Internet Enhanced Service

In This Chapter

This chapter provides information about Internet Enhanced Service (IES), process overview, and implementation notes.

Topics in this chapter include:

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.

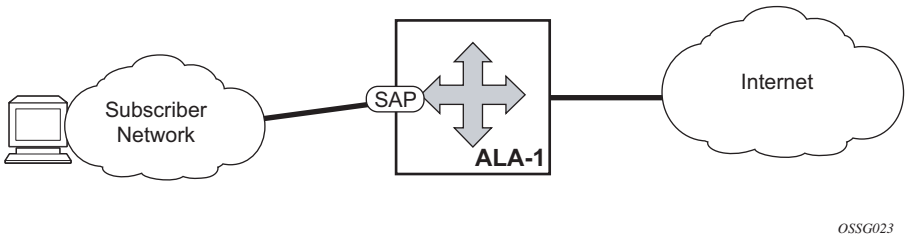

Figure 1: Internet Enhanced Service

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.

Note: Refer to the 7750 SR Triple Play Guide for information about how subscriber group-interfaces function in the Routed Central Office model.

IES Features

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

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) on page 19
- 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

QoS Policy Propagation Using BGP (QPPB)

This section discusses QPPB as it applies to VPRN, IES, and router interfaces. Refer to the Internet Enhanced Service section on page 15 and the IP Router Configuration section in the 7x50 OS 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 routers, 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, 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 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. 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.

QPPB Applications

There are two typical applications of QPPB:

1. Coordination of QoS policies between different administrative domains.
2. Traffic differentiation within a single domain, based on route characteristics.

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 certain 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 towards network XYZ into the QoS class implied by the BGP community value.

QPPB may also be used to request that traffic sourced from certain 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.

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 certain AS_PATH, or destined for a particular next-hop (remaining on-net vs. off-net).

Figure 2 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. Note however, that the DSCP or other COS markings could be left unchanged in the ISP's network and QPPB used on every node.

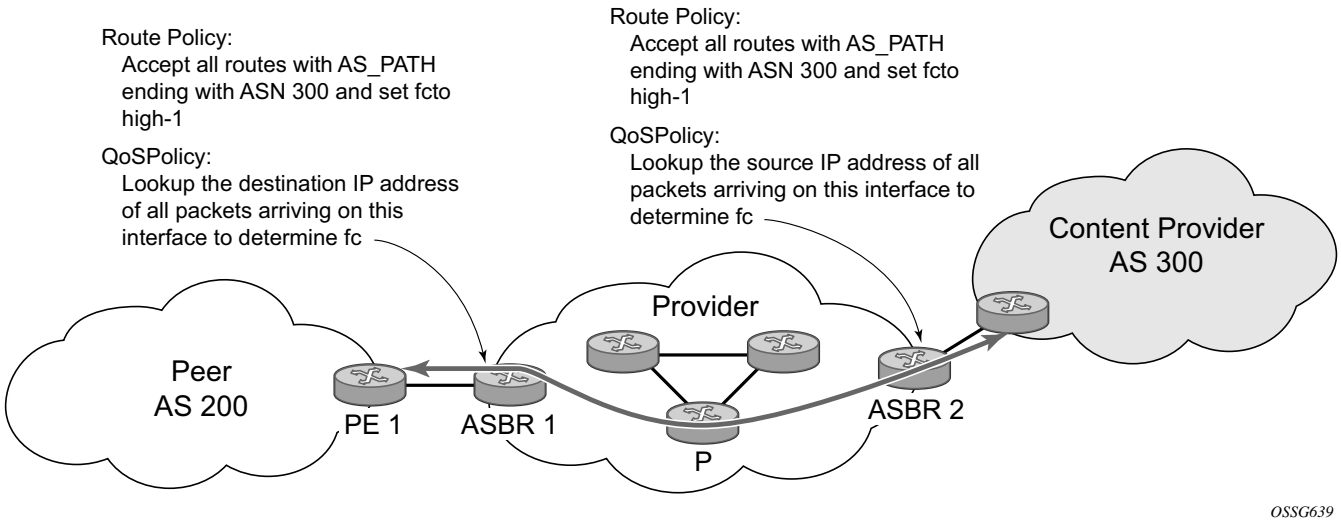

Figure 2: Use of QPPB to Differentiate Traffic in an ISP Network

QPPB

There are two main aspects of the QPPB feature:

- The ability to associate a forwarding-class and priority with certain 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.

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)

Note that 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 using the following modified versions of the static-route commands:

- static-route {*ip-prefix*/*prefix-length* | *ip-prefix netmask*} [fc *fc-name* [priority {low | high}]] next-hop *ip-int-name*|*ip-address*
- static-route {*ip-prefix*/*prefix-length* | *ip-prefix netmask*} [fc *fc-name* [priority {low | high}]] indirect *ip-address*

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.

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#
```

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. Note however, 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 (see section 5.7 for further details). 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.

QPPB When Next-Hops are Resolved by QPPB Routes

In some circumstances (IP VPN inter-AS model C, Carrier Supporting Carrier, indirect static routes, etc.) 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, etc. In release 9.0 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, etc. have forwarding-class and priority associations.

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 given flow on the same path. The QPPB configuration model described in Associating an FC and Priority with a Route on page 22 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 on page 22 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.

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

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.

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 will be 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 or not **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 will be 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 will be 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 will be based on QPPB (unless DE=1), but if no priority information is associated with the route then the packet priority will be 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 summarizes these 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
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

Table 2: QPPB Interactions with SAP Ingress QoS (Continued)

Original FC object mapping	New FC object mapping	Profile	Priority (drop preference)	DE=1 override	In/out of profile marking
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

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 certain 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 certain actions as a result of status transitions. For example, if the operational group goes down, the monitoring objects will be brought down.

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 member(s).

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 is example uses an R-VPLS context. The VPLS instance includes the **allow-ip-int-binding** and the **service-name v1**. The IES interface links to the VPLS using the **vpls v1** option. All commands are under the configuration service hierarchy.

To further explain 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 will derive its state from the state of oper-group g1.

```
oper-group g1 create

vpls 1 customer 1 create
       allow-ip-int-binding
       stp
           shutdown
       exit
       service-name "v1"
       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 21.1.1.1/24
           monitor-oper-group "g1"
           vpls "v1"
       exit
     no shutdown
     exit
```

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 interface — An interface that allows the sharing of a subnet among one or many group interfaces in the routed CO model.
- Group interface — Aggregates multiple SAPs on the same port.
- Redundant interfaces — 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 given subscriber.

IPv6 Enhanced Subscriber Management (ESM)

All IPv6 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).

RADIUS Accounting

In the 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 thus 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 will 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 will be 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 new behavior means certain AVP may be in either host, sla-profile instance, or both accounting records.

Note that 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	No
NAS-Identifier	Yes	Yes
NAS-IP-Address	Yes	Yes
Nas-Port-Id	Yes	No
Nas-Port	Yes	No
Nas-Port-Type	Yes	No
Service-Type	Yes	No
Framed-Protocol	Yes	No
Framed-Ip-Address	Yes	No
Framed-Ip-Netmark	Yes	No
Framed-Route	Yes	No
Class	Yes	No
Session-Timeout	Yes	Yes
Circuit-Id VSA	Yes	No
Called-Station-Id	Yes	No
Calling-Station-Id	Yes	No
MAC-Addr VSA	Yes	No
Remote-Id VSA	Yes	No
Acct-Input-Octets	No	Yes
Acct-Output-Octets	No	Yes
Acct-Input-Gigawords	No	Yes
Acct-Output-Gigawords	No	Yes
Acct-Session-Id	Yes	Yes
Acct-Session-Time	Yes	Yes

Table 3: RADIUS Accounting Table (Continued)

RADIUS Accounting AVP	Host Accounting	SLA-Profile Accounting
Acct-Input-Packets	No	Yes
Acct-Output-Packets	No	Yes
Agent-Circuit-Id	Yes	No
Agent-Remote-Id	Yes	No
Actual-Data-Rate-Upstream	Yes	No
Actual-Data-Rate-Downstream	Yes	No
Access-Loop-Encapsulation	Yes	No
Alc-”Accounting”	No	Yes
Alc-Subscriber-Id	Yes	Yes
Alc-Subscriber-Profile-String	Yes	Yes
Alc-Sla-Profile-String	Yes	Yes

SAPs

Encapsulations

The following SAP encapsulations are supported on IES services:

- Ethernet null
- Ethernet dot1q
- SONET/SDH IPCP
- SONET/SDH BCP-null
- SONET/SDH BCP-dot1q
- SONET/SDH ATM
- ATM - LLC SNAP or VC-MUX

ATM SAP Encapsulations for IES

The 7750 SR series supports ATM PVC service encapsulation for IES SAPs. Both UNI and NNI cell formats are supported. The format is configurable on a SONET/SDH path basis. A path maps to an ATM VC. All VCs on a path must use the same cell format.

The following ATM encapsulation and transport modes are supported:

- RFC 2684, *Multiprotocol Encapsulation over ATM Adaptation Layer 5*:
  - → AAL5 LLC/SNAP IPv4 routed
  - → AAL5 VC mux IPv4 routed
  - → AAL5 LLC/SNAP IPv4 bridged
  - → AAL5 VC mux IPv4 bridged

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 SR OS Triple Play Service Delivery Architecture user 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 given access-port on the Access-PE can be backhauled over a single MPLS pseudowire towards 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 (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 and on the HSMDAv2.

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

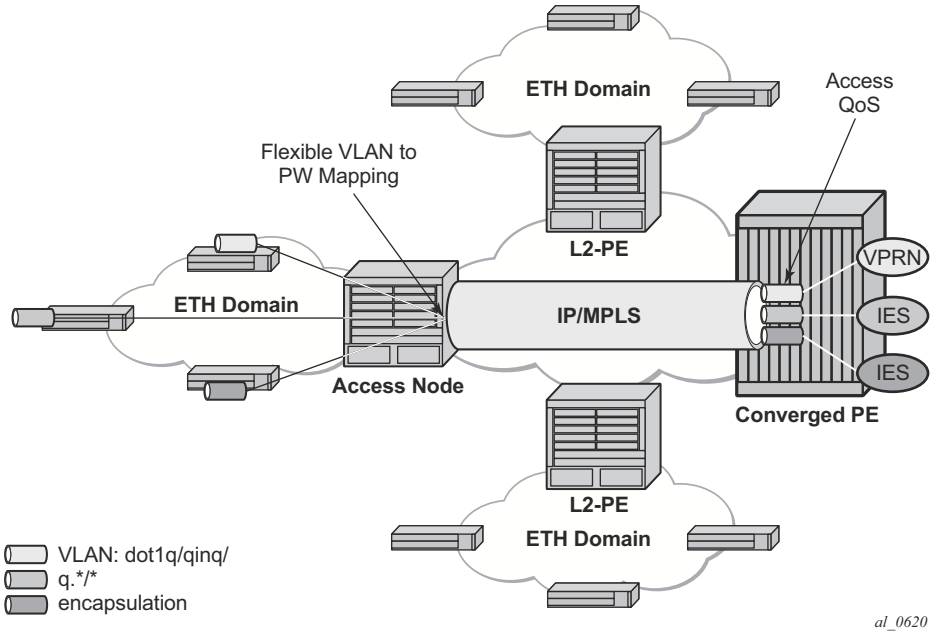

Figure 3: Network Architecture using Pseudowire SAPs

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 optional control word is not supported. The SDP could use an LDP LSP, RSVP LSP, BGP RFC3107 tunnel, or LDP over RSVP tunnel. Hash labels are not supported. The SDP may be bound to a port or a LAG, although note that 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.

Pseudowire SAP Configuration

The following steps 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 following steps are used to configure a pseudowire SAP on the IES or VPRN service at the Layer 3 PE:

1. Define a pseudowire port

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

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
```

3. Perform one of the following steps:
   a. 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 30.1.1.1/24
mac 00:00:00:00:00:ff
   static-arp 30.1.1.2 00:00:00:00:00:aa
   sap pw-1:1 create
      exit
   exit
   no shutdown
      exit
```

b. 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
```

The PW SAP may be mated to an Ethernet SAP or an Ethernet spoke-sdp in the Epipe VLL service

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.

  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 IOM3 with Ethernet MDA or HSMDAv2, and policy based redirection for the IOM3 with Ethernet MDA, as applicable.
- Policing and H-POL

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 will inherit the vport used by the pw-port egress shaping.

Note that the vport is represented by a secondary shaper on an HSMDAv2. 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(s) 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 (**configure>port>ethernet>access>egress>vport** *vport-name* **create**).
2. Explicitly assign (via 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.

As in the ESM on pseudowire case, vport scheduling on the HSMDAv2 is implemented using an exp-secondary-shaper. This is referred to as a pw-sap-secondary-shaper in the new CLI below. If an 'hsmda-queue-override secondary-shape' is defined for the pw-sap, then the system will use the override, else:

- If a named pw-sap-secondary-shaper is defined for the pw-port, then that is used,
- Else, the default exp-secondary-shaper for the port is used.

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) or secondary shaper (on HSMDAv2).

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. See the SR OS Quality of Service guide for more details.

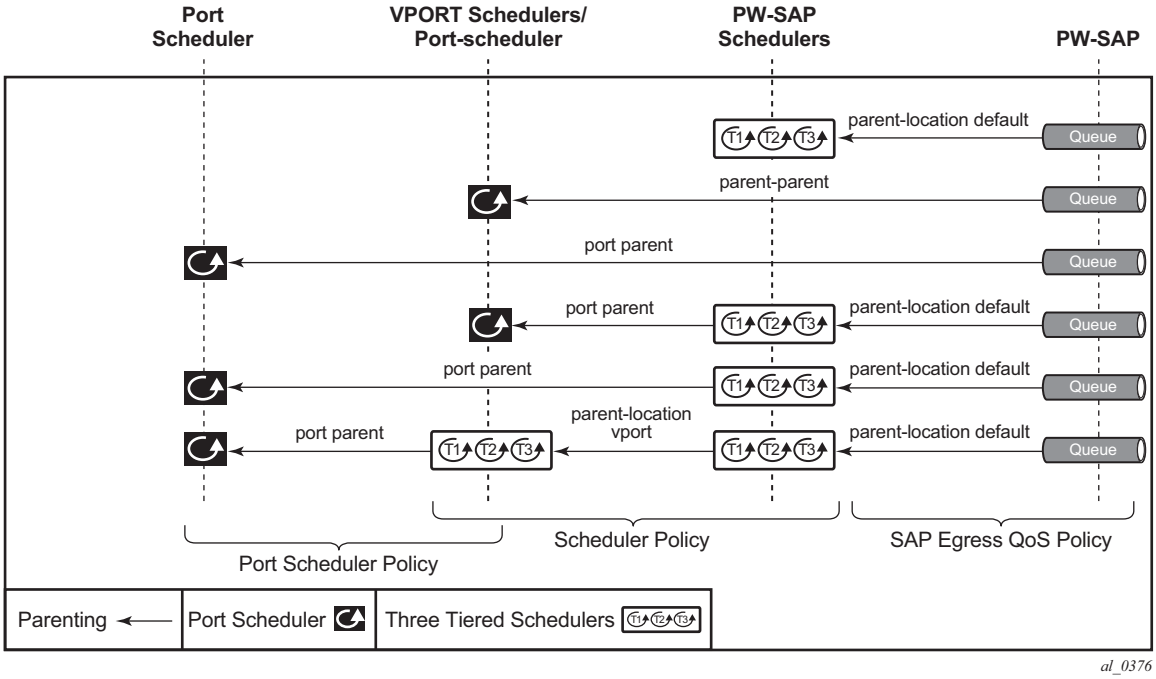

Figure 4: PW SAP Egress Scheduling Hierarchy Options

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.

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.

Note that the exp-secondary-shaper used on the HSMDAv2 always assumes MPLS overhead and does not account for the packet-byte-offset. In all other cases, 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 summarizes the default packet sizes used at each of the schedulers on the IOM/Ethernet MDA and HSMDAv2, assuming a 1000byte customer packet.

Table 4: Packet Sizes Used for Pseudowire SAPs

Type	Size
exp-secondary-shaper	20B preamble + 26 MPLS + 1000B pkt
queue/policer rate on hsmdav2	1000B customer 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

Redundancy with Pseudowire SAPs

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.

In order 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.

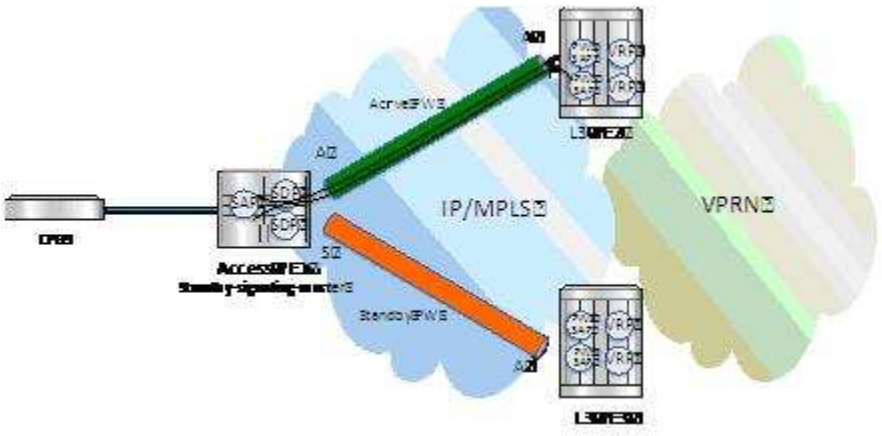

Figure 5: Dual Homing into Multiple Layer 3 PEs

Dual homing operates in a similar manner to spoke-sdp termination on IES/VPRN. Figure 5 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 given spoke-sdp and the local configuration does not allow this bit to be ignored, the PE will take 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 thus will stop forwarding packets towards 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 towards 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. Support for existing master-slave PW redundancy mechanisms for dual homing of the access PE into separate converged PEs using active/standby PW status is required as shown in Figure 6.

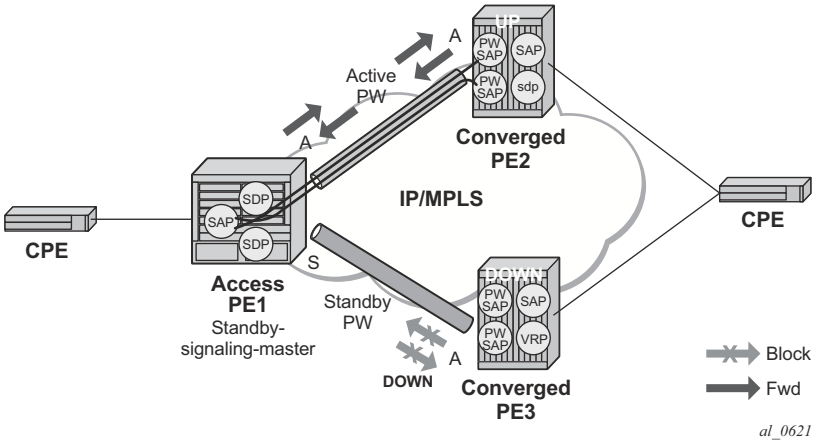

Figure 6: Master-Slave PW Redundancy

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 for Epipe spoke-sdps. That is, when enabled, if a PW Forwarding Standby (0x20) LDP status message is received for the PW, 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 e.g. 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.

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 will also go operationally down, and thus the corresponding PW status bit signaled (0x00000001 - Pseudowire Not Forwarding). Note that, if a status of 0x00000001 is signaled for a currently active PW, and active/standby dual homing is in use then the access PE will fail 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 i.e. 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"
```

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 that 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.

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 will be 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 ICPM ping mechanism is used to test the connectivity.

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

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. Note that both Layer 2 or Layer 3 criteria can be used in the QoS policies for traffic classification in an IES.

Filter Policies

Only IP filter policies can be applied to IES services.

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. Refer to the *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 not access QoS policies.

Figure 7 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.

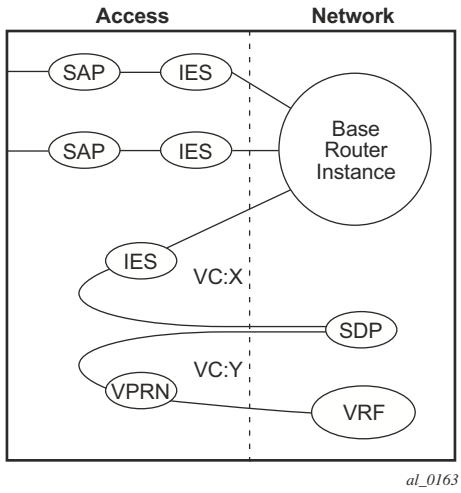

Figure 7: SDP-ID and VC Label Service Identifiers

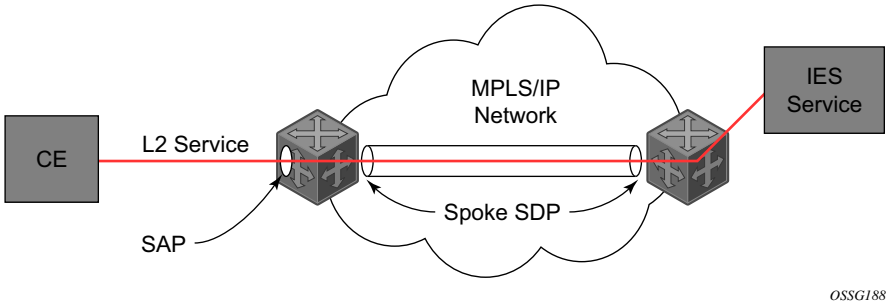

Figure 8: IES Spoke-SDP Termination

Figure 8 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.

Note that, 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.

SRRP

Subscriber Router Redundancy Protocol (SRRP) 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. Once 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. Once 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. Since 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 proper 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 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:

- 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).
- Create a redundant IP interface and associate it with the SRRP instances group IP interface for shunting traffic to the remote router when master.
- 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:

- 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).
- Define a description string for the SRRP instance.
- 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).
- Change the base priority for the SRRP instance.
- Specify one or more VRRP policies to dynamically manage the SRRP instance base priority.
- Specify a new keep alive interval for the SRRP instance.

Table 5 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 will 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 will use the native group IP interface MAC address. • The group IP interface redundant IP interface will not be used. • Will not accept packets destined to the SRRP gateway MAC received on the group IP interface.
Becoming Master (In order 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 use the native group IP interface MAC address. • Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface. • Will not accept packets destined to the SRRP gateway MAC received on the group IP interface.

Table 5: SRRP State Effect on Subscriber Hosts Associated with Group IP Interface (Continued)

SRRP State	ARP	Local Proxy ARP Enabled	Remote Proxy ARP Enabled	Subscriber Host Routing
Master	• 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 will use the SRRP gateway MAC address. • Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface. • Will accept packets destined to 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). • Will not respond to ARP for subscriber subnet SRRP gateway IP addresses	• Will not respond to ARP for any subscriber hosts on the subscriber subnet.	• Will not respond to ARP for any remote IP hosts.	• Will 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. • Will accept packets destined to the SRRP gateway MAC received on the group IP interface.

Table 5: SRRP State Effect on Subscriber Hosts Associated with Group IP Interface (Continued)

SRRP State	ARP	Local Proxy ARP Enabled	Remote Proxy ARP Enabled	Subscriber Host Routing
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). • Will not respond to ARP for subscriber subnet SRRP gateway IP addresses.	• Will not respond to ARP for any subscriber hosts on the subscriber subnet.	• Will 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 remapped to the group IP interface. • Will accept packets destined to 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.	• Will not respond to ARP for any subscriber hosts on the subscriber subnet	• Will not respond to ARP for any remote IP hosts	• Will 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. • Will not accept packets destined to the SRRP gateway MAC received on the group IP interface.

Table 5: SRRP State Effect on Subscriber Hosts Associated with Group IP Interface (Continued)

SRRP State	ARP	Local Proxy ARP Enabled	Remote Proxy ARP Enabled	Subscriber Host Routing
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). • Will not respond to ARP for subscriber subnet SRRP gateway IP addresses.	• Will not respond to ARP for any subscriber hosts on the subscriber subnet.	• Will 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 remapped to the group IP interface. • Will not accept packets destined to the SRRP gateway MAC received on the group IP interface.

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.

Due to 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. Since only two nodes are participating, the VRRP skew timer is not utilized 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 will 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 will be used to choose the master. If a master already exists, despite having the lower IP address, the system will 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.

SRRP and Multi-Chassis Synchronization

In order 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 will synchronize 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 SRRP instance MCS key
- Containing service type and ID
- Containing subscriber IP interface name
- Subscriber subnet information
- Containing group IP interface information
- The SRRP group IP interface redundant IP interface name, IP address and mask
- The SRRP advertisement message SAP
- The local system IP address (SRRP advertisement message source IP address)
- The Group IP interface MAC address
- The SRRP gateway MAC address
- The SRRP instance administration state (up / down)
- The SRRP instance operational state (disabled / becoming-backup / backup / becoming-master / master)
- The current SRRP priority
- Remote redundant IP interface availability (available / unavailable)
- Local receive SRRP advertisement SAP availability (available / unavailable)

SRRP Instance

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

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. Once 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.

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.

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.

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. Once the subscriber subnet information matches, an SRRP Subscriber Subnet Mismatch Clear event is generated.

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.

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.

Remote Sending Redundant IP Interface Unavailable

If the remote node is sending redundant IP interface unavailable, the local node will treat 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.

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. Since SRRP advertisement messages cannot be received, the local node will immediately become master if it has the lower system IP address.

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 will be 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. Since the remote node cannot receive SRRP advertisement messages, the local node will immediately become master if it has the lower system IP address.

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.

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 will be 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 given subscriber will always flow through one node.

Subscriber Subnet SRRP Gateway IP Address Connectivity

The SRRP gateway IP addresses on the subscriber subnets cannot be advertised as /32 host routes since they may be active (master) on multiple group IP interfaces on multiple SRRP routers. Without a /32 host route path, the network will forward any packet destined to 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.

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. Since the SRRP advertisement source IP addresses are not subscriber hosts, an anti-spoof entry will 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.

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.