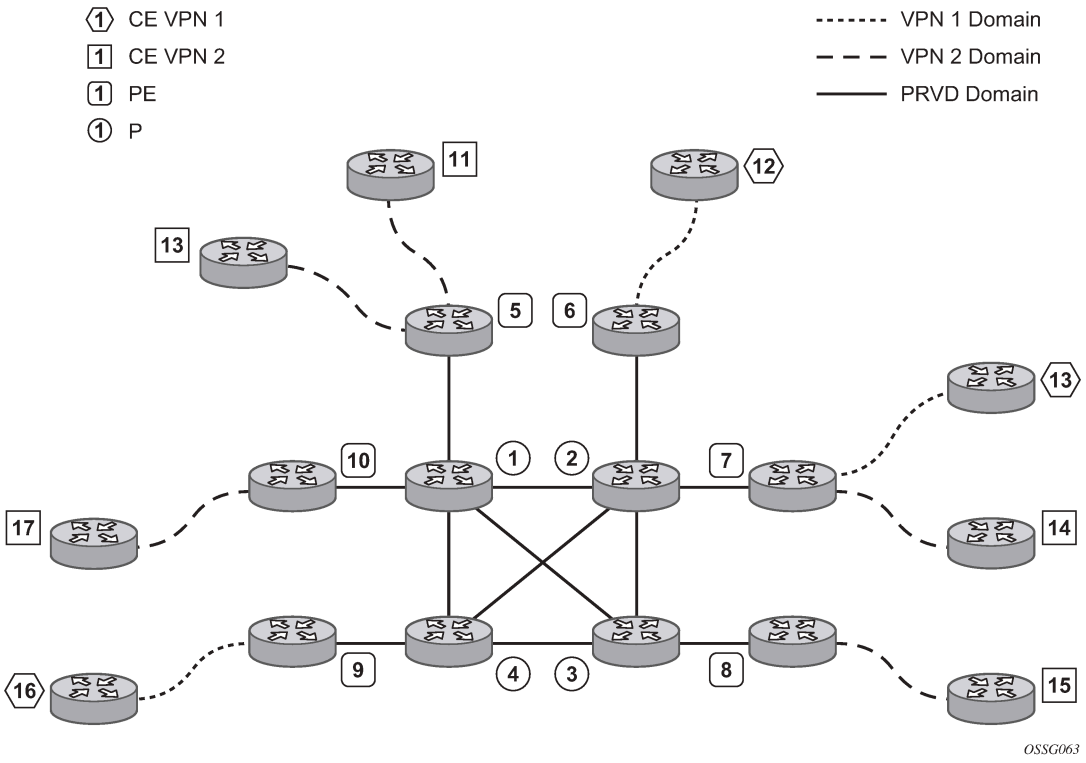


Figure 27: Multicast in IP-VPN applications depicts an example of multicast in an IP-VPN application. The provider's domain encompasses the core routers (1 through 4) and the edge routers (5 through 10). The various IP-VPN customers each have their own multicast domain, VPN-1 (CE routers 12, 13 and 16) and VPN-2 (CE Routers 11, 14, 15, 17 and 18). Multicast in this VPRN example, the VPN-1 data generated by the customer behind router 16 is multicast only by PE 9 to PE routers 6 and 7 for delivery to CE routers 12

and 13 respectively. Data generated for VPN-2 generated by the customer behind router 15 is forwarded by PE 8 to PE routers 5, 7 and 10 for delivery to CE routers 18, 11, 14 and 17 respectively.

The demarcation of these domains is in the PE's (routers 5 through 10). The PE router participates in both the customer multicast domain and the provider's multicast domain. The customer's CEs are limited to a multicast adjacency with the multicast instance on the PE specifically created to support that specific customer's IP-VPN. This way, customers are isolated from the provider's core multicast domain and other customer multicast domains while the provider's core routers only participate in the provider's multicast domain and are isolated from all customers' multicast domains.

The PE for a specific customer's multicast domain becomes adjacent to the CE routers attached to that PE and to all other PE's that participate in the IP-VPN (or customer) multicast domain. This is achieved by the PE who encapsulates the customer multicast control data and multicast streams inside the provider's multicast packets. These encapsulated packets are forwarded only to the PE nodes that are attached to the same customer's edge routers as the originating stream and are part of the same customer VPRN. This prunes the distribution of the multicast control and data traffic to the PEs that participate in the customer's multicast domain. The Rosen draft refers to this as the default multicast domain for this multicast domain; the multicast domain is associated with a unique multicast group address within the provider's network.

3.4.1 Use of data MDTs

Using the above method, all multicast data offered by a specific CE is always delivered to all other CEs that are part of the same multicast. It is possible that a number of CEs do not require the delivery of a particular multicast stream because they have no downstream receivers for a specific multicast group. At low traffic volumes, the impact of this is limited. However, at high data rates this could be optimized by devising a mechanism to prune PEs from the distribution tree that although forming part of the customer multicast have no need to deliver a specific multicast stream to the CE attached to them. To facilitate this optimization, the Rosen draft specifies the use of data MDTs. These data MDTs are signaled after the bandwidth for a specific SG exceeds the configurable threshold.

When a PE detects it is transmitting data for the SG in excess of this threshold, it sends an MDT join TLV (at 60 second intervals) over the default MDT to all PEs. All PEs that require the SG specified in the MDT join TLV join the data MDT that is used by the transmitting PE to send the specific SG. PEs that do not require the SG do not join the data MDT, therefore pruning the multicast distribution tree to just the PEs requiring the SG. After providing sufficient time for all PEs to join the data MDT, the transmitting PE switches the specific multicast stream to the data MDT.

PEs that do not require the SG to be delivered, keep state to allow them to join the data MDT as required.

When the bandwidth requirement no longer exceeds the threshold, the PE stops announcing the MDT join TLV. At this point the PEs using the data MDT leave this group and transmission resumes over the default MDT.

Sampling to check if an s,g has exceeded the threshold occurs every ten seconds. If the rate has exceeded the configured rate in that sample period then the data MDT is created. If during that period the transmission rate has not exceeded the configured threshold then the data MDT is not created. If the data MDT is active and the transmission rate in the last sample period has not exceeded the configured rate then the data MDT is torn down and the multicast stream resumes transmission over the default MDT.

3.4.2 Multicast protocols supported in the provider network

When MVPN auto-discovery is disabled, PIM-SM can be used for I-PMSI, and PIM-SSM or PIM-SM (Draft-Rosen Data MDT) can be used for S-PMSI; When MVPN S-PMSI auto-discovery is enabled, both PIM-SM and PIM SSM can be used for I-PMSI, and PIM-SSM can be used for S-PMSI. In the customer network, both PIM-SM and PIM-SSM are supported.

An MVPN is defined by two sets of sites: sender sites set and receiver sites set, with the following properties:

- Hosts within the sender sites set could originate multicast traffic for receivers in the receiver sites set.
- Receivers not in the receiver sites set should not be able to receive this traffic.
- Hosts within the receiver sites set could receive multicast traffic originated by any host in the sender sites set.
- Hosts within the receiver sites set should not be able to receive multicast traffic originated by any host that is not in the sender sites set.

A site could be both in the sender sites set and receiver sites set, which implies that hosts within such a site could both originate and receive multicast traffic. An extreme case is when the sender sites set is the same as the receiver sites set, in which case all sites could originate and receive multicast traffic from each other.

Sites within a specific MVPN may be either within the same, or in different organizations, which implies that an MVPN can be either an intranet or an extranet. A site may be in more than one MVPN, which implies that MVPNs may overlap. Not all sites of a specific MVPN have to be connected to the same service provider, which implies that an MVPN can span multiple service providers.

Another way to look at MVPN is to say that an MVPN is defined by a set of administrative policies. Such policies determine both sender sites set and receiver site set. Such policies are established by MVPN customers, but implemented by MVPN service providers using the existing BGP/MPLS VPN mechanisms, such as route targets, with extensions, as necessary.

3.4.3 MVPN membership auto-discovery using BGP

BGP-based auto-discovery is performed by a multicast VPN address family. Any PE that attaches to an MVPN must issue a BGP update message containing an NLRI in this address family, along with a specific set of attributes.

The PE router uses route targets to specify MVPN route import and export. The route target may be the same as the one used for the corresponding unicast VPN, or it may be different. The PE router can specify separate import route targets for sender sites and receiver sites for a specific MVPN.

The route distinguisher (RD) that is used for the corresponding unicast VPN can also be used for the MVPN.

When BGP auto-discovery is enabled, PIM peering on the I-PMSI is disabled, so no PIM hellos are sent on the I-PMSI. C-trees to P-tunnels bindings are also discovered using BGP S-PMSI AD routes, instead of PIM join TLVs. Configure PIM join TLVs when **c-mcast-signaling** is set to **pim** in the **config>service>vprn>mvpn>provider-tunnel>selective>auto-discovery-disable** context.

Table 10: Supported configuration combinations and Table 11: Supported configuration combinations describe the supported configuration combinations. If the CLI combination is not allowed, the system returns an error message. If the CLI command is marked as "ignored" in the table, the configuration is not blocked, but its value is ignored by the software.

Table 10: Supported configuration combinations

Auto-discovery	Inclusive PIM SSM	Action
Yes	Yes	Allowed
MDT-SAFI	Yes	Allowed
No	Yes	Not Allowed
Yes or No	No	Allowed
MDT-SAFI	No	Ignored
MDT-SAFI	No (RSVP and MLDP)	Not Allowed

Table 11: Supported configuration combinations

Auto-discovery	C-mcast-signaling	s-PMSI auto-discovery	Action
Yes	BGP	Ignored	Allowed
Yes	PIM	Yes	Allowed
Yes	PIM	No	Allowed
No	BGP	Ignored	Not Allowed
No	PIM	Ignored	Allowed
MDT-SAFI	Ignored (PIM behavior)	Ignored ("No" behavior)	Allowed

For example, if **auto-discovery** is disabled, the **c-mcast-signaling bgp** command fails with an error message stating:

C-multicast signaling in BGP requires auto-discovery to be enabled

If **c-mcast-signaling** is set to **bgp** then **no auto-discovery** fails with an error message stating

C-multicast signaling in BGP requires auto-discovery to be enabled

When **c-mcast-signaling** is set to **bgp**, S-PMSI A-D is always enabled (its configuration is ignored);

When **auto-discovery** is disabled, S-PMSI A-D is always disabled (its configuration is ignored).

When **auto-discovery** is enabled and **c-multicast-signaling** is set to **pim**, the S-PMSI A-D configuration value is used.

mdt-safi uses **pim c-mcast-signaling** and **s-pmsi-signaling** regardless of what is configured. A **c-mcast-signaling** or **s-pmsi-signaling** configuration is ignored, but both **pim** and **bgp** values are allowed.

mdt-safi is only applicable to PIM-SSM I-PMSI. PIM-SM (ASM) I-PMSI is configurable but is ignored. RSVP and MLDP I-PMSI are not allowed.

MVPN implementation based on RFC 6037, *Cisco Systems' Solution for Multicast in MPLS/BGP IP VPNs* can support membership auto-discovery using BGP MDT-SAFI. A CLI option is provided per MVPN

instance to enable auto-discovery either using BGP MDT-SAFI or NG-MVPN. Only PIM-MDT is supported with BGP MDT-SAFI method.

3.4.4 PE-PE transmission of C-multicast routing using BGP

MVPN c-multicast routing information is exchanged between PEs by using c-multicast routes that are carried using MCAST-VPN NLRI.

3.4.5 VRF route import extended community

VRF route import is an IP address-specific extended community, of an extended type, and is transitive across AS boundaries (RFC 4360, *BGP Extended Communities Attribute*).

To support MVPN, in addition to the import/export route target extended communities used by the unicast routing, each VRF on a PE must have an import route target extended community that controls imports of C-multicast routes into a particular VRF.

The c-multicast import RT uniquely identifies a VRF, and is constructed as follows:

- The Global Administrator field of the c-multicast import RT must be set to an IP address of the PE. This address should be common for all the VRFs on the PE (this address may be the PE's loopback address).
- The Local Administrator field of the c-multicast import RT associated with a specific VRF contains a 2 octets long number that uniquely identifies that VRF within the PE that contains the VRF.

A PE that has sites of a specific MVPN connected to it communicates the value of the c-multicast import RT associated with the VRF of that MVPN on the PE to all other PEs that have sites of that MVPN. To accomplish this, a PE that originates a (unicast) route to VPN-IP addresses includes in the BGP updates message that carries this route the VRF route import extended community that has the value of the c-multicast import RT of the VRF associated with the route, except if it is known a priori that none of these addresses act as multicast sources or RP, or both, in which case the (unicast) route need not carry the VRF Route Import extended community.

All c-multicast routes with the c-multicast import RT specific to the VRF must be accepted. In this release, vrf-import and vrf-target policies do not apply to C-multicast routes.

The decision flow path is shown below.

```
if (route-type == c-mcast-route)
   if (route_target_list includes C-multicast_Import_RT){
      else
         drop;
   else
   Run vrf_import or vrf-target, or both;
```

3.4.6 Provider tunnel support

3.4.6.1 Point-to-Multipoint Inclusive (I-PMSI) and Selective (S-PMSI) Provider Multicast Service Interface

BGP C-multicast signaling must be enabled for an MVPN instance to use P2MP RSVP-TE or LDP as I-PMSI (equivalent to 'Default MDT', as defined in draft Rosen MVPN) and S-PMSI (equivalent to 'Data MDT', as defined in draft Rosen MVPN).

By default, all PE nodes participating in an MVPN receive data traffic over I-PMSI. Optionally, (C-*, C-*) wildcard S-PMSI can be used instead of I-PMSI. See section Wildcard (C-*, C-*) P2MP LSP S-PMSI for more information. For efficient data traffic distribution, one or more S-PMSIs can be used, in addition to the default PMSI, to send traffic to PE nodes that have at least one active receiver connected to them. For more information, see P2MP LSP S-PMSI.

Only one unique multicast flow is supported over each P2MP RSVP-TE or P2MP LDP LSP S-PMSI. Number of S-PMSI that can be initiated per MVPN instance is restricted by CLI command **maximum-p2mp-spmsi**. P2MP LSP S-PMSI cannot be used for more than one (S,G) stream (that is, multiple multicast flow) as number of S-PMSI per MVPN limit is reached. Multicast flows that cannot switch to S-PMSI remain on I-PMSI.

3.4.6.2 P2MP RSVP-TE I-PMSI and S-PMSI

Point-to-Multipoint RSVP-TE LSP as inclusive or selective provider tunnel is available with BGP NG-MVPN only. P2MP RSVP-TE LSP is dynamically setup from root node on auto discovery of leaf PE nodes that are participating in multicast VPN. Each RSVP-TE I-PMSI or S-PMSI LSP can be used with a single MVPN instance only.

RSVP-TE LSP template must be defined (see *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide*) and bound to MVPN as inclusive or selective (S-PMSI is for efficient data distribution and is optional) provider tunnel to dynamically initiate P2MP LSP to the leaf PE nodes learned via NG-MVPN auto-discovery signaling. Each P2MP LSP S2L is signaled based on parameters defined in LSP template.

3.4.6.3 P2MP LDP I-PMSI and S-PMSI

Point-to-Multipoint LDP LSP as inclusive or selective provider tunnel is available with BGP NG-MVPN only. P2MP LDP LSP is dynamically setup from leaf nodes on auto discovery of leaf node PE nodes that are participating in multicast VPN. Each LDP I-PMSI or S-PMSI LSP can be used with a single MVPN instance only.

The **multicast-traffic** CLI command must be configured per LDP interface to enable P2MP LDP setup. P2MP LDP must also be configured as inclusive or selective (S-PMSI is for efficient data distribution and is optional) provider tunnel per MVPN to dynamically initiate P2MP LSP to leaf PE nodes learned via NG-MVPN auto-discovery signaling.

3.4.6.4 Wildcard (C-*, C-*) P2MP LSP S-PMSI

Wildcard S-PMSI allows usage of selective tunnel as a default tunnel for a specific MVPN. By using wildcard S-PMSI, operators can avoid full mesh of LSPs between MVPN PEs, reducing related signaling, state, and BW consumption for multicast distribution (no traffic is sent to PEs without any receivers active on the default PMSI).

The SR OS allows an operator to configure wildcard S-PMSI for ng-MVPN (**config>service>vprn>mvpn>pt>inclusive>wildcard-spmsi**), using LDP and RSVP-TE in P-instance. Support includes:

- IPv4 and IPv6
- PIM ASM and SSM
- directly attached receivers

The SR OS (C-*, C-*) wildcard implementation uses wildcard S-PMSI instead of I-PMSI for a specific MVPN. To switch MVPN from I-PMSI to (C-*, C-*) S-PMSI a VPRN shutdown is required. ISSU and Upstream Multicast Hop (UMH) redundancy can be used to minimize the impact.

To minimize outage, the following upgrade order is recommended:

1. Route Reflector
2. Receiver PEs
3. backup UMH
4. active UMH

RSVP-TE/mLDP configuration under inclusive provider tunnel (**config>service>vprn>mvpn>pt>inclusive**) apply to wildcard S-PMSI when enabled.

Wildcard C-S and C-G values are encoded as defined in RFC6625: using zero for Multicast Source Length and Multicast Group Length and omitting Multicast Source and Multicast Group values respectively in MCAST_VPN_NLRI. For example, a (C-*, C-*) is advertised as: RD, 0x00, 0x00, and originating router's IP address.

Note: All SR OSs with BGP peering session to the PE with RFC6625 support enabled must be upgraded to SR OS Release 13.0 before the feature is enabled. Failure to do so results in the following processing on a router with BGP peering session to an RFC6625-enabled PE:

- BGP peer running Release 12.0 version R4 or newer accepts 0-length address and it keeps encoding length 4 with all zeros for the address
- BGP peer running Release 12 version R3 or older does not accept 0-length address and keeps restarting BGP session

The procedures implemented by SR OS are compliant to section 3 and 4 of RFC, 6625. Wildcards encoded as described above are carried in NLRI field of MP_REACH_NLRF_ATTRIBUTE. Both IPv4 and IPv6 are supported: (AFI) of 1 or 2 and a Subsequent AFI (SAFI) of MCAST-VPN.

The (C-*, C-*) S-PMSI is established as follows:

- UMH PEs advertise I-PMSI A-D routes without tunnel information present (empty PTA) - encoded as per RFC6513/6514 before advertising wildcard S-PMSI. I-PMSI needs to be signaled and installed on receiver PEs, because (C-*, C-*) S-PMSI is only installed when a first receiver is added. However, no LSP is established for I-PMSI).
- UMH PEs advertise S-PMSI A-D route whose NLRI contains (C-*, C-*) with tunnel information encoded as per RFC 6625.
- Receiver PEs join wildcard S-PMSI if there are any receivers present.

Note: If UMH PE does not encode I-PMSI/S-PMSI A-D routes as per the above, or advertises both I-PMSI and wildcard S-PMSI with the tunnel information present, no interoperability can be achieved.

To ensure correct operation of BSR between PEs with (C-*, C-*) S-PMSI signaling, SR OS implements two modes of operations for BSR.

By default (bsr unicast):

- BSR PDUs are sent/forwarded as unicast PDUs to neighbor PEs when I-PMSI with Pseudo-tunnel interface is installed.
- At every BSR interval timer BSR Unicast PDU are sent to all I-PMSI interfaces when this is an elected BSR.
- BSMs received as multicast from C-instance interfaces are flooded as unicast in the P-instance.
- All PEs process BSR PDU's received on I-PMSI Pseudo-tunnel interface as unicast packets.
- BSR PDU's are not forwarded to PE's management control interface.
- BSR unicast PDU's use PE's System IP address as destination IP and sender PE's System address as Source IP.
- The BSR unicast functionality ensures that no special state needs to be created for BSR when (C-*, C-*) S-PMSI is enabled, which is beneficiary considering low volume of BSR traffic.

Note:

- For bsr unicast, the base IPv4 system address (IPv4) or the mapped version of the base IPv4 system address (IPv6) must be configured under the VPRN to ensure bsr unicast messages can reach the VPRN.
- For bsr spmsi, the base IPv4/IPv6 system address must be configured under the VPRN to ensure B-SR S-PMSI's are established.

BSR S-PMSI mode can be enabled to allow interoperability with other vendors. In that mode full mesh S-PMSI is required and created between all PEs in MVPN to exchange BSR PDUs between all PEs in MVPN. To operate as expected, the BSR S-PMSI mode requires a selective P-tunnel configuration. For IPv6 support (including dual-stack) of BSR S-PMSI mode, the IPv6 default system interface address must be configured as a loopback interface address under the VPRN and VPRN PIM contexts. Changing BSR signaling requires a VPRN shutdown.

Other key feature interactions and restrictions for (C-*, C-*) include the following:

- Extranet is fully supported with wildcard S-PMSI trees.
- (C-S, C-G) S-PMSIs are supported when (C-*, C-*) S-PMSI is configured (including both BW and receiver PE driven thresholds).
- Geo-redundancy is supported (deploying with geo-redundancy eliminates traffic duplication when geo-redundant source has no active receivers at a cost of slightly increased outage upon a switch because wildcard S-PMSI may need to be re-establish).
- PIM in P-instance is not supported.
- SR OS implementation requires wildcard encoding as per RFC6625 and I-PMSI/S-PMSI signaling as defined above (I-PMSI signaled with empty PTA then S-PMSI signaled with P-tunnel PTA) for interoperability. Implementations that do not adhere to RFC6625 encoding, or signal both I-PMSI and S-PMSI with P-tunnel PTA do not inter-operate with SR OS implementation).

3.4.6.5 P2MP LSP S-PMSI

NG-MVPN support P2MP RSVP-TE and P2MP LDP LSPs as selective provider multicast service interface (S-PMSI). S-PMSI is used to avoid sending traffic to PEs that participate in multicast VPN, but do not have any receivers for a specific C-multicast flow. This allows more-BW efficient distribution of multicast traffic over the provider network, especially for high bandwidth multicast flows. S-PMSI is spawned dynamically based on configured triggers as described in S-PMSI trigger thresholds section.

In MVPN, 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. In the scenario of S-PMSI:

1. The head-end PE sends an S-PMSI A-D route for a specific C-flow with the Leaf Information Required bit set.
2. The PEs who are interested in the C-flow respond with Leaf A-D routes.
3. The head-end PE then signals the P2MP LSP to all the leaf PEs using RSVP-TE.

Also, because the receivers may come and go, the implementation supports dynamically adding and pruning leaf nodes to and from the P2MP LSP.

When the tunnel type in the PMSI attribute is set to RSVP-TE P2MP LSP, the tunnel identifier is <Extended Tunnel ID, Reserved, Tunnel ID, P2MP ID>, as carried in the RSVP-TE P2MP LSP SESSION Object.

The PE can also learn via an A-D route that it needs to receive traffic on a particular RSVP-TE P2MP LSP before the LSP is actually setup. In this case, the PE needs to wait until the LSP is operational before it can modify its forwarding tables as directed by the A-D route.

Because of the way that LDP normally works, mLDP P2MP LSPs are setup without solicitation from the leaf PEs toward the head-end PE. The leaf PE discovers the head-end PE via I-PMSI or S-PMSI A-D routes. The tunnel identifier carried in the PMSI attribute is used as the P2MP FEC element. The tunnel identifier consists of the head-end PE's address, along with a Generic LSP identifier value. The Generic LSP identifier value is automatically generated by the head-end PE.

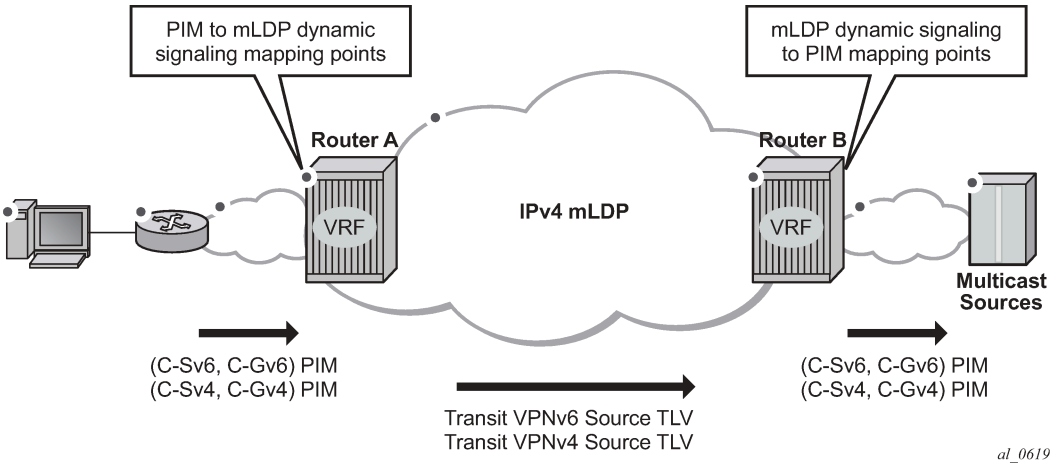
3.4.6.6 Dynamic multicast signaling over P2MP LDP in VRF

This feature provides a multicast signaling solution for IP-VPNs, allowing the connection of IP multicast sources and receivers in C-instances, which are running PIM multicast protocol using Rosen MVPN with BGP SAFI and P2MP mLDP in P-instance. The solution dynamically maps each PIM multicast flow to a P2MP LDP LSP on the source and receiver PEs.

The feature uses procedures defined in RFC 7246: *Multipoint Label Distribution Protocol In-Band Signaling in Virtual Routing and Forwarding (VRF) Table Context*. On the receiver PE, PIM signaling is dynamically mapped to the P2MP LDP tree setup. On the source PE, signaling is handed back from the P2MP mLDP to the PIM. Because of dynamic mapping of multicast IP flow to P2MP LSP, provisioning and maintenance overhead is eliminated as multicast distribution services are added and removed from the VRF. Per (C-S, C-G) IP multicast state is also removed from the network, because P2MP LSPs are used to transport multicast flows.

Figure 28: Dynamic mLDP signaling for IP multicast in VPRN illustrates dynamic mLDP signaling for IP multicast in VPRN.

Figure 28: Dynamic mLDP signaling for IP multicast in VPRN



As illustrated in Figure 28: Dynamic mLDP signaling for IP multicast in VPRN, P2MP LDP LSP signaling is initiated from the receiver PE that receives PIM JOIN from a downstream router (Router A). To enable dynamic multicast signaling, the **p2mp-ldp-tree-join** must be configured on PIM customer-facing interfaces for the specific VPRN of Router A. This enables handover of multicast tree signaling from the PIM to the P2MP LDP LSP. Being a leaf node of the P2MP LDP LSP, Router A selects the upstream-hop as the root node of P2MP LDP FEC, based on a routing table lookup. If an ECMP path is available for a specific route, then the number of trees are equally balanced toward multiple root nodes. The PIM joins are carried in the Transit Source PE (Router B), and multicast tree signaling is handed back to the PIM and propagated upstream as native-IP PIM JOIN toward C-instance multicast source.

The feature is supported with IPv4 and IPv6 PIM SSM and IPv4 mLDP. Directly connected IGMP/MLD receivers are also supported, with PIM enabled on outgoing interfaces and SSM mapping configured, if required.

The following are feature restrictions:

- Dynamic mLDP signaling in a VPRN instance and Rosen or NG-MVPN are mutually exclusive.
- A single instance of P2MP LDP LSP is supported between the receiver PE and Source PE per multicast flow; there is no stitching of dynamic trees.
- Extranet functionality is not supported.
- The router LSA link ID or the advertising router ID must be a routable IPv4 address (including IPv6 into IPv4 mLDP use cases).
- IPv6 PIM with dynamic IPv4 mLDP signaling is not supported with EBGP or IBGP with IPv6 next-hop.
- Inter-AS and IGP inter-area scenarios where the originating router is altered at the ASBR and ABR respectively, (therefore PIM has no way to create the LDP LSP toward the source), are not supported.
- When dynamic mLDP signaling is deployed, a change in Route Distinguisher (RD) on the Source PE is not acted upon for any (C-S, C-G)s until the receiver PEs learn about the new RD (via BGP) and send explicit delete and create with the new RD.
- Procedures of Section 2 of RFC 7246 for a case where UMH and the upstream PE do not have the same IP address are not supported.

3.4.6.7 MVPN sender-only/receiver-only

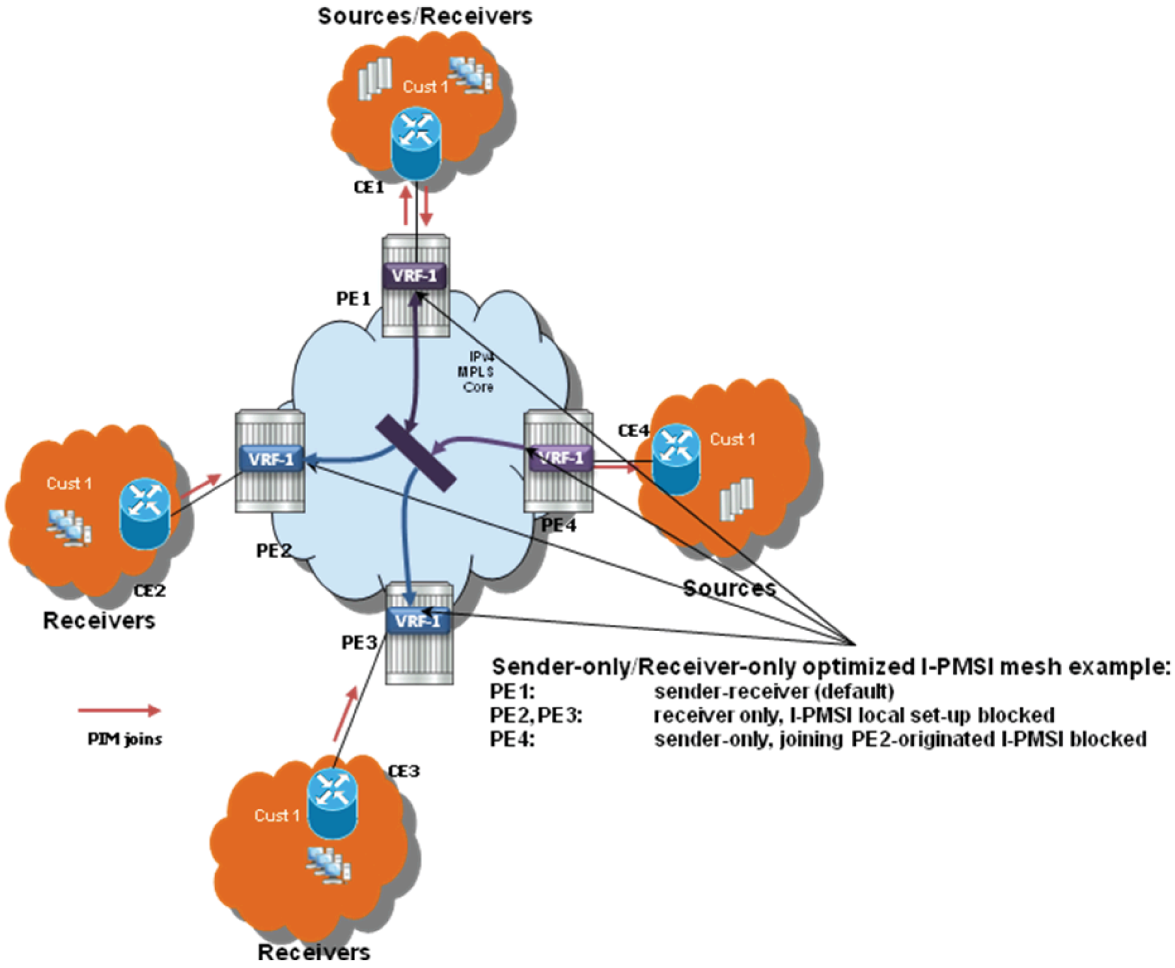
In multicast MVPN, by default, if multiple PE nodes form a peering with a common MVPN instance then each PE node originates a multicast tree locally toward the remaining PE nodes that are member of this MVPN instance. This behavior creates a mesh of I-PMSI across all PE nodes in the MVPN. It is often a case, that a specific VPN has many sites that host multicast receivers, but only few sites that either host both receivers and sources or sources only.

MVPN Sender-only/Receiver-only allows optimization of control and data plane resources by preventing unnecessary I-PMSI mesh when a specific PE hosts multicast sources only or multicast receivers only for a specific MVPN.

For PE nodes that host only multicast sources for a specific VPN, operators can now block those PEs, through configuration, from joining I-PMSIs from other PEs in this MVPN. For PE nodes that host only multicast receivers for a specific VPN, operators can now block those PEs, through configuration, to set-up a local I-PMSI to other PEs in this MVPN.

MVPN Sender-only/Receiver-only is supported with ng-MVPN using IPv4 RSVP-TE or IPv4 LDP provider tunnels for both IPv4 and IPv6 customer multicast. Figure 29: MVPN sender-only/receiver-only example depicts 4-site MVPN with sender-only, receiver-only and sender-receiver (default) sites.

Figure 29: MVPN sender-only/receiver-only example



Extra attention needs to be paid to BSR/RP placement when Sender-only/Receiver-only is enabled. Source DR sends unicast encapsulated traffic toward RP, therefore, RP shall be at sender-receiver or sender-only site, so that *G traffic can be sent over the tunnel. BSR shall be deployed at the sender-receiver site. BSR can be at sender-only site if the RPs are at the same site. BSR needs to receive packets from other candidate-BSR and candidate-RPs and also needs to send BSM packets to everyone.

3.4.6.8 S-PMSI trigger thresholds

The mLDP and RSVP-TE S-PMSIs support two types of data thresholds: bandwidth-driven and receiver-PE-driven. The threshold evaluation and bandwidth driven threshold functionality are described in Use of data MDTs.

In addition to the bandwidth threshold functionality, an operator can enable receiver-PE-driven threshold behavior. Receiver PE thresholds ensure that S-PMSI is only created when BW savings in P-instance justify extra signaling required to establish a new S-PMSI. For example, the number of receiver PEs interested in a specific C-multicast flow is meaningfully smaller than the number of receiver PEs for default PMSI (I-PMSI or wildcard S-PMSI). To ensure that S-PMSI is not constantly created/deleted, two

thresholds need to be specified: receiver PE add threshold and receiver PE delete threshold (expected to be significantly higher).

When a (C-S, C-G) crosses a data threshold to create S-PMSI, instead of regular S-PMSI signaling, sender PE originates S-PMSI explicit tracking procedures to detect how many receiver PEs are interested in a specific (C-S, C-G). When receiver PEs receive an explicit tracking request, each receiver PE responds, indicating whether there are multicast receivers present for that (C-S, C-G) on the specific PE (PE is interested in a specific (C-S, C-G)). If the geo-redundancy feature is enabled, receiver PEs do not respond to explicit tracking requests for suppressed sources and therefore only Receiver PEs with an active join are counted against the configured thresholds on Source PEs.

Upon regular sampling and check interval, if the previous check interval had a non-zero receiver PE count (one interval delay to not trigger S-PMSI prematurely) and current count of receiver PEs interested in the specific (C-S, C-G) is non-zero and is less than the configured receiver PE add threshold, Source PE sets up S-PMSI for this (C-S, C-G) following standard ng-MVPN procedures augmented with explicit tracking for S-PMSI being established.

Data threshold timer should be set to ensure enough time is given for explicit tracking to complete (for example, setting the timer to value that is too low may create S-PMSI prematurely).

Upon regular data-delay-interval expiry processing, when BW threshold validity is being checked, a current receiver PE count is also checked (for example, explicit tracking continues on the established S-PMSI). If BW threshold no longer applies or the receiver PEs exceed receiver PE delete threshold, the S-PMSI is torn down and (C-S, C-G) joins back the default PMSI.

Changing of thresholds (including enabling disabling the thresholds) is allowed in service. The configuration change is evaluated at the next periodic threshold evaluation.

The explicit tracking procedures follow RFC 6513/6514 with clarification and wildcard S-PMSI explicit tracking extensions as described in IETF Draft: *draft-dolganow-l3vpn-expl-track-00*.

3.4.6.9 Migration from existing Rosen implementation

The existing Rosen implementation is compatible to provide an easy migration path.

The following migration procedures are supported:

- Upgrade all the PE nodes that need to support MVPN to the newer release.
- The old configuration is converted automatically to the new style.
- Node by node, MCAST-VPN address-family for BGP is enabled. Enable auto-discovery using BGP.
- Change PE-to-PE signaling to BGP.

3.4.6.10 Policy-based S-PMSI

SR OS creates a single Selective P-Multicast Service Interface (S-PMSI) per multicast stream: (S,G) or (*,G). To better manage bandwidth allocation in the network, multiple multicast streams are often bundled starting from the same root node into a single, multi-stream S-PMSI. Network bandwidth is usually managed per package or group of packages, instead of per channel.

Multi-stream S-PMSI supports a single S-PMSI for one or more IPv4 (C-S, C-G) or IPv6 (C-S, C-G) streams. Multiple multicast streams starting from the same VPRN going to the same set of leafs can use a single S-PMSI. Multi-stream S-PMSIs can:

- carry exclusively IPv4 or exclusively IPv6, or a mix of channels

- coexist with a single group S-PMSI

To create a multi-stream S-PMSI, an S-PMSI policy needs to be configured in the VPN context on the source node. This policy maps multiple (C-S, C-G) streams to a single S-PMSI. Because this configuration is done per MVPN, multiple VPNs can have identical policies, each configured for its own VPN context.

When mapping a multicast stream to a multi-stream S-PMSI policy, the data traverses the S-PMSI without first using the I-PMSI. (Before this feature, when a multicast stream was sourced, the data used the I-PMSI first until a configured threshold was met. After this, if the multicast data exceeded the threshold, it would signal the S-PMSI tunnel and switch from I-PMSI to S-PMSI.)

For multi-stream S-PMSI, if the policy is configured and the multicast data matches the policy rules, the multi-stream S-PMSI is used from the start without using the default I-PMSI.

Multiple multi-stream S-PMSI policies could be assigned to a specific S-PMSI configuration. In this case, the policy acts as a link list. The first (lowest index) that matches the multi-stream S-PMSI policy is used for that specific stream.

The rules for matching a multi-stream S-PMSI on the source node are listed here.

S-PMSI to (C-S, C-G) mapping on Source-PE, in sequence:

1. The multi-stream S-PMSI policy is evaluated, starting from the lowest numerical policy index. This allows the feature to be enabled in the service when per-(C-S, C-G) stream configuration is present. Only entries that are not shut down are evaluated. First, the multi-stream S-PMSI (the lowest policy index) that the (C-S, C-G) stream maps to is selected.
2. If (C-S, C-G) does not map to any of multi-stream S-PMSIs, per-(C-S, C-G) S-PMSIs are used for transmission if a (C-S, C-G) maps to an existing S-PMSI (based on data-thresholds).
3. If S-PMSI cannot be used, the default PMSI is used.

To address multi-stream S-PMSI P-tunnel failure, if an S-PMSI P-tunnel is not available, a default PMSI tunnel is used. When an S-PMSI tunnel fails, all (C-S, C-G) streams using this multi-stream S-PMSI move to the default PMSI. The groups move back to S-PMSI after the S-PMSI tunnel is restored.

3.4.6.10.1 Supported MPLS tunnels

Multi-stream S-PMSI is configured in the context of an **auto-discovery** default (that is, NG-MVPN). It supports all existing per-mLDP/RSVP-TE P2MP S-PMSI tunnel functionality for multi-stream S-PMSI LSP templates (RSVP-TE P-instance).

Note:

- Per-multicast group statistics are not available for multi-stream S-PMSIs on S-PMSI level.
- GRE tunnels are not supported for multi-stream S-PMSI.

3.4.6.10.2 Supported multicast features

S-PMSI is supported with PIM ASM and PIM SSM in C-instances.

Note:

- The multi-stream S-PMSI model uses BSR RP co-located with the source PE or an RP between the source PE and multicast source, that is, upstream of receivers. Both **bsr unicast** and **bsr spmsi** can be deployed as applicable.

- The model also supports other RP types.

3.4.6.10.3 In-service changes to multi-stream S-PMSI

The operator can change the mapping in service; that is, the operator can move active streams (C-S, C-G) from one S-PMSI to another using the configuration, or from the default PMSI to the S-PMSI, without having to stop data transmission or having to disable a PMSI.

The change is performed by moving a (C-S, C-G) stream from a per-group S-PMSI to a multi-stream S-PMSI and the other way around, and moving a (C-S, C-G) stream from one multi-stream S-PMSI to another multi-stream S-PMSI.

Note:

- During re-mapping, a changed (C-S, C-G) stream is first moved to the default PMSI before it is moved to a new S-PMSI, regardless of the type of move. Unchanged (C-S, C-G) streams must remain on an existing PMSIs.
- Any change to a multi-stream S-PMSI policy or to a preferred multi-stream S-PMSI policy (for example, an index change equivalent to less than or equal to the current policy) should be performed during a maintenance window. Failure to perform these types of changes in a maintenance window could potentially cause a traffic outage.

3.4.6.10.4 Configuration example

In this example, two policies are created on the source node: multi-stream S-PMSI 1 and multi-stream-S-PMSI 10.

A multicast stream with group 224.0.0.x and source 192.0.2.0/24 maps to the first multi-stream policy. The group in the range of 224.0.0.0/24 and source 192.0.2.1/24 maps to policy 10.

```
*A:SwSim14>config>service>vprn# info
mvpn
                auto-discovery default
                c-mcast-signaling bgp
                provider-tunnel
                    inclusive
                        mldp
                            no shutdown
                        exit
                    exit
                    selective
                        mldp
                            no shutdown
                        exit
                        no auto-discovery-disable
                        data-threshold 239.70.1.0/24 1
                        data-threshold 239.0.0.0/8 1
                        multistream-spmsi 1 create
                            group 224.0.0.0/24
                                source 192.0.2.0/24
                            exit
                        exit
                        multistream-spmsi 10 create
                            group 224.0.0.0/24
                                source 192.0.2.1/24
                            exit
```

```
                    exit
                exit
            exit
```

3.4.6.11 Policy-based data MDT

A single data MDT can transport one or more IPv4 (C-S, C-G) streams. This allows multiple multicast streams starting from the same VPRN going to the same set of leafs to use a single data MDT. Characteristics of an MDT include:

- a multi-stream data MDT can carry IPv4 only
- a multi-stream data MDT can coexist with a single-group data MDT
- a default MDT must be configured

To create a multi-stream MDT data, an MDA data policy must be configured in the context of MVPN on the source node. This policy maps multiple (C-S, C-G) streams to a single data MDT. Because this configuration is per MVPN, multiple VPNs can have identical policies configured, each for its own VPN context.

When a multicast stream is mapped to a multi-stream data MDT policy, the data traverses the default MDT first. The data delay timer is used to switch the data from the default MDT to the multi-stream data MDT.

When the multi-stream data MDT is deleted, the traffic switches back to the default MDT.

In some cases when a new multi-stream data MDT is configured which is better suited, some streams may prefer this new multi-stream data MDT. To switch, the traffic switches to the default MDT and then to the new multi-stream data MDT.

MDT data can be configured as SSM or ASM.

There can be multiple multi-stream policies assigned to a single data MDT configuration. In this case, the policy acts as a link list, the first (lowest index) matched multi-stream policy is used for that specific stream. The following are the rules of matching a multi-stream data MDT to (C-S, C-G) mapping on a source PE (in order):

1. The multi-stream policy is evaluated to enable the feature in service when per-(C-S, C-G) configuration is present, starting from the lowest numerical policy index (only entries that are not shutdown are evaluated). The first multi-stream data MDT (the lowest policy index) the (C-S, C-G) maps to is selected.
2. If (C-S, C-G) does not map to any of the multi-stream data MDTs, per-(C-S, C-G) single data MDTs are used if a (C-S, C-G) maps to an existing MDT based on data-thresholds.
3. The default MDT is used if no policy matches to a data MDT.
4. When a (C-S, C-G) arrives, it start on the default MDT but can switch to a data MDT when the data delay interval expires. It does not check the data-threshold before switching.
5. When going from one multi-stream data MDT to a more suitable one, the traffic first switches to the default MDT and then switch to the new multi-stream data MDT based on the data-delay-interval.

When a data MDT tunnel fails, all (C-S, C-G)s using this multi-stream data MDT move to the default MDT, and the groups move back to the data MDT when it is restored.

3.4.7 MVPN (NG-MVPN) upstream multicast hop fast failover

MVPN upstream PE or P node fast failover detection method is supported with RSVP P2MP I-PMSI only. A receiver PE achieves fast upstream failover based on the capability to subscribe multicast flow from multiple UMH nodes and the capability to monitor the health of the upstream PE and intermediate P nodes using an unidirectional multi-point BFD session running over the provider tunnel.

A receiver PE subscribes multicast flow from multiple upstream PE nodes to have active redundant multicast flow available during failure of primary flow. Active redundant multicast flow from standby upstream PE allows instant switchover of multicast flow during failure of primary multicast flow.

Faster detection of multicast flow failure is achieved by keeping track of unidirectional multi-point BFD sessions enabled on the provider tunnel. Multi-point BFD sessions must be configured with 10 ms transmit interval on sender (root) PE to achieve sub-50ms fast failover on receiver (leaf) PE.

UMH **tunnel-status** selection option must be enabled on the receiver PE for upstream fast failover. Primary and standby upstream PE pairs must be configured on the receiver PE to enable active redundant multicast flow to be received from the standby upstream PE.

3.4.8 Multicast VPN extranet

Multicast VPN extranet distribution allows multicast traffic to flow across different routing instances. A routing instance that received a PIM/IGMP JOIN but cannot reach source of multicast source directly within its own instance is selected as receiver routing instance (receiver C-instance). A routing instance that has source of multicast stream and accepts PIM/IGMP JOIN from other routing instances is selected as source routing instance (source C-instance). A routing instance that does not have either source or receivers but is used in the core is selected as a transit instance (transit P-instance). The following subsections detail supported functionality.

3.4.8.1 Multicast extranet for Rosen MVPN for PIM SSM

Multicast extranet is supported for Rosen MVPN with MDT SAFI. Extranet is supported for IPv4 multicast stream for default and data MDTs (PIM and IGMP).

The following extranet cases are supported:

- local replication into a receiver VRF from a source VRF on a source PE
- transit replication from a source VRF onto a tunnel of a transit core VRF on a source PE. A source VRF can replicate its streams into multiple core VRFs as long as any specific stream from source VRF uses a single core VRF (the first tunnel in any core VRF on which a join for the stream arrives). Streams with overlapping group addresses (same group address, different source address) are supported in the same core VRF.
- remote replication from source or transit VRF into one or more receiver VRFs on receiver PEs
- multiple replications from multiple source or transit VRFs into a receiver VRF on receiver PEs

Rosen MVPN extranet requires routing information exchange between the source VRF and the receiver VRF based on route export or import policies:

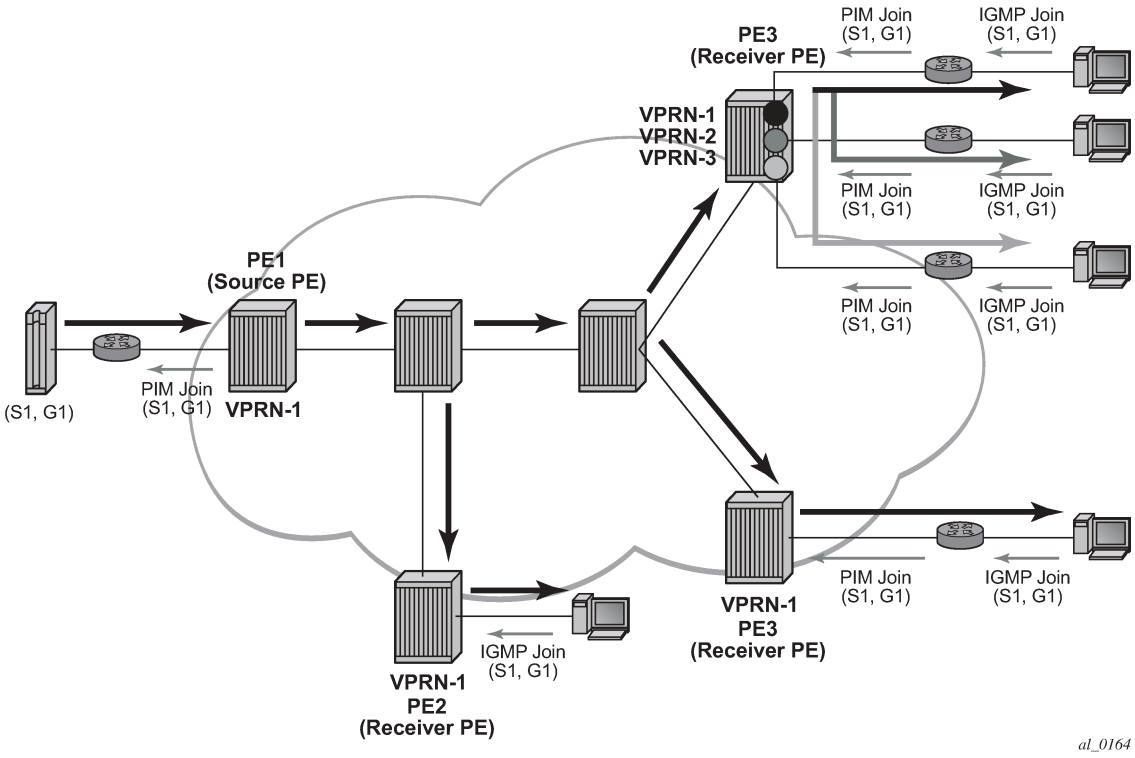
- Routing information for multicast sources must be exported using an RT export policy from the source VRF instance.
- Routing information must be imported into the receiver or transit VRF instance using an RT import policy.

The following are restrictions:

- The source VRF instance and receiver VRF instance of extranet must exist on a common PE node (to allow local extranet mapping).
- SSM translation is required for IGMP (C-*, C-G).
- An I-PMSI route cannot be imported into multiple VPRNs, and NG-MVPN routes do not need to be imported into multiple VPRNs.

In Figure 30: Multicast VPN traffic flow, VPRN-1 is the source VPRN instance and VPRN-2 and VPRN-3 are receiver VPRN instances. The PIM/IGMP JOIN received on VPRN-2 or VPRN-3 is for (S1, G1) multicast flow. Source S1 belongs to VPRN-1. Because of the route export policy in VPRN-1 and the import policy in VPRN-2 and VPRN-3, the receiver host in VPRN-2 or VPRN-3 can subscribe to the stream (S1, G1).

Figure 30: Multicast VPN traffic flow



3.4.8.2 Multicast extranet for NG-MVPN for PIM SSM

Multicast extranet is supported for ng-MVPN with IPv4 RSVP-TE and mLDP I-PMSIs and S-PMSIs including (C-*, C-*) S-PMSI support where applicable. Extranet is supported for IPv4 C-multicast traffic (PIM/IGMP joins).

The following extranet cases are supported:

- local replication into a receiver C-instance MVPNs on a source PE from a source P-instance MVPN
- remote replication from P-instance MVPN into one or more receiver C-instance MVPNs on receiver PEs

- multiple replications from multiple source/transit P-instance MVPNs into a receiver C-instance MVPN on receiver PEs

Note: Transit replication on Source PE is not supported.

Multicast extranet for ng-MVPN, similarly to extranet for Rosen MVPN, requires routing information exchange between source ng-MVPN and receiver ng-MVPN based on route export and import policies. Routing information for multicast sources is exported using an RT export policy from a source ng-MVPN instance and imported into a receiver ng-MVPN instance using an RT import policy. S-PMSI/I-PMSI establishment and C-multicast route exchange occurs in a source ng-MVPN P-instance only (import and export policies are not used for MVPN route exchange). Sender-only functionality must not be enabled for the source/transit ng-MVPN on the receiver PE. It is recommended to enable receiver-only functionality on a receiver ng-MVPN instance.

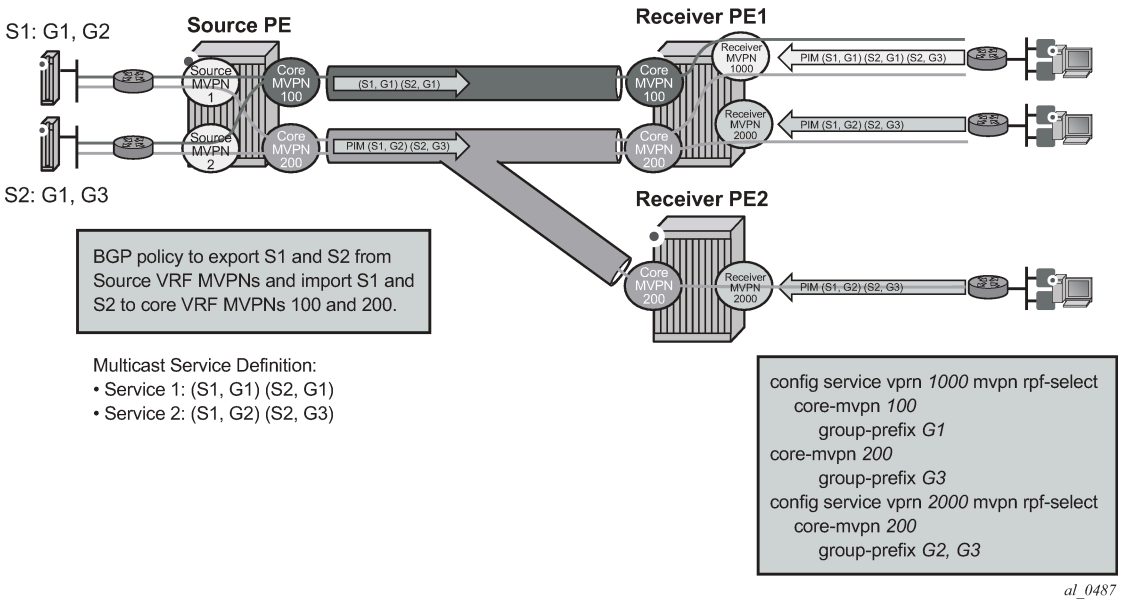
The following are restrictions:

- Source P-instance MVPN and receiver C-instance MVPN must reside on the receiver PE (to allow local extranet mapping).
- SSM translation is required for IGMP (C-*, C-G).

3.4.8.3 Multicast extranet with per-group mapping for PIM SSM

In some deployments, such as IPTV or wholesale multicast services, it may be desirable to create one or more transit MVPNs to optimize delivery of multicast streams in the provider core. Figure 31: Source PE transit replication and receiver PE per-group SSM extranet mapping represents an example deployment model.

Figure 31: Source PE transit replication and receiver PE per-group SSM extranet mapping



The architecture displayed in Figure 31: Source PE transit replication and receiver PE per-group SSM extranet mapping requires a source routing instance MVPN to place its multicast streams into one or more

transit core routing instance MVPNs (each stream mapping to a single transit core instance only). It also requires receivers within each receiver routing instance MVPN to know which transit core routing instance MVPN they need to join for each of the multicast streams. To achieve this functionality, transit replication from a source routing instance MVPN onto a tunnel of a transit core routing instance MVPN on a source PE (see earlier sub-sections for MVPN topologies supporting transit replication on source PEs) and per-group mapping of multicast groups from receiver routing instance MVPNs to transit core routing instance MVPNs (as defined below) are required.

For per-group mapping on a receiver PE, the operator must configure a receiver routing instance MVPN per-group mapping to one or more source/transit core routing instance MVPNs. The mapping allows propagation of PIM joins received in the receiver routing instance MVPN into the core routing MVPN instance defined by the map. All multicast streams sourced from a single multicast source address are always mapped to a single core routing instance MVPN for a specific receiver routing instance MVPN (multiple receiver MVPNs can use different core MVPNs for groups from the same multicast source address). If per-group map in receiver MVPN maps multicast streams sourced from the same multicast source address to multiple core routing instance MVPNs, then the first PIM join processed for those streams selects the core routing instance MVPN to be used for all multicast streams from a specific source address for this receiver MVPN. PIM joins for streams sourced from the source address not carried by the selected core VRF MVPN instance remains unresolved. When a PIM join or prune is received in a receiver routing instance MVPN with per-group mapping configured, if no mapping is defined for the PIM join's group address, non-extranet processing applies when resolving how to forward the PIM join/prune.

The main attributes for per-group SSM extranet mapping on receiver PE include support for:

- Rosen MVPN with MDT SAFI. RFC6513/6514 NG-MVPN with IPv4 RSVP-TE/mLDP in P-instance (a P-instance tunnel must be in the same VPRN service as multicast source)
- IPv4 PIM SSM
- IGMP (C-S, C-G), and for IGMP (C-*, C-G) using SSM translation
- a receiver routing instance MVPN to map groups to multiple core routing instance MVPNs
- in-service changes of the map to a different transit/source core routing instance (this is service affecting)

The following are restrictions:

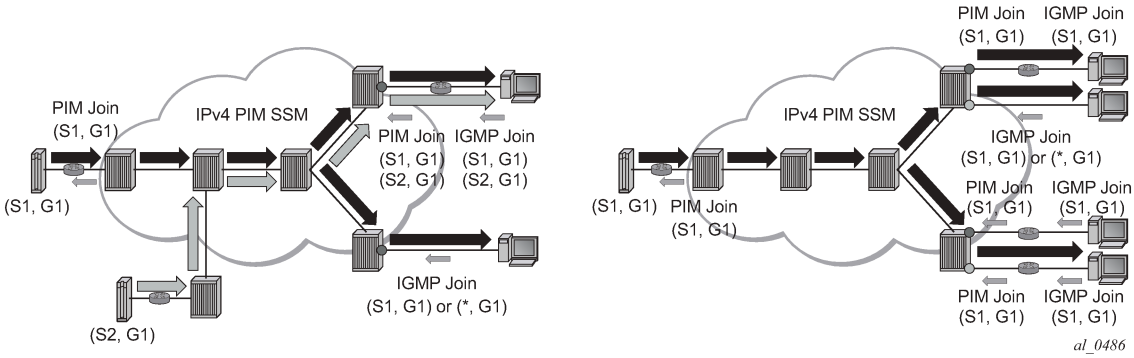
- When a receiver routing instance MVPN is on the same PE as a source routing instance MVPN, basic extranet functionality and not per-group (C-S, C-G) mapping must be configured (extranet from receiver routing instance to core routing instance to source routing instance on a single PE is not supported).
- Local receivers in the core routing instance MVPN are not supported when per-group mapping is deployed.
- Receiver routing instance MVPN that has per-group mapping enabled cannot have tunnels in its OIF lists.
- Per-group mapping is blocked if GRT/VRF extranet is configured.

3.4.8.4 Multicast GRT-source/VRF-receiver extranet with per group mapping for PIM SSM

Multicast GRT-source/VRF-receiver (GRT/VRF) extranet with per-group mapping allows multicast traffic to flow from GRT into VRF MVPN instances. A VRF routing instance that received a PIM/IGMP join but cannot reach the source multicast stream directly within its own instance is selected as receiver routing instance. A GRT instance that has sources of multicast streams and accepts PIM joins from other VRF MVPN instances is selected as source routing instance.

Figure 32: GRT/VRF extranet shows an example deployment.

Figure 32: GRT/VRF extranet



Routing information is exchanged between GRT and VRF receiver MVPN instances of extranet by enabling grt-extranet under a receiver MVPN PIM configuration for all or a subset of multicast groups. When enabled, multicast receivers in a receiver routing instances can subscribe to streams from any multicast source node reachable in GRT source instance.

The main feature attributes are:

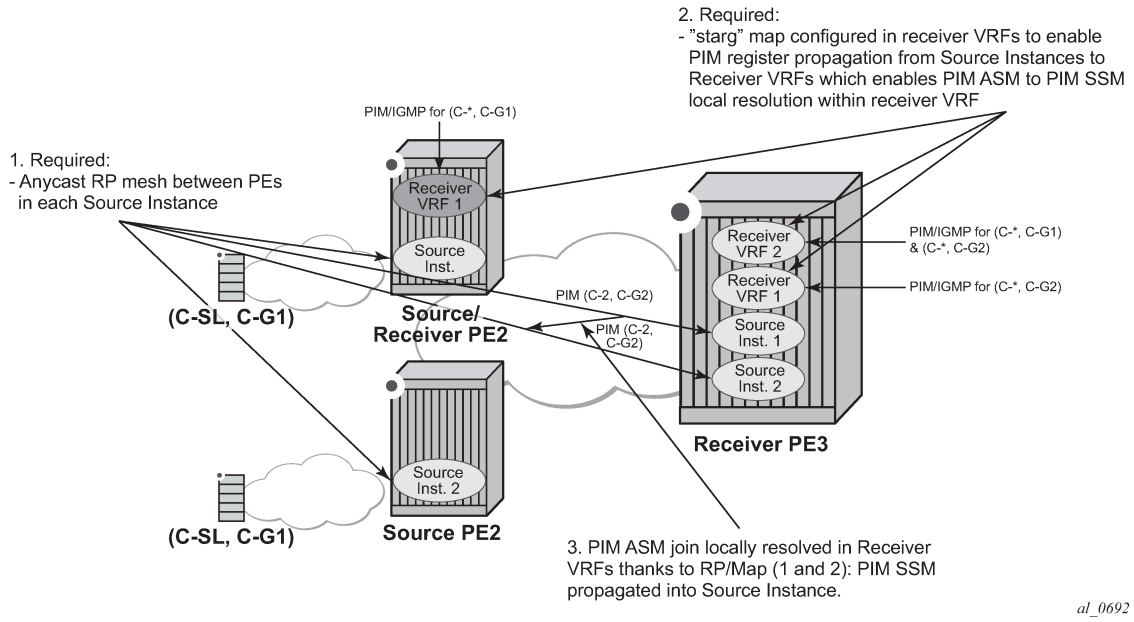
- GRT/VRF extranet can be performed on all streams or on a configured group of prefixes within a receiver routing instance.
- GRT instance requires Classic Rosen multicast.
- IPv4 PIM joins are supported in receiver VRF instances.
- Local receivers using IGMP: (C-S, C-G) and (C-*, C-G) using SSM translation are supported.
- The feature is blocked if a per-group mapping extranet is configured in receiver VRF.

3.4.8.5 Multicast extranet with per-group mapping for PIM ASM

Multicast extranet with per-group mapping for PIM ASM allows multicast traffic to flow from a source routing instance to a receiver routing instance when a PIM ASM join is received in the receiver routing instance.

Figure 33: Multicast extranet with per group PIM ASM mapping depicts PIM ASM extranet map support.

Figure 33: Multicast extranet with per group PIM ASM mapping



PIM ASM extranet is achieved by local mapping from the receiver to source routing instances on a receiver PE. The mapping allows propagation of anycast RP PIM register messages between the source and receiver routing instances over an auto-created internal extranet interface. This PIM register propagation allows the receiver routing instance to resolve PIM ASM joins to multicast sources and to propagate PIM SSM joins over an auto-created extranet interface to the source routing instance. PIM SSM joins are then propagated toward the multicast source within the source routing instance.

The following MVPN topologies are supported:

- Rosen MVPN with MDT SAFI: a local replication on a source PE and multiple-source/multiple-receiver replication on a receiver PE
- RFC 6513/6514 NG-MVPN (including RFC 6625 (C-*, C-*) wildcard S-PMSI): a local replication on a source PE and a multiple source/multiple receiver replication on a receiver PE
- Extranet for GRT-source/VRF receiver with a local replication on a source PE and a multiple-receiver replication on a receiver PE
- Locally attached receivers are supported without SSM mapping.

To achieve the extranet replication, the operator must configure:

- local PIM ASM mapping on a receiver PE from a receiver routing instance to a source routing instance (**config>service>vprn>mvpn>rpf-select>core-mvpn** or **config>service>vprn>pim>grt-extranet** as applicable)
- an anycast RP mesh between source and receiver PEs in the source routing instance

The following are restrictions:

- This feature is supported for IPv4 multicast.
- The multicast source must reside in the source routing instance the ASM map points to on a receiver PE (the deployment of transit replication extranet from source instance to core instance on Source PE with ASM map extranet from receiver instance to core instance on a receiver PE is not supported).

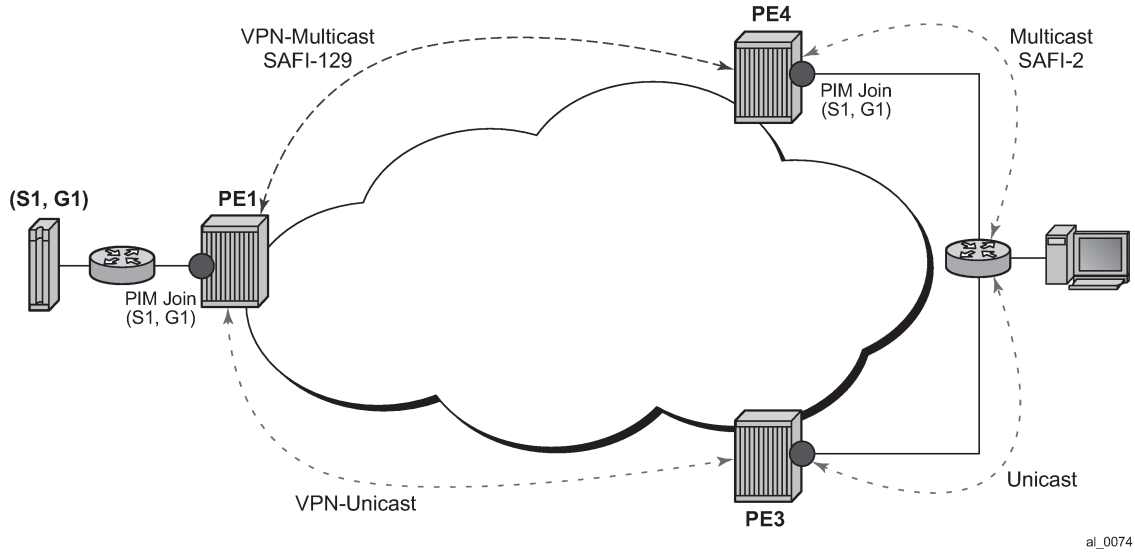
- A specific multicast group can be mapped in a receiver routing instance using either PIM SSM mapping or PIM ASM mapping but not both.
- A specific multicast group cannot map to multiple source routing instances.

3.4.9 Non-congruent unicast and multicast topologies for multicast VPN

Operators that prefer to keep unicast and multicast traffic on separate links in a network have the option to maintain two separate instances of the route table (unicast and multicast) per VPRN.

Multicast BGP can be used to advertise separate multicast routes using Multicast NLRI (SAFI 2) on PE-CE link within VPRN instance. Multicast routes maintained per VPRN instance can be propagated between PE-PE using BGP Multicast-VPN NLRI (SAFI 129).

Figure 34: Incongruent multicast and unicast topology for non-overlapping traffic links



SR OS supports option to perform RPF check per VPRN instance using multicast route table, unicast route table or both.

Non-congruent unicast and multicast topology is supported with NG-MVPN. Draft Rosen is not supported.

3.4.10 Automatic discovery of Group-to-RP mappings (auto-RP)

Auto-RP is a proprietary group discovery and mapping mechanism for IPv4 PIM that is described in cisco-ipmulticast/pim-autorp-spec, *Auto-RP: Automatic discovery of Group-to-RP mappings for IP multicast*. The functionality is similar to the IETF standard bootstrap router (BSR) mechanism that is described in RFC 5059, *Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)*, to dynamically learn about the availability of Rendezvous Points (RPs) in a network. When a router is configured as an RP-mapping agent with the **pim>rp>auto-rp-discovery** command, it listens to the CISCO-RP-ANNOUNCE (224.0.1.39) group and caches the announced mappings. The RP-mapping agent then periodically sends out RP-mapping packets to the CISCO-RP-DISCOVERY (224.0.1.40) group. SR OS supports version 1 of

the auto-RP specification, so the ability to deny RP-mappings by advertising negative group prefixes is not supported.

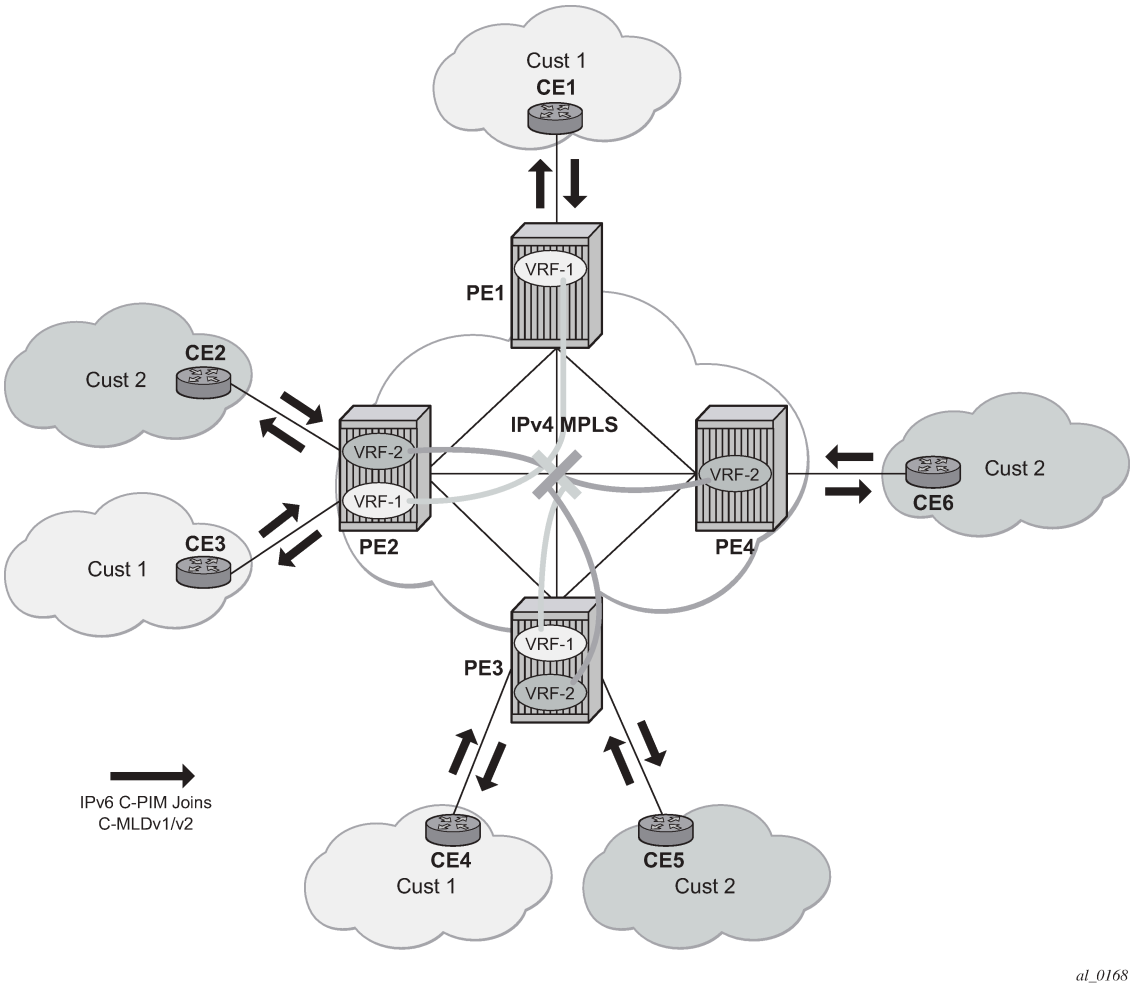
PIM dense-mode (PIM-DM) as described in RFC 3973, *Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)*, is used for the auto-RP groups to support multihoming and redundancy. The RP-mapping agent supports announcing, mapping, and discovery functions; candidate RP functionality is not supported.

Auto-RP is supported for IPv4 in multicast VPNs and in the global routing instance. Either BSR or auto-RP for IPv4 can be configured; the two mechanisms cannot be enabled together. BSR for IPv6 and auto-RP for IPv4 can be enabled together. In a multicast VPN, auto-RP cannot be enabled together with sender-only or receiver-only multicast distribution trees (MDTs), or wildcard S-PMSI configurations that could block flooding.

3.4.11 IPv6 MVPN support

IPv6 multicast support in SR OS allows operators to offer customers IPv6 multicast MVPN service. An operator uses IPv4 mLDP or RSVP-TE core to carry IPv6 c-multicast traffic inside IPv4 mLDP or RSVP-TE provider tunnels (p-tunnels). The IPv6 customer multicast on a specific MVPN can be blocked, enabled on its own or in addition to IPv4 multicast per PE or per interface. When both IPv4 and IPv6 multicast is enabled for a specific MVPN, a single tree is used to carry both IPv6 and IPv4 traffic. Figure 35: IPv6 MVPN example shows an example of an operator with IPv4 MPLS backbone providing IPv6 MVPN service to Customer 1 and Customer 2.

Figure 35: IPv6 MVPN example



SR OS IPv6 MVPN multicast implementation provides the following functionality:

- IPv6 C-PIM-SM (ASM and SSM)
- MLDv1 and MLDv2
- SSM mapping for MLDv1
- I-PMSI and S-PMSI using IPv4 P2MP mLDP p-tunnels
- I-PMSI and S-PMSI using IPv4 P2MP RSVP p-tunnels
- BGP auto-discovery
- PE-PE transmission of C-multicast routing using BGP mvpn-ipv6 address family
- IPv6 BSR/RP functions on functional par with IPv4 (auto-RP using IPv4 only)
- Embedded RP
- Inter-AS Option A

The following known restrictions exist for IPv6 MVPN support:

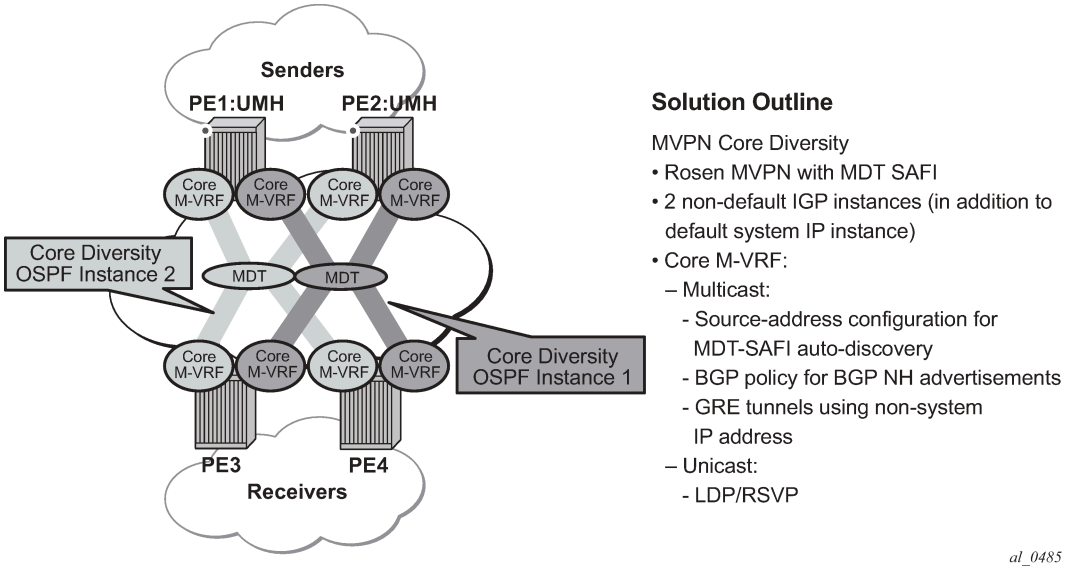
- Non-congruent topologies are not supported.

- IPv6 is not supported in MCAC.
- If IPv4 and IPv6 multicast is enabled, per-MVPN multicast limits apply to entire IPv4 and IPv6 multicast traffic as it is carried in a single PMSI. For example IPv4 AND IPv6 S-PMSIs are counted against a single S-PMSI maximum per MVPN.
- IPv6 Auto-RP is not supported.

3.4.12 Multicast core diversity for Rosen MDT_SAFI MVPNs

Figure 36: Multicast core diversity depicts Rosen MVPN core diversity deployment:

Figure 36: Multicast core diversity



Core diversity allows operator to optionally deploy multicast MVPN in either default IGP instance or one of two non-default IGP instances to provide, for example, topology isolation or different level of services. The following describes main feature attributes:

- Rosen MVPN IPv4 multicast with MDT SAFI is supported with default and data MDTs.
- Rosen MVPN can use a non-default OSPF or ISIS instance (using their loopback addresses instead of a system address).
- Up to 3 distinct core instances are supported: system + 2 non-default OSPF instances – referred as "red" and "blue" below.
- The BGP Connector also uses non-default OSPF loopback as NH, allowing Inter-AS Option B/C functionality to work with Core diversity as well.
- The feature is supported with CSC-VPRN.

On source PEs (PE1: UMH, PE2: UMH in Figure 36: Multicast core diversity), an MVPN is assigned to a non-default IGP core instance as follows:

1. MVPN is statically pointed to use one of the non-default "red"/"blue" IGP instances loopback addresses as source address instead of system loopback IP.

2. MVPN export policy is used to change unicast route next-hop VPN address (no longer required as of SR OS Release 12.0.R4 - BGP Connector support for non-default instances).

The above configuration ensures that MDT SAFI and IP-VPN routes for the non-default core instance use non-default IGP loopback instead of system IP. This ensures PIM advertisement/joins run in the correct core instance and GRE tunnels for multicast can be set-up using and terminated on non-system IP.

If BGP export policy is used to change unicast route next-hop VPN address, unicast traffic must be forwarded in non-default "red" or "blue" core instance LDP or RSVP (terminating on non-system IP) must be used. GRE unicast traffic termination on non-system IP is not supported, and any GRE traffic arriving at the PE in "blue", "red" instances destined for non-default IGP loopback IP is forwarded to CPM (ACL or CPM filters can be used to prevent the traffic from reaching the CPM). This limitation does not apply if BGP connector attribute is used to resolve the multicast route.

No configuration is required on non-source PEs.

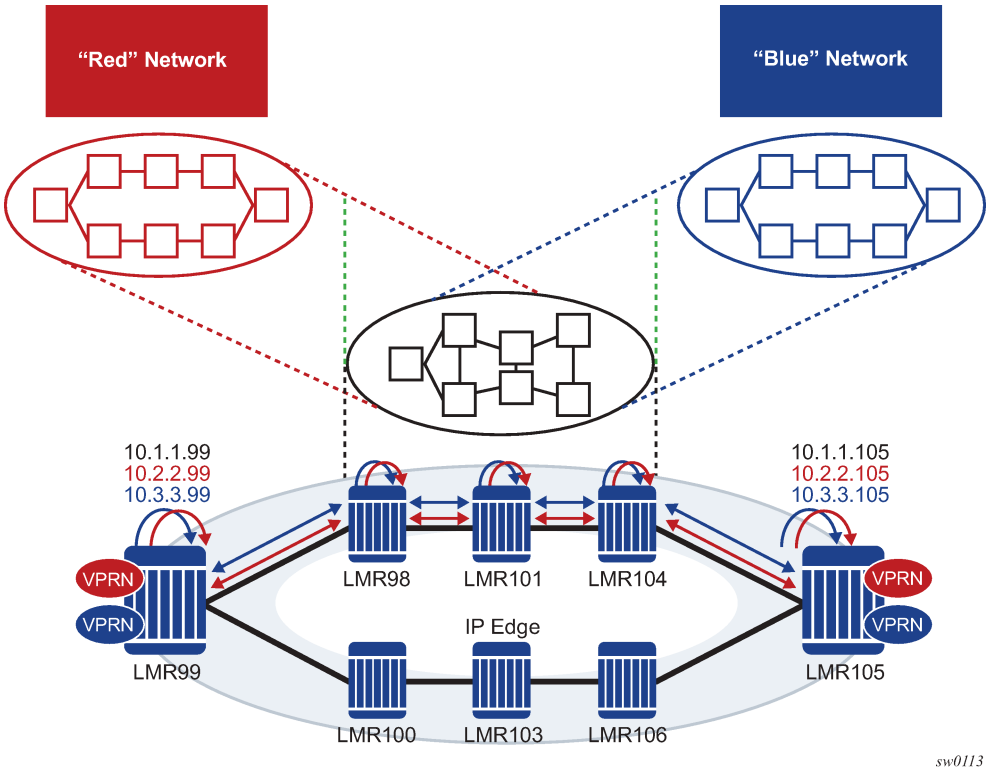
The following are feature restrictions:

- VPRN instance must be shutdown to change the mdt-safi source-address. The CLI rollback that includes change of the auto-discovery is therefore service impacting.
- To reset mdt-safi source-address to system IP, operator must first execute no auto-discovery (or auto-discovery default) then auto-discovery mdt-safi
- Configuring system IP as a source-address consumes one of the 2 IP addresses allowed, therefore it should not be done.
- Operators must configure correct IGP instance loopback IP addresses within Rosen MVPN context and must configure correct BGP policies (Before Release 12.0.R4) for the feature to operate as expected. There is no verification that the address entered for MVPN provider tunnel source-address is such an address or is not a system IP address.

3.4.13 NG-MVPN core diversity

See Figure 37: Logical networks using multi-instance IGP for an operational example of logical networks using Multi-Instance IGP.

Figure 37: Logical networks using multi-instance IGP



SR OS is being positioned in multi-instance IGP as a virtualization or migration strategy in numerous cases. One specific application is as a virtual LSR core whereby various topologies are created using separate IGP instances. Specifically, the migration to SR solutions requires the deployment of multi-instance IGP with service migrations. The objective is to more cleanly segregate protocols and service bindings to specific routing instances. NG-MVPN does not currently allow non-system loopbacks to be used for PMSI (for example, MVPN address family).

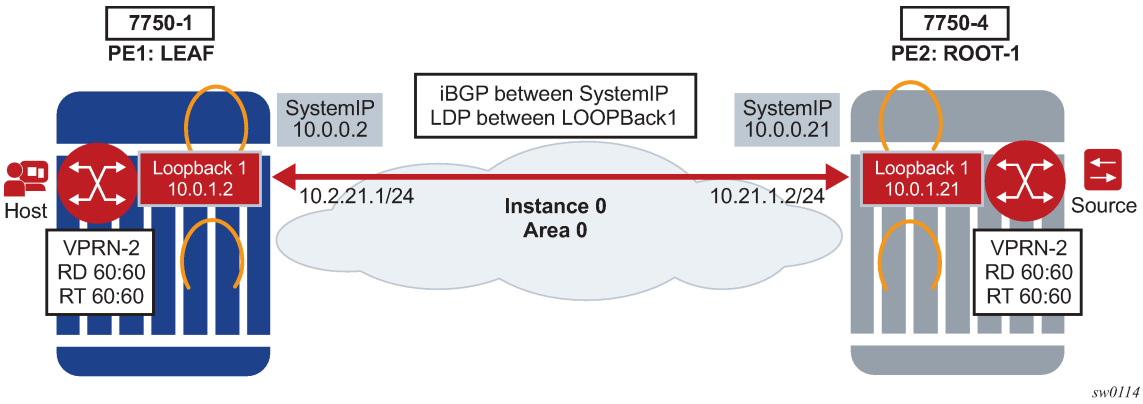
The ability to support binding MVPN with mLDP tunnels to different loopback interfaces is one of the main requirements. In addition, assigning these loopback interfaces to different IGP instances creates parallel NG-MVPN services, each running over a separate IGP instance.

This feature has two main components to it:

- The ability of advertising a MVPN route, via a loopback interface, and generating an mLDP FEC with that loopback interface.
- The ability of creating multiple IGP instances and assigning a loopback to each IGP instance. This, combined with the component above, creates parallel NG-MVPN services on different IGP instances.

3.4.13.1 NG-MVPN to loopback interface

Figure 38: NG-MVPN setup via loopback interface



SystemIP is usually used for management purposes and, therefore, it may be needed to create services to a separate loopback interface. Because of security concerns, many operators do not want to create services to the systemIP address. See Figure 38: NG-MVPN setup via loopback interface and Figure 39: Intra-AS basic opaque FEC to loopback interface.

The first portion of this feature allows MVPN routes be advertised with a nexthop specific loopback.

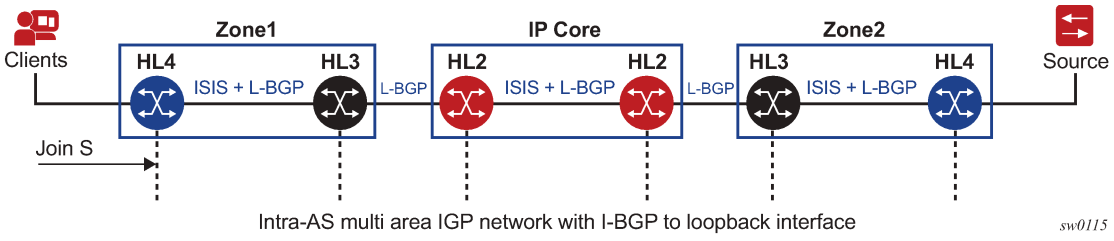
This can be achieved via two methods:

- Having IBGP session to the loopback interface of the peer, and using the corresponding local loopback for local-address.
- Having IBGP session to the systemIP but using policies to change the nexthop of the AD route to the corresponding loopback interface.

After the AD routes are advertised via the loopback interface as nexthop, PIM generates an mLDP FEC for the loopback interface.

The preceding configuration allows an NG-MVPN to be established via a specific loopback.

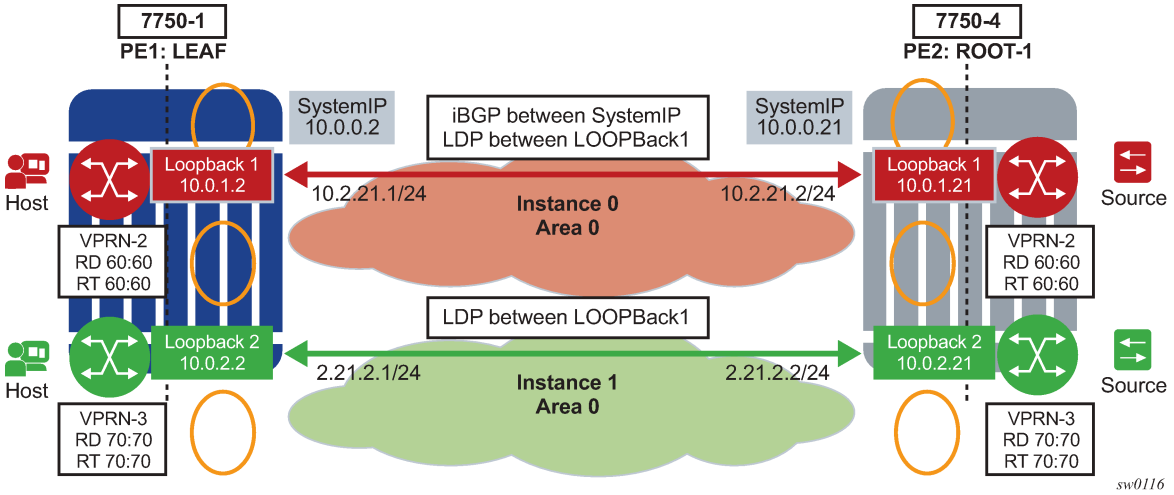
Figure 39: Intra-AS basic opaque FEC to loopback interface



It should be noted that this setup works in a single area or multi-area through an ABR.

3.4.13.2 NG-MVPN core diversity

Figure 40: Core diversity with parallel NG-MVPN services on parallel IGP instances



In Figure 40: Core diversity with parallel NG-MVPN services on parallel IGP instances, Red and Blue networks correspond to separate IGP instances tied to separate loopback interfaces. In core diversity, MVPN services on the LER are bound to a domain by advertising MP-BGP routes, with next hop set to the appropriate loopback address, and having the receiving LERs resolve those BGP next hops, based on the corresponding IGP instance. All subsequent MVPN BGP routing exchanges must set and resolve the BGP next hop with the configured non-system loopback.

With this feature, an MP-BGP next hop would resolve to a link LDP label which is indirectly associated with a specific IGP instance. Services which advertise BGP routes with next hop Red loopback would result in traffic flowing over the Red network using LDP and Blue loopback would result in traffic flowing over Blue network. See Figure 40: Core diversity with parallel NG-MVPN services on parallel IGP instances.

Most importantly, the common routes in each IGP instance must have a unique and active label. This is required to ensure that the same route advertised in two different IGP domains do not resolve to the same label. LDP *local-lsr-id* can be used to ensure FECs and label mapping be advertised via the right instance of IGP.

Core diversity allows an operator to optionally deploy multicast NG-MVPN in either default IGP instance or one of the non-default IGP instances to provide, for example, topology isolation or different level of services. The following describes the main feature attributes:

- NG- MVPN can use IPv4/IPv6 multicast.
- NG-MVPN can use a non-default OSPF or ISIS instance.

  Note: This is accomplished by using their loopback addresses instead of a system address.

- The BGP Connector also uses non-default OSPF loopback as NH, via two methods:
  - Setting the BGP local-address to a loopback interface IP address
  - Creating a routing policy to set the NH of a MVPN AD route to the corresponding loopback IP

- RSVP-TE/LDP transport tunnels also use non-systemIP loopback for session creation. As an example, RSVP-TE p2mp LSPs would be created to non systemIP loopbacks and mLDP creates a session via *local-lsr-id*.
- Need to add the source-address for mvpn auto-discover default.

On the source PEs, a NG-MVPN is assigned to a non-default IGP core instance as follows:

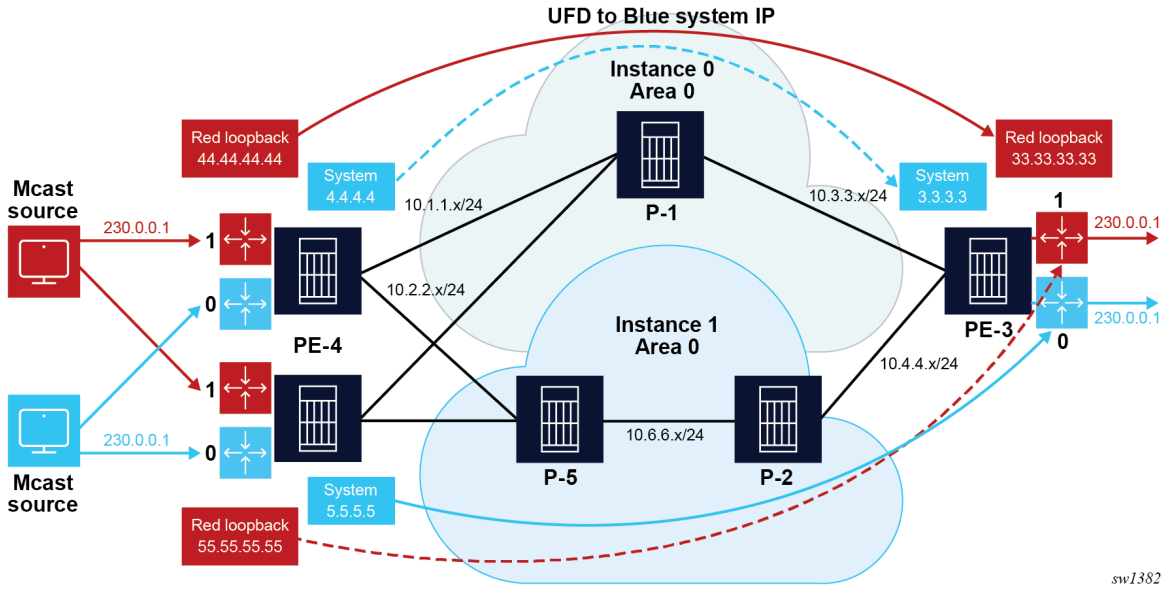
1. NG-MVPN is statically pointed to use one of the non-default *red*/*blue* IGP instances loopback addresses as source address instead of system loopback IP.
2. MVPN export policy is used to change unicast route next-hop VPN address (BGP Connector support for non-default instances).
3. Alternatively, the BGP local-address can be set to the correct loopback interface assigned for that specific instance.

The preceding configuration ensures that MVPN-IPv4/v6 and IP-VPN routes for the non-default core instance use non-default IGP loopback instead of system IP. This ensures MVPN-IPv4/v6 advertisement/ joins run in the correct core instance and mLDP and P2MP RSVP tunnels (I-PMSI and S-PMSI) for multicast can be set-up using and terminating on non-system IP.

If BGP export policy is used to change unicast route next-hop VPN address, then unicast traffic must be forwarded in non-default *red* or *blue* core instance LDP or RSVP (terminating on non-system IP) must be used.

3.4.13.3 P2MP RSVP-TE core diversity with UFD for UMH redundancy

Figure 41: P2MP RSVP-TE core diversity with UFD for UMH redundancy



Each NG-MVPN can be established over a separate IGP instance for core diversity support. In Figure 41: P2MP RSVP-TE core diversity with UFD for UMH redundancy, there are two IGP instances, Blue and Red, each having its own dedicated NG-MVPN service. The Blue plane is identified with the system IP address, and Red plane is identified with a loopback IP address. MVPN Autodiscovery (AD) routes are advertised and installed accordingly based on their NLRI in each instance. To advertise these MVPN AD routes in the

corresponding planes, route policies are used to change the next hops of the AD routes. For example, for the Blue plane the AD routes are advertised by default with the system IP address as their next hop, so no route policy is necessary. In the Red plane, a route policy can change the next hop of the corresponding NG-MVPN AD routes to use the loopback IP address.

Core diversity also supports UMH redundancy solutions. In the case of P2MP RSVP-TE, which uses UFD for source UMH failure, the UFD packets are generated for each IGP instance (Blue, Red) accordingly, with the correct source and destination IP address that is part of that IGP instance.

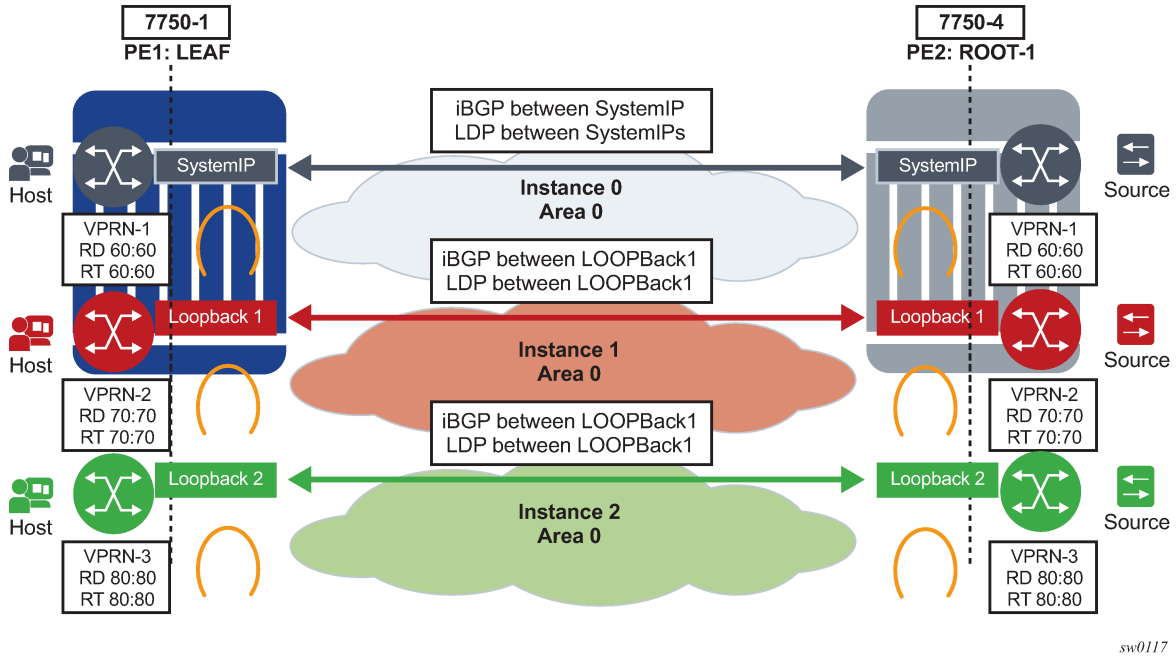
3.4.13.4 UFD packet generation

SR OS generates the UFD packets with the correct source and destination IP addresses for each IGP instance used in the core diversity configuration. SR OS uses the NLRI next-hop information of the AD route for the destination IP address and derives the source IP address from the MVPN autodiscovery default source-address loopback for the UFD packets. This ensures the UFD packets are traversing the appropriate core corresponding to their UMH reachability.

For example, in Figure 41: P2MP RSVP-TE core diversity with UFD for UMH redundancy, the Blue plane UFD packets are generated from the root PE to the leaf PE with the source IP address as the local system IP on the root and the destination IP address as the system IP of the leaf. The Red plane generates the UFD packet with the loopback IP address to which it is corresponding to as its source and destination IP address.

3.4.13.5 Configuration example

Figure 42: Core diversity with parallel NG-MVPN services on parallel IGP instances



In Figure 42: Core diversity with parallel NG-MVPN services on parallel IGP instances, there are three IGP instances, the default Instance 0, Instance 1, and Instance 2. Each instance binds to its own loopback

interface. For Instance 0, it is the systemIP *system Interface* used as loopback. LDP, RSVP and MP-BGP need to run between the corresponding loopback associated with each instance.

For example, for the blue Instance *2,* both MP-BGP and LDP need to be configured to its corresponding loopback *loopback2* and the next-hop for BGP MVPN-IP4/IPv6 and the VPN-IPv4/IPv6 need to be *loopback2*.

From a configuration point of view, the following steps need to be performed:

1. For MLDP, configure LDP with **local-lsr-id** with loopback interface of instance 2 *loopback2*:

```
*A:SwSim14>config>router>ldp# info

      interface-parameters
         interface “2ROOT” dual-stack
           ipv4
             local-lsr-id interface-name “loopback2”
              no shutdown
           exit
           no shutdown
         exit
```

2. Enable the source-address for default auto discover as follows:

```
*A:Dut-A>config>service>vprn 2
      mvpn auto-discovery default source-address LOOPBack1
      vrf-export “vprnexp100”
*A:Dut-A>config>service>vprn 3
      mvpn auto-discovery default source-address LOOPBack2
      vrf-export “vprnexp101”
```

3. Define community **vprnXXXX** for each VPRN using non-default core-instance and define a policy to tag each VPRN with either a *blue* or *red* standard community attribute:

```
*A:Dut-A>config>router>policy-options# info
        community "vprn2" members "target:70:70"
        community "vprn3" members "target:80:80"
        policy-statement "vprnexp2"
            entry 10
                from
                   protocol direct
                exit
                action accept
                   community add "vprn2" "red"
                exit
            exit
        policy-statement "vprnexp3"
            entry 10
                from
                   protocol direct
                exit
                action accept
                   community add "vprn3" "blue"
                exit
            exit
        exit
```

4. Define a single global BGP policy to change next hop for red/blue MVPNs:

```
*A:Dut-A>config>router>policy-options# info
        policy-statement "MVPN_CoreDiversity_Exp"
```

```
            entry 10
                from
                   community "red"
                exit
                to
                   protocol bgp-vpn
                exit
                action accept
                     next-hop loopback1
                exit
            exit
            entry 20
                from
                   community "blue"
                exit
                to
                   protocol bgp-vpn
                exit
                action accept
                     next-hop loopback2
                exit
            exit
        exit
```

5. Configure BGP default MVPN export in the group as required:

```
configure router bgp group "mvpn" export "MVPN_CoreDiveristy_Exp"
```

6. Configure each VPRN to use correct IGP source address and correct VRF export policy:

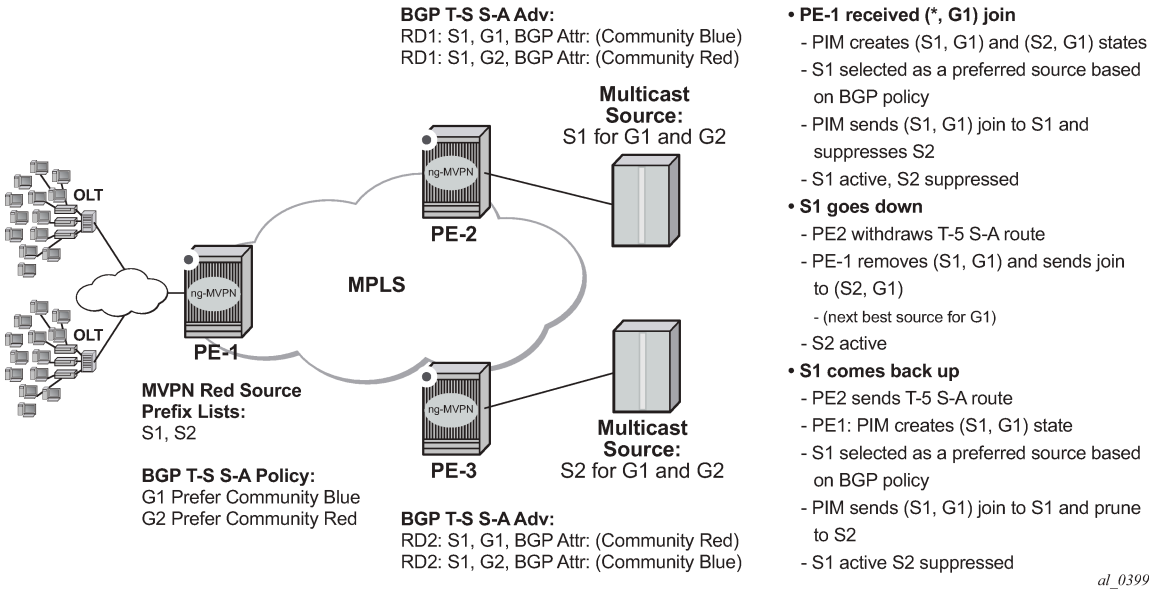
```
*A:Dut-A>config>service>vprn 2
      mvpn auto-discovery default source-address loopback1
      vrf-export "vprnexp2"
*A:Dut-A>config>service>vprn 3
      mvpn auto-discovery default source-address loopback2
      vrf-export "vprnexp3"
```

3.4.14 NG-MVPN multicast source geo-redundancy

Multicast source geo-redundancy is targeted primarily for MVPN deployments for multicast delivery services like IPTV. The solutions allows operators to configure a list of geographically dispersed redundant multicast sources (with different source IPs) and then, using configured BGP policies, ensure that each Receiver PE (a PE with receivers in its C-instance) selects only a single, most-preferred multicast source for a specific group from the list. Although the data may still be replicated in P-instance (each multicast source sends (C-S, C-G) traffic onto its I-PMSI tree or S-PMSI tree), each Receiver PE only forwards data to its receivers from the preferred multicast source. This allows operators to support multicast source geo-redundancy without the replication of traffic for each (C-S, C-G) in the C-instance while allowing fast recovery of service when an active multicast source fails.

Figure 43: Preferred source selection for multicast source geo-redundancy shows an operational example of multicast source geo-redundancy:

Figure 43: Preferred source selection for multicast source geo-redundancy



Operators can configure a list of prefixes for multicast source redundancy per MVPN on Receiver PEs:

- Up to 8 multicast source prefixes per VPRN are supported.
- Any multicast source that is not part of the source prefix list is treated as a unique source and automatically joined in addition to joining the most preferred source from the redundant multicast source list.

A Receiver PE selects a single, most-preferred multicast source from the list of pre-configured sources for a specific MVPN during (C-*, C-G) processing as follows:

- A single join for the group is sent to the most preferred multicast source from the operator-configured multicast source list. Joins to other multicast sources for a specific group are suppressed. Operator can see active and suppressed joins on a Receiver PE. Although a join is sent to a single multicast source only, (C-S, C-G) state is created for every source advertising Type-5 S-A route on the Receiver PE.
- The most preferred multicast source is a reachable source with the highest local preference for Type-5 SA route based on the BGP policy, as described later in this section.
- On a failure of the preferred multicast source or when a new multicast source with a better local preference is discovered, Receiver PE joins the new most-preferred multicast source. The outage experienced depends on how quickly Receiver PE receives Type-5 S-A route withdrawal or loses unicast route to multicast source, and how quickly the network can process joins to the newly selected preferred multicast sources.
- Local multicast sources on a Receiver PE are not subject to the most-preferred source selection, regardless of whether they are part of redundant source list or not.

BGP policy on Type-5 SA advertisements is used to determine the most preferred multicast source based on the best local preference as following:

- Each Source PE (a PE with multicast sources in its C-instance) tags Type-5 SA routes with a unique standard community attribute using global BGP policy or MVPN vrf-export policy. Depending on multicast topology, the policy may require source-aware tagging in the policy. Either all MVPN routes or Type 5 SA routes only can be tagged in the policy (new attribute **mvpn-type 5**).

- Each receiver PE has a BGP VRF import policy that sets local preference using match on Type-5 SA routes (new attribute *mvpn-type 5*) and standard community attribute value (as tagged by the Source PEs). Using policy statements that also include group address match, allows receiver PEs to select the best multicast source per group. The BGP VRF import policy must be applied as **vrf-import** under **config>service>vprn>mvpn** context. It must have default-action *accept* specified, or all MVPN routes other than those matched by specified entries are rejected. In addition, it must have **vrf-target** as a community match condition, because **vrf-target mvpn** configuration is ignored when **vrf-import** policy is defined.

Operators can change redundant source list or BGP policy affecting source selection in service. If such a change of the list/policy results in a new preferred multicast source election, make-before-break is used to join the new source and prune the previously best source.

For the correct operations, MVPN multicast source geo-redundancy requires the router:

- To maintain the list of eligible multicast sources on Receiver PEs, Source PE routers must generate Type-5 S-A route even if the Source PE sees no active joins from any receiver for a specific group.
- To trigger a switch from a currently active multicast source on a Receiver PE, Source PE routers must withdraw Type-5 S-A route when the multicast source fails or alternatively unicast route to multicast source must be withdrawn or go down on a Receiver PE.

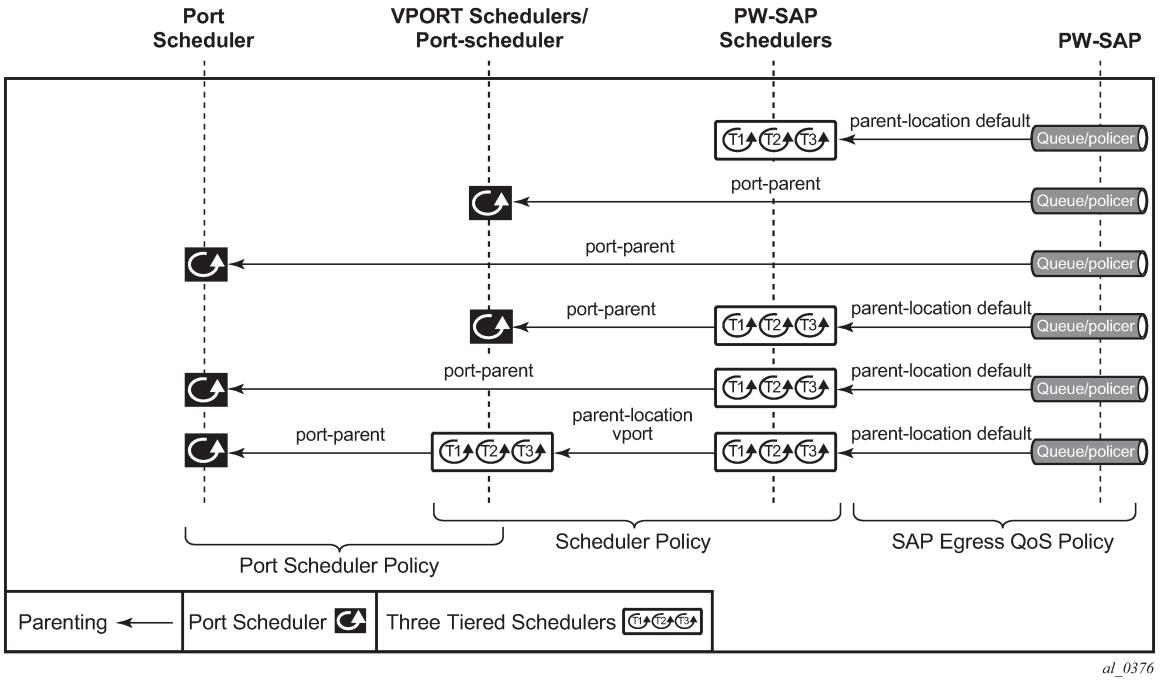
MVPN multicast source redundancy solutions is supported for the following configurations only. Enabling the feature in unsupported configuration must be avoided.

- NG-MVPN with RSVP-TE or mLDP or PIM with BGP c-multicast signaling in P-instance. Both I-PMSI and S-PMSI trees are supported.
- IPv4 and IPv6 (C-*, C-G) PIM ASM joins in the C-instance.
- Both **intersite-shared** *enabled* and *disabled* are supported. For **intersite-shared enabled**, operators must enable generation of Type-5 S-A routes even in the absence of receivers seen on Source PEs (intersite-shared persistent-type5-adv must be enabled).
- The Source PEs must be configured as a sender-receiver, the Receiver PEs can be configured as a sender-receiver or a receiver-only.
- The RPs must be on the Source PEs side. Static RP, anycast-RP, embedded-RP types are supported.
- UMH redundancy can be deployed to protect Source PE to any multicast source. When deployed, UMH selection is executed independently of source selection after the most preferred multicast source had been chosen. Supported **umh-selection** options include: **highest-ip**, **hash-based**, **tunnel-status** (not supported for IPv6), and **unicast-rt-pref**.

3.4.15 Multicast core diversity for Rosen MDT SAFI MVPNs

Figure 44: Multicast core diversity shows a Rosen MVPN core diversity deployment.

Figure 44: Multicast core diversity



Core diversity allows operators to optionally deploy multicast MVPN in either default IGP instance. or one of two non-default IGP instances to provide; for example, topology isolation or different level of services. The following describes the main feature attributes:

- Rosen MVPN IPv4 multicast with MDT SAFI is supported with default and data MDTs.
- Rosen MVPN can use a non-default OSPF or ISIS instance (using their loopback addresses instead of a system address).
- Up to 3 distinct core instances are supported: system + 2 non-default OSPF instances shown in Figure 44: Multicast core diversity.
- The BGP Connector also uses non-default OSPF loopback as NH, allowing Inter-AS Option B/C functionality to work with Core diversity as well.
- The feature is supported with CSC-VPRN.

On source PEs (PE1: UMH, PE2: UMH in the above picture), an MVPN is assigned to a non-default IGP core instance as follows:

- MVPN is statically pointed to use one of the non-default IGP instances loopback addresses as source address instead of system loopback IP.
- MVPN export policy is used to change unicast route next-hop VPN address.
- BGP Connector support for non-default instances.

The configuration shown above ensures that MDT SAFI and IP-VPN routes for the non-default core instance use non-default IGP loopback instead of system IP. This ensures PIM advertisement/joins run in the correct core instance and GRE tunnels for multicast can be set-up using and terminated on non-system IP. If BGP export policy is used to change unicast route next-hop VPN address instead of BGP Connector attribute-based processing and unicast traffic must be forwarded in non-default core instances 1 or 2, LDP or RSVP (terminating on non-system IP) must be used. GRE unicast traffic termination on non-system IP

is not supported and any GRE traffic arriving at the PE in instances 1 or 2, destined for non-default IGP loopback IP is forwarded to CPM (ACL or CPM filters can be used to prevent the traffic from reaching the CPM).

No configuration is required on non-source PEs.

Known feature restrictions include:

- VPRN instance must be shutdown to change the mdt-safi source-address. The CLI rollback that includes change of the auto-discovery is therefore service impacting.
- To reset mdt-safi source-address to system IP, operator must first execute no auto-discovery (or auto-discovery default) then auto-discovery mdt-safi
- Configuring system IP as a source-address consumes one of the 2 IP addresses allowed, therefore it should not be done.
- Operators must configure correct IGP instance loopback IP addresses within Rosen MVPN context and must configure correct BGP policies (before Release 12.0R4) for the feature to operate as expected. There is no verification that the address entered for MVPN provider tunnel source-address is such an address or is not a system IP address.

3.4.16 Inter-AS MVPN

The Inter-AS MVPN feature allows set-up of Multicast Distribution Trees (MDTs) that span multiple Autonomous Systems (ASes).This section focuses on multicast aspects of the Inter-AS MVPN solution.

To support Inter-AS option for MVPNs, a mechanism is required that allows setup of Inter-AS multicast tree across multiple ASes. Because of limited routing information across AS domains, it is not possible to setup the tree directly to the source PE. Inter-AS VPN Option A does not require anything specific to inter-AS support as customer instances terminate on ASBR and each customer instance is handed over to the other AS domain via a unique instance. This approach allows operators to provide full isolation of ASes, but the solution is the least scalable case, as customer instances across the network have to exist on ASBR.

Inter-AS MVPN Option B allows operators to improve upon the Option A scalability while still maintaining AS isolation, while Inter-AS MVPN Option C further improves Inter-AS scale solution but requires exchange of Inter-AS routing information and therefore is typically deployed when a common management exists across all ASes involved in the Inter-AS MVPN. The following sub-sections provide further details on Inter-AS Option B and Option C functionality.

3.4.16.1 BGP connector attribute

BGP connector attribute is a transitive attribute (unchanged by intermediate BGP speaker node) that is carried with VPNv4 advertisements. It specifies the address of source PE node that originated the VPNv4 advertisement.

With Inter-AS MVPN Option B, BGP next-hop is modified by local and remote ASBR during re-advertisement of VPNv4 routes. On BGP next-hop change, information about the originator of prefix is lost as the advertisement reaches the receiver PE node.

BGP connector attribute allows source PE address information to be available to receiver PE, so that a receiver PE is able to associate VPNv4 advertisement to the corresponding source PE.

3.4.16.2 PIM RPF vector

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 ASes. Routers in an AS other than that of a source PE, have no routes available to reach the source PE and therefore PIM JOINs would never be sent upstream. To enable setup of MDT toward a source PE, BGP next-hop (ASBR) information from that PE's MDT-SAFI advertisement is used to fake a route to the PE. If the BGP next-hop is a PIM neighbor, the PIM JOINs would be sent upstream. Otherwise, the PIM JOINs would be sent to the immediate IGP next-hop (P) to reach the BGP next-hop. Because the IGP next-hop does not have a route to source PE, the PIM JOIN would not be propagated forward unless it carried extra information contained in RPF Vector.

In case of Inter-AS MVPN Option C, unicast routing information toward the source PE is available in a remote AS PEs/ASBRs as BGP 8277 tunnels, but unavailable at remote P routers. If the tunneled next-hop (ASBR) is a PIM neighbor, the PIM JOINs would be sent upstream. Otherwise, the PIM JOINs would be sent to the immediate IGP next-hop (P) to reach the tunneled next-hop. Because the IGP next-hop does not have a route to source PE, the PIM JOIN would not be propagated forward unless it carried extra information contained in RPF Vector.

To enable setup of MDT toward a source PE, PIM JOIN therefore carries BGP next hop information in addition to source PE IP address and RD for this MVPN. For option-B, both these pieces of information are derived from MDT-SAFI advertisement from the source PE. For option-C, both these pieces of information are obtained from the BGP tunneled route.

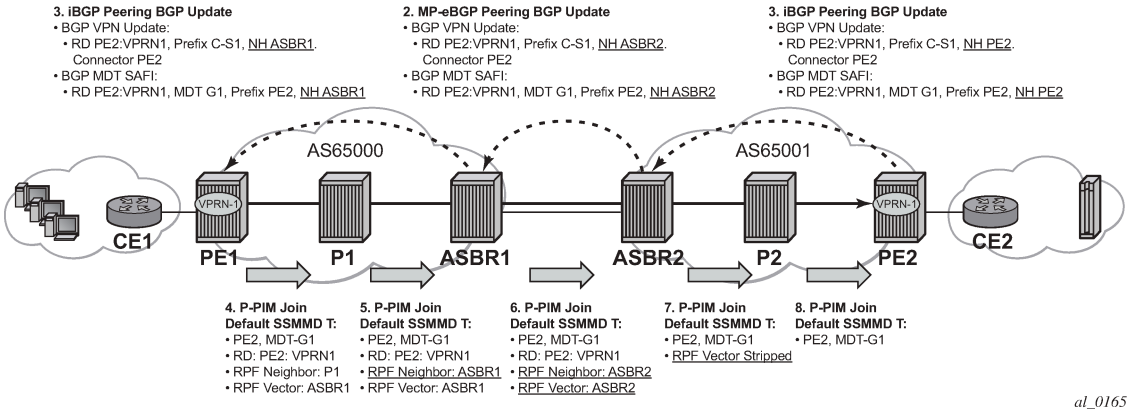
The RPF vector is added to a PIM join at a PE router when configure router **pim rpfv** option is enabled. P routers and ASBR routers must also have the option enabled to allow RPF Vector processing. If the option is not enabled, the RPF Vector is dropped and the PIM JOIN is processed as if the PIM Vector were not present.

For further details about RPF Vector processing please see [RFCs 5496, 5384 and 6513]

3.4.16.3 Inter-AS MVPN Option B

Inter-AS Option B is supported for Rosen MVPN PIM SSM using BGP MDT SAFI, PIM RPF Vector and BGP Connector attribute. The Figure 45: Inter-AS Option B default MDT setup depict set-up of a default MDT:

Figure 45: Inter-AS Option B default MDT setup



SR OS inter-AS Option B is designed to be standard compliant based on the following RFCs:

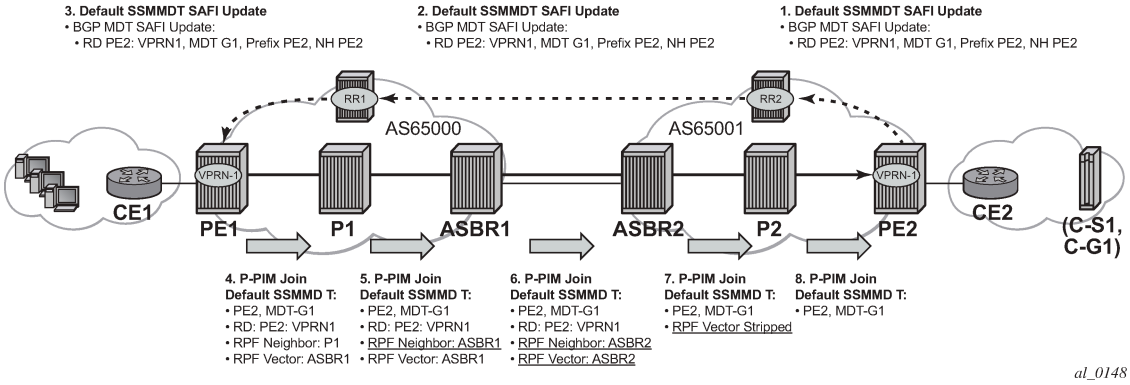
RFC 5384	The Protocol Independent Multicast (PIM) Join Attribute Format
RFC 5496	The Reverse Path Forwarding (RPF) Vector TLV
RFC 6513	Multicast in MPLS/BGP IP VPNs

The SR OS implementation was designed also to interoperate with older routers Inter-AS implementations that do not comply with the RFC 5384 and RFC 5496.

3.4.16.4 Inter-AS MVPN Option C

Inter-AS Option C is supported for Rosen MVPN PIM SSM using BGP MDT SAFI and PIM RPF Vector. Figure 46: Inter-AS Option C default MDT setup depicts a default MDT setup:

Figure 46: Inter-AS Option C default MDT setup



Additional restrictions for Inter-AS MVPN Option B and C support are the following:

- Inter-AS MVPN Option B is not supported with duplicate PE addresses.
- For Inter-AS Option C, BGP 8277 routes are installed into unicast rtm (rtable-u), unless routes are installed by some other means into multicast rtm (rtable-m), and Option C does not build core MDTs, therefore, rpf-table is configured to rtable-u or both.

Additional Cisco interoperability notes are the following:

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 SR OS implementation was designed to inter-operate with Cisco routers Inter-AS implementations that do not comply with the RFC5384 and RFC5496.

When **configure router pim rpfv mvpn** option is enabled, Cisco routers need to be configured to include RD in an RPF vector using the following command: **ip multicast vrf vrf-name rpf proxy rd vector** for interoperability. When Cisco routers are not configured to include RD in an RPF vector, operator should configure SR OS router (if supported) using **configure router pim rpfv core mvpn**: PIM joins received can be a mix of core and mvpn RPF vectors.

3.4.16.5 NG-MVPN non-segmented inter-AS solution

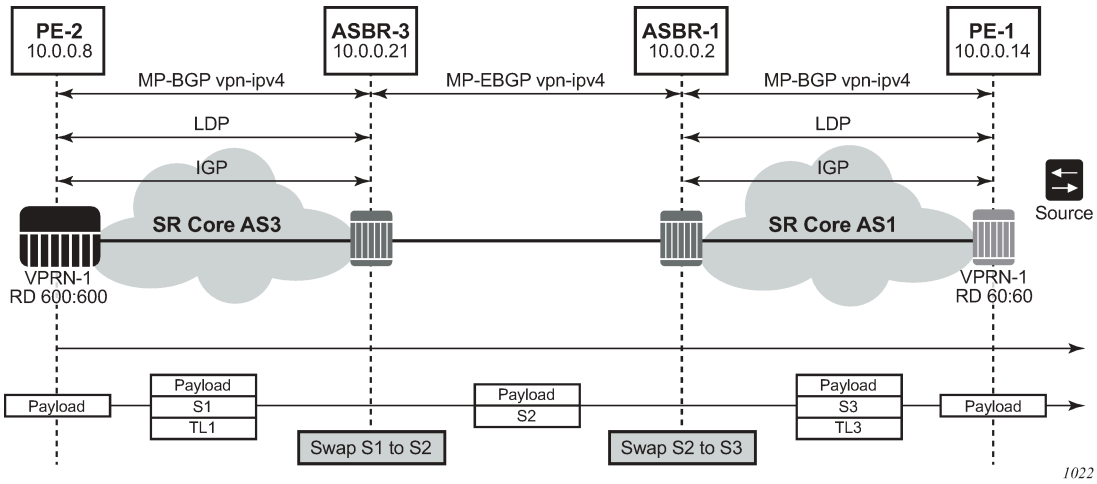
This feature allows multicast services to use segmented protocols and span them over multiple autonomous systems (ASs), as done in unicast services. As IP VPN or GRT services span multiple IGP areas or multiple ASs, either because of a network designed to deal with scale or as result of commercial acquisitions, operators may require Inter-AS VPN (unicast) connectivity. For example, an Inter-AS VPN can break the IGP, MPLS and BGP protocols into access segments and core segments, allowing higher scaling of protocols by segmenting them into their own islands. SR OS also allows for similar provision of multicast services and for spanning these services over multiple IGP areas or multiple ASs.

For multicast VPN (MVPN), SR OS previously supported Inter-AS Model A/B/C for Rosen MVPN; however, when MPLS was used, only Model A was supported for Next Generation Multicast VPN (NG-MVPN) and d-mLDP signaling.

For unicast VPRNs, the Inter-AS or Intra-AS Option B and C breaks the IGP, BGP and MPLS protocols at ABR routers (in case of multiple IGP areas) and ASBR routers (in case of multiple ASs). At ABR and ASBR routers, a stitching mechanism of MPLS transport is required to allow transition from one segment to next, as shown in Figure 47: Unicast VPN Option B with segmented MPLS and Figure 48: Unicast VPN Option C with segmented MPLS.

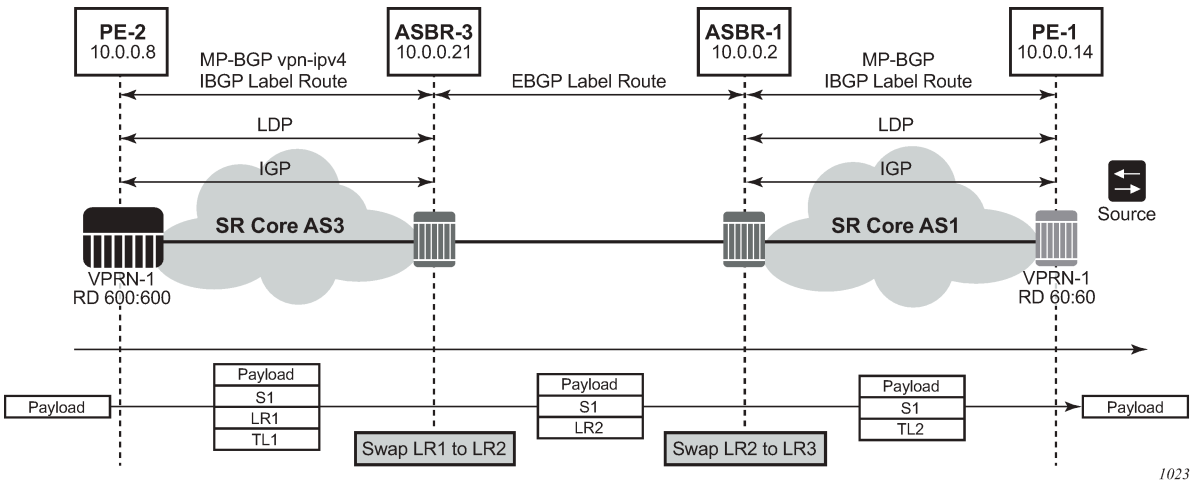
In Figure 47: Unicast VPN Option B with segmented MPLS, the Service Label (S) is stitched at the ASBR routers.

Figure 47: Unicast VPN Option B with segmented MPLS



In Figure 48: Unicast VPN Option C with segmented MPLS, the 8277 BGP Label Route (LR) is stitched at ASBR1 and ASBR3. At ASBR1, the LR1 is stitched with LR2, and at ASBR3, the LR2 is stitched with TL2.

Figure 48: Unicast VPN Option C with segmented MPLS



Previously, in case of NG-MVPN, segmenting an LDP MPLS tunnel at ASBRs or ABRs was not possible. As such, RFC 6512 and 6513 used a non-segmented mechanism to transport the multicast data over P-tunnels end-to-end through ABR and ASBR routers. The signaling of LDP needed to be present and possible between two ABR routers or two ASBR routers in different ASs.

Note: For unicast VPNs, it was usually preferred to only have EBGP between ASBR routers. The non-segmented behavior of d-mLDP would have broken this by requiring LDP signaling between ASBR routers.

SR OS now has d-mLDP non-segmented intra-AS and inter-AS signaling for NG-MVPN and GRT multicast. The non-segmented solution for d-mLDP is possible for inter-ASs as Option B and C.

3.4.16.5.1 Non-segmented d-mLDP and inter-AS VPN

There are three types of VPN Inter-AS solutions:

- Inter-AS Option A
- Inter-AS Option B
- Inter-AS Option C

Options B and C use recursive opaque types 8 and 7 respectively, from Table 12: Recursive opaque types
.

Table 12: Recursive opaque types

Opaque type	Opaque name	RFC	SR OS use
1	Basic Type	RFC 6388	VPRN Local AS
3	Transit IPv4	RFC 6826	IPv4 multicast over mLDP in GRT

Opaque type	Opaque name	RFC	SR OS use
4	Transit IPv6	RFC 6826	IPv6 multicast over mLDP in GRT
7	Recursive Opaque (Basic Type)	RFC 6512	Inter-AS Option C MVPN over mLDP
8	Recursive Opaque (VPN Type)	RFC 6512	Inter-AS Option B MVPN over mLDP

3.4.16.5.1.1 Inter-AS Option A

In Inter-AS Option A, ASBRs communicate using VPN access interfaces, which need to be configured under PIM for the two ASBRs to exchange multicast information.

3.4.16.5.1.2 Inter-AS Option B

The recursive opaque type used for Inter-AS Option B is the Recursive Opaque (VPN Type), shown as opaque type 8 in Table 12: Recursive opaque types .

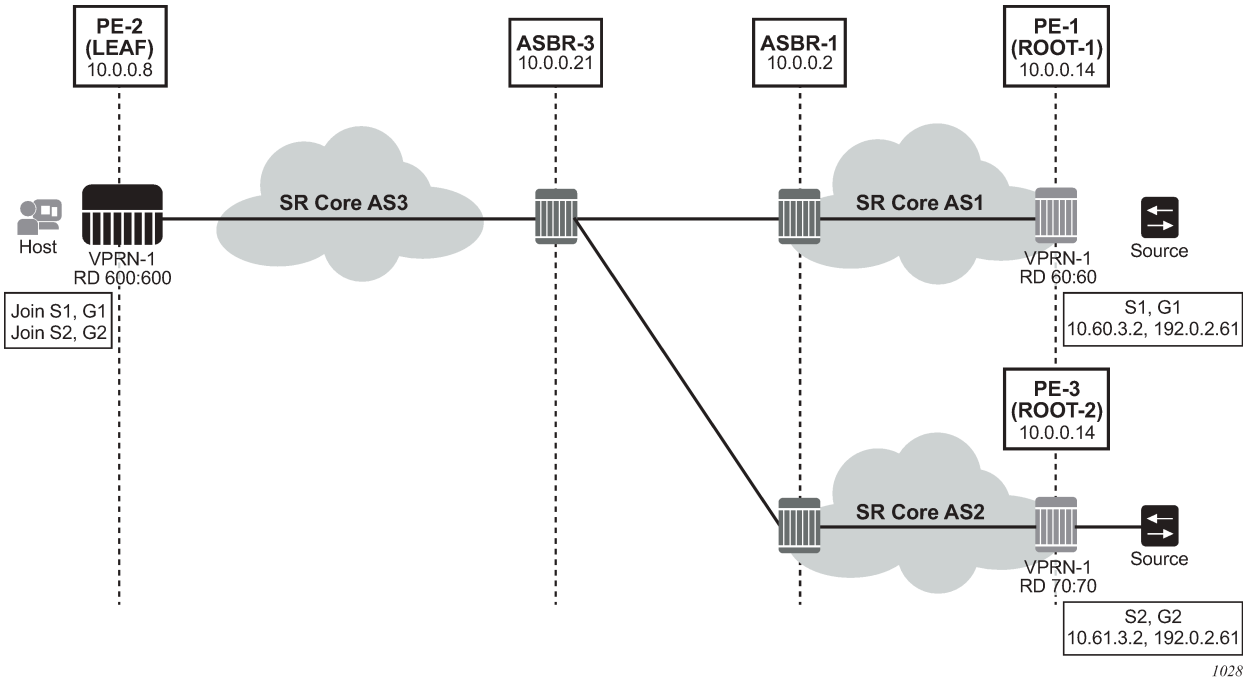
Inter-AS Option B Support for NG-MVPN

Inter-AS Option B requires additional processing on ASBR routers and recursive FEC encoding than that of Inter-AS Option A. Because BGP adjacency is not e2e, ASBRs must cache and use a PMSI route to build the tree. For that, mLDP recursive FEC must carry RD information—therefore, VPN recursive FEC is required (opaque type 8).

In Inter-AS Option B, the PEs in two different ASs do not have their system IP address in the RTM. As such, for NG-MVPN, a recursive opaque value in mLDP FEC is required to signal the LSP to the first ASBR in the local AS path.

Because the system IPs of the peer PEs (Root-1 and Root-2) are not installed on the local PE (leaf), it is possible to have two PEs in different ASs with same system IP address, as shown in Figure 49: Identical system IP on multiple PEs (Option B). However, SR OS does not support this topology. The system IP address of all nodes (root or leaf) in different ASs must be unique.

Figure 49: Identical system IP on multiple PEs (Option B)



For inter-AS Option B and NG-MVPN, SR OS as a leaf does not support multiple roots in multiple ASs with the same system IP and different RDs; however, the first root that is advertised to an SR OS leaf is used by PIM to generate an MLDP tunnel to this actual root. Any dynamic behavior after this point, such as removal of the root and its replacement by a second root in a different AS, is not supported and the SR OS behavior is nondeterministic.

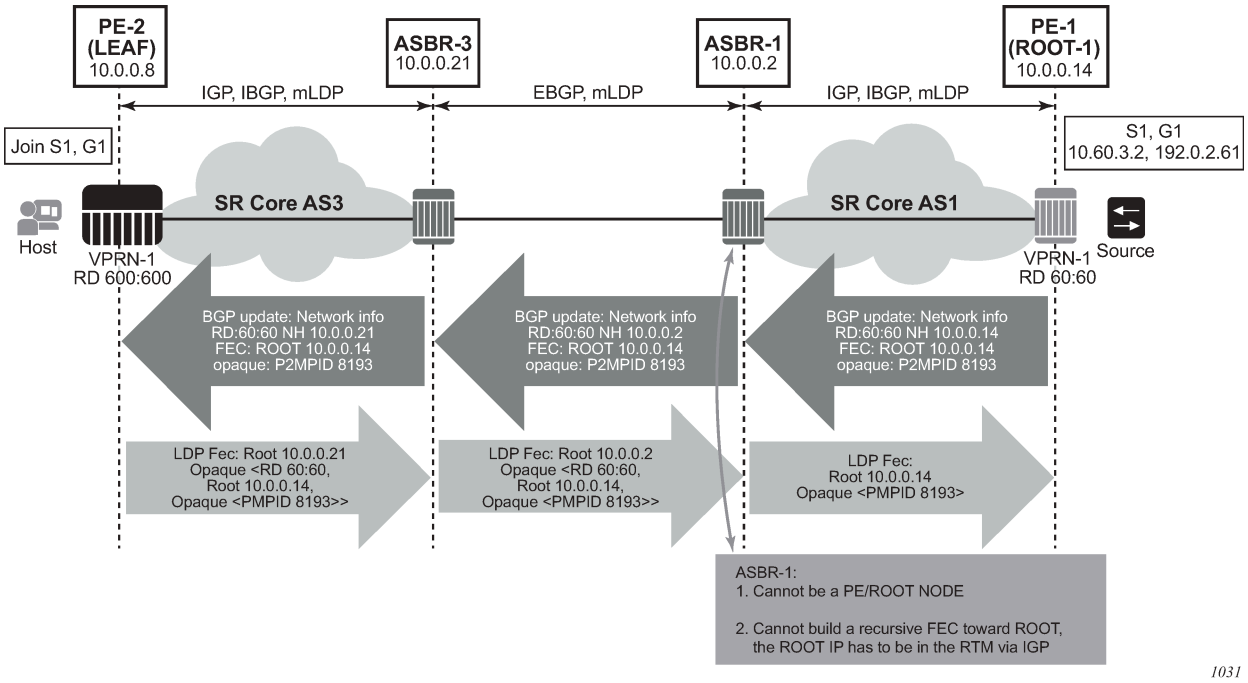
I-PMSI and S-PMSI Establishment

I-PMSI and S-PMSI functionality follows RFC 6513 section 8.1.1 and RFC 6512 sections 3.1 and 3.2.1. For routing, the same rules as for GRT d-mLDP use case apply, but the VRR Route Import External community now encodes the VRF instance in the local administrator field.

Option B uses an outer opaque of type 8 and inter opaque of type 1 (see Table 12: Recursive opaque types).

Figure 50: Non-segmented mLDP PMSI establishment (Option B) depicts the processing required for I-PMSI and S-PMSI Inter-AS establishment.

Figure 50: Non-segmented mLDP PMSI establishment (Option B)



For non-segmented mLDP trees, A-D procedures follow those of the Intra-AS model, with the exception that NO EXPORT community must be excluded; LSP FEC includes mLDP VPN-recursive FEC.

For I-PMSI on Inter-AS Option B:

- A-D routes must be installed by ASBRs and next-hop information is changed as the routes are propagated, as shown in Figure 50: Non-segmented mLDP PMSI establishment (Option B).
- PMSI A-D routes are used to provide inter-domain connectivity on remote ASBRs.

On a receipt of an Intra-AS PMSI A-D route, PE2 resolves PE1's address (next-hop in PMSI route) to a labeled BGP route with a next-hop of ASBR3, because PE1 (Root-1) is not known via IGP. Because ASBR3 is not the originator of the PMSI route, PE2 sources an mLDP VPN recursive FEC with a root node of ASBR3, and an opaque value containing the information advertised by Root-1 (PE-1) in the PMSI A-D route, shown below, and forwards the FEC to ASBR 3 using IGP.

PE-2 LEAF FEC: (Root ASBR3, Opaque value {Root: ROOT-1, RD 60:60, Opaque Value: P2MPLSP-ID xx}}

When the mLDP VPN-recursive FEC arrives at ASBR3, it notes that it is the identified root node, and that the opaque value is a VPN-recursive opaque value. Because Root-1 PE1 Is not known via IGP, ASBR3 resolves the root node of the VPN-Recursive FEC using PMSI A-D (I or S) matching the information in the VPN-recursive FEC (the originator being PE1 (Root-1), RD being 60:60, and P2MP LSP ID xx). This yields ASBR1 as next hop. ASBR3 creates a new mLDP FEC element with a root node of ASBR1, and an opaque value being the received recursive opaque value, as shown below. ASBR then forwards the FEC using IGP.

ASBR-3 FEC: {Root ASBR 1, Opaque Value {Root: ROOT-1, RD 60:60, Opaque Value: P2MPLSP-ID xx}}

When the mLDP FEC arrives at ASBR1, it notes that it is the root node and that the opaque value is a VPN-recursive opaque value. As PE1's ROOT-1 address is known to ASBR1 through the IGP, no further recursion is required. Regular processing begins, using received Opaque mLDP FEC information.

ASBR-1 FEC: {Root: ROOT-1, Opaque Value: P2MP LSP-ID xx

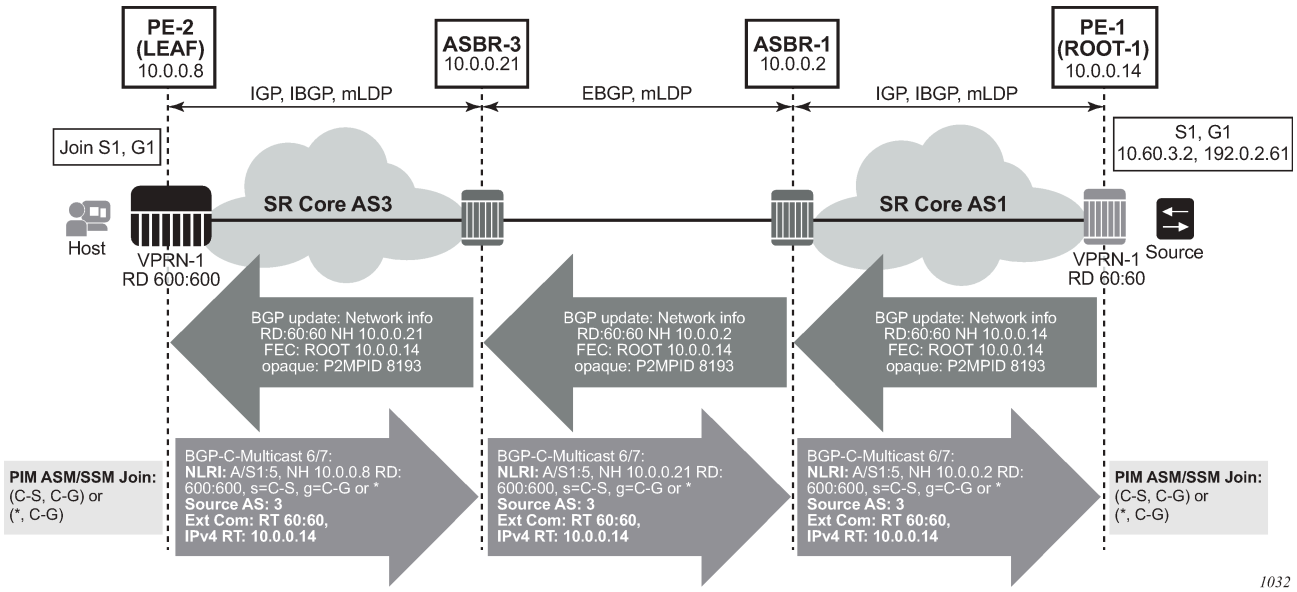
Note: VPN-Recursive FEC carries P2MPLSP ID. The P2MPLSP ID is used in addition to PE RD and Root to select a route to the mLDP root using the correct I-PMSI or S-PMSI route.

The functionality as described above for I-PMSI applies also to S-PMSI and (C-*, C-*) S-PMSI.

C-multicast Route Processing

C-multicast route processing functionality follows RFC 6513 section 8.1.2 (BGP used for route exchange). The processing is analogous to BGP Unicast VPN route exchange described in Figure 47: Unicast VPN Option B with segmented MPLS and Figure 48: Unicast VPN Option C with segmented MPLS. Figure 51: Non-segmented mLDP C-multicast exchange (Option B) shows C-multicast route processing with non-segmented mLDP PMSI details.

Figure 51: Non-segmented mLDP C-multicast exchange (Option B)



3.4.16.5.1.3 Inter-AS Option C

In Inter-AS Option C, the PEs in two different ASs have their system IP address in the RTM, but the intermediate nodes in the remote AS do not have the system IP of the PEs in their RTM. As such, for NG-MVPN, a recursive opaque value in mLDP FEC is needed to signal the LSP to the first ASBR in the local AS path.

The recursive opaque type used for Inter-AS Option B is the Recursive Opaque (Basic Type), shown as opaque type 7 in Table 12: Recursive opaque types .

3.4.16.5.1.3.1 Inter-AS Option C support for NG-MVPN

For Inter-AS Option C, on a leaf PE, a route exists to reach root PE's system IP and, as ASBRs can use BGP unicast routes, recursive FEC processing using BGP unicast routes, and not VPN recursive FEC processing using PMSI routes, is required.

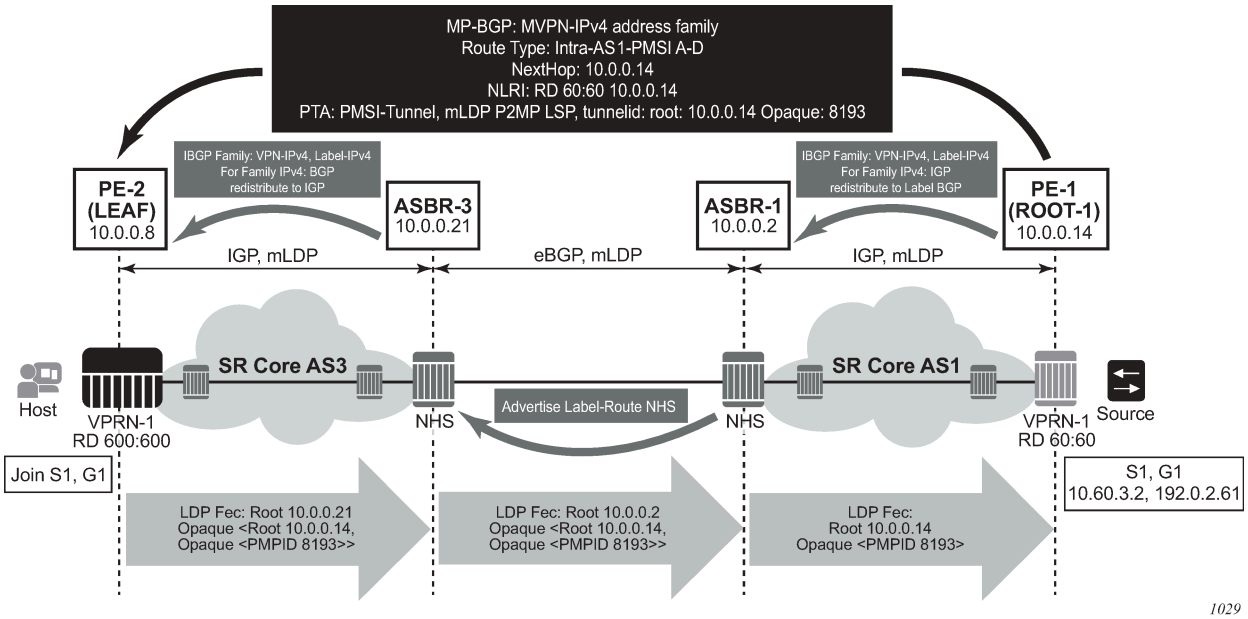
I-PMSI and S-PMSI establishment

I-PMSI and S-PMSI functionality follows RFC 6513 section 8.1.1 and RFC 6512 Section 2. The same rules as per the GRT d-mLDP use case apply, but the VRR Route Import External community now encodes the VRF instance in the local administrator field.

Option C uses an outer opaque of type 7 and inter opaque of type 1.

Figure 52: Non-segmented mLDP PMSI establishment (Option C) shows the processing required for I-PMSI and S-PMSI Inter-AS establishment.

Figure 52: Non-segmented mLDP PMSI establishment (Option C)



For non-segmented mLDP trees, A-D procedures follow those of the Intra-AS model, with the exception that NO EXPORT Community must be excluded; LSP FEC includes mLDP recursive FEC (and not VPN recursive FEC).

For I-PMSI on Inter-AS Option C:

- A-D routes are not installed by ASBRs and next-hop information is not changed in MVPN A-D routes.
- BGP-labeled routes are used to provide inter-domain connectivity on remote ASBRs.

On a receipt of an Intra-AS I-PMSI A-D route, PE2 resolves PE1's address (N-H in PMSI route) to a labeled BGP route with a next-hop of ASBR3, because PE1 is not known via IGP. PE2 sources an mLDP FEC with a root node of ASBR3, and an opaque value, shown below, containing the information advertised by PE1 in the I-PMSI A-D route.

PE-2 LEAF FEC: {root = ASBR3, opaque value: {Root: ROOT-1, opaque value: P2MP-ID xx}}

When the mLDP FEC arrives at ASBR3, it notes that it is the identified root node, and that the opaque value is a recursive opaque value. ASBR3 resolves the root node of the Recursive FEC (ROOT-1) to a labeled BGP route with the next-hop of ASBR1, because PE-1 is not known via IGP. ASBR3 creates a new mLDP FEC element with a root node of ASBR1, and an opaque value being the received recursive opaque value.

ASBR3 FEC: {root: ASBR1, opaque value: {root: ROOT-1, opaque value: P2MP-ID xx}}

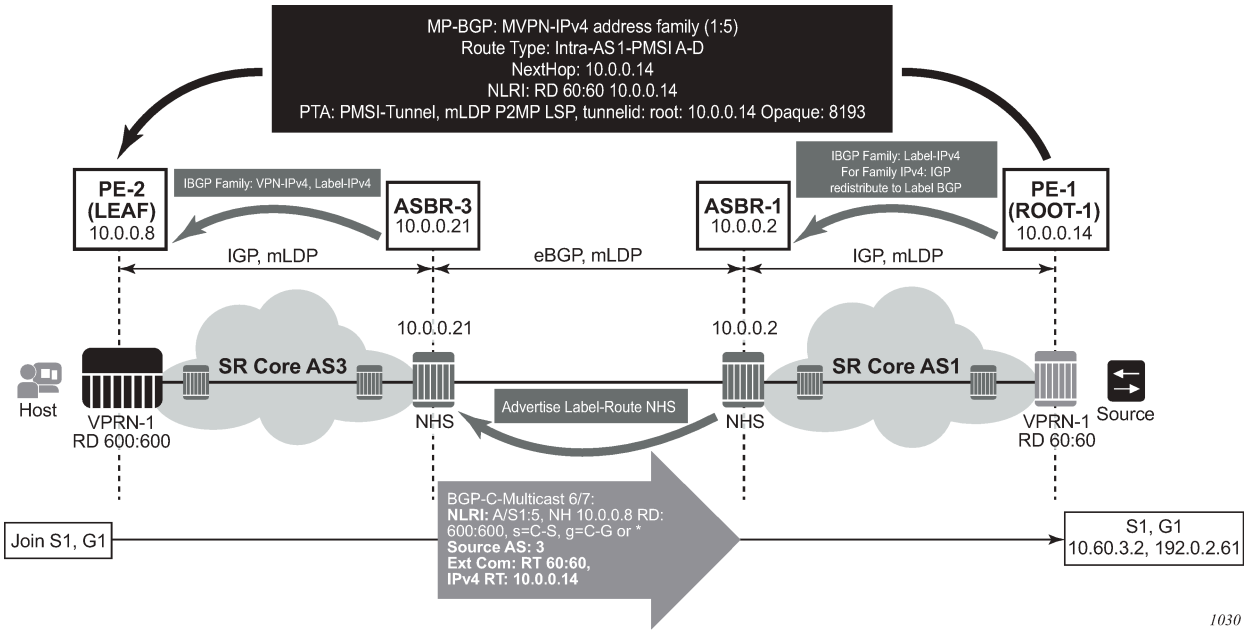
When the mLDP FEC arrives at ASBR1, it notes that it is the root node and that the opaque value is a recursive opaque value. As PE-1's address is known to ASBR1 through the IGP, no further recursion is required. Regular processing begins, using the received Opaque mLDP FEC information.

The functionality as described above for I-PMSI applies to S-PMSI and (C-*, C-*) S-PMSI.

C-multicast route processing

C-multicast route processing functionality follows RFC 6513 section 8.1.2 (BGP used for route exchange). The processing is analogous to BGP Unicast VPN route exchange. Figure 53: Non-segmented mLDP C-multicast exchange (Option C) shows C-multicast route processing with non-segmented mLDP PMSI details.

Figure 53: Non-segmented mLDP C-multicast exchange (Option C)



LEAF node cavities

Caution: The SR OS ASBR does not currently support receiving a non-recursive opaque FEC (opaque type 1).

The LEAF (PE-2) has to have the ROOT-1 system IP installed in RTM via BGP. If the ROOT-1 is installed in RTM via IGP, the LEAF does not generate the recursive opaque FEC. As such, the ASBR 3 does not process the LDP FEC correctly.

3.4.16.5.2 Configuration example

No configuration is required for Option B or Option C on ASBRs, although for Option B, **config>router>bgp>enable-inter-as-vpn** is required to enable inter-as-non-segmented MLDP through the ASBR router.

Policy is required for a root or leaf PE for removing the NO_EXPORT community from MVPN routes, which can be configured using an export policy on the PE.

The following is an example for configuring a policy on PEs to remove the **no-export**:

```
*A:Dut-A>config>router>policy-options# info
----------------------------------------------
            community "no-export" members "no-export"
            policy-statement "remNoExport"
                default-action accept
                    community remove "no-export"
                exit
            exit
----------------------------------------------
*A:Dut-A>config>router>policy-options#
```

The following is an example for configuring in BGP the policy in a global, group, or peer context:

```
*A:Dut-A>config>router>bgp# info
----------------------------------------------
            vpn-apply-export
            export "remNoExport"
```

3.4.16.5.3 Inter-AS non-segmented MLDP

See the "Inter-AS Non-segmented MLDP" section of the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide* for more information.

3.4.16.5.4 ECMP

See the "ECMP" section of the *7450 ESS, 7750 SR, 7950 XRS, and VSR MPLS Guide* for more information about ECMP.

3.4.17 mLDP non-Segmented intra-AS (inter-area) MVPN solution

SR OS now supports intra-AS (Inter-Area) option B and C. The following interaction between inter and intra is as follows:

- Intra-AS option B with inter-AS option B
- Intra-AS option C with inter-AS option C

3.4.17.1 Intra-AS and inter-AS Option B

For intra/inter-as option B, the root is not visible on the leaf. LDP is responsible for building the recursive FEC and signaling the FEC to ABR/ASBR on the leaf. ABR/ASBR must have the PMSI AD router to re-build the FEC (recursive or basic) depending on whether they are connected to another ABR/ASBR or root node. As such, LDP must import the MVPN PMSI AD routes. To save resources, importing MVPN PMSI AD routes are performed manually by the operator using configuration commands. When **mvpn-no-export-community** command is enabled, LDP requests BGP to provide the LDP task with all the MVPN PMSI AD routes and LDP internally caches these routes. If this knob is disabled and to save resources, MVPN discards the catch routes.

In a scenario where a node running an older image that does not support the **mvpn-no-export-community** command and that node upgrades to a new image that does support the **mvpn-no-export-community** command, and if the older image had an inter-as MVPN configuration, then after the upgrade to the newer image, the **mvpn-no-export-community** command is enabled by default, to ensure a smooth upgrade. This is to verify that all the routes are imported to mLDP so the inter-as functionality works after the upgrade.

In addition, SR OS supports two major upgrades to enable the MVPN command if an upgrade to a load supporting this knob.

3.4.17.2 MVPN next hop self on ABRs

For option B, the ABR routers must change the next-hop of MVPN AD routes to be the ABR systemIP or the loopback IP for core diversity. Currently, the **next-hop-self** BGP command does not change the next hop of the MVPN AD routes. This functionality will be available in a future release.

In the meantime, a BGP policy can be used to change the MVPN AD routes next hop at the ABR.

3.4.17.2.1 MVPN next-hop-self policy example

MVPN Type 1 route (intra-AS IPMSI AD route) and MVPN Type 3 (S-PMSI AD route) must have a policy to set their next hop to be the ABR systemIP. In the following example, the ABR systemIP is 10.20.1.4 with the same token as the unicast vpn-ipv4 family and can be configured within the policy to have the next hop changed to the ABR systemIP.

Configure three policies on all ABRs:

- a policy to change mvpn-ipv4 IntraAD Route Type 1 next hop to next-hop-self
- a policy to change vpn-ipv4 next hop to next-hop-self
- a policy to change mvpn-ipv4 IntraAD Route Type 3 to next-hop-self

```
*A:Dut-D>config>router>policy-options# info
----------------------------------------------
            policy-statement "mod_nh_10.20.1.4"
                entry 1
                    from
                        mvpn-type 1
                    exit
                    action accept
                        next-hop 10.20.1.4
                    exit
                exit
```

```
                default-action next-policy
                exit
            exit
            policy-statement "mod_nh_vpn_10.20.1.4"
                entry 1
                    from
                        family vpn-ipv4
                    exit
                    action accept
                        next-hop 10.20.1.4
                    exit
                exit
                default-action next-policy
                exit
            exit
            policy-statement "mod_nh_spmsi_10.20.1.4"
                entry 1
                    from
                        mvpn-type 3
                    exit
                    action accept
                        next-hop 10.20.1.4
                    exit
                exit
                default-action next-policy
                exit
            exit
-------------------------------------------
```

3.4.17.2.2 LDP configuration example

Under the LDP configuration, all ABR and non-ABR routers must enable **import- pmsi-routes mvpn** and **mvpn-no-export-community** to import all inter-AS and intra-AS (inter-area) routes.

```
A:Dut-B>config>router>ldp# info
-------------------------------------------
            fast-reroute
            mp-mbb-time 10
            generate-basic-fec-only
            import-pmsi-routes
                mvpn
                mvpn-no-export-community
            exit
```

3.4.17.2.3 BGP configuration example

The MVPN AD **next-hop-self** policies and the **vpn-ipv4** policy must be imported under BGP on ABR routers.

In addition, for unicast vpn-ipv4 connectivity, the **enable-inter-as-vpn** command must be configured under BGP.

```
A:Dut-B>config>router>bgp# info
-------------------------------------------
            family ipv4 ipv6 vpn-ipv4 vpn-ipv6 mvpn-ipv4 mcast-vpn-ipv4 mvpn-ipv6 mcast-vpn-
ipv6
            vpn-apply-import
            vpn-apply-export
```

```
            connect-retry 10
            keepalive 10
            hold-time 30
            enable-inter-as-vpn
            rapid-update vpn-ipv4 vpn-ipv6 mvpn-ipv4 mcast-vpn-ipv4 mvpn-ipv6 mcast-vpn-ipv6
            group "ibgp_A"
                next-hop-self
                cluster 10.20.1.2
                export "mod_nh_10.20.1.2" "mod_nh_spmsi_10.20.1.2" "mod_nh_vpn_10.20.1.2"
                neighbor 10.20.1.1
                    local-address 10.20.1.2
                    med-out 100
                    peer-as 100
                exit
            exit
            group "ibgp_D"
                next-hop-self
                cluster 10.180.4.2
                export "mod_nh_10.20.1.2" "mod_nh_spmsi_10.20.1.2" "mod_nh_vpn_10.20.1.2"
```

3.4.18 Weighted ECMP and ECMP for VPRN IPv4 and IPv6 over MPLS LSPs

ECMP over MPLS LSPs for VPRN services refers to spraying packets across multiple named RSVP and SR-TE LSPs within the same ECMP set.

The ECMP-like spraying consists of hashing the relevant fields in the header of a labeled packet and selecting the next-hop tunnel based on the modulo operation of the output of the hash and the number of ECMP tunnels. The maximum number of ECMP tunnels selected from the TTM matches the value of the user-configured **ecmp** option. Only LSPs with the same lowest LSP metric can be part of the ECMP set. If the number of these LSPs is higher than the value configured in the **ecmp** option, the LSPs with the lowest tunnel IDs are selected first.

In weighted ECMP, the load-balancing weight of the LSP is normalized by the system and then used to bias the amount of traffic forwarded over each LSP. The weight of the LSP is configured using the **config**>**router**>**mpls**>**lsp**>**load-balancing-weight weight** and **config**>**router**>**mpls**>**lsp-template**>**load-balancing-weight weight** commands.

If one or more LSPs in the ECMP set have no **load-balancing-weight** configured, and the ECMP is set to a specific next hop, regular ECMP spraying is used.

Weighted ECMP is configured for VPRN services with SDP auto-bind by using the **config**>**service**>**vprn**>**auto-bind-tunnel**>**ecmp** *max-ecmp-routes* and **config**>**service**>**vprn**>**auto-bind-tunnel**>**weighted-ecmp** commands. Weighted ECMP is disabled by default.

The **ecmp max-ecmp-routes** command allows explicit configuration of the number of tunnels that **auto-bind-tunnel** can use to resolve for a VPRN. The **max-ecmp-routes** parameter range is 1 to 32.

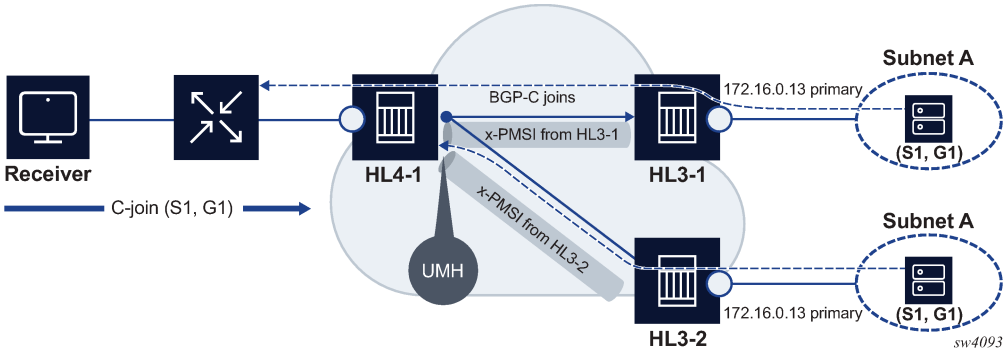
If weighted ECMP is enabled, then a path is selected based on the output of the hashing algorithm. Packet paths are then mapped to LSPs in the SDP in proportion to the configured load-balancing weight of the LSP. The hash is based on the system load-balancing configuration.

3.4.19 UMH redundancy using bandwidth monitoring

Bandwidth monitoring is used in MVPN for NG-MVPN and mLDP transport. It is used for multicast source redundancy, where both sources have the same IP address but are connected to two different root nodes. Bandwidth monitoring can be used with basic or recursive mLDP FEC. Upstream Multicast Hop (UMH)

redundancy for bandwidth monitoring is supported for mLDP basic FEC and recursive FEC type 7 and 8 only.

Figure 54: Bandwidth monitoring



With bandwidth monitoring, the leaf node sends a single (S1,G1) join to both root nodes. PIM SSM and ASM can be used between the receiver and the leaf, or between the UMH and the source. For ASM, bandwidth monitoring works only when traffic is switched from <*,G> to <S,G>.

After the source starts the multicast flow toward the root PEs, both root nodes transport the traffic to the leaf node on the PMSI (I-PMSI or S-PMSI).

The leaf listens to the active PMSI, blocks the other PMSI, and monitors the traffic rate on both the active and inactive PMSI. For faster than 50 ms switchover, both the active and the inactive PMSIs must arrive on the same IOM, because a single IOM must make the decision about which PMSI the leaf listens to and which PMSI to block.

The threshold for the rate of traffic lost between the active PMSI and the inactive PMSI is configured on the leaf PE. If the rate exceeds the configured value, the traffic switches from the active PMSI to the inactive PMSI. Rate monitoring is per PMSI, and not per (C-S,C-G).

After the active PMSI traffic rate is stored, there is a revertive behavior, which has a configurable timer. The revertive timer starts after the active PMSI traffic is recovered. When the timer expires and the primary PMSI traffic is stable, the traffic is switched back to the primary path. If the traffic goes below the threshold while the timer is decrementing, the timer is reset. This feature supports 1K of PMSI switchovers within 50 ms.

3.4.19.1 Fault recovery mitigation at PMSI switchover time

<S,G> switching between I-PMSI and S-PMSI is not symmetrical (synchronized in time) on the active and the inactive UMH. While the active UMH attempts to switch an <S,G> between I-PMSI and S-PMSI, the active PMSI traffic rate arriving from the active UMH may be different from that arriving from the inactive UMH. This asymmetrical behavior can generate a premature switch from the active PMSI to the inactive PMSI.

The traffic rate delta can be set to account for this behavior. For example, if a 1080P channel uses 5 Mb/s, the traffic rate delta can be set to 15 Mb/s to avoid the switchover from the primary to the secondary PMSI if one or two 1080 <S,G>s are switched between I-PMSI and S-PMSI. This provides a 10 Mb/s tolerance of asymmetric traffic.

3.4.19.2 S-PMSI behavior

If the network FDV is large or the sources are not synchronized, switching from I-PMSI to S-PMSI can happen at a different time on the primary and backup UMHs. This can cause asymmetric traffic on the I-PMSI and S-PMSI, resulting in a switch from the active UMH. The <S,G> traffic can arrive for the I-PMSI from the backup UMH and for S-PMSI from the active UMH, which causes temporary duplicate traffic until both UMHs switch to S-PMSI.

Multistream S-PMSI provides a solution for this case by mapping an <S,G> to an S-PMSI. The <S,G> is locked to the multistream S-PMSI, which is always configured and never torn down, even if the traffic goes down to 0, so the multistream S-PMSI is not susceptible to S-PMSI traffic drops.

The number of <S,G>s must be less than, or equal to, the **maximum-p2mp-spmsi** value configured under the MVPN selective provider tunnel. Otherwise, different <S,G>s can switch to the S-PMSI and when the S-PMSI limit is exhausted, the 2 UMHs become out of sync.

3.4.19.3 Bandwidth monitoring on single IOMs

Bandwidth monitoring is supported on single IOMs, that is, both the active and the backup PMSI terminate on the same IOM. The IOM monitors the statistics of both PMSIs and makes the switchover decision. The IOM does not include the CPM in any of the bandwidth monitoring decisions, which ensures fast detection times and switchover times under 50 ms.

All leaf and bud nodes must be configured with the same UMH PEs, I-PMSI and S-PMSI, and bandwidth threshold configuration to avoid traffic drops.

All LAG members must be in the same IOM that is performing the bandwidth monitoring function. The LAG interfaces spanning between multiple port members that belong to different IOMs have an unpredictable behavior, including traffic duplication.

3.4.19.4 ASM behavior

Bandwidth monitoring is supported with ASM only after the traffic is switched from <*,G> to <S,G>. Traffic arriving on separate IOMs from the active UMH and the inactive UMH results in traffic duplication because the pairing of the active and inactive UMH is per IOM, and the IOMs do not have a view of the pair. If the active traffic and the backup traffic arrive on different IOMs, each IOM treats the flow as the active flow and processes the traffic accordingly.

3.4.19.5 Low traffic rate

At low traffic rates, the packet on the active PMSI and the packet on the inactive PMSI can arrive at different times. If the packets arrive at the point that the statistics are read, there can be an inconsistency, resulting in a switchover. To avoid a switchover, UMH redundancy using bandwidth monitoring should be used only when the traffic rate is higher than 2 or 3 packets per second.

3.4.19.6 Revertive timer

If the active PMSI traffic goes below the threshold while the revertive timer (**configure service vprn mvpn provider-tunnel inclusive umh-rate-monitoring revertive-timer** and **configure service vprn mvpn**

provider-tunnel selective umh-rate-monitoring group source revertive-timer) is running, the timer is reset.

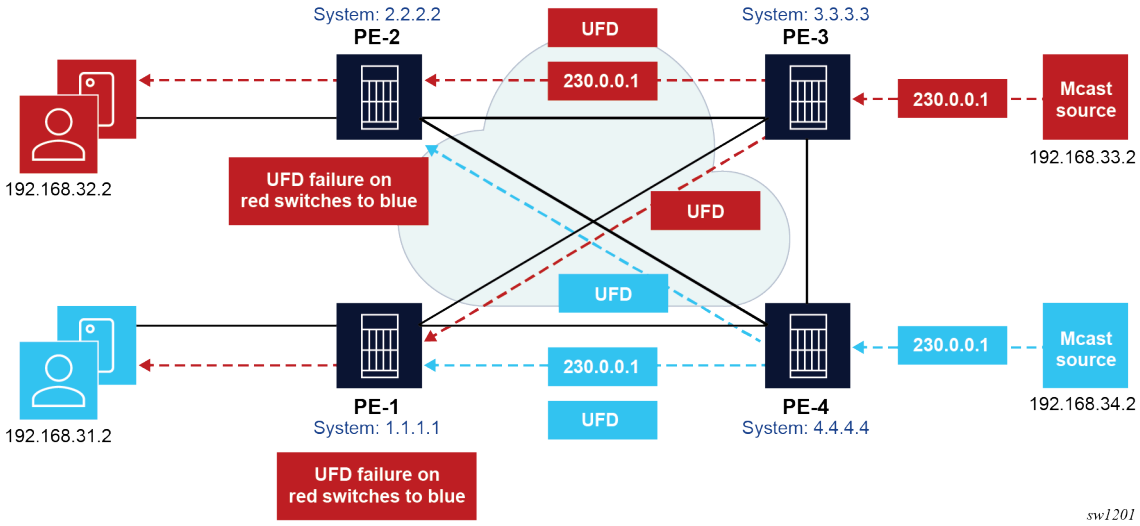
3.4.19.7 MVPN upstream PE fast failover

MVPN upstream UMH fast failover is supported for P2MP SR PMSI, as illustrated in Figure 55: Example of MVPN upstream PE fast failover. The upstream PE failure detection uses the Unidirectional Forwarding Detection (UFD) method. A downstream PE (receiver PE) supports fast upstream failover using the capability:

- to select two UMH nodes
- to monitor the upstream PE health using UFD

The tunnel status is monitored using the UFD session status received over the P-tunnel; the C-flow source is declared to be active based on this monitoring information. If multiple nodes are sourcing C-flows, the receiver PE node can choose to receive traffic from a primary and a standby source, but forwards only the multicast stream received from the primary source. If the detected P-tunnel status is down, the multicast receiver PE forwards the traffic received over the standby P-tunnel.

Figure 55: Example of MVPN upstream PE fast failover



3.4.19.7.1 MVPN upstream PE fast failover for tree SID

MVPN upstream PE fast failover for tree SID is supported as follows:

- Only inclusive PMSIs are supported.
- SM and SSM modes are supported. For SM, only fast switchover is supported on SPT. Fast protection is not supported on a shared tree.
- UFD sessions with 10-millisecond interval are supported on the CPM.
- Traffic duplication for restoration of the primary stream from the standby can occur for up to a second or more, depending on the CPU load during switchover.

- The P2MP policy tunnel ID is used as the UFD discriminator. Consequently, the solution is not interoperable with other vendors.

3.4.19.8 Multicast-only Fast Reroute

Multicast-only Fast Reroute (MoFRR) is not supported when UMH redundancy with bandwidth monitoring is enabled.

3.5 FIB prioritization

The RIB processing of specific routes can be prioritized through the use of the **rib-priority** command. This command allows specific routes to be prioritized through the protocol processing so that updates are propagated to the FIB as quickly as possible.

The **rib-priority** command can be configured within the VPRN instance of the OSPF or IS-IS routing protocols. For OSPF, a prefix list can be specified that identifies which route prefixes should be considered high priority. If the rib-priority high command is configured under an **VPRN>OSPF>area>interface** context then all routes learned through that interface is considered high priority. For the IS-IS routing protocol, RIB prioritization can be either specified though a prefix-list or an IS-IS tag value. If a prefix list is specified than route prefixes matching any of the prefix list criteria is considered high priority. If instead an IS-IS tag value is specified then any IS-IS route with that tag value is considered high priority.

The routes that have been designated as high priority are the first routes processed and then passed to the FIB update process so that the forwarding engine can be updated. All known high priority routes should be processed before the routing protocol moves on to other standard priority routes. This feature has the most impact when there are a large number of routes being learned through the routing protocols.

3.6 Configuring a VPRN service with CLI

This section provides information to configure Virtual Private Routed Network (VPRN) services using the command line interface.

3.6.1 Basic configuration

The following fields require specific input (there are no defaults) to configure a basic VPRN service:

- customer ID (see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Services Overview Guide*)
- interface parameters
- spoke SDP parameters

The following example displays a configuration of a VPRN service.

```
*A:ALA-1>config>service>vprn# info
----------------------------------------------
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
```

```
            auto-bind-tunnel
                resolution filter
                resolution-filter ldp
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                proxy-arp
                exit
                sap 1/1/10:1 create
                    ingress
                        qos 100
                    exit
                    egress
                        qos 1010
                        filter ip 10
                    exit
                exit
                dhcp
                    description "DHCP test"
                exit
                vrrp 1
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            bgp
                router-id 10.0.0.1
                group "to-cel"
                    export "vprnBgpExpPolCust1"
                    peer-as 65101
                    neighbor 10.1.1.2
                    exit
                exit
            exit
            pim
                apply-to all
                rp
                    static
                    exit
                    bsr-candidate
                        shutdown
                    exit
                    rp-candidate
                        shutdown
                    exit
                exit
            exit
            rip
                export "vprnRipExpPolCust1"
                group "cel"
                    neighbor "to-ce1"
                    exit
                exit
            exit
            no shutdown
-------------------------------------------
*A:ALA-1>config>service>vprn#
```

3.6.2 Common configuration tasks

About this task

This section provides a brief overview of the tasks that must be performed to configure a VPRN service and provides the CLI commands.

Procedure

Step 1. Associate a VPRN service with a customer ID.

Step 2. Optionally define an autonomous system.

Step 3. Define a route distinguisher (mandatory).

Step 4. Define VRF route-target associations or VRF import/export policies.

Step 5. Optionally define PIM parameters.

Step 6. Create a subscriber interface (applies to the 7750 SR only and is optional).

Step 7. Create an interface.

Step 8. Define SAP parameters on the interface.

- Select nodes and ports.
- Optionally select QoS policies other than the default (configured in config>qos context).
- Optionally select filter policies (configured in config>filter context).
- Optionally select accounting policy (configured in config>log context).
- Optionally configure DHCP features. (applies to the 7450 ESS and 7750 SR)

Step 9. Optionally define BGP parameters.
BGP must be enabled in the **config>router>bgp** context.

Step 10. Optionally define RIP parameters.

Step 11. Optionally spoke SDP parameters.

Step 12. Optionally create confederation autonomous systems within an AS.

Step 13. Enable the service.

3.6.3 Configuring VPRN components

This section provides VPRN configuration examples.

3.6.3.1 Creating a VPRN service

Use the following CLI syntax to create a VPRN service. A route distinguisher must be defined and the VPRN service must be administratively up in order for VPRN to be operationally active.

```
config>service# vprn service-id [customer customer-id]
    – route-distinguisher [ip-address:number1 | asn:number2]
    – description description-string
    – no shutdown
```

The following example displays a VPRN service configuration.

```
*A:ALA-1>config>service# info
----------------------------------------------
...
        vprn 1 customer 1 create
            route-distinguisher 10001:0
            no shutdown
        exit
...
----------------------------------------------
*A:ALA-1>config>service>vprn#
```

3.6.3.2 Configuring global VPRN parameters

The following example displays a VPRN service with configured parameters.

```
*A:ALA-1>config>service# info
----------------------------------------------
...
        vprn 1 customer 1 create
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            autonomous-system 10000
            route-distinguisher 10001:1
            spoke-sdp 2 create
            exit
            no shutdown
        exit
...
----------------------------------------------
*A:ALA-1>config>service#
```

3.6.3.3 Configuring VPRN log parameters

The following output displays a VPRN log configuration example.

```
B:Dut-C>config>service>vprn# info
----------------------------------------------
            dhcp
                local-dhcp-server "vprn_1" create
                    use-pool-from-client
                    force-renews
                    no shutdown
                exit
            exit
            snmp
                 community "YsMv96H2KZVKQeakNAq.38gvyr.MH9vA" hash2 r version both
                 community "gkYL94l90FFgu91PiRNvn3Rnl0edkMU1" hash2 rw version v2c
                 access
            log
                filter 1
                    default-action forward
                    entry 1
                        action forward
                    exit
```

```
                exit
                syslog 1
                    address 3ffe::e01:403
                    log-prefix "vprn1"
                exit
                snmp-trap-group 32001
                    trap-target "3" address 3ffe::e01:403 port 9000 snmpv2c notify-
community "vprn1"
                exit
                log-id 1
                    filter 1
                    from main change
                    to syslog 1
                exit
                log-id 3
                    filter 1
                    from main change
                    to snmp
                exit
            exit
...
-----------------------------------------------
B:Dut-C>config>service>vprn#
```

3.6.3.3.1 Configuring a spoke-SDP

Use the following CLI syntax to configure spoke SDP parameters:

```
config>service# vprn service-id [customer customer-id]
    – spoke-sdp sdp-id
        – no shutdown
    – interface ip-int-name
        – spoke-sdp sdp-id:vc-id [vc-type {ether | vlan}]
            – egress
                – filter {ip ip-filter-id}
                – vc-label egress-vc-label
            – ingress
                – filter {ip ip-filter-id}
                – vc-label ingress-vc-label
            – tos-marking-state {trusted | untrusted}
            – no shutdown
```

Use the following CLI syntax to configure spoke SDP parameters for the 7750 SR:

```
    – config>service# vprn service-id [customer customer-id]
    – spoke-sdp sdp-id
        – no shutdown
    – interface ip-int-name
        – spoke-sdp sdp-id:vc-id [vc-type {ether | vlan | vpls}]
            – egress
                – filter {ip ip-filter-id}
                – vc-label egress-vc-label
            – ingress
                – filter {ip ip-filter-id}
                – vc-label ingress-vc-label
            – tos-marking-state {trusted | untrusted}
            – no shutdown
```

The following output displays a spoke SDP configuration.

```
A:ALA-48>config>service>vprn# info
----------------------------------------------
...
            interface "SpokeSDP" create
                spoke-sdp 3:4 create
                    ingress
                        vc-label 3000
                        filter ip 10
                    exit
                    egress
                        vc-label 2000
                        filter ip 10
                    exit
                exit
            exit
...
            spoke-sdp 3 create
            exit
            no shutdown
----------------------------------------------
A:ALA-48>config>service>vprn#
```

3.6.3.4 Configuring VPRN protocols - PIM

The following example displays a VPRN PIM configuration for the 7750 SR:

```
config>service# info
#------------------------------------------
...
        vprn 1 customer 2 create
            route-distinguisher 1:11
            interface "if1" create
                address 10.13.14.15/32
                loopback
            exit
            interface "if2" create
                address 10.13.14.1/24
                sap 1/1/2:0 create
                exit
            exit
            pim
                interface "if1"
                exit
                interface "if2"
                exit
                rp
                    static
                    exit
                    bsr-candidate
                        shutdown
                    exit
                    rp-candidate
                        shutdown
                    exit
                exit
            exit
            no shutdown
        exit
```

```
    exit
#------------------------------------------
config>service#
```

3.6.3.4.1 Configuring router interfaces

See the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide* for command descriptions and syntax information to configure router interfaces.

The following example displays a router interface configurations:

```
ALA48>config>router# info
#------------------------------------------
echo "IP Configuration"
#------------------------------------------
...
        interface "if1"
            address 10.2.2.1/24
            port 1/1/33
        exit
        interface "if2"
            address 10.49.1.46/24
            port 1/1/34
        exit
        interface "if3"
            address 10.11.11.1/24
            port 1/1/35
        exit
...
#------------------------------------------
ALA48>config>router#
```

3.6.3.4.2 Configuring VPRN protocols - BGP

The autonomous system number and router ID configured in the VPRN context only applies to that particular service.

The minimal parameters that should be configured for a VPRN BGP instance are:

- Specify an autonomous system number for the router. See Configuring global VPRN parameters.
- Specify a router ID. If a new or different router ID value is entered in the BGP context, then the new values takes precedence and overwrites the VPRN-level router ID. See Configuring global VPRN parameters.
- Specify a VPRN BGP peer group.
- Specify a VPRN BGP neighbor with which to peer.
- Specify a VPRN BGP peer-AS that is associated with the above peer.

VPRN BGP is administratively enabled upon creation. Minimally, to enable VPRN BGP in a VPRN instance, you must associate an autonomous system number and router ID for the VPRN service, create a peer group, neighbor, and associate a peer AS number. There are no default VPRN BGP groups or neighbors. Each VPRN BGP group and neighbor must be explicitly configured.

All parameters configured for VPRN BGP are applied to the group and are inherited by each peer, but a group parameter can be overridden on a specific basis. VPRN BGP command hierarchy consists of three levels:

- the global level
- the group level
- the neighbor level

For example:

```
config>service>vprn>bgp#  (global level)
        – group  (group level)
    – neighbor      (neighbor level)
```

The local-address must be explicitly configured if two systems have multiple BGP peer sessions between them for the session to be established.

For more information about the BGP protocol, see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

3.6.3.4.2.1 Configuring VPRN BGP group and neighbor parameters

A group is a collection of related VPRN BGP peers. The group name should be a descriptive name for the group. Follow your group, name, and ID naming conventions for consistency and to help when troubleshooting faults.

All parameters configured for a peer group are applied to the group and are inherited by each peer (neighbor), but a group parameter can be overridden on a specific neighbor-level basis.

After a group name is created and options are configured, neighbors can be added within the same autonomous system to create IBGP connections or neighbors in different autonomous systems to create EBGP peers. All parameters configured for the peer group level are applied to each neighbor, but a group parameter can be overridden on a specific neighbor basis.

3.6.3.4.2.2 Configuring route reflection

Route reflection can be implemented in autonomous systems with a large internal BGP mesh to reduce the number of IBGP sessions required. One or more routers can be selected to act as focal points, for internal BGP sessions. Several BGP-speaking routers can peer with a route reflector. A route reflector forms peer connections to other route reflectors. A router assumes the role as a route reflector by configuring the **cluster cluster-id** command. No other command is required unless you want to disable reflection to specific peers.

If you configure the cluster command at the global level, then all subordinate groups and neighbors are members of the cluster. The route reflector cluster ID is expressed in dotted-decimal notation. The ID should be a significant topology-specific value. No other command is required unless you want to disable reflection to specific peers.

If a route reflector client is fully meshed, the **disable-client-reflect** command can be enabled to stop the route reflector from reflecting redundant route updates to a client.

3.6.3.4.2.3 Configuring BGP confederations

A VPRN can be configured to belong to a BGP confederation. BGP confederations are one technique for reducing the degree of IBGP meshing within an AS. When the confederation command is in the configuration of a VPRN the type of BGP session formed with a VPRN BGP neighbor is determined as follows:

- The session is of type IBGP if the peer AS is the same as the local AS.
- The session is of type confed-EBGP if the peer AS is different than the local AS AND the peer AS is listed as one of the members in the confederation command.
- The session is of type EBGP if the peer AS is different than the local AS AND the peer AS is not listed as one of the members in the confederation command.

3.6.3.4.2.4 VPRN BGP CLI syntax

Use the CLI syntax to configure VPRN BGP parameters.

The following example displays a VPRN BGP configuration:

```
*A:ALA-1>config>service# info
----------------------------------------------
...
        vprn 1 customer 1 create
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
            auto-bind-tunnel
                resolution filter
                resolution-filter ldp
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                sap 1/1/10:1 create
                    ingress
                        scheduler-policy "SLA2"
                        qos 100
                    exit
                    egress
                        scheduler-policy "SLA1"
                        qos 1010
                        filter ip 6
                    exit
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            bgp
                router-id 10.0.0.1
                group "to-cel"
                    export "vprnBgpExpPolCust1"
                    peer-as 65101
                    neighbor 10.1.1.2
                    exit
                exit
            exit
            spoke-sdp 2 create
            exit
```

```
            no shutdown
        exit
...
-------------------------------------------
*A:ALA-1>config>service#
```

3.6.3.4.3 Configuring VPRN protocols - RIP

PE routers attached to a specific VPN must learn the set of addresses for each site in that VPN. There are several ways for a PE router to obtain this information, one of which is the Routing Information Protocol (RIP). RIP sends routing update messages that include entry changes to the routing table, which is updated with the new information.

RIP can be used as a PE/CE distribution technique. PE and CE routers can be configured as RIP peers, and the CE router can transmit RIP updates to inform the PE router about the set of address prefixes which are reachable at the CE router's site. When RIP is configured in the CE router, care must be taken to ensure that address prefixes from other sites (address prefixes learned by the CE router from the PE router) are never advertised to the PE router. Specifically, if a PE router receives a VPN-IPv4 route and, as a result, distributes an IPv4 route to a CE router, then that route must not be distributed back from that CE's site to a PE router (either the same router or different routers).

To enable a VPRN RIP instance, enable the RIP protocol in the **config>service> >vprn>rip** context of the VPRN. VPRN RIP is administratively enabled upon creation. Configuring other RIP commands and parameters is optional.

Caution: Careful planning is essential to implement commands that can affect the behavior of VPRN RIP global, group, and neighbor levels. Because the RIP commands are hierarchical, analyze the values that can disable features on a particular level.

The parameters configured at the VPRN RIP global level are inherited by the group and neighbor levels. Several hierarchical VPRN RIP commands can be modified on different levels; the most specific value is used. That is, a VPRN RIP group-specific command takes precedence over a global VPRN RIP command. A neighbor-specific command takes precedence over a global VPRN RIP and group-specific command. For example, if you modify a VPRN RIP neighbor-level command default, the new value takes precedence over VPRN RIP group- and global-level settings. VPRN RIP groups and neighbors are not created by default. Each VPRN RIP group and neighbor must be explicitly configured.

The minimal parameters that should be configured for a VPRN instance are:

- Specify a VPRN RIP peer group.
- Specify a VPRN RIP neighbor with which to peer.
- Specify a VPRN RIP peer-AS that is associated with the above peer.

The VPRN RIP command hierarchy consists of three levels:

- the global level
- the group level
- the neighbor level

For example:

```
config>service>vprn>rip#  (global level)
    – group (group level)
        – neighbor  (neighbor level)
```

3.6.3.4.3.1 VPRN RIP CLI syntax

The following example displays a VPRN RIP configuration:

```
*A:ALA-1>config>service# info
----------------------------------------------
...
        vprn 1 customer 1 create
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
            auto-bind-tunnel
                resolution filter
                resolution-filter ldp
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                sap 1/1/10:1 create
                    ingress
                        scheduler-policy "SLA2"
                        qos 100
                    exit
                    egress
                        scheduler-policy "SLA1"
                        qos 1010
                        filter ip 6
                    exit
                exit
            exit
            static-route-entry 10.5.0.0/24
                       next-hop 10.1.1.2
            bgp
                router-id 10.0.0.1
                group "to-cel"
                    export "vprnBgpExpPolCust1"
                    peer-as 65101
                    neighbor 10.1.1.2
                    exit
                exit
            exit
            rip
                export "vprnRipExpPolCust1"
                group "cel"
                    neighbor "to-ce1"
                    exit
                exit
            exit
            spoke-sdp 2 create
            exit
            no shutdown
        exit
...
----------------------------------------------
*A:ALA-1>config>service# info
```

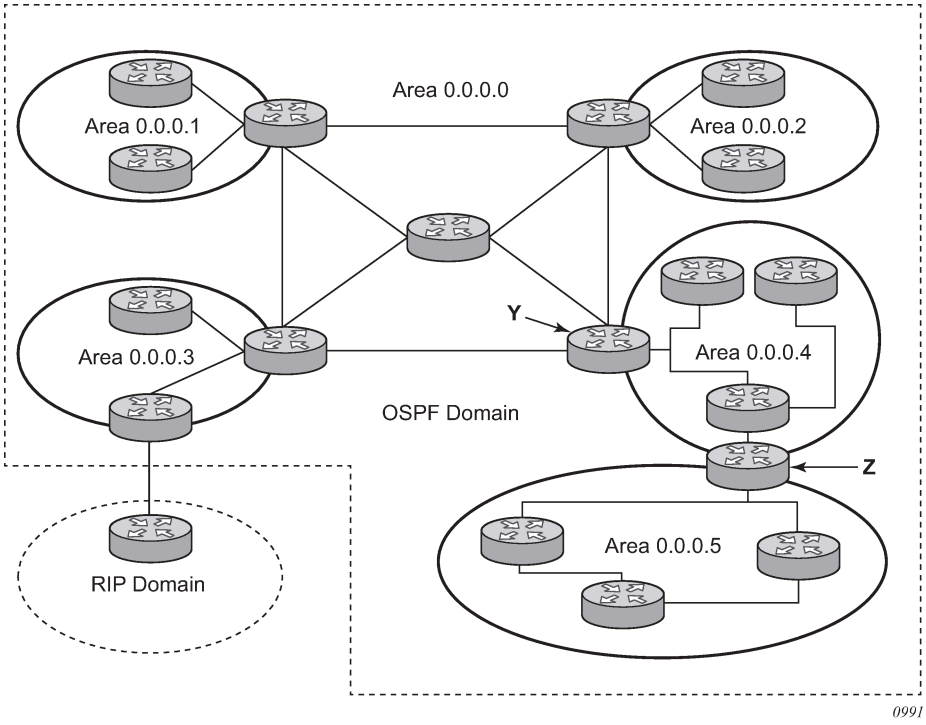
For more information about the RIP protocol, see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

3.6.3.4.4 Configuring VPRN protocols - OSPF

Each VPN routing instance is isolated from any other VPN routing instance, and from the routing used across the backbone. OSPF can be run with any VPRN, independently of the routing protocols used in other VPRNs, or in the backbone itself. For more information about the OSPF protocol, see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

The OSPF backbone area, area 0.0.0.0, must be contiguous and all other areas must be connected to the backbone area. The backbone distributes routing information between areas. If it is not practical to connect an area to the backbone (see Area 0.0.0.5 in Figure 56: OSPF areas), the area border routers (such as routers Y and Z) must be connected via a virtual link. The two area border routers form a point-to-point-like adjacency across the transit area (see Area 0.0.0.4). A virtual link can only be configured while in the area 0.0.0.0 context.

Figure 56: OSPF areas




```
config>service>vprn>ospf#
```

3.6.3.4.4.1 VPRN OSPF CLI syntax

The following example displays the VPRN OSPF configuration shown above:

```
*A:ALA-48>config>service# info
----------------------------------------------
 vprn 2 customer 1 create
           interface "test" create
           exit
           no shutdown
       exit
```

```
            area 0.0.0.0
                virtual-link 1.2.3.4 transit-area 1.2.3.4
                    hello-interval 9
                    dead-interval 40
                exit
    exit
-------------------------------------------
*A:ALA-48>config>service#
```

For more information about the OSPF protocol, see the *7450 ESS, 7750 SR, 7950 XRS, and VSR Router Configuration Guide*.

3.6.3.4.5 Configuring a VPRN interface

Interface names associate an IP address to the interface, and then associate the IP interface with a physical port. The logical interface can associate attributes like an IP address, port, Link Aggregation Group (LAG) or the system.

There are no default interfaces.

You can configure a VPRN interface as a loopback interface by issuing the loopback command instead of the **sap sap-id** command. The loopback flag cannot be set on an interface where a SAP is already defined and a SAP cannot be defined on a loopback interface.

When using mtrace or mstat in a Layer 3 VPN context then the configuration for the VPRN should have a loopback address configured which has the same address as the core instance's system address (BGP next-hop).

The following example displays a VPRN interface configuration:

```
*A:ALA-1>config>service>vprn# info
-----------------------------------------------
...
        vprn 1 customer 1 create
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
            auto-bind-tunnel
                resolution filter
                resolution-filter ldp
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            spoke-sdp 2 create
            exit
            no shutdown
        exit
...
-----------------------------------------------
*A:ALA-1>config>service#
```

3.6.3.4.6 Configuring overload state on a single SFM

When a router has fewer than the full set of SFMs functioning, the forwarding capacity can be reduced. Some scenarios include:

- fewer than the maximum number of SFMs installed in the system
- one or more SFMs have failed
- the system is in the ISSU process and the SFM is co-located on the CPM

An overload condition can be set for IS-IS and OSPF to enable the router to still participate in exchanging routing information, but route all traffic away from it when insufficient SFMs are active. This is achieved using the following CLI commands:

```
config>router>single-sfm-overload [holdoff-time hold-off-time]
    config>service>vprn>single-sfm-overload [holdoff-time hold-off-time]
    tools>perform>redundancy>forced-single-sfm-overload
```

These cause an overload state in the IGP to trigger the traffic reroute by setting the overload bit in IS-IS or setting the metric to maximum in OSPF. When PIM uses IS-IS or OSPF to find out the upstream router, a next-hop change in the IS-IS or OSPF causes PIM to join the new path and prune the old path, which effectively also reroutes the multicast traffic downstream as well as the unicast traffic.

When the problem is resolved, and the required compliment of SFMs become active in the router, the overload condition is cleared, which causes the traffic to be routed back to the router.

The conditions to set overload are:

- For 7950 XRS-20, 7750 SR-12/SR-7 and 7450 ESS-12/ESS-7 platforms, if an SF/CPM fails, the protocol sets the overload.
- For 7950-40 XRS and 7750 SR-12e platforms, if two SFMs fail (a connected pair on the XRS-40) the protocol sets the overload.

3.6.3.4.7 Configuring a VPRN interface SAP

A SAP is a combination of a port and encapsulation parameters which identifies the service access point on the interface and within the SR. Each SAP must be unique within a router. A SAP cannot be defined if the interface **loopback** command is enabled.

When configuring VPRN interface SAP parameters, a default QoS policy is applied to each ingress and egress SAP. Additional QoS policies and scheduler policies must be configured in the **config>qos** context. Filter policies are configured in the **config>filter** context and must be explicitly applied to a SAP. There are no default filter policies.

The following example displays a VPRN interface SAP configuration:

```
*A:ALA-1>config>service# info
----------------------------------------------
...
        vprn 1 customer 1 create
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
            auto-bind-tunnel
                resolution filter
```

```
                resolution-filter ldp
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                sap 1/1/10:1 create
                    ingress
                        scheduler-policy "SLA2"
                        qos 100
                    exit
                    egress
                        scheduler-policy "SLA1"
                        qos 1010
                        filter ip 6
                    exit
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            spoke-sdp 2 create
            exit
            no shutdown
        exit
...
-----------------------------------------------
*A:ALA-1>config>service#
```

3.7 Service management tasks

This section discusses VPRN service management tasks.

3.7.1 Modifying VPRN service parameters

Use the CLI syntax to modify VPRN parameters. The following example displays the VPRN service creation output.

```
*A:ALA-1>config>service# info
-----------------------------------------------
...
vprn 1 customer 1 create
            shutdown
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            maximum-routes 2000
            autonomous-system 10000
            route-distinguisher 10001:1
            interface "to-ce1" create
                address 10.1.1.1/24
                sap 1/1/10:1 create
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            bgp
                router-id 10.0.0.1
                group "to-ce1"
                    export "vprnBgpExpPolCust1"
                    peer-as 65101
```

```
                    neighbor 10.1.1.2
                    exit
                exit
            exit
            spoke-sdp 2 create
            exit
        exit
...
-------------------------------------------
*A:ALA-1>config>service>vprn#
```

3.7.2 Deleting a VPRN service

A VPRN service cannot be deleted until SAPs and interfaces are shut down and deleted. If protocols or a spoke-SDP, or both are defined, they must be shut down and removed from the configuration as well.

Use the following CLI syntax to delete a VPRN service:

```
config>service#
    – [no] vprn service-id [customer customer-id]
        – shutdown
        – [no] interface ip-int-name
            – shutdown
            – [no] sap sap-id]
        – [no] bgp
            – shutdown
        – [no] rip
            – shutdown
        – [no] spoke-sdp sdp-id
            – [no] shutdown
```

3.7.3 Disabling a VPRN service

A VPRN service can be shut down without deleting any service parameters.

```
config>service#
    – vprn service-id [customer customer-id]
        – shutdown
```

```
config>service# vprn 1
    – config>service>vprn# shutdown
    – config>service>vprn# exit
```

```
*A:ALA-1>config>service# info
-------------------------------------------
...
        vprn 1 customer 1 create
            shutdown
            vrf-import "vrfImpPolCust1"
            vrf-export "vrfExpPolCust1"
            ecmp 8
            autonomous-system 10000
            route-distinguisher 10001:1
            auto-bind-tunnel
                resolution filter
                resolution-filter ldp
```

```
            vrf-target target:10001:1
            interface "to-ce1" create
                address 10.1.0.1/24
                sap 1/1/10:1 create
                    ingress
                        scheduler-policy "SLA2"
                        qos 100
                    exit
                    egress
                        scheduler-policy "SLA1"
                        qos 1010
                        filter ip 6
                    exit
                exit
            exit
            static-route-entry 10.5.0.0/24
                next-hop 10.1.1.2
            bgp
                router-id 10.0.0.1
                group "to-cel"
                    export "vprnBgpExpPolCust1"
                    peer-as 65101
                    neighbor 10.1.1.2
                    exit
                exit
            exit
            rip
                export "vprnRipExpPolCust1"
                group "cel"
                    neighbor "to-ce1"
                    exit
                exit
            exit
            spoke-sdp 2 create
            exit
        exit
...
----------------------------------------------
*A:ALA-1>config>service#
```

3.7.4 Re-enabling a VPRN service

To re-enable a VPRN service that was shut down.

```
config>service#
    – vprn service-id [customer customer-id]
        – no shutdown
        –
```

4 Standards and protocol support

Note:
The information provided in this chapter is subject to change without notice and may not apply to all platforms.

Nokia assumes no responsibility for inaccuracies.

4.1 Access Node Control Protocol (ANCP)

draft-ietf-ancp-protocol-02, *Protocol for Access Node Control Mechanism in Broadband Networks*

RFC 5851, *Framework and Requirements for an Access Node Control Mechanism in Broadband Multi-Service Networks*

4.2 Bidirectional Forwarding Detection (BFD)

draft-ietf-idr-bgp-ls-sbfd-extensions-01, *BGP Link-State Extensions for Seamless BFD*

draft-ietf-lsr-ospf-bfd-strict-mode-10, *OSPF BFD Strict-Mode*

RFC 5880, *Bidirectional Forwarding Detection (BFD)*

RFC 5881, *Bidirectional Forwarding Detection (BFD) IPv4 and IPv6 (Single Hop)*

RFC 5882, *Generic Application of Bidirectional Forwarding Detection (BFD)*

RFC 5883, *Bidirectional Forwarding Detection (BFD) for Multihop Paths*

RFC 7130, *Bidirectional Forwarding Detection (BFD) on Link Aggregation Group (LAG) Interfaces*

RFC 7880, *Seamless Bidirectional Forwarding Detection (S-BFD)*

RFC 7881, *Seamless Bidirectional Forwarding Detection (S-BFD) for IPv4, IPv6, and MPLS*

RFC 7883, *Advertising Seamless Bidirectional Forwarding Detection (S-BFD) Discriminators in IS-IS*

RFC 7884, *OSPF Extensions to Advertise Seamless Bidirectional Forwarding Detection (S-BFD) Target Discriminators*

4.3 Border Gateway Protocol (BGP)

draft-gredler-idr-bgplu-epe-14, *Egress Peer Engineering using BGP-LU*

draft-hares-idr-update-attrib-low-bits-fix-01, *Update Attribute Flag Low Bits Clarification*

draft-ietf-idr-add-paths-guidelines-08, *Best Practices for Advertisement of Multiple Paths in IBGP*

draft-ietf-idr-best-external-03, *Advertisement of the best external route in BGP*

draft-ietf-idr-bgp-flowspec-oid-03, *Revised Validation Procedure for BGP Flow Specifications*

draft-ietf-idr-bgp-gr-notification-01, *Notification Message support for BGP Graceful Restart*

draft-ietf-idr-bgp-ls-app-specific-attr-16, *Application-Specific Attributes Advertisement with BGP Link-State*

draft-ietf-idr-bgp-ls-flex-algo-06, *Flexible Algorithm Definition Advertisement with BGP Link-State*

draft-ietf-idr-bgp-optimal-route-reflection-10, *BGP Optimal Route Reflection (BGP-ORR)*

draft-ietf-idr-error-handling-03, *Revised Error Handling for BGP UPDATE Messages*

draft-ietf-idr-flowspec-interfaceset-03, *Applying BGP flowspec rules on a specific interface set*

draft-ietf-idr-flowspec-path-redirect-05, *Flowspec Indirection-id Redirect* – localised ID

draft-ietf-idr-flowspec-redirect-ip-02, *BGP Flow-Spec Redirect to IP Action*

draft-ietf-idr-link-bandwidth-03, *BGP Link Bandwidth Extended Community*

draft-ietf-idr-long-lived-gr-00, *Support for Long-lived BGP Graceful Restart*

RFC 1772, *Application of the Border Gateway Protocol in the Internet*

RFC 1997, *BGP Communities Attribute*

RFC 2385, *Protection of BGP Sessions via the TCP MD5 Signature Option*

RFC 2439, *BGP Route Flap Damping*

RFC 2545, *Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing*

RFC 2858, *Multiprotocol Extensions for BGP-4*

RFC 2918, *Route Refresh Capability for BGP-4*

RFC 4271, *A Border Gateway Protocol 4 (BGP-4)*

RFC 4360, *BGP Extended Communities Attribute*

RFC 4364, *BGP/MPLS IP Virtual Private Networks (VPNs)*

RFC 4456, *BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)*

RFC 4486, *Subcodes for BGP Cease Notification Message*

RFC 4659, *BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN*

RFC 4684, *Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)*

RFC 4724, *Graceful Restart Mechanism for BGP* – helper mode

RFC 4760, *Multiprotocol Extensions for BGP-4*

RFC 4798, *Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)*

RFC 5004, *Avoid BGP Best Path Transitions from One External to Another*

RFC 5065, *Autonomous System Confederations for BGP*

RFC 5291, *Outbound Route Filtering Capability for BGP-4*

RFC 5396, *Textual Representation of Autonomous System (AS) Numbers* – asplain

RFC 5492, *Capabilities Advertisement with BGP-4*

RFC 5668, *4-Octet AS Specific BGP Extended Community*

RFC 6286, *Autonomous-System-Wide Unique BGP Identifier for BGP-4*

RFC 6793, *BGP Support for Four-Octet Autonomous System (AS) Number Space*

RFC 6810, *The Resource Public Key Infrastructure (RPKI) to Router Protocol*

RFC 6811, *Prefix Origin Validation*

RFC 6996, *Autonomous System (AS) Reservation for Private Use*

RFC 7311, *The Accumulated IGP Metric Attribute for BGP*

RFC 7606, *Revised Error Handling for BGP UPDATE Messages*

RFC 7607, *Codification of AS 0 Processing*

RFC 7674, *Clarification of the Flowspec Redirect Extended Community*

RFC 7752, *North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP*

RFC 7854, *BGP Monitoring Protocol (BMP)*

RFC 7911, *Advertisement of Multiple Paths in BGP*

RFC 7999, *BLACKHOLE Community*

RFC 8092, *BGP Large Communities Attribute*

RFC 8097, *BGP Prefix Origin Validation State Extended Community*

RFC 8212, *Default External BGP (EBGP) Route Propagation Behavior without Policies*

RFC 8277, *Using BGP to Bind MPLS Labels to Address Prefixes*

RFC 8571, *BGP - Link State (BGP-LS) Advertisement of IGP Traffic Engineering Performance Metric Extensions*

RFC 8950, *Advertising IPv4 Network Layer Reachability Information (NLRI) with an IPv6 Next Hop*

RFC 8955, *Dissemination of Flow Specification Rules*

RFC 8956, *Dissemination of Flow Specification Rules for IPv6*

RFC 9086, *Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing BGP Egress Peer Engineering*

4.4 Broadband Network Gateway (BNG) Control and User Plane Separation (CUPS)

3GPP 23.007, *Restoration procedures*

3GPP 29.244, *Interface between the Control Plane and the User Plane nodes*

3GPP 29.281, *General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U)*

BBF TR-459, *Control and User Plane Separation for a Disaggregated BNG*

BBF TR-459.2, *Multi-Service Disaggregated BNG with CUPS: Integrated Carrier Grade NAT function*

RFC 8300, *Network Service Header (NSH)*

4.5 Certificate management

RFC 4210, *Internet X.509 Public Key Infrastructure Certificate Management Protocol (CMP)*

RFC 4211, *Internet X.509 Public Key Infrastructure Certificate Request Message Format (CRMF)*

RFC 5280, *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*

RFC 6712, *Internet X.509 Public Key Infrastructure -- HTTP Transfer for the Certificate Management Protocol (CMP)*

RFC 7030, *Enrollment over Secure Transport*

RFC 7468, *Textual Encodings of PKIX, PKCS, and CMS Structures*

4.6 Circuit emulation

RFC 4553, *Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)*

RFC 5086, *Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)*

RFC 5287, *Control Protocol Extensions for the Setup of Time-Division Multiplexing (TDM) Pseudowires in MPLS Networks*

4.7 Ethernet

IEEE 802.1AB, *Station and Media Access Control Connectivity Discovery*

IEEE 802.1ad, *Provider Bridges*

IEEE 802.1ag, *Connectivity Fault Management*

IEEE 802.1ah, *Provider Backbone Bridges*

IEEE 802.1ak, *Multiple Registration Protocol*

IEEE 802.1aq, *Shortest Path Bridging*

IEEE 802.1ax, *Link Aggregation*

IEEE 802.1D, *MAC Bridges*

IEEE 802.1p, *Traffic Class Expediting*

IEEE 802.1Q, *Virtual LANs*

IEEE 802.1s, *Multiple Spanning Trees*

IEEE 802.1w, *Rapid Reconfiguration of Spanning Tree*

IEEE 802.1X, *Port Based Network Access Control*

IEEE 802.3ac, *VLAN Tag*

IEEE 802.3ad, *Link Aggregation*

IEEE 802.3ah, *Ethernet in the First Mile*

IEEE 802.3x, *Ethernet Flow Control*

ITU-T G.8031/Y.1342, *Ethernet Linear Protection Switching*

ITU-T G.8032/Y.1344, *Ethernet Ring Protection Switching*

ITU-T Y.1731, *OAM functions and mechanisms for Ethernet based networks*

4.8 Ethernet VPN (EVPN)

draft-ietf-bess-evpn-ipvpn-interworking-06, *EVPN Interworking with IPVPN*

draft-ietf-bess-evpn-irb-mcast-04, *EVPN Optimized Inter-Subnet Multicast (OISM) Forwarding* – ingress replication

draft-ietf-bess-evpn-pref-df-06, *Preference-based EVPN DF Election*

draft-ietf-bess-evpn-unequal-lb-16, *Weighted Multi-Path Procedures for EVPN Multi-Homing* – section 9

draft-ietf-bess-evpn-virtual-eth-segment-06, *EVPN Virtual Ethernet Segment*

draft-ietf-bess-pbb-evpn-isid-cmacflush-00, *PBB-EVPN ISID-based CMAC-Flush*

draft-sajassi-bess-evpn-ip-aliasing-05, *EVPN Support for L3 Fast Convergence and Aliasing/Backup Path* – IP Prefix routes

RFC 7432, *BGP MPLS-Based Ethernet VPN*

RFC 7623, *Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)*

RFC 8214, *Virtual Private Wire Service Support in Ethernet VPN*

RFC 8317, *Ethernet-Tree (E-Tree) Support in Ethernet VPN (EVPN) an Provider Backbone Bridging EVPN (PBB-EVPN)*

RFC 8365, *A Network Virtualization Overlay Solution Using Ethernet VPN (EVPN)*

RFC 8560, *Seamless Integration of Ethernet VPN (EVPN) with Virtual Private LAN Service (VPLS) and Their Provider Backbone Bridge (PBB) Equivalents*

RFC 8584, *DF Election and AC-influenced DF Election*

RFC 9047, *Propagation of ARP/ND Flags in an Ethernet Virtual Private Network (EVPN)*

RFC 9135, *Integrated Routing and Bridging in Ethernet VPN (EVPN)* – Asymmetric IRB Procedures and Mobility Procedure

RFC 9136, *IP Prefix Advertisement in Ethernet VPN (EVPN)*

RFC 9161, *Operational Aspects of Proxy ARP/ND in Ethernet Virtual Private Networks*

RFC 9251, *Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Proxies for Ethernet VPN (EVPN)*

4.9 gRPC Remote Procedure Calls (gRPC)

cert.proto version 0.1.0, *gRPC Network Operations Interface (gNOI) Certificate Management Service*

file.proto version 0.1.0, *gRPC Network Operations Interface (gNOI) File Service*

gnmi.proto version 0.8.0, *gRPC Network Management Interface (gNMI) Service Specification*

PROTOCOL-HTTP2, *gRPC over HTTP2*

system.proto Version 1.0.0, *gRPC Network Operations Interface (gNOI) System Service*

4.10 Intermediate System to Intermediate System (IS-IS)

draft-ietf-isis-mi-02, *IS-IS Multi-Instance*

draft-kaplan-isis-ext-eth-02, *Extended Ethernet Frame Size Support*

ISO/IEC 10589:2002 Second Edition, *Intermediate system to Intermediate system intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)*

RFC 1195, *Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*

RFC 2973, *IS-IS Mesh Groups*

RFC 3359, *Reserved Type, Length and Value (TLV) Codepoints in Intermediate System to Intermediate System*

RFC 3719, *Recommendations for Interoperable Networks using Intermediate System to Intermediate System (IS-IS)*

RFC 3787, *Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)*

RFC 4971, *Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information*

RFC 5120, *M-ISIS: Multi Topology (MT) Routing in IS-IS*

RFC 5130, *A Policy Control Mechanism in IS-IS Using Administrative Tags*

RFC 5301, *Dynamic Hostname Exchange Mechanism for IS-IS*

RFC 5302, *Domain-wide Prefix Distribution with Two-Level IS-IS*

RFC 5303, *Three-Way Handshake for IS-IS Point-to-Point Adjacencies*

RFC 5304, *IS-IS Cryptographic Authentication*

RFC 5305, *IS-IS Extensions for Traffic Engineering TE*

RFC 5306, *Restart Signaling for IS-IS* – helper mode

RFC 5308, *Routing IPv6 with IS-IS*

RFC 5309, *Point-to-Point Operation over LAN in Link State Routing Protocols*

RFC 5310, *IS-IS Generic Cryptographic Authentication*

RFC 6119, *IPv6 Traffic Engineering in IS-IS*

RFC 6213, *IS-IS BFD-Enabled TLV*

RFC 6232, *Purge Originator Identification TLV for IS-IS*

RFC 6233, *IS-IS Registry Extension for Purges*

RFC 6329, *IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging*

RFC 7775, *IS-IS Route Preference for Extended IP and IPv6 Reachability*

RFC 7794, *IS-IS Prefix Attributes for Extended IPv4 and IPv6 Reachability* – sections 2.1 and 2.3

RFC 7987, *IS-IS Minimum Remaining Lifetime*

RFC 8202, *IS-IS Multi-Instance* – single topology

RFC 8570, *IS-IS Traffic Engineering (TE) Metric Extensions* – Min/Max Unidirectional Link Delay metric for flex-algo, RSVP, SR-TE

RFC 8919, *IS-IS Application-Specific Link Attributes*

4.11 Internet Protocol (IP) Fast Reroute (FRR)

draft-ietf-rtgwg-lfa-manageability-08, *Operational management of Loop Free Alternates*

RFC 5286, *Basic Specification for IP Fast Reroute: Loop-Free Alternates*

RFC 7431, *Multicast-Only Fast Reroute*

RFC 7490, *Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)*

RFC 8518, *Selection of Loop-Free Alternates for Multi-Homed Prefixes*

4.12 Internet Protocol (IP) general

draft-grant-tacacs-02, *The TACACS+ Protocol*

RFC 768, *User Datagram Protocol*

RFC 793, *Transmission Control Protocol*

RFC 854, *Telnet Protocol Specifications*

RFC 1350, *The TFTP Protocol (revision 2)*

RFC 2347, *TFTP Option Extension*

RFC 2348, *TFTP Blocksize Option*

RFC 2349, *TFTP Timeout Interval and Transfer Size Options*

RFC 2428, *FTP Extensions for IPv6 and NATs*

RFC 2617, *HTTP Authentication: Basic and Digest Access Authentication*

RFC 2784, *Generic Routing Encapsulation (GRE)*

RFC 2818, *HTTP Over TLS*

RFC 2890, *Key and Sequence Number Extensions to GRE*

RFC 3164, *The BSD syslog Protocol*

RFC 4250, *The Secure Shell (SSH) Protocol Assigned Numbers*

RFC 4251, *The Secure Shell (SSH) Protocol Architecture*

RFC 4252, *The Secure Shell (SSH) Authentication Protocol* – publickey, password

RFC 4253, *The Secure Shell (SSH) Transport Layer Protocol*

RFC 4254, *The Secure Shell (SSH) Connection Protocol*

RFC 4511, *Lightweight Directory Access Protocol (LDAP): The Protocol*

RFC 4513, *Lightweight Directory Access Protocol (LDAP): Authentication Methods and Security Mechanisms* – TLS

RFC 4632, *Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan*

RFC 5082, *The Generalized TTL Security Mechanism (GTSM)*

RFC 5246, *The Transport Layer Security (TLS) Protocol Version 1.2* – TLS client, RSA public key

RFC 5425, *Transport Layer Security (TLS) Transport Mapping for Syslog* – RFC 3164 with TLS

RFC 5656, *Elliptic Curve Algorithm Integration in the Secure Shell Transport Layer* – ECDSA

RFC 5925, *The TCP Authentication Option*

RFC 5926, *Cryptographic Algorithms for the TCP Authentication Option (TCP-AO)*

RFC 6398, *IP Router Alert Considerations and Usage* – MLD

RFC 6528, *Defending against Sequence Number Attacks*

RFC 7011, *Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information*

RFC 7012, *Information Model for IP Flow Information Export*

RFC 7230, *Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing*

RFC 7231, *Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content*

RFC 7232, *Hypertext Transfer Protocol (HTTP/1.1): Conditional Requests*

RFC 7301, *Transport Layer Security (TLS) Application Layer Protocol Negotiation Extension*

RFC 7616, *HTTP Digest Access Authentication*

RFC 8446, *The Transport Layer Security (TLS) Protocol Version 1.3*

4.13 Internet Protocol (IP) multicast

cisco-ipmulticast/pim-autorp-spec01, *Auto-RP: Automatic discovery of Group-to-RP mappings for IP multicast* – version 1

draft-ietf-bier-pim-signaling-08, *PIM Signaling Through BIER Core*

draft-ietf-idmr-traceroute-ipm-07, *A "traceroute" facility for IP Multicast*

draft-ietf-l2vpn-vpls-pim-snooping-07, *Protocol Independent Multicast (PIM) over Virtual Private LAN Service (VPLS)*

RFC 1112, *Host Extensions for IP Multicasting*

RFC 2236, *Internet Group Management Protocol, Version 2*

RFC 2365, *Administratively Scoped IP Multicast*

RFC 2375, *IPv6 Multicast Address Assignments*

RFC 2710, *Multicast Listener Discovery (MLD) for IPv6*

RFC 3306, *Unicast-Prefix-based IPv6 Multicast Addresses*

RFC 3376, *Internet Group Management Protocol, Version 3*

RFC 3446, *Anycast Rendevous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)*

RFC 3590, *Source Address Selection for the Multicast Listener Discovery (MLD) Protocol*

RFC 3618, *Multicast Source Discovery Protocol (MSDP)*

RFC 3810, *Multicast Listener Discovery Version 2 (MLDv2) for IPv6*

RFC 3956, *Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address*

RFC 3973, *Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)* – auto-RP groups

RFC 4541, *Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches*

RFC 4604, *Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast*

RFC 4607, *Source-Specific Multicast for IP*

RFC 4608, *Source-Specific Protocol Independent Multicast in 232/8*

RFC 4610, *Anycast-RP Using Protocol Independent Multicast (PIM)*

RFC 4611, *Multicast Source Discovery Protocol (MSDP) Deployment Scenarios*

RFC 5059, *Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)*

RFC 5186, *Internet Group Management Protocol Version 3 (IGMPv3) / Multicast Listener Discovery Version 2 (MLDv2) and Multicast Routing Protocol Interaction*

RFC 5384, *The Protocol Independent Multicast (PIM) Join Attribute Format*

RFC 5496, *The Reverse Path Forwarding (RPF) Vector TLV*

RFC 6037, *Cisco Systems' Solution for Multicast in MPLS/BGP IP VPNs*

RFC 6512, *Using Multipoint LDP When the Backbone Has No Route to the Root*

RFC 6513, *Multicast in MPLS/BGP IP VPNs*

RFC 6514, *BGP Encodings and Procedures for Multicast in MPLS/IP VPNs*

RFC 6515, *IPv4 and IPv6 Infrastructure Addresses in BGP Updates for Multicast VPNs*

RFC 6516, *IPv6 Multicast VPN (MVPN) Support Using PIM Control Plane and Selective Provider Multicast Service Interface (S-PMSI) Join Messages*

RFC 6625, *Wildcards in Multicast VPN Auto-Discover Routes*

RFC 6826, *Multipoint LDP In-Band Signaling for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Path*

RFC 7246, *Multipoint Label Distribution Protocol In-Band Signaling in a Virtual Routing and Forwarding (VRF) Table Context*

RFC 7385, *IANA Registry for P-Multicast Service Interface (PMSI) Tunnel Type Code Points*

RFC 7716, *Global Table Multicast with BGP Multicast VPN (BGP-MVPN) Procedures*

RFC 7761, *Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)*

RFC 8279, *Multicast Using Bit Index Explicit Replication (BIER)*

RFC 8296, *Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks* – MPLS encapsulation

RFC 8401, *Bit Index Explicit Replication (BIER) Support via IS-IS*

RFC 8444, *OSPFv2 Extensions for Bit Index Explicit Replication (BIER)*

RFC 8487, *Mtrace Version 2: Traceroute Facility for IP Multicast*

RFC 8534, *Explicit Tracking with Wildcard Routes in Multicast VPN* – (C-*,C-*) wildcard

RFC 8556, *Multicast VPN Using Bit Index Explicit Replication (BIER)*

4.14 Internet Protocol (IP) version 4

RFC 791, *Internet Protocol*

RFC 792, *Internet Control Message Protocol*

RFC 826, *An Ethernet Address Resolution Protocol*

RFC 951, *Bootstrap Protocol (BOOTP)* – relay

RFC 1034, *Domain Names - Concepts and Facilities*

RFC 1035, *Domain Names - Implementation and Specification*

RFC 1191, *Path MTU Discovery* – router specification

RFC 1519, *Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy*

RFC 1534, *Interoperation between DHCP and BOOTP*

RFC 1542, *Clarifications and Extensions for the Bootstrap Protocol*

RFC 1812, *Requirements for IPv4 Routers*

RFC 1918, *Address Allocation for Private Internets*

RFC 2003, *IP Encapsulation within IP*

RFC 2131, *Dynamic Host Configuration Protocol*

RFC 2132, *DHCP Options and BOOTP Vendor Extensions*

RFC 2401, *Security Architecture for Internet Protocol*

RFC 3021, *Using 31-Bit Prefixes on IPv4 Point-to-Point Links*

RFC 3046, *DHCP Relay Agent Information Option (Option 82)*

RFC 3768, *Virtual Router Redundancy Protocol (VRRP)*

RFC 4884, *Extended ICMP to Support Multi-Part Messages* – ICMPv4 and ICMPv6 Time Exceeded

4.15 Internet Protocol (IP) version 6

RFC 2464, *Transmission of IPv6 Packets over Ethernet Networks*

RFC 2529, *Transmission of IPv6 over IPv4 Domains without Explicit Tunnels*

RFC 3122, *Extensions to IPv6 Neighbor Discovery for Inverse Discovery Specification*

RFC 3315, *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*

RFC 3587, *IPv6 Global Unicast Address Format*

RFC 3596, *DNS Extensions to Support IP version 6*

RFC 3633, *IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6*

RFC 3646, *DNS Configuration options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*

RFC 3736, *Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6*

RFC 3971, *SEcure Neighbor Discovery (SEND)*

RFC 3972, *Cryptographically Generated Addresses (CGA)*

RFC 4007, *IPv6 Scoped Address Architecture*

RFC 4193, *Unique Local IPv6 Unicast Addresses*

RFC 4291, *Internet Protocol Version 6 (IPv6) Addressing Architecture*

RFC 4443, *Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification*

RFC 4861, *Neighbor Discovery for IP version 6 (IPv6)*

RFC 4862, *IPv6 Stateless Address Autoconfiguration* – router functions

RFC 4890, *Recommendations for Filtering ICMPv6 Messages in Firewalls*

RFC 4941, *Privacy Extensions for Stateless Address Autoconfiguration in IPv6*

RFC 5007, *DHCPv6 Leasequery*

RFC 5095, *Deprecation of Type 0 Routing Headers in IPv6*

RFC 5722, *Handling of Overlapping IPv6 Fragments*

RFC 5798, *Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6* – IPv6

RFC 5952, *A Recommendation for IPv6 Address Text Representation*

RFC 6092, *Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service* – Internet Control and Management, Upper-Layer Transport Protocols, UDP Filters, IPsec and Internet Key Exchange (IKE), TCP Filters

RFC 6106, *IPv6 Router Advertisement Options for DNS Configuration*

RFC 6164, *Using 127-Bit IPv6 Prefixes on Inter-Router Links*

RFC 6437, *IPv6 Flow Label Specification*

RFC 6603, *Prefix Exclude Option for DHCPv6-based Prefix Delegation*

RFC 8021, *Generation of IPv6 Atomic Fragments Considered Harmful*

RFC 8200, *Internet Protocol, Version 6 (IPv6) Specification*

RFC 8201, *Path MTU Discovery for IP version 6*

4.16 Internet Protocol Security (IPsec)

draft-ietf-ipsec-isakmp-mode-cfg-05, *The ISAKMP Configuration Method*

draft-ietf-ipsec-isakmp-xauth-06, *Extended Authentication within ISAKMP/Oakley (XAUTH)*

RFC 2401, *Security Architecture for the Internet Protocol*

RFC 2403, *The Use of HMAC-MD5-96 within ESP and AH*

RFC 2404, *The Use of HMAC-SHA-1-96 within ESP and AH*

RFC 2405, *The ESP DES-CBC Cipher Algorithm With Explicit IV*

RFC 2406, *IP Encapsulating Security Payload (ESP)*

RFC 2407, *IPsec Domain of Interpretation for ISAKMP (IPsec DoI)*

RFC 2408, *Internet Security Association and Key Management Protocol (ISAKMP)*

RFC 2409, *The Internet Key Exchange (IKE)*

RFC 2410, *The NULL Encryption Algorithm and Its Use With IPsec*

RFC 2560, *X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP*

RFC 3526, *More Modular Exponential (MODP) Diffie-Hellman group for Internet Key Exchange (IKE)*

RFC 3566, *The AES-XCBC-MAC-96 Algorithm and Its Use With IPsec*

RFC 3602, *The AES-CBC Cipher Algorithm and Its Use with IPsec*

RFC 3706, *A Traffic-Based Method of Detecting Dead Internet Key Exchange (IKE) Peers*

RFC 3947, *Negotiation of NAT-Traversal in the IKE*

RFC 3948, *UDP Encapsulation of IPsec ESP Packets*

RFC 4106, *The Use of Galois/Counter Mode (GCM) in IPsec ESP*

RFC 4109, *Algorithms for Internet Key Exchange version 1 (IKEv1)*

RFC 4301, *Security Architecture for the Internet Protocol*

RFC 4303, *IP Encapsulating Security Payload*

RFC 4307, *Cryptographic Algorithms for Use in the Internet Key Exchange Version 2 (IKEv2)*

RFC 4308, *Cryptographic Suites for IPsec*

RFC 4434, *The AES-XCBC-PRF-128 Algorithm for the Internet Key Exchange Protocol (IKE)*

RFC 4543, *The Use of Galois Message Authentication Code (GMAC) in IPsec ESP and AH*

RFC 4754, *IKE and IKEv2 Authentication Using the Elliptic Curve Digital Signature Algorithm (ECDSA)*

RFC 4835, *Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)*

RFC 4868, *Using HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512 with IPsec*

RFC 4945, *The Internet IP Security PKI Profile of IKEv1/ISAKMP, IKEv2 and PKIX*

RFC 5019, *The Lightweight Online Certificate Status Protocol (OCSP) Profile for High-Volume Environments*

RFC 5282, *Using Authenticated Encryption Algorithms with the Encrypted Payload of the IKEv2 Protocol*

RFC 5903, *ECP Groups for IKE and IKEv2*

RFC 5996, *Internet Key Exchange Protocol Version 2 (IKEv2)*

RFC 5998, *An Extension for EAP-Only Authentication in IKEv2*

RFC 6379, *Suite B Cryptographic Suites for IPsec*

RFC 6380, *Suite B Profile for Internet Protocol Security (IPsec)*

RFC 6960, *X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP*

RFC 7296, *Internet Key Exchange Protocol Version 2 (IKEv2)*

RFC 7321, *Cryptographic Algorithm Implementation Requirements and Usage Guidance for Encapsulating Security Payload (ESP) and Authentication Header (AH)*

RFC 7383, *Internet Key Exchange Protocol Version 2 (IKEv2) Message Fragmentation*

RFC 7427, *Signature Authentication in the Internet Key Exchange Version 2 (IKEv2)*

4.17 Label Distribution Protocol (LDP)

draft-pdutta-mpls-ldp-adj-capability-00, *LDP Adjacency Capabilities*

draft-pdutta-mpls-ldp-v2-00, *LDP Version 2*

draft-pdutta-mpls-mldp-up-redundancy-00, *Upstream LSR Redundancy for Multi-point LDP Tunnels*

draft-pdutta-mpls-multi-ldp-instance-00, *Multiple LDP Instances*

draft-pdutta-mpls-tldp-hello-reduce-04, *Targeted LDP Hello Reduction*

RFC 3037, *LDP Applicability*

RFC 3478, *Graceful Restart Mechanism for Label Distribution Protocol* – helper mode

RFC 5036, *LDP Specification*

RFC 5283, *LDP Extension for Inter-Area Label Switched Paths (LSPs)*

RFC 5443, *LDP IGP Synchronization*

RFC 5561, *LDP Capabilities*

RFC 5919, *Signaling LDP Label Advertisement Completion*

RFC 6388, *Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths*

RFC 6512, *Using Multipoint LDP When the Backbone Has No Route to the Root*

RFC 6826, *Multipoint LDP in-band signaling for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths*

RFC 7032, *LDP Downstream-on-Demand in Seamless MPLS*

RFC 7473, *Controlling State Advertisements of Non-negotiated LDP Applications*

RFC 7552, *Updates to LDP for IPv6*

4.18 Layer Two Tunneling Protocol (L2TP) Network Server (LNS)

draft-mammoliti-l2tp-accessline-avp-04, *Layer 2 Tunneling Protocol (L2TP) Access Line Information Attribute Value Pair (AVP) Extensions*

RFC 2661, *Layer Two Tunneling Protocol "L2TP"*

RFC 2809, *Implementation of L2TP Compulsory Tunneling via RADIUS*

RFC 3438, *Layer Two Tunneling Protocol (L2TP) Internet Assigned Numbers: Internet Assigned Numbers Authority (IANA) Considerations Update*

RFC 3931, *Layer Two Tunneling Protocol - Version 3 (L2TPv3)*

RFC 4719, *Transport of Ethernet Frames over Layer 2 Tunneling Protocol Version 3 (L2TPv3)*

RFC 4951, *Fail Over Extensions for Layer 2 Tunneling Protocol (L2TP) "failover"*

4.19 Multiprotocol Label Switching (MPLS)

draft-ietf-mpls-lsp-ping-ospfv3-codepoint-02, *OSPFv3 CodePoint for MPLS LSP Ping*

RFC 3031, *Multiprotocol Label Switching Architecture*

RFC 3032, *MPLS Label Stack Encoding*

RFC 3270, *Multi-Protocol Label Switching (MPLS) Support of Differentiated Services* – E-LSP

RFC 3443, *Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks*

RFC 4023, *Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)*

RFC 4182, *Removing a Restriction on the use of MPLS Explicit NULL*

RFC 5332, *MPLS Multicast Encapsulations*

RFC 5884, *Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)*

RFC 6374, *Packet Loss and Delay Measurement for MPLS Networks* – Delay Measurement, Channel Type 0x000C

RFC 6424, *Mechanism for Performing Label Switched Path Ping (LSP Ping) over MPLS Tunnels*

RFC 6425, *Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS) - Extensions to LSP Ping*

RFC 6790, *The Use of Entropy Labels in MPLS Forwarding*

RFC 7308, *Extended Administrative Groups in MPLS Traffic Engineering (MPLS-TE)*

RFC 7510, *Encapsulating MPLS in UDP*

RFC 7746, *Label Switched Path (LSP) Self-Ping*

RFC 7876, *UDP Return Path for Packet Loss and Delay Measurement for MPLS Networks* – Delay Measurement

RFC 8029, *Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures*

4.20 Multiprotocol Label Switching - Transport Profile (MPLS-TP)

RFC 5586, *MPLS Generic Associated Channel*

RFC 5921, *A Framework for MPLS in Transport Networks*

RFC 5960, *MPLS Transport Profile Data Plane Architecture*

RFC 6370, *MPLS Transport Profile (MPLS-TP) Identifiers*

RFC 6378, *MPLS Transport Profile (MPLS-TP) Linear Protection*

RFC 6426, *MPLS On-Demand Connectivity and Route Tracing*

RFC 6427, *MPLS Fault Management Operations, Administration, and Maintenance (OAM)*

RFC 6428, *Proactive Connectivity Verification, Continuity Check and Remote Defect indication for MPLS Transport Profile*

RFC 6478, *Pseudowire Status for Static Pseudowires*

RFC 7213, *MPLS Transport Profile (MPLS-TP) Next-Hop Ethernet Addressing*

4.21 Network Address Translation (NAT)

draft-ietf-behave-address-format-10, *IPv6 Addressing of IPv4/IPv6 Translators*

draft-ietf-behave-v6v4-xlate-23, *IP/ICMP Translation Algorithm*

draft-miles-behave-l2nat-00, *Layer2-Aware NAT*

draft-nishitani-cgn-02, *Common Functions of Large Scale NAT (LSN)*

RFC 4787, *Network Address Translation (NAT) Behavioral Requirements for Unicast UDP*

RFC 5382, *NAT Behavioral Requirements for TCP*

RFC 5508, *NAT Behavioral Requirements for ICMP*

RFC 6146, *Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers*

RFC 6333, *Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion*

RFC 6334, *Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite*

RFC 6887, *Port Control Protocol (PCP)*

RFC 6888, *Common Requirements For Carrier-Grade NATs (CGNs)*

RFC 7753, *Port Control Protocol (PCP) Extension for Port-Set Allocation*

RFC 7915, *IP/ICMP Translation Algorithm*

4.22 Network Configuration Protocol (NETCONF)

RFC 5277, *NETCONF Event Notifications*

RFC 6020, *YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)*

RFC 6022, *YANG Module for NETCONF Monitoring*

RFC 6241, *Network Configuration Protocol (NETCONF)*

RFC 6242, *Using the NETCONF Protocol over Secure Shell (SSH)*

RFC 6243, *With-defaults Capability for NETCONF*

RFC 8342, *Network Management Datastore Architecture (NMDA)* – Startup, Candidate, Running and Intended datastores

RFC 8525, *YANG Library*

RFC 8526, *NETCONF Extensions to Support the Network Management Datastore Architecture* – <get-data> operation

4.23 Open Shortest Path First (OSPF)

RFC 1765, *OSPF Database Overflow*

RFC 2328, *OSPF Version 2*

RFC 3101, *The OSPF Not-So-Stubby Area (NSSA) Option*

RFC 3509, *Alternative Implementations of OSPF Area Border Routers*

RFC 3623, *Graceful OSPF Restart Graceful OSPF Restart* – helper mode

RFC 3630, *Traffic Engineering (TE) Extensions to OSPF Version 2*

RFC 4222, *Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance*

RFC 4552, *Authentication/Confidentiality for OSPFv3*

RFC 4576, *Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks (VPNs)*

RFC 4577, *OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)*

RFC 5185, *OSPF Multi-Area Adjacency*

RFC 5187, *OSPFv3 Graceful Restart* – helper mode

RFC 5243, *OSPF Database Exchange Summary List Optimization*

RFC 5250, *The OSPF Opaque LSA Option*

RFC 5309, *Point-to-Point Operation over LAN in Link State Routing Protocols*

RFC 5340, *OSPF for IPv6*

RFC 5642, *Dynamic Hostname Exchange Mechanism for OSPF*

RFC 5709, *OSPFv2 HMAC-SHA Cryptographic Authentication*

RFC 5838, *Support of Address Families in OSPFv3*

RFC 6549, *OSPFv2 Multi-Instance Extensions*

RFC 6987, *OSPF Stub Router Advertisement*

RFC 7471, *OSPF Traffic Engineering (TE) Metric Extensions* – Min/Max Unidirectional Link Delay metric for flex-algo, RSVP, SR-TE

RFC 7684, *OSPFv2 Prefix/Link Attribute Advertisement*

RFC 7770, *Extensions to OSPF for Advertising Optional Router Capabilities*

RFC 8362, *OSPFv3 Link State Advertisement (LSA) Extensibility*

RFC 8920, *OSPF Application-Specific Link Attributes*

4.24 OpenFlow

TS-007 Version 1.3.1, *OpenFlow Switch Specification* – OpenFlow-hybrid switches

4.25 Path Computation Element Protocol (PCEP)

draft-ietf-pce-binding-label-sid-15, *Carrying Binding Label/Segment Identifier (SID) in PCE-based Networks.* – MPLS binding SIDs

draft-alvarez-pce-path-profiles-04, *PCE Path Profiles*

draft-dhs-spring-pce-sr-p2mp-policy-00, *PCEP extensions for p2mp sr policy*

RFC 5440, *Path Computation Element (PCE) Communication Protocol (PCEP)*

RFC 8231, *Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE*

RFC 8253, *PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)*

RFC 8281, *PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model*

RFC 8408, *Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages*

RFC 8664, *Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing*

4.26 Point-to-Point Protocol (PPP)

RFC 1332, *The PPP Internet Protocol Control Protocol (IPCP)*

RFC 1990, *The PPP Multilink Protocol (MP)*

RFC 1994, *PPP Challenge Handshake Authentication Protocol (CHAP)*

RFC 2516, *A Method for Transmitting PPP Over Ethernet (PPPoE)*

RFC 4638, *Accommodating a Maximum Transit Unit/Maximum Receive Unit (MTU/MRU) Greater Than 1492 in the Point-to-Point Protocol over Ethernet (PPPoE)*

RFC 5072, *IP Version 6 over PPP*

4.27 Policy management and credit control

3GPP TS 29.212 Release 11, *Policy and Charging Control (PCC); Reference points* – Gx support as it applies to wireline environment (BNG)

RFC 4006, *Diameter Credit-Control Application*

RFC 6733, *Diameter Base Protocol*

4.28 Pseudowire (PW)

draft-ietf-l2vpn-vpws-iw-oam-04, *OAM Procedures for VPWS Interworking*

MFA Forum 12.0.0, *Multiservice Interworking - Ethernet over MPLS*

MFA Forum 13.0.0, *Fault Management for Multiservice Interworking v1.0*

MFA Forum 16.0.0, *Multiservice Interworking - IP over MPLS*

RFC 3916, *Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)*

RFC 3985, *Pseudo Wire Emulation Edge-to-Edge (PWE3)*

RFC 4385, *Pseudo Wire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN*

RFC 4446, *IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)*

RFC 4447, *Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)*

RFC 4448, *Encapsulation Methods for Transport of Ethernet over MPLS Networks*

RFC 5085, *Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires*

RFC 5659, *An Architecture for Multi-Segment Pseudowire Emulation Edge-to-Edge*

RFC 5885, *Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)*

RFC 6073, *Segmented Pseudowire*

RFC 6310, *Pseudowire (PW) Operations, Administration, and Maintenance (OAM) Message Mapping*

RFC 6391, *Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network*

RFC 6575, *Address Resolution Protocol (ARP) Mediation for IP Interworking of Layer 2 VPNs*

RFC 6718, *Pseudowire Redundancy*

RFC 6829, *Label Switched Path (LSP) Ping for Pseudowire Forwarding Equivalence Classes (FECs) Advertised over IPv6*

RFC 6870, *Pseudowire Preferential Forwarding Status bit*

RFC 7023, *MPLS and Ethernet Operations, Administration, and Maintenance (OAM) Interworking*

RFC 7267, *Dynamic Placement of Multi-Segment Pseudowires*

RFC 7392, *Explicit Path Routing for Dynamic Multi-Segment Pseudowires* – ER-TLV and ER-HOP IPv4 Prefix

RFC 8395, *Extensions to BGP-Signaled Pseudowires to Support Flow-Aware Transport Labels*

4.29 Quality of Service (QoS)

RFC 2430, *A Provider Architecture for Differentiated Services and Traffic Engineering (PASTE)*

RFC 2474, *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*

RFC 2597, *Assured Forwarding PHB Group*

RFC 3140, *Per Hop Behavior Identification Codes*

RFC 3246, *An Expedited Forwarding PHB (Per-Hop Behavior)*

4.30 Remote Authentication Dial In User Service (RADIUS)

draft-oscca-cfrg-sm3-02, *The SM3 Cryptographic Hash Function*

RFC 2865, *Remote Authentication Dial In User Service (RADIUS)*

RFC 2866, *RADIUS Accounting*

RFC 2867, *RADIUS Accounting Modifications for Tunnel Protocol Support*

RFC 2868, *RADIUS Attributes for Tunnel Protocol Support*

RFC 2869, *RADIUS Extensions*

RFC 3162, *RADIUS and IPv6*

RFC 4818, *RADIUS Delegated-IPv6-Prefix Attribute*

RFC 5176, *Dynamic Authorization Extensions to RADIUS*

RFC 6613, *RADIUS over TCP* – with TLS

RFC 6614, *Transport Layer Security (TLS) Encryption for RADIUS*

RFC 6929, *Remote Authentication Dial-In User Service (RADIUS) Protocol Extensions*

RFC 6911, *RADIUS attributes for IPv6 Access Networks*

4.31 Resource Reservation Protocol - Traffic Engineering (RSVP-TE)

draft-newton-mpls-te-dynamic-overbooking-00, *A Diffserv-TE Implementation Model to dynamically change booking factors during failure events*

RFC 2702, *Requirements for Traffic Engineering over MPLS*

RFC 2747, *RSVP Cryptographic Authentication*

RFC 2961, *RSVP Refresh Overhead Reduction Extensions*

RFC 3097, *RSVP Cryptographic Authentication -- Updated Message Type Value*

RFC 3209, *RSVP-TE: Extensions to RSVP for LSP Tunnels*

RFC 3477, *Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)*

RFC 3564, *Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering*

RFC 3906, *Calculating Interior Gateway Protocol (IGP) Routes Over Traffic Engineering Tunnels*

RFC 4090, *Fast Reroute Extensions to RSVP-TE for LSP Tunnels*

RFC 4124, *Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering*

RFC 4125, *Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering*

RFC 4127, *Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering*

RFC 4561, *Definition of a Record Route Object (RRO) Node-Id Sub-Object*

RFC 4875, *Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)*

RFC 4950, *ICMP Extensions for Multiprotocol Label Switching*

RFC 5712, *MPLS Traffic Engineering Soft Preemption*

RFC 5817, *Graceful Shutdown in MPLS and Generalized MPLS Traffic Engineering Networks*

4.32 Routing Information Protocol (RIP)

RFC 1058, *Routing Information Protocol*

RFC 2080, *RIPng for IPv6*

RFC 2082, *RIP-2 MD5 Authentication*

RFC 2453, *RIP Version 2*

4.33 Segment Routing (SR)

draft-bashandy-rtgwg-segment-routing-uloop-06, *Loop avoidance using Segment Routing*

draft-filsfils-spring-net-pgm-extension-srv6-usid-13, *Network Programming extension: SRv6 uSID instruction*

draft-filsfils-spring-srv6-net-pgm-insertion-04, *SRv6 NET-PGM extension: Insertion*

draft-ietf-6man-spring-srv6-oam-10, *Operations, Administration, and Maintenance (OAM) in Segment Routing Networks with IPv6 Data plane (SRv6)*

draft-ietf-idr-bgp-ls-segment-routing-ext-16, *BGP Link-State extensions for Segment Routing*

draft-ietf-idr-bgpls-srv6-ext-13, *BGP Link State Extensions for SRv6*

draft-ietf-idr-segment-routing-te-policy-11, *Advertising Segment Routing Policies in BGP*

draft-ietf-isis-mpls-elc-10, *Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS* – advertising ELC

draft-ietf-lsr-flex-algo-16, *IGP Flexible Algorithm*

draft-ietf-lsr-isis-srv6-extensions-14, *IS-IS Extension to Support Segment Routing over IPv6 Dataplane*

draft-ietf-ospf-mpls-elc-12, *Signaling Entropy Label Capability and Entropy Readable Label-stack Depth Using OSPF* – advertising ELC

draft-ietf-rtgwg-segment-routing-ti-lfa-01, *Topology Independent Fast Reroute using Segment Routing*

draft-ietf-spring-conflict-resolution-05, *Segment Routing MPLS Conflict Resolution*

draft-ietf-spring-segment-routing-policy-08, *Segment Routing Policy Architecture*

draft-ietf-teas-sr-rsvp-coexistence-rec-02, *Recommendations for RSVP-TE and Segment Routing LSP co-existence*

draft-voyer-6man-extension-header-insertion-10, *Deployments With Insertion of IPv6 Segment Routing Headers*

draft-voyer-pim-sr-p2mp-policy-02, *Segment Routing Point-to-Multipoint Policy*

draft-voyer-spring-sr-p2mp-policy-03, *SR Replication Policy for P2MP Service Delivery*

RFC 8287, *Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes*

RFC 8476, *Signaling Maximum SID Depth (MSD) Using OSPF* – node MSD

RFC 8491, *Signaling Maximum SID Depth (MSD) Using IS-IS* – node MSD

RFC 8660, *Segment Routing with the MPLS Data Plane*

RFC 8661, *Segment Routing MPLS Interworking with LDP*

RFC 8663, *MPLS Segment Routing over IP* – BGP SR with SR-MPLS-over-UDP/IP

RFC 8665, *OSPF Extensions for Segment Routing*

RFC 8666, *OSPFv3 Extensions for Segment Routing*

RFC 8667, *IS-IS Extensions for Segment Routing*

RFC 8669, *Segment Routing Prefix Segment Identifier Extensions for BGP*

RFC 8754, *IPv6 Segment Routing Header (SRH)*

RFC 8814, *Signaling Maximum SID Depth (MSD) Using the Border Gateway Protocol - Link State*

RFC 8986, *Segment Routing over IPv6 (SRv6) Network Programming*

RFC 9252, *BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)*

4.34 Simple Network Management Protocol (SNMP)

draft-blumenthal-aes-usm-04, *The AES Cipher Algorithm in the SNMP's User-based Security Model* – CFB128-AES-192 and CFB128-AES-256

draft-ietf-isis-wg-mib-06, *Management Information Base for Intermediate System to Intermediate System (IS-IS)*

draft-ietf-mboned-msdp-mib-01, *Multicast Source Discovery protocol MIB*

draft-ietf-mpls-ldp-mib-07, *Definitions of Managed Objects for the Multiprotocol Label Switching, Label Distribution Protocol (LDP)*

draft-ietf-mpls-lsr-mib-06, *Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base Using SMIv2*

draft-ietf-mpls-te-mib-04, *Multiprotocol Label Switching (MPLS) Traffic Engineering Management Information Base*

draft-ietf-ospf-mib-update-08, *OSPF Version 2 Management Information Base*

draft-ietf-vrrp-unified-mib-06, *Definitions of Managed Objects for the VRRP over IPv4 and IPv6* – IPv6

ESO-CONSORTIUM-MIB revision 200406230000Z, *esoConsortiumMIB*

IANA-ADDRESS-FAMILY-NUMBERS-MIB revision 200203140000Z, *ianaAddressFamilyNumbers*

IANAifType-MIB revision 200505270000Z, *ianaifType*

IANA-RTPROTO-MIB revision 200009260000Z, *ianaRtProtoMIB*

IEEE8021-CFM-MIB revision 200706100000Z, *ieee8021CfmMib*

IEEE8021-PAE-MIB revision 200101160000Z, *ieee8021paeMIB*

IEEE8023-LAG-MIB revision 200006270000Z, *lagMIB*

LLDP-MIB revision 200505060000Z, *lldpMIB*

RFC 1157, *A Simple Network Management Protocol (SNMP)*

RFC 1212, *Concise MIB Definitions*

RFC 1215, *A Convention for Defining Traps for use with the SNMP*

RFC 1724, *RIP Version 2 MIB Extension*

RFC 1901, *Introduction to Community-based SNMPv2*

RFC 2021, *Remote Network Monitoring Management Information Base Version 2 using SMIv2*

RFC 2206, *RSVP Management Information Base using SMIv2*

RFC 2213, *Integrated Services Management Information Base using SMIv2*

RFC 2494, *Definitions of Managed Objects for the DS0 and DS0 Bundle Interface Type*

RFC 2578, *Structure of Management Information Version 2 (SMIv2)*

RFC 2579, *Textual Conventions for SMIv2*

RFC 2580, *Conformance Statements for SMIv2*

RFC 2787, *Definitions of Managed Objects for the Virtual Router Redundancy Protocol*

RFC 2819, *Remote Network Monitoring Management Information Base*

RFC 2856, *Textual Conventions for Additional High Capacity Data Types*

RFC 2863, *The Interfaces Group MIB*

RFC 2864, *The Inverted Stack Table Extension to the Interfaces Group MIB*

RFC 2933, *Internet Group Management Protocol MIB*

RFC 3014, *Notification Log MIB*

RFC 3165, *Definitions of Managed Objects for the Delegation of Management Scripts*

RFC 3231, *Definitions of Managed Objects for Scheduling Management Operations*

RFC 3273, *Remote Network Monitoring Management Information Base for High Capacity Networks*

RFC 3410, *Introduction and Applicability Statements for Internet Standard Management Framework*

RFC 3411, *An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks*

RFC 3412, *Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)*

RFC 3413, *Simple Network Management Protocol (SNMP) Applications*

RFC 3414, *User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)*

RFC 3415, *View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)*

RFC 3416, *Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)*

RFC 3417, *Transport Mappings for the Simple Network Management Protocol (SNMP)* – SNMP over UDP over IPv4

RFC 3418, *Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)*

RFC 3419, *Textual Conventions for Transport Addresses*

RFC 3498, *Definitions of Managed Objects for Synchronous Optical Network (SONET) Linear Automatic Protection Switching (APS) Architectures*

RFC 3584, *Coexistence between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework*

RFC 3592, *Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type*

RFC 3593, *Textual Conventions for MIB Modules Using Performance History Based on 15 Minute Intervals*

RFC 3635, *Definitions of Managed Objects for the Ethernet-like Interface Types*

RFC 3637, *Definitions of Managed Objects for the Ethernet WAN Interface Sublayer*

RFC 3826, *The Advanced Encryption Standard (AES) Cipher Algorithm in the SNMP User-based Security Model*

RFC 3877, *Alarm Management Information Base (MIB)*

RFC 3895, *Definitions of Managed Objects for the DS1, E1, DS2, and E2 Interface Types*

RFC 3896, *Definitions of Managed Objects for the DS3/E3 Interface Type*

RFC 4001, *Textual Conventions for Internet Network Addresses*

RFC 4022, *Management Information Base for the Transmission Control Protocol (TCP)*

RFC 4113, *Management Information Base for the User Datagram Protocol (UDP)*

RFC 4220, *Traffic Engineering Link Management Information Base*

RFC 4273, *Definitions of Managed Objects for BGP-4*

RFC 4292, *IP Forwarding Table MIB*

RFC 4293, *Management Information Base for the Internet Protocol (IP)*

RFC 4631, *Link Management Protocol (LMP) Management Information Base (MIB)*

RFC 4878, *Definitions and Managed Objects for Operations, Administration, and Maintenance (OAM) Functions on Ethernet-Like Interfaces*

RFC 7420, *Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module*

RFC 7630, *HMAC-SHA-2 Authentication Protocols in the User-based Security Model (USM) for SNMPv3*

SFLOW-MIB revision 200309240000Z, *sFlowMIB*

4.35 Timing

GR-1244-CORE Issue 3, *Clocks for the Synchronized Network: Common Generic Criteria*

GR-253-CORE Issue 3, *SONET Transport Systems: Common Generic Criteria*

IEEE 1588-2008, *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*

ITU-T G.781, *Synchronization layer functions*

ITU-T G.813, *Timing characteristics of SDH equipment slave clocks (SEC)*

ITU-T G.8261, *Timing and synchronization aspects in packet networks*

ITU-T G.8262, *Timing characteristics of synchronous Ethernet equipment slave clock (EEC)*

ITU-T G.8262.1, *Timing characteristics of an enhanced synchronous Ethernet equipment slave clock (eEEC)*

ITU-T G.8264, *Distribution of timing information through packet networks*

ITU-T G.8265.1, *Precision time protocol telecom profile for frequency synchronization*

ITU-T G.8275.1, *Precision time protocol telecom profile for phase/time synchronization with full timing support from the network*

RFC 3339, *Date and Time on the Internet: Timestamps*

RFC 5905, *Network Time Protocol Version 4: Protocol and Algorithms Specification*

4.36 Two-Way Active Measurement Protocol (TWAMP)

RFC 5357, *A Two-Way Active Measurement Protocol (TWAMP)* – server, unauthenticated mode

RFC 5938, *Individual Session Control Feature for the Two-Way Active Measurement Protocol (TWAMP)*

RFC 6038, *Two-Way Active Measurement Protocol (TWAMP) Reflect Octets and Symmetrical Size Features*

RFC 8545, *Well-Known Port Assignments for the One-Way Active Measurement Protocol (OWAMP) and the Two-Way Active Measurement Protocol (TWAMP)* – TWAMP

RFC 8762, *Simple Two-Way Active Measurement Protocol* – unauthenticated

RFC 8972, *Simple Two-Way Active Measurement Protocol Optional Extensions* – unauthenticated

4.37 Virtual Private LAN Service (VPLS)

RFC 4761, *Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling*

RFC 4762, *Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling*

RFC 5501, *Requirements for Multicast Support in Virtual Private LAN Services*

RFC 6074, *Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)*

RFC 7041, *Extensions to the Virtual Private LAN Service (VPLS) Provider Edge (PE) Model for Provider Backbone Bridging*

RFC 7117, *Multicast in Virtual Private LAN Service (VPLS)*

4.38 Voice and video

DVB BlueBook A86, *Transport of MPEG-2 TS Based DVB Services over IP Based Networks*

ETSI TS 101 329-5 Annex E, *QoS Measurement for VoIP - Method for determining an Equipment Impairment Factor using Passive Monitoring*

ITU-T G.1020 Appendix I, *Performance Parameter Definitions for Quality of Speech and other Voiceband Applications Utilizing IP Networks - Mean Absolute Packet Delay Variation & Markov Models*

ITU-T G.107, *The E Model - A computational model for use in planning*

ITU-T P.564, *Conformance testing for voice over IP transmission quality assessment models*

RFC 3550, *RTP: A Transport Protocol for Real-Time Applications* – Appendix A.8

RFC 4585, *Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)*

RFC 4588, *RTP Retransmission Payload Format*

4.39 Wireless Local Area Network (WLAN) gateway

3GPP TS 23.402, *Architecture enhancements for non-3GPP accesses* – S2a roaming based on GPRS

4.40 Yet Another Next Generation (YANG)

RFC 6991, *Common YANG Data Types*

RFC 7950, *The YANG 1.1 Data Modeling Language*

RFC 7951, *JSON Encoding of Data Modeled with YANG*

4.41 Yet Another Next Generation (YANG) OpenConfig Modules

openconfig-aaa.yang version 0.4.0, *OpenConfig AAA Module*

openconfig-aaa-radius.yang version 0.3.0, *OpenConfig AAA RADIUS Module*

openconfig-aaa-tacacs.yang version 0.3.0, *OpenConfig AAA TACACS+ Module*

openconfig-acl.yang version 1.0.0, *OpenConfig ACL Module*

openconfig-bfd.yang version 0.2.2, *OpenConfig BFD Module*

openconfig-bgp.yang version 6.1.0, *OpenConfig BGP Module*

openconfig-bgp-common.yang version 6.0.0, *OpenConfig BGP Common Module*

openconfig-bgp-common-multiprotocol.yang version 6.0.0, *OpenConfig BGP Common Multiprotocol Module*

openconfig-bgp-common-structure.yang version 6.0.0, *OpenConfig BGP Common Structure Module*

openconfig-bgp-global.yang version 6.0.0 *OpenConfig BGP Global Module*

openconfig-bgp-neighbor.yang version 6.1.0, *OpenConfig BGP Neighbor Module*

openconfig-bgp-peer-group.yang version 6.1.0, *OpenConfig BGP Peer Group Module*

openconfig-bgp-policy.yang version 4.0.1, *OpenConfig BGP Policy Module*

openconfig-if-aggregate.yang version 2.0.0, *OpenConfig Interfaces Aggregated Module*

openconfig-if-ethernet.yang version 2.0.0, *OpenConfig Interfaces Ethernet Module*

openconfig-if-ip.yang version 2.0.0, *OpenConfig Interfaces IP Module*

openconfig-if-ip-ext.yang version 2.0.0, *OpenConfig Interfaces IP Extensions Module*

openconfig-igmp.yang version 0.2.0, *OpenConfig IGMP Module*

openconfig-interfaces.yang version 2.0.0, *OpenConfig Interfaces Module*

openconfig-isis.yang version 0.3.2, *OpenConfig IS-IS Module*

openconfig-isis-policy.yang version 0.3.2, *OpenConfig IS-IS Policy Module*

openconfig-isis-routing.yang version 0.3.2, *OpenConfig IS-IS Routing Module*

openconfig-lacp.yang version 1.1.0, *OpenConfig LACP Module*

openconfig-lldp.yang version 0.1.0, *OpenConfig LLDP Module*

openconfig-local-routing.yang version 1.2.0, *OpenConfig Local Routing Module*

openconfig-mpls.yang version 2.3.0, *OpenConfig MPLS Module*

openconfig-mpls-ldp.yang version 3.0.2, *OpenConfig MPLS LDP Module*

openconfig-mpls-rsvp.yang version 2.3.0, *OpenConfig MPLS RSVP Module*

openconfig-mpls-te.yang version 2.3.0, *OpenConfig MPLS TE Module*

openconfig-network-instance.yang version 1.1.0, *OpenConfig Network Instance Module*

openconfig-network-instance-l3.yang version 0.11.1, *OpenConfig L3 Network Instance Module* – static routes

openconfig-packet-match.yang version 1.0.0, *OpenConfig Packet Match Module*

openconfig-pim.yang version 0.2.0 *OpenConfig PIM Module*

openconfig-platform.yang version 0.15.0, *OpenConfig Platform Module*

openconfig-platform-fan.yang version 0.1.1, *OpenConfig Platform Fan Module*

openconfig-platform-linecard.yang version 0.1.2, *OpenConfig Platform Linecard Module*

openconfig-platform-port.yang version 0.4.2, *OpenConfig Port Module*

openconfig-platform-transceiver.yang version 0.9.0, *OpenConfig Transceiver Module*

openconfig-procmon.yang version 0.4.0, *OpenConfig Process Monitoring Module*

openconfig-relay-agent.yang version 0.1.0, *OpenConfig Relay Agent Module*

openconfig-routing-policy.yang version 3.0.0, *OpenConfig Routing Policy Module*

openconfig-rsvp-sr-ext.yang version 0.1.0, *OpenConfig RSVP-TE and SR Extensions Module*

openconfig-system.yang version 0.10.1, *OpenConfig System Module*

openconfig-system-grpc.yang version 1.0.0, *OpenConfig System gRPC Module*

openconfig-system-logging.yang version 0.3.1, *OpenConfig System Logging Module*

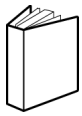
openconfig-system-terminal.yang version 0.3.0, *OpenConfig System Terminal Module*

openconfig-telemetry.yang version 0.5.0, *OpenConfig Telemetry Module*

openconfig-terminal-device.yang version 1.9.0, *OpenConfig Terminal Optics Device Module*

openconfig-vlan.yang version 2.0.0, *OpenConfig VLAN Module*

Customer document and product support



Customer documentation

Customer documentation welcome page

Technical support

Product support portal

Documentation feedback

Customer documentation feedback