# Triple Play Enhanced Subscriber Management

## In This Section

This section describes features which provide Enhanced Subscriber Management functions for Triple Play services.

Topics in this section include:

# Uniform RADIUS Server Configuration

In the past, RADIUS servers for Enhanced Subscriber Management were configured in the authentication and accounting policies. Because more applications are using RADIUS servers today (WLAN gateway radius proxy for example), a new uniform RADIUS server configuration is introduced. It is now recommended to define RADIUS server destinations in subscriber authentication and accounting policies through a **radius-server-policy**.

Functionality, that is only available via the uniform RADIUS server configuration includes that includes:

- Accounting on/off:
- The accounting on/off behavior is controlled from within the radius-server-policy. The operational state of the **radius-server-policy** can be changed based on the reachability of the RADIUS server (reception of an accounting response for the Accounting On request).
- An Accounting On message is sent at power on, after a node reboot, when the **acct-on-off** command is configured in a radius-server-policy and user triggered with a CLI command.
- An Accounting Off message is sent before an admin initiated node reboot, when the **acct-on-off** command is removed from a **radius-server-policy** and user triggered with a CLI command.
- Buffering of accounting messages: When all servers in a radius-server-policy are unreachable, it is possible to buffer the acct-stop and acct-interim-update messages for up to 25 hrs. When a RADIUS server becomes reachable again then the messages in the buffer are retransmitted.
- A configurable hold down time for accounting servers that are marked down and during which no new communication attempts will be made (**hold-down-time**).
- A configurable maximum number of outstanding RADIUS requests for accounting servers (**pending-requests-limit**). Before, an internal limit restricted the number of pending accounting request messages. This internal limit has now been removed for both RADIUS server configuration methods.
- Increased retry and timeout values for unsuccessful RADIUS communication.
- Enhanced RADIUS server statistics

# RADIUS Server Configuration Method

Following two configuration methods co-exist but are mutually exclusive:

- Uniform RADIUS Server Configuration (Preferred) on page 829
- Legacy RADIUS Server Configuration on page 831

## Uniform RADIUS Server Configuration (Preferred)

To attach a RADIUS server policy to an authentication policy:

```
configure
    subscriber-mgmt
        authentication-policy "auth-policy-1" create
            radius-server-policy "aaa-server-policy-1"
        exit
    exit
```

**Notes:**

- To avoid conflicts, the following CLI commands are ignored in the authentication policy when a **radius-server-policy** is attached:
  - All commands in the **radius-authentication-server** context
  - **accept-authorization-change**
  - **coa-script-policy**
  - **accept-script-policy**
  - **request-script-policy**
- The **fallback-action** command specifies the action when no RADIUS server is available is configured direct in the **config>subscr-mgmt>auth-plcy** CLI context.

To attach a RADIUS server policy to a RADIUS accounting policy:

For example:

```
configure
    subscriber-mgmt
        radius-accounting-policy "acct-policy-1" create
            radius-server-policy "aaa-server-policy-1"
        exit
    exit
```

**Note:** To avoid conflicts, the following CLI commands are ignored in the RADIUS accounting policy when a **radius-server-policy** is attached:

- All commands in the **radius-accounting-server** context
- **acct-request-script-policy**

To configure the RADIUS servers in a RADIUS server policy:

For example:

```
configure
    aaa
        radius-server-policy "aaa-server-policy-1" create
            description "Radius AAA server policy"
            accept-script-policy "script-policy-2"
            acct-request-script-policy "script-policy-3"
            auth-request-script-policy "script-policy-1"
            acct-on-off oper-state-change
            servers
                access-algorithm direct
                retry 3
                router "Base"
                no source-address
                hold-down-time sec 30
                timeout sec 5
                buffering
                    acct-interim min 60 max 3600 lifetime 5
                    acct-stop min 60 max 3600 lifetime 5
                exit
                server 1 name "server-1"
                server 2 name "server-2"
            exit
        exit
    exit
```

To configure the RADIUS servers in the routing instance:

- In the Base routing instance: **configure>router>radius-server**.
- In a VPRN routing instance: **configure>service>vprn 10>radius-server**.
- In the management routing instance (out of band): **configure>router management>radius-server**.

For example:

```
configure
    router
        radius-server
            server "server-1" address 172.16.1.1 secret <shared secret> hash2 create
                accept-coa
                coa-script-policy "script-policy-4"
```

```
                description "Radius server 1"
                pending-requests-limit 4096
                acct-port 1813
                auth-port 1812
            exit
            server "server-2" address 172.16.1.2 secret <shared secret> hash2 create
                accept-coa
                coa-script-policy "script-policy-4"
                description "Radius server 2"
                pending-requests-limit 4096
                acct-port 1813
                auth-port 1812
            exit
        exit
    exit
```

**Note:** To configure RADIUS CoA servers for use in Enhanced Subscriber Management, the server must be configured in the corresponding routing instance with the **accept-coa** command enabled.

## Legacy RADIUS Server Configuration

**Note:** It is recommended to migrate to the uniform RADIUS server configuration as described above to have additional functionality enabled.

To configure a RADIUS server in an authentication policy:

```
configure
    subscriber-mgmt
        authentication-policy "auth-policy-1" create
            radius-authentication-server
                access-algorithm direct
                hold-down-time 30
                retry 3
                no source-address
                timeout 5
                router "Base"
               server 1 address 172.16.1.1 secret <shared secret> hash2 port 1812 pending-
requests-limit 4096
                server 2 address 172.16.1.2 secret <shared secret> hash2 port 1812 pending-
requests-limit 4096
            exit
            accept-authorization-change
            accept-script-policy "script-policy-2"
            coa-script-policy "script-policy-4"
            request-script-policy "script-policy-1"
        exit
    exit
```

**Note:** In legacy RADIUS server configuration, to configure RADIUS CoA servers for use in Enhanced Subscriber Management, the server must be configured in the authentication policy

with the **accept-authorization-change** command enabled. A CoA only server can be configured with the optional coa-only flag.

To configure a RADIUS server in a RADIUS accounting policy:

```
configure
    subscriber-mgmt
        radius-accounting-policy "acct-policy-1" create
            radius-accounting-server
                access-algorithm direct
                retry 3
                timeout 5
                no source-address
                router "Base"
                server 1 address 172.16.1.1 secret <shared secret> hash2 port 1813
                server 2 address 172.16.1.2 secret <shared secret> hash2 port 1813
            exit
            acct-request-script-policy "script-policy-3"
        exit
    exit
```

# RADIUS Authentication of Subscriber Sessions

This section describes the Alcatel-Lucent router acting as a Broadband Subscriber Aggregator (BSA).

Note that in the 7750 and 7710 TPSDA solutions, the Alcatel-Lucent 5750 Subscriber Services Controller (SSC) serves as the policy manager, DHCP and RADIUS server.

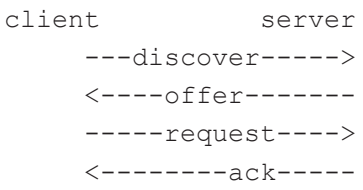
In this application, one of the required functions can be to authenticate users trying to gain access to the network.While sometimes the DHCP server (an SSC) can perform authentication, in most cases a RADIUS server (an SSC) is used to check the customer's credentials.

Note: Refer to section DHCP Principles on page 336 for an explanation of DHCP and DHCP Snooping on page 344 for an explanation of DHCP snooping.

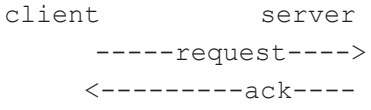
For information about the RADIUS server selection algorithm, refer to the Security chapter in the OS System Management Guide.

If authentication is enabled, the router will temporarily hold any received DHCP discover message and will send a access-request message to a configured RADIUS server containing the client's MAC address and/or Circuit-ID (from the Option 82 field) as the user name. If and when access is granted by the RADIUS server, the router will then forward or relay the DHCP discover message to the DHCP server and thus allow an IP address to be assigned. If the RADIUS authentication request is denied, the DHCP message is dropped and an event is generated.

A typical initial DHCP scenario (after client bootup) is:




But, when the client already knows its IP address (when an existing lease is being renewed), it can skip straight to the request/ack phase:




In the first scenario, the DHCP discover triggers an authentication message to RADIUS and the DHCP request also triggers RADIUS authentication. The previous reply is cached for 10 seconds, the second DHCP packet will not result in a RADIUS request.
In the second scenario, the DHCP request triggers an authentication message to RADIUS.

If the optional subscriber management authentication policy **re-authentication** command is enabled, DHCP authentication is performed at every DHCP lease renew request. Only dynamic DHCP sessions are subject to remote authentication. Statically provisioned hosts are not authenticated.

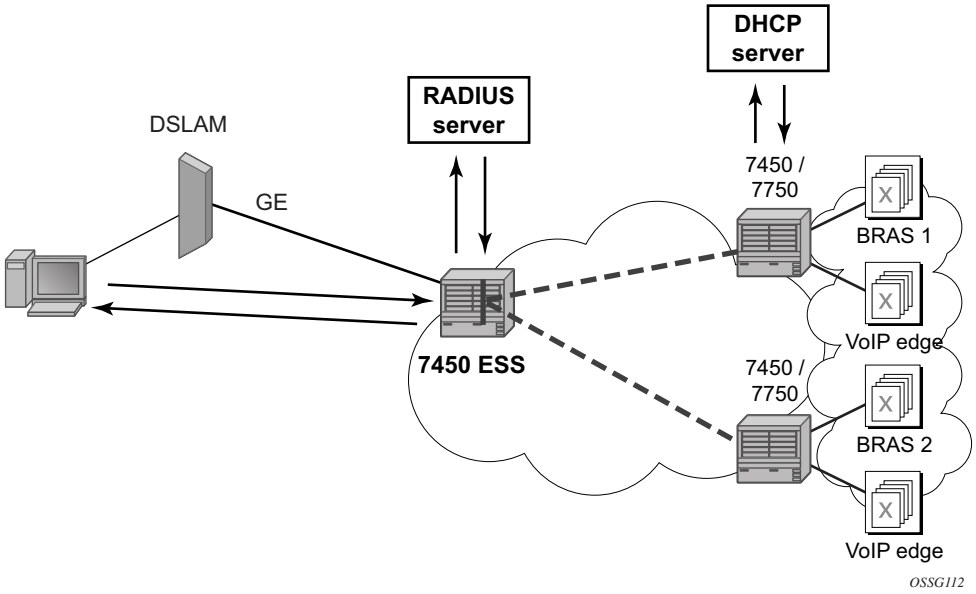
# RADIUS Authentication Extensions

This section describes an extension to RADIUS functionality in the subscriber management context. As part of subscriber host authentication, RADIUS can respond with access-response message, which, in the case of an accept, can include several RADIUS attributes (standard and vendor-specific) that allow proper provisioning of a given subscriber-host.

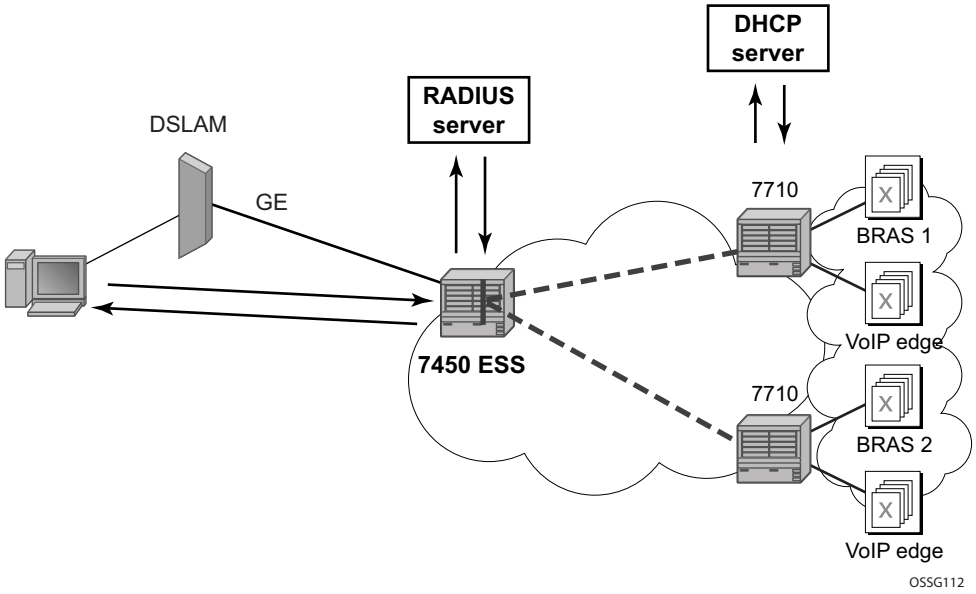
Change-of-Authorization (CoA) messages as defined by RFC 3576, *Dynamic Authorization Extensions to Remote Authentication Dial In User Service (RADIUS)*, are supported. The goal of CoA messages is to provide a mechanism for "mid-session change" support through RADIUS.

## Triple Play Network with RADIUS Authentication

7450 ESS:




7710 SR:



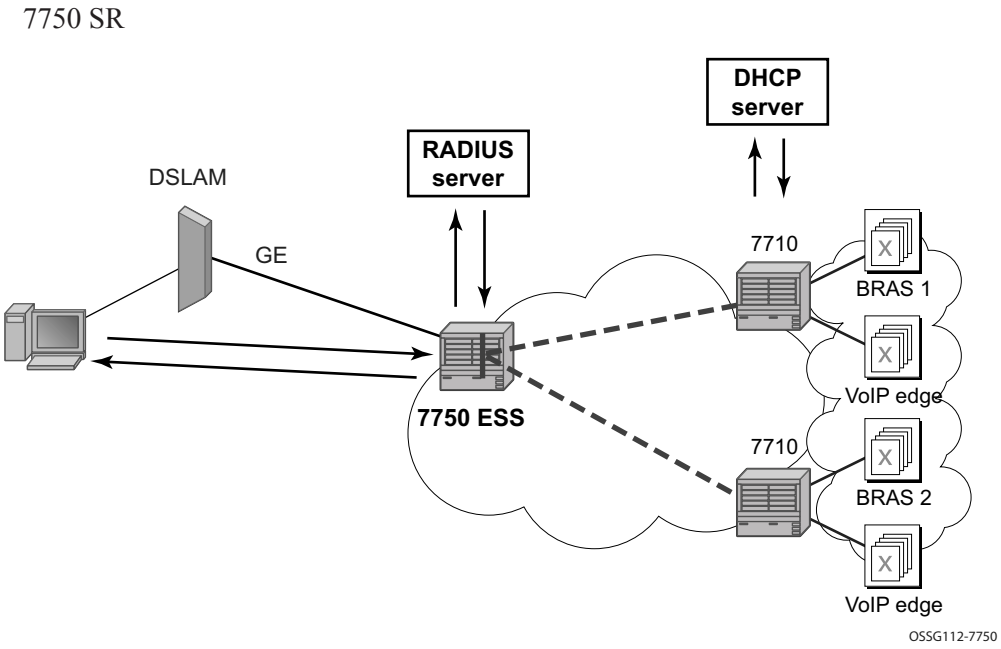



**Figure 56: Triple Play Aggregation Network with RADIUS-Based DHCP Host Authentication**

Figure 56 shows a flow of RADIUS authentication of DHCP hosts in the triple play aggregation environment. Besides granting the authentication of given DHCP host, the RADIUS server can include RADIUS attributes (standard and/or Vendor-Specific Attributes (VSAs)) which are then used by the network element to provision objects related to a given DHCP host.

RADIUS is a distributed client/server concept that is used to protect networks against unauthorized access. In the context of the router's subscriber management in TPSDA, the RADIUS client running on nodes sends authentication requests to the SSC.

RADIUS can be used to perform three distinct services:

- Authentication determines whether or not a given subscriber-host is allowed to access a specific service.
- Authorization associates connection attributes or characteristics with a specific subscriber host.
- Accounting tracks service use by individual subscribers.

The RADIUS protocol uses "attributes" to describe specific authentication, authorization and accounting elements in a user profile (which are stored on the RADIUS server). RADIUS messages contain RADIUS attributes to communicate information between network elements running a RADIUS client and a RADIUS server.

RADIUS divides attributes into two groups, standard attributes and Vendor-Specific Attributes (VSAs). VSA is a concept allowing conveying vendor specific configuration information in a RADIUS messages, discussed in RFC 2865, *Remote Authentication Dial In User Service*

*(RADIUS)*. It is up to the vendor to specify the exact format of the VSAs. Alcatel-Lucent-specific VSAs are identified by vendor-id 6527.

# RADIUS Authorization Extensions

The following sections define different functional extensions and list relevant RADIUS attributes.

## Basic Provisioning of Authentication Extensions

In order to comply with RFC 4679, *DSL Forum Vendor-Specific RADIUS Attributes*, the software includes the following attributes in the authentication-request message:

- agent-circuit-id (as defined by DSL forum)
- agent-remote-id (as defined by DSL forum)

The following attributes can also be included if configured and provided by downstream equipment:

- actual-data-rate-upstream
- actual-data-rate-downstream
- minimum-data-rate-upstream
- minimum-data-rate-downstream
- access-loop-encapsulation

When the node is configured to insert (or replace) Option 82, the above mentioned attributes will have the content after this operation has been performed by the software.

In addition, the following standard RADIUS attributes will be included in authentication request messages (subject to configuration):

- NAS-identifier — string containing system-name
- NAS-port-id
- NAS-port-type — Values: 32 (null encap), 33 (dot1q), 34 (qinq), 15 (DHCP hosts), specified value (0 — 255)
- MAC-address (Alcatel-Lucent VSA 27)
- dhcp-vendor-class-id (Alcatel-Lucent VSA 36)
- calling-station-id

These will only be included in the access-request if they have been configured.

In order to provide the possibility to push new policies for currently active subscribers, the routers support commands to force re-authentication of the given subscriber-host. After issuing such a command, the router will send a DHCP force-renew packet, which causes the subscriber to renew its lease (provided it supports force-renew). The DHCP request and ACK are then authenticated and processed by the routers as they would be during a normal DHCP renew.

## Calling-Station-ID

A **calling-station-id** can be configured at SAP level and can be included in the RADIUS authentication and accounting messages. This attribute is used in legacy BRAS to identify the user (typically phone number used for RAS connection). In the broadband networks this was replaced by circuit-id in Option 82. However, the Option 82 format is highly dependent on access-node vendor, which makes interpretation in management servers (such as RADIUS) troublesome. Some operators use the calling-station-id attribute as an attribute indicating the way the circuit-id should be interpreted. The calling-station-id attribute can be configured as a string which will be configured on the SAP. It can also be configured to use the **sap-id**, **remote-id** or **mac-address**.

## Subscriber Session Timeout

To limit the lifetime of a PPP session or DHCPv4 host to a fixed time interval, a timeout can be specified from RADIUS. By default, a PPP session or DHCPv4 host has no session timeout (infinite).

For PPP sessions, a session-timeout can be configured in the ppp-policy. A RADIUS specified session-timeout overrides the CLI configured value.

```
subscriber-mgmt
    ppp-policy "ppp-policy-1" create
        session-timeout 86400
    exit
exit
```

When the session timeout expires a PPP session is terminated and a DHCPv4 host deleted. For a DHCPv4 host, a DHCP release message is also sent to the server .

The following two attributes can be used in RADIUS Access-Accept and CoA messages to limit the PPP session or DHCPv4 host session time (Table 13):

**Table 13: Subscriber Session Timeout**

| Attribute ID | Attribute Name | Type | Limits | Purpose and Format |
|---|---|---|---|---|
| 27 | Session-Timeout | integer | 2147483647 seconds | 0 = infinite (no session-timeout)<br>[1.. 2147483647] in seconds<br>For example:<br>Session-Timeout = 3600 |
| 26-6527-160 | Alc-Relative-Session-Timeout | integer | [0..2147483647 ] seconds | 0 = infinite (no session-timeout)<br>[1..2147483647] in seconds<br>For example:<br>Alc-Relative-Session-Timeout = 3600 |

When specified in a RADIUS Access-Accept message, both attributes specify an absolute value for session timeout. When specified in a RADIUS CoA message, attribute [26-6527-160] Alc-Relative-Session-Timeout specifies a relative session timeout value in addition to the current session time while attribute [27] Session-Timeout specifies an absolute session timeout value. If the current session time is greater than the received Session-Timeout, a CoA NAK is sent with error cause “Invalid Attribute Value (407)”.

Only one of the above attributes to specify a session timeout can be present in a single RADIUS message. An event is raised when both are specified in a single message.

The output of the “show service id <service-id> ppp session detail” CLI command contains following fields related to session timeout for PPP sessions:

- Up Time: the PPP session uptime
- Session Time Left: the remaining time before the session is terminated
- RADIUS Session-TO: the RADIUS received session timeout value.

The output of the “show service id <service-id> dhcp lease-state detail” CLI command contains following fields related to session timeout for DHCPv4 hosts:

- Up Time: the DHCPv4 host uptime
- Remaining Lease Time: the remaining time before the lease expires in the DHCP server. The client should renew its lease before this time.
- Remaining SessionTime: the remaining time before the DHCPv4 host is deleted
- Session-Timeout: the DHCPv4 host is deleted when its uptime reaches the Session-Timeout value.
- Lease-Time: the lease time specified by the DHCPv4 server

**Note**:

In a radius-proxy scenario or when a DHCPv4 host is created with a RADIUS CoA message, the RADIUS attribute [26-6527-174] Alc-Lease-Time attribute must be used to specify the lease time. If the [26-6527-174] Alc-Lease-Time is not present in these scenarios, then the RADIUS attribute [27] Session-Timeout is interpreted as DHCPv4 lease time.

### Domain Name in Authentication

In many networks, the user name has specific meaning with respect to the domain (ISP) where the user should be authenticated. In order to identify the user correctly, the user name in an authentication-request message should contain a domain-name. The domain-name can be derived from different places. In PPPoE authentication the domain name is given by the PPPoE client with the user name used in PAP or CHAP authentication. For DHCP hosts similar functionality is implemented by a “pre-authentication” lookup in a local user database before performing the RADIUS request.

For example, it can be derived from option60 which contains the vendor-specific string identifying the ISP the set-box has been commissioned by.

To append a domain name to a DHCP host, the following configuration steps should be taken:

- Under the (group or IP) interface of the service, a local user database should be configured in the DHCP node and no authentication policy should be configured.
- In the local user database, there should be a host entry containing both the domain name to be appended and an authentication policy that should be used for RADIUS authentication of the host. The host entry should contain no other information needed for setting up the host (IP address, ESM string), otherwise the DHCP request will be dropped.
- In the authentication policy, the **user-name-format** command should contain the parameter **append** *domain-name*.

---

## RADIUS Reply Message for PPPoE PAP/CHAP

The string returned in a [18] Reply-Message attribute in a RADIUS Access-Accept is passed to the PPPoE client in the CHAP Success or PAP Authentication-Ack message.

The string returned in a [18] Reply-Message attribute in a RADIUS Access-Reject is passed to the PPPoE client in the CHAP Failure or PAP Authentication-Nak message.

When no [18] Reply-Message attribute is available, the SROS default messages are used instead: “CHAP authentication success” or ”CHAP authentication failure” for CHAP and “Login ok” or ” Login incorrect” for PAP.

# Provisioning of Enhanced Subscriber Management (ESM) Objects

In the ESM concept on network elements, a subscriber host is described by the following aspects:

- subscriber-id-string
- subscriber-profile-string
- sla-profile-string
- ancp-string
- intermediate-destination-identifier-string
- application-profile-string

This information is typically extracted from DHCP-ACK message using a Python script, and is used to provision subscriber-specific resources such as queues and filter entries. As an alternative to extracting this information from DHCP-ACK packet, provisioning from RADIUS server is supported.

As a part of this feature, the following VSAs have been defined:

- alc-subscriber-id-string — Contains a string which is interpreted as a subscriber-id.
- alc-subscriber-profile-string — Contains a string which is interpreted as a subscriber profile
- alc-sla-profile-string — Contains string which is interpreted as an SLA profile.
- alc-ancp-string — Contains string which is interpreted as an ANCP string.
- alc-int-dest-id-string — Contains a string which is interpreted as an intermediate destination ID
- alc-app-profile-string — Contains a string which is interpreted as an application profile

  Note that these strings can be changed in a CoA request.

When RADIUS authentication response messages contain the above VSAs, the information is used during processing of DHCP-ACK message as an input for the configuration of subscriber-host parameters, such as QoS and filter entries.

If ESM is not enabled on a given SAP, information in the VSAs is ignored.

If ESM is enabled and the RADIUS response does not include all ESM-related VSAs (an ANCP string is not considered as a part of ESM attributes), only the subscriber-id is mandatory (the other ESM-related VSAs are not included). The remaining ESM information (sub-profile, sla-profile) will be extracted from DHCP-ACK message according to “normal” flow (Python script, etc.). If the profiles are missing from RADIUS, they are not extracted from the DHCP data with Python to prevent inconsistent information. Instead, he data will revert to the configured default values.

However, if the above case, a missing subscriber ID will cause the DHCP request to be dropped. The DHCP server will not be queried in that case.

When no DHCP server is configured, DHCP-discover/request messages are discarded.

## Provisioning IP Configuration of the Host

The other aspect of subscriber-host authorization is providing IP configuration (ip-address, subnet-mask, default gateway and dns) through RADIUS directory rather then using centralized DHCP server. In this case, the node receiving following RADIUS attributes will assume role of DHCP server in conversation with the client and provide the IP configuration received from RADIUS server.

These attributes will be accepted only if the system is explicitly configured to perform DHCP-server functionality on a given interface.

The following RADIUS attributes will be accepted from authentication-response messages:

- framed-ip-address — The IP address to be configured for the subscriber-host.
- framed-ip-netmask — The IP network to be configured for the subscriber host. If RADIUS does not return a netmask, the DHCP request is dropped.
- framed-pool — The pool on a local DHCP server from which a DHCP-provided IP address should be selected.
- alc-default-router — The address of the default gateway to be configured on the DHCP client.
- alc-primary-dns — The DNS address to be provided in DHCP configuration.
  - → Juniper VSA for primary DNS.
  - → Redback VSA for primary DNS.
- alc-secondary-dns
  - → Juniper VSA for secondary DNS.
  - → Redback VSA for secondary DNS.
- alc-lease-time — Defines the lease time.
- session-timeout — Defines the lease time in absence of the alc-lease-time attribute.
- NetBIOS
  - → alc-primary-nbns
  - → alc-secondary-nbns

---

## RADIUS Based Authentication in Wholesale Environment

In order to support VRF selection, the following attributes are supported:

- alc-retail-serv-id — Indicates the service-id of the required retail VPRN service configured on the system.

## Change of Authorization and Disconnect-Request

In a typical RADIUS environment, the network element serves as a RADIUS client, which means the messages are originated by a routers. In some cases, such as “mid-session” changes, it is desirable that the RADIUS server initiates a CoA request to impose a change in policies applicable to the subscriber, as defined by RFC 3576.

To configure a RADIUS server to accept CoA and Disconnect Messages is achieved in one of the following ways:

1. Configure up to 64 RADIUS CoA servers per routing instance:

```
config>router>radius-server#
config>service>vprn>radius-server#

 server "coa-1" address 10.1.1.1 secret <shared-secret> hash2 create
      accept-coa
 exit
```

   This is the preferred method.

2. Configure up to 16 RADIUS CoA servers per authentication policy.

```
config>subscr-mgmt>auth-plcy#

 accept-authorization-change
```

The UDP port for CoA and Disconnect Messages is configurable per system with the command:

```
config>aaa#

radius-coa-port {1647|1700|1812|3799}
```

Note that there is a priority in the functions that can be performed by CoA. The first matching one will be performed:

- If the CoA packet contains a force-renew attribute, the subscriber gets a force-renew DHCP packet. This function is not supported for PPPoE or ARP hosts.
- If the CoA packet contains a create-host attribute, a new lease-state is created. Only DHCP lease-states can be created by a CoA message. PPPoE sessions and ARP hosts cannot be created.
- Otherwise, the ESM strings are updated.

There are several reasons for using RADIUS initiated CoA messages:

1. Changing ESM attributes (SLA or subscriber profiles) or queues/policers/schedulers rates of the given subscriber host — CoA messages containing the identification of the given subscriber-host along with new ESM attributes.

2. Changing (or triggering the change) of IP configuration of the given subscriber-host — CoA messages containing the identification of the given subscriber-host along with VSA indicating request of forcerenew generation.
3. Configuring new subscriber-host — CoA messages containing the full configuration for the given host.

If the changes to ESM attributes are required, the RADIUS sever will send CoA messages to the network element requesting the change in attributes included in the CoA request:

- attribute(s) to identify a single or multiple subscriber host(s): "NAS-Port-Id + IP address/prefix" or "Acct-Session-Id" or "Alc-Subsc-ID-Str"
  - → Nas-Port-Id attribute + single IP address/prefix attribute:
    - – Framed-IP-Address
    - – Alc-Ipv6-Address
    - – Framed-Ipv6-Prefix
    - – Delegated-Ipv6-Prefix
  - → Acct-Session-Id (number format)
  - → Alc-Subsc-ID-Str
- alc-subscriber-profile-string
- alc-sla-profile-string
- alc-ancp-string
- alc-app-profile-string
- alc-int-dest-id-string
- alc-subscriber-id-string
- alc-subscriber-qos-override

  Note that if the subscriber-id-string is changed while the ANCP string is explicitly set, the ANCP-string **must** be changed simultaneously. When changing the alc-subscriber-id-string, the lease state is temporarily duplicated, causing two identical ANCP-strings to be in the system at the same time. This is not allowed.

As a reaction to such message, the router changes the ESM settings applicable to the given host.

If changes to the IP configuration (including the VRF-id in the case of wholesaling) of the given host are needed, the RADIUS server may send a CoA message containing VSA indicating request for forcerenew generation:

- attribute(s) to identify a single or multiple subscriber host(s): "NAS-Port-Id + IP address/prefix" or "Acct-Session-Id" or "Alc-Subsc-ID-Str":
  - → Nas-Port-Id attribute + single IP address/prefix attribute:
    - – Framed-IP-Address
    - – Alc-Ipv6-Address

- Framed-Ipv6-Prefix
- Delegated-Ipv6-Prefix

→ Acct-Session-Id (number format)

→ Alc-Subsc-ID-Str

- alc-force-renew
- alc-force-nak

As a reaction to such message, router will generate a DHCP forcerenew message for the given subscriber host. Consequently, during the re-authentication, new configuration parameters can be populated based on attributes included in Authentication-response message. The force-NAK attribute has the same function as the force-renew attribute, but will cause the ESR to reply with a NAK to the next DHCP renew. This will invalidate the lease state on the ESR and force the client to completely recreate its lease, making it possible to update parameters that cannot be updated through normal CoA messages, such as IP address or address pool.

If the configuration of the new subscriber-host is required, RADIUS server will send a CoA message containing VSA request new host generation along with VSAs specifying all required parameters.

- alc-create-host
- alc-subscriber-id-string — This attribute is mandatory in case ESM is enabled, and optional for new subscriber host creation otherwise.
- NAS-port-id — This attribute indicates the SAP where the host should be created.
- framed-ip-address —
- alc-client-hw-address — A string in the xx:xx:xx:xx:xx:xx format. This attribute is mandatory for new subscriber-host creation.
- alc-lease-time — Specifies the lease time. If both session-timeout and alc-lease-time are not present, then a default lease time of 7 days is used.
- session-timeout — Specifies the lease time in absence of the alc-lease-time attribute. If both session-timeout and alc-lease-time are not present, then a default lease time of 7 days is used.
- alc-retail-svc-id — This is only used in case of wholesaling for selection of the retail service
- Optionally other VSAs describing given subscriber host. Obviously, if the ESM is enabled, but the CoA message does not contain ESM attributes the new host will not be created.

After executing the requested action, the router element responds with an ACK or NAK message depending on the success/failure of the operation. In case of failure (and hence NAK response), the element will include the error code in accordance with RFC 3576 definitions if an appropriate error code is available.

Supporting CoA messages has security risks as it essentially requires action to unsolicited messages from the RADIUS server. This can be primarily the case in an environment where RADIUS servers from multiple ISPs share the same aggregation network. To minimize the security risks, the following rules apply:

- Support of CoA messages is disabled by default. They can be enabled on a per RADIUS server or authentication-policy basis.
- When CoA is enabled, the node will listen and react only to CoA messages received from RADIUS servers. In addition, CoA messages must be protected with the key corresponding to the given RADIUS server. All other CoA messages will be silently discarded.

In all cases (creation, modification, forcerenew) subscriber host identification attributes are mandatory in the CoA request: “NAS-Port-Id + IP” or “Acct-Session-Id” or “Alc-Subsc-ID-Str”

- Nas-Port-Id + single IP address/prefix:
  - → Nas-Port-Id
  - → Framed-IP-Address
  - → Alc-Ipv6-Address
  - → Framed-Ipv6-Prefix
  - → Delegated-Ipv6-Prefix
- Acct-Session-Id (number format)
- Alc-Subsc-ID-Str

When there are no subscriber host identification attributes present in the CoA, the message will be NAK’d with corresponding error code.

- hosts, meaning only subscriber host to which the given authentication-policy is applicable.
- Receiving CoA message with the same attributes as currently applicable to the given host will be responded to with an ACK message.
- In case of dual homing (through SRRP), the RADIUS server should send CoA messages to both redundant nodes and this with all corresponding attributes (NAS-port-id with its local meaning to corresponding node).
- In the case of change requests, the node which has the given host active (active-sap for master-sap for SRRP) will process the RADIUS message and reply to RADIUS. The standby node will always reply with a NAK.
- In the case of create requests, the active node (the SAP described by NAS-port-id is “active” or “master”). Both nodes will reply, but the standby will NAK the request.

The properties of an existing RADIUS-authenticated PPPoE session can be changed by sending a Change of Authorization (CoA) message from the RADIUS server. Processing of a CoA is done in the same way as for DHCP hosts, with the exception that only the ESM settings can be changed for a PPPoE session (the force-renew attribute is not supported for PPPoE sessions and a Create-Host CoA will always generate a DHCP host.)

For terminating PPPoE sessions from the RADIUS server, the disconnect-request message can be sent from the RADIUS server. This message triggers a shutdown of the PPPoE session. The attributes needed to identify the PPPoE session are the same as for DHCP hosts.

## RADIUS-Based Accounting

When a router is configured to perform RADIUS-based accounting, at the creation of a subscriber-host, it will generate an accounting-start packet describing the subscriber-host and send it to the RADIUS accounting server. At the termination of the session, it will generate an accounting-stop packet including accounting statistics for a given host. The router can also be configured to send an interim-accounting message to provide updates for a subscriber-host.

The exact format of accounting messages, their types, and communication between client running on the routers and RADIUS accounting server is described in RFC 2866, *RADIUS Accounting*. The following describes a few specific configurations.

In order to identify a subscriber-host in accounting messages different RADIUS attributes can be included in the accounting-start, interim-accounting, and accounting-stop messages. The inclusion of the individual attributes is controlled by configuration commands. Following RADIUS attributes will be supported in accounting messages:

- framed-ip-address
- framed-ip-netmask
- agent-circuit-id (as defined by DSL forum)
- agent-remote-id (as defined by DSL forum)
- calling-station-id
- alc-subscriber-id-string
- alc-subscriber-profile-string
- alc-sla-profile-string
- user-name
- NAS-identifier
- NAS-port-id
- NAS-port-type — Values: 32 (null encap), 33 (dot1q), 34 (qinq), 15 (DHCP hosts), specified value (0 — 255)
- MAC-address

In the subscriber management concept, a single subscriber-host can have multiple associated queues. Every queue represents a separate set of accounting statistics and therefore, the single accounting message will contain statistics for all queues associated with the given subscriber host.

RADIUS accounting defines accounting statistics in terms of the following attributes which contain the queue-id (the first 2 bytes of a 10B word) and the counter (the remaining 8B):

- alc-acct-input-inprof-octets-64 — ingress-in-profile-forwarded-bytes.
- alc-acct-input-outprof-octets-64 — ingress-out-of-profile-forwarded-bytes.
- alc-acct-input-inprof-packets-64 — ingress-in-profile-forwarded-packets.
- alc-acct-input-outprof-packets-64 — ingress-out-of-profile-forwarded-packets.

- alc-acct-output-inprof-octets-64 — egress-in-profile-forwarded-bytes.
- alc-acct-output-outprof-octets-64 — egress-out-of-profile-forwarded-bytes.
- alc-acct-output-inprof-packets-64 — egress-in-profile-forwarded-packets.
- alc-acct-output-outprof-packets-64 — egress-out-of-profile-forwarded-packets.

In addition to accounting-start, interim-accounting, and accounting-stop messages, a RADIUS client on a routers will send also accounting-on and accounting-off messages. An accounting-on message will be sent when a given RADIUS accounting-policy is applied to a given subscriber-profile, or the first server is defined in the context of an already applied policy. The following attributes will be included in such message:

- NAS-identifier
- alc-subscriber-profile-string
- Accounting-session-id
- Event-timestamp

Accounting-off messages will be sent at following events:

- An accounting policy has been removed from a sub-profile.
- The last RADIUS accounting server has been removed from an already applied accounting policy.

These messages contain following attributes:

- NAS-identifier
- alc-subscriber-profile-string
- Accounting-session-id
- Accounting-terminate-cause
- Event-timestamp

In case of dual homing, both nodes will send RADIUS accounting messages for the host, with all attributes as it is locally configured. The RADIUS log files on both boxes need to be parsed to get aggregate accounting data for the given subscriber host regardless the node used for forwarding.

For RADIUS-based accounting, a custom record can be defined to refine the data that is sent to the RADIUS server. Refer to the Configuring an Accounting Custom Record in the OS System Management Guide for further information.

## Accounting Modes Of Operation

This section is applicable to the 7750 SR or the 7450 ESS in mixed mode
There are three basic accounting models in 7750 SR or the 7450 ESS in mixed mode:

- Per queue-instance
- Per Host
- Per Session

Each of the basic models can optionally be enabled to send interim-updates. Inclusion/exclusion of interim-updates will depend on whether volume based (start/interim-updates/stop) or time based (start/stop) accounting is required.

The difference between the three basic accounting models is in its core related to the processing of the acc-session-id for each model. The differences are related to:

- acct-session-id generation within each model.
- outcome in response to the CoA action relative to the targeted acct-session-id.

The counters for volume-based accounting are collected from queues or policers that are instantiated per sla-profile instance (SPI) on non-HSMDA based hardware or per subscriber on HSMDA based hardware. This is true irrespective of which model of accounting (or combination of models) is deployed. Within accounting context, the SPI on non-HSMDA or subscriber on HSMDA equates to queue-instance.

Table 14 summarizes the key differences between various accounting modes of operation that are supported. Interim-updates for each individual mode can be enabled/disabled via configuration (interim-updates keyword as an extension to the commands that enable three basic modes of accounting). This is denoted by the IU-Config keyword under the 'I-U' column in the table. The table also shows that any two combinations of the three basic models (including their variants for volume/time based accounting) can be enabled simultaneously.

**Table 14: Accounting Modes of Operation**

| Accounting Mode | Accounting Entity | START | I-U | STOP | Acct-session-id | Acct-multi-session-id |
|---|---|---|---|---|---|---|
| queue-instance-accounting | queue-instance | X | IU-config | X | X | |
| | session | | | | | |
| | host | | | | | |

**Table 14: Accounting Modes of Operation (Continued)**

| Accounting Mode | Accounting Entity | START | I-U | STOP | Acct-session-id | Acct-multi-session-id |
|---|---|---|---|---|---|---|
| session-accounting | queue-instance | | | | | |
| | session | X | IU-config | X | X | q-instance |
| | host | | | | | |
| host-accounting | queue-instance | | | | | |
| | session | | | | | |
| | host | X | IU-config | X | X | queue-instance |
| queue-instance-accounting + host-accounting | queue-instance | X | IU-config | X | X | queue-instance |
| | session | | | | | |
| | host | X | IU-config | X | X | queue-instance |
| queue-instance-accounting + session-accounting | queue-instance | X | IU-config | X | X | queue-instance |
| | session | X | IU-config | X | X | queue-instance |
| | host | | | | | |
| session- accounting + host-accounting | queue-instance | | | | | queue-instance |
| | session | X | IU-config | X | X | |
| | host | X | | X | X | SESSION |

Note that hosts within the targeted CoA entity will be affected as follows:

- If the CoA target is the session, then both constituting members (IPv4 and IPv6) of the dual-stack host will be affected.
- If the CoA target is the queuing-instance, then up to 32 hosts that are sharing that SPI will be affected.

The same principle applies to LI.

The accounting behavior (accounting messages and accounting attributes) in case that the SPI is changed via CoA depends on the accounting mode of operation. On non-HSMDA hardware, the behavior is the following:

- SPI change in conjunction with per queuing instance accounting will trigger a STOP for the old SPI and a START for the new SPI with corresponding counters. Acct-session-id/Acct-Multi-Session-Id will be unique per SPI. Note that Acct-Multi-Session-Id is only generated if per queuing-instance accounting mode of operation is combined with some other mode of operation (host or session).
- SPI change in conjunction with per host or per session accounting (no interim updates for either method) will NOT trigger any new accounting messages. In other words, SPI change will go unnoticed from the perspective of the accounting server until the host/session is terminated. When the host/session is terminated a STOP will be sent with the VSA carrying the latest SPI name and the acct-multi-session-id attribute of the latest SPI. Acct-session-id will stay the same during the lifetime of the host. Counters are not included in STOP (interim-update not enabled).

  SPI change in conjunction with per host accounting with interim-updates or per session accounting with interim-updates will trigger two interim-update messages:

  → One with the old counters (terminated queues) and the old SPI name VSA. This behavior is similar to the triggered STOP message in per queuing-instance accounting upon SPI change.

  → One with the new counters (new queues instantiated), the VSA carrying the new SPI name and the new acct-multi-session-id referencing the new SPI. This behavior is similar to the triggered START message in per queuing-instance when SPI is changed in per queuing-instance accounting.

On HSMDA, no START/STOPS are sent since queues are not re-instantiated on ingress or egress.

## Per Session Accounting

In the per session accounting mode of operation the accounting message stream[1] (START/INTERIM-UPDATE/STOP) is generated per session.

- A session is defined for PPPoE hosts for which a state is maintained. The state of the host (single stack or dual-stack) is normally refreshed via PPP keepalives. Each PPPoE host of the same address family (v4 or v6) corresponds to a unique session which is identified by the <session-ID, mac> combination.

In dual-stack PPPoE case, IPv4 and IPv6 hosts are tied to the same (LCP) session. A single authentication request is initiated for such session (triggered by the first host that initiates the session).

For a single stack PPPoE host, the behavior defined in the per session accounting model is indistinguishable from the per host accounting model. The per session accounting model makes difference in behavior only for dual stack PPPoE hosts.

The following are the properties of the Per Session Accounting model:

- A single accounting session ID (acct-session-id) is generated per (PPPoE) session and it can optionally be sent in RADIUS Access-Request message.
- This acct-session-id is synchronized via MCS in dual homing environment.
- The accounting messages (START, INTERIM-UPDATE, STOP) carry the acct-multi-session-id attribute denoting the sla-profile instance with which the session is associated.
- The counters are collected from the queues instantiated through the sla-profile instance. If multiple sessions are sharing the same sla-profile instance, the counters are aggregated. In other words, counters per individual session cannot be extracted from the aggregated count.
- RADIUS triggered changes and LI are applicable per session:
  - → Queue/policer RADIUS overrides — Parameters for the referenced queue/policer within the session are changed accordingly.
  - → Subscriber aggregate rate limits, scheduler rates and arbiter rates are changed accordingly.
  - → CoA DISCONNECT brings down the entire session.
  - → LI — Activation based on the session acct-session-id affects the hosts within the session (dual-stack).
  - → SLA profile instance change affects all hosts (or sessions) sharing the same sla-profile instance (SPI). If the SPI is changed on a non-HSMDA based MDA, then queues are re-instantiated and counters are reset.
- All applicable IP addresses (v4 and v6 – including all v6 attributes – alc-ipv6-address, framed-ipv6-prefix, delegated-ipv6-prefix) are present in accounting messages for the session.

---

1. The accounting message stream refers to a collection of accounting messages (START/INTERIM-UPDATE/STOP) sharing the same acct-session-id.

## Caveats

Per session accounting is supported for entities that have concept of a session. Currently only PPPoE hosts (single or dual-stack) fall into this category.

## RADIUS Per Host Accounting

In 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 SR-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.

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

Under accounting-policy, a host-accounting command allows configurable behavior.

---

## No Host-Accounting

In prior releases and when no host-accounting is configured, the 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-accting messages.

| RADIUS Accounting AVP | Include-radius-attrs Acct/ Auth | Host Accounting | SLA-Profile Accounting |
|---|---|---|---|
| User-Name | Yes/No | Yes | No |
| NAS-Identifier | Yes/No | Yes | Yes |
| NAS-Ip-Address | No/No | Yes | Yes |
| NAS-Port-Id | Yes/Yes | Yes | No |
| Nas-Port | Yes/No | Yes | No |
| NAS-Port-Type | Yes/Yes | Yes | No |
| Service-Type | No/No | Yes | No |
| Framed-Protocol | No/No | Yes | No |
| Framed-Ip-Address | Yes/No | Yes | No |
| Framed-Ip-Netmask | Yes/No | Yes | No |
| Framed-Route | No/No | Yes | No |
| Class | No/No | Yes | No |
| Session-Timeout | No/No | Yes | Yes |
| Circuit-Id VSA | Yes/Yes | Yes | No |
| Called-Station-Id | Yes/Yes | Yes | No |
| Calling-Station-Id | Yes/Yes | Yes | No |
| MAC-Addr VSA | Yes/Yes | Yes | No |
| Remote-Id VSA | Yes/Yes | Yes | No |
| Acct-Input-Octets | No/No | No | Yes |
| Acct-Output-Octets | No/No | No | Yes |
| Acct-Input-Gigawords | No/No | No | Yes |
| Acct-Output-Gigawords | No/No | No | Yes |
| Acct-Session-Id | No/No | Yes | Yes |
| Acct-Session-Time | No/No | Yes | Yes |

| RADIUS Accounting AVP | Include-radius-attrs Acct/ Auth | Host Accounting | SLA-Profile Accounting |
|---|---|---|---|
| Acct-Input-Packets | No/No | No | Yes |
| Acct-Output-Packets | No/No | No | Yes |
| Acct-Multi-Session-Id | No/No | Yes | Yes |
| Actual-Data-Rate-Upstream | No/No | Yes | No |
| Actual-Data-Rate-Downstream | No/No | Yes | No |
| Access-Loop-Encapsulation | No/Yes | Yes | No |
| Alc-Accounting | No/No | No | Yes |
| Alc-Subscriber-Id | Yes/No | Yes | Yes |
| Alc-Subscriber-Profile-String | Yes/No | Yes | Yes |
| Alc-Sla-Profile-String | Yes/No | Yes | Yes |

## Accounting Interim Update Message Interval

The interval between two RADIUS Accounting Interim Update messages can be configured in the RADIUS accounting policy with the **update-interval** command, for example:

```
configure
    subscriber-mgmt
        radius-accounting-policy "acct-policy-1" create
            update-interval 60
            update-interval-jitter absolute 600
```

A RADIUS specified interim interval (attribute [85] Acct-Interim-Interval) overrides the CLI configured value.

By default, a random delay of 10% of the configured **update-interval** is added to the update-interval between two Accounting Interim Update messages. This jitter value can be configured with the **update-interval-jitter** to an absolute value in seconds between zero and 3600. The effective maximum random delay value is the minimum value of the configured absolute jitter value and 10% of the configured **update-interval**.

A value of zero will send the Accounting Interim Update message without introducing an additional random delay.

---

## Class Attribute

The RADIUS class-attribute helps to aid in user identification

User identification is used to correlate RADIUS accounting messages with the given user. During authentication process, the RADIUS authentication server inserts a class-attribute into the RADIUS authenticate response message and then the router echoes this class attribute in all RADIUS accounting messages.

---

## User Name

The user-name, which is used for user authentication (user-name attribute in RADIUS authentication request), can be included in RADIUS accounting messages. Per RFC 2865, when a RADIUS server returns a (different) user-name attribute, the changed user name will be used in accounting and not the originally sent user name.

## Accounting-On and Accounting Off

For RADIUS servers configured in a RADIUS server policy, the accounting on/off behavior is controlled via the **acct-on-off** command in the radius-server-policy.

By default, no Accounting-On or Accounting-Off messages are sent (**no acct-on-off**).

With the **acct-on-off** command configured in the radius-server-policy:

- An Accounting-On is sent for the following:
  - When the system is powered on.
  - After a system reboots.
  - When the **acct-on-off** command is added to the **radius-server-policy** configuration.
  - User triggered via CLI: tools perform aaa acct-on
- An Accounting-Off is sent for the following:
  - Before a user initiated system reboot.
  - When the **acct-on-off** command is removed from the **radius-server-policy** configuration.
  - User triggered via CLI: tools perform aaa acct-off.

The Accounting-On or Accounting-Off message is sent to the servers configured in the radius-server-policy, following the configured access-algorithm until an Accounting Response is received. If the first server responds, no message is sent to the other servers.

The Accounting-On message is repeated until an Accounting Response message is received from a RADIUS server: If after the configured retry/timeout timers for each RADIUS server in the radius-server-policy no response is received then the process starts again after a fixed one minute wait interval.

The Accounting-Off message is attempted once: If after the configured retry/timeout timers for each RADIUS server in the radius-server-policy no response is received then no new attempt is made.

It is possible to block a radius-server-policy until an Accounting Response is received from one of the RADIUS servers in the radius-server-policy that acknowledges the reception of an Accounting-On. The radius-server-policy cannot be used by applications for sending RADIUS messages until the state becomes “Not Blocked”. This is achieved with the optional “oper-state-change” flag, for example:

```
configure
    aaa
        radius-server-policy "aaa-server-policy-1" create
            acct-on-off oper-state-change
            servers
                router "Base"
                server 1 name "server-1"
            exit
        exit
    exit
```

If multiple radius-server-policies are in use for different applications (for example, authentication and accounting) and an Accounting-On must be send for only one radius-server-policy, it is possible to tie the acct-on-off states of both policies together using an acct-on-off-group. With this configuration, it is possible to block the authentication servers until the accounting servers are available. An acct-on-off-group can be referenced by:

- a single radius-server-policy as controller: the acct-on-off oper-state of the acct-on-off-group is set to the acct-on-off oper-state of the radius-server-policy (acts as master)
- multiple radius-server-policies as monitor: the acct-on-off oper-state of the radius-server-policy is inherited from the acct-on-off oper-state of the acct-on-off group. (acts as a slave)

```
configure
    aaa
        acct-on-off-group "group-1" create
            description "Grouping of radius-server-policies acct-on-off"
        exit
        radius-server-policy "aaa-server-policy-1" create
            acct-on-off oper-state-change group "group-1"
            servers
                router "Base"
                server 1 name "server-1"
            exit
        exit
        radius-server-policy "aaa-server-policy-2" create
            acct-on-off monitor-group "group-1"
            servers
                router "Base"
                server 1 name "server-2"
            exit
        exit
```

It is possible to force an Accounting-On or Accounting-Off message for a radius-server-policy with acct-on-off enabled using following CLI commands:

**tools perform aaa acct-on** [**radius-server-policy** *policy-name*] [**force**]

**tools perform aaa acct-off** [**radius-server-policy** *policy-name*] [**force**] [**acct-terminate-cause** *number*]

If an Accounting-On was sent to the radius-server-policy and it was acknowledged with an Accounting Response then a new Accounting-On can only be sent with the “force” flag.

If an Accounting-Off was sent to the radius-server-policy and it was acknowledged with an Accounting Response then a new Accounting-Off can only be sent with the “force” flag. The Acct-Terminate-Cause value in the Accounting-Off can be overwritten.

Use the following CLI command to display the Accounting On/Off information for a radius-server-policy:

```
# show aaa radius-server-policy "aaa-server-policy-3" acct-on-off
===============================================================================
RADIUS server policy "aaa-server-policy-3" AcctOnOff info
===============================================================================
Oper state                  : on
Session Id                  : 242FFF0000008F512A3985
Last state change           : 02/24/2013 16:06:41
Trigger                     : startUp
Server                      : "server-1"
===============================================================================
```

The operational state provides following state information: The sending of the Accounting-On or Accounting-Off message is ongoing (sendAcctOn, SendAcctOff), is successfully responded (on, off) or no response received (OffNoResp).

The Session-Id is a unique identifier for each RADIUS server policy accounting Accounting-On/ Accounting-Off sequence.

The Trigger field shows what triggered the Accounting On or Accounting Off message. If the radius-server-policy is part of an acct-on-off group then the group name is shown in brackets.

The Server field shows which server in the RADIUS server policy responded to the Accounting-On or Accounting-Off message.

To display the  acct-on-off state of a radius-server-policy, use the command, for example:

```
# show aaa radius-server-policy "aaa-server-policy-3"
===============================================================================
RADIUS server policy "aaa-server-policy-3"
===============================================================================
Description                 : (Not Specified)
Acct Request script policy  : script-policy-1
Auth Request script policy  : script-policy-1
Accept script policy        : script-policy-1
Acct-On-Off                 : Enabled (state Blocked)
-------------------------------------------------------------------------------
RADIUS server settings
-------------------------------------------------------------------------------
Router                      : "Base"
Source address              : (Not Specified)
Access algorithm            : direct
Retry                       : 3
Timeout (s)                 : 5
Hold down time (s)          : 30
Last management change      : 02/20/2013 13:32:05
===============================================================================
===============================================================================
Servers for "aaa-server-policy-3"
===============================================================================
Idx Name                            Address         Port        Oper State
                                                    Auth/Acct
-------------------------------------------------------------------------------
1   server-3                        172.16.1.10     1812/1813   unknown
===============================================================================
```

The Acct-On-Off field indicates if the sending of Accounting-On and Accounting-Off messages is enabled or disabled. If enabled, the oper-state is displayed: state Blocked or state Not Blocked. When Blocked, the radius-server-policy cannot be used to send RADIUS messages.

To display acct-on-off-group information, use following command, for example:

```
# show aaa acct-on-off-group "group-1"
===============================================================================
Acct-On-Off-Group Information
===============================================================================
acct on off group name               : group-1
  - controlling Radius-Server-policy :
        aaa-server-policy-1
  - monitored by Radius-Serer-policy :
        aaa-server-policy-2

-------------------------------------------------------------------------------
Nbr of Acct-on-off-groups displayed : 1
-------------------------------------------------------------------------------
===============================================================================
```

## RADIUS Accounting Message Buffering

When all servers in a radius-server-policy are unreachable, it is possible to buffer the Accounting Stop and Accounting Interim-Update messages for up to 25 hours. When a RADIUS server becomes reachable again then the messages in the buffer are retransmitted.

RADIUS Accounting message buffering parameters can be configured per message type, for example:

```
configure
    aaa
        radius-server-policy "aaa-server-policy-1" create
            servers
                router "Base"
                buffering
                    acct-interim min 60 max 3600 lifetime 12
                    acct-stop min 60 max 3600 lifetime 12
                exit
                server 1 name "server-1"
            exit
        exit
    exit
```

When RADIUS accounting message buffering is enabled:

1. The message is stored in the buffer, a lifetime timer is started and the message is sent to the RADIUS server.
2. If after retry*timeout seconds no RADIUS accounting response is received for the Accounting Interim Update or Accounting Stop then a new attempt to send the message is started after minimum [(min-val*2n), max-val] seconds.
3. Repeat step 2 until:

    a. RADIUS accounting response is received, or

    b. the lifetime of the buffered message expires, or

    c. (if the buffered message is an Accounting Interim-Update only) A new Accounting Interim-Update or an Accounting Stop or for the same accounting session-id and radius-server-policy is stored in the buffer, or

    d. the message is manually purged from the message buffer via a clear command
4. The message is purged from the buffer as shown in Figure 57.

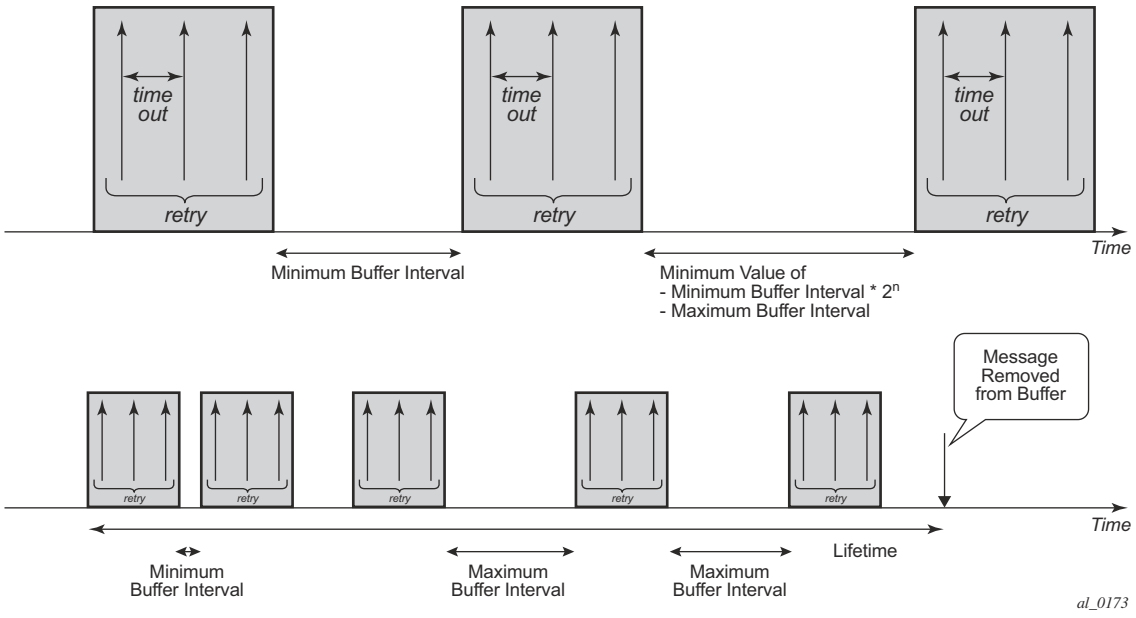



**Figure 57: Purging Message from Buffer**

When Accounting Interim-Update message buffering is enabled, it is recommended to also enable Accounting Stop message buffering. This will guarantee the message ordering per accounting session.

Use following clear command to manually delete messages from the RADIUS accounting message buffer:

**# clear aaa radius-server-policy** *policy-name* **msg-buffer** [**acct-session-id** *acct-session-id*]

When specifying the Acct-Session-Id, only that specific message will be deleted from the message buffer. If no Acct-Session-Id is specified, all messages for that radius-server-policy are deleted from the message buffer.

Use the following show commands to display the RADIUS accounting message buffer statistics:

```
# show aaa radius-server-policy "aaa-server-policy-1" msg-buffer-stats
===============================================================================
RADIUS server policy "aaa-server-policy-1" message buffering stats
===============================================================================
buffering acct-interim      : enabled
  min interval (s)          : 60
  max interval (s)          : 3600
  lifetime (hrs)            : 12
buffering acct-stop         : enabled
  min interval (s)          : 60
  max interval (s)          : 3600
  lifetime (hrs)            : 12

Statistics
-------------------------------------------------------------------------------
```

```
Total acct-stop messages in buffer                        : 0
Total acct-interim messages in buffer                     : 5
Total acct-stop messages dropped (lifetime expired)       : 0
Total acct-interim messages dropped (lifetime expired)    : 0
Last buffer clear time                                    : N/A
Last buffer statistics clear time                         : N/A
-------------------------------------------------------------------------------
===============================================================================
```

Use following clear command to reset the RADIUS accounting message buffer statistics:

# **clear aaa radius-server-policy** *policy-name* **statistics msg-buffer-only**

Use following tools commands to display the RADIUS accounting message buffer content:

**# tools dump aaa radius-server-policy** *policy-name* **msg-buffer [session-id** *acct-session-id***]**

For example:

```
# tools dump aaa radius-server-policy "aaa-server-policy-1" msg-buffer
===============================================================================
RADIUS server policy "aaa-server-policy-1" message buffering
===============================================================================
message type Acct-Session-Id                                remaining lifetime
-------------------------------------------------------------------------------
acct-interim 242FFF0000009A512B36FC                         0d 11:58:54
acct-interim 242FFF0000009B512B36FC                         0d 11:58:48
acct-interim 242FFF0000009C512B36FC                         0d 11:58:30
acct-interim 242FFF0000009D512B36FC                         0d 11:58:29
acct-interim 242FFF0000009E512B36FC                         0d 11:59:05
-------------------------------------------------------------------------------
No. of messages in buffer: 5
===============================================================================
```

When specifying the Acct-Session-Id, the message details are displayed.

## Sending an Accounting Stop Message upon a RADIUS Authentication Failure of a PPPoE Session

In scenarios where Radius authentication is used for PPPoE sessions, an accounting stop message can be generated to notify the Radius servers in case of an authentication failure.

The failure events are categorized in three categories:

- "**on-request-failure**" — All failure conditions between the sending of an Access-Request and the reception of an Access-Accept or Access-Reject.
- "**on-reject**" — When an Access-Reject is received.
- "**on-accept-failure**" — All failure conditions that appear after receiving an Access-Accept and before successful instantiation of the host or session.

Each of the categories can be enabled separately in the RADIUS authentication policy.

In the Enhanced Subscriber Management (ESM) model, the RADIUS accounting server is found after authentication and host identification as part of the subscriber profile configuration. To report authentication failures to accounting servers, an alternative RADIUS accounting policy configuration is required: local user database pre-authentication can provide the RADIUS authentication policy to be used for authentication and the RADIUS accounting policy to be used for authentication failure reporting. A duplicate RADIUS accounting policy can be specified if the accounting stop resulting from a RADIUS authentication failure must also be sent to a second RADIUS destination.

```
configure
    subscriber-mgmt
        local-user-db "ludb-1" create
            ppp
                match-list username
                host "default" create
                    auth-policy "auth-policy-1"
                    password ignore
                    acct-policy "acct-policy-1" duplicate "acct-policy-2"
                    no shutdown
                exit
            exit
            no shutdown
        exit
        authentication-policy "auth-policy-1" create
            pppoe-access-method pap-chap
            include-radius-attribute
               - - - snip - - -
            exit
            send-acct-stop-on-fail on-request-failure on-reject on-accept-failure
            radius-server-policy "aaa-server-policy-1"
        exit
        radius-accounting-policy "acct-policy-1" create
            - - - snip - - -
            radius-server-policy "aaa-server-policy-1"
        exit
        radius-accounting-policy "acct-policy-2" create
            - - - snip - - -
```

```
            radius-server-policy "aaa-server-policy-2"
        exit
```

To enable local user database pre-authentication, use the user-db configuration in the capture SAP and in the group-interface. For example:

```
configure
    service
        vpls 10 customer 1 create
            sap 1/1/1:1.* capture-sap create
                trigger-packet pppoe
                pppoe-policy "ppp-policy-1"
                pppoe-user-db "ludb-1"
            exit
            no shutdown
        exit
        ies 1000 customer 1 create
            subscriber-interface "sub-int-1" create
               - - - snip - - -
                group-interface "group-int-1-1" create
                    - - - snip - - -
                    pppoe
                        policy "ppp-policy-1"
                        user-db "ludb-1"
                        no shutdown
                    exit
                exit
            exit
            no shutdown
        exit
```

# Enhanced Subscriber Management Overview

Topics in this section:



## Enhanced Subscriber Management Basics

In residential broadband networks numerous subscribers can be provisioned that can require significant changes on a daily basis. Manually configuring the applicable parameters for each subscriber would be prohibitive. The Alcatel-Lucent 7750 SR has been designed to support fully dynamic provisioning of access, QoS and security aspects for residential subscribers using DHCP to obtain an IP address. Enabling Enhanced Subscriber Management drastically reduces the configuration burden.

Enhanced Subscriber Management in the 7750 SR supports many vendor's access nodes and network aggregation models, including VLAN per customer, per service or per access node.

### Standard and Enhanced Subscriber Management

The system can switch between standard and enhanced subscriber management modes on a per SAP basis. The Enhanced Subscriber Management mode is supported on the SR-7 and SR-12 chassis and on the ESS-7 chassis.

Some functions are common between the standard and enhanced modes. These include DHCP lease management, static subscriber host definitions and anti-spoofing. While the functions of these features may be similar between the two modes, the behavior is considerably different.

- Standard mode — The system performs SLA enforcement functions on a per SAP basis, that is, the attachment to a SAP with DHCP lease management capabilities. The node can authenticate a subscriber session with RADIUS based on the MAC address, the circuit-id (from Option 82) or both. It will then maintain the lease state in a persistent manner. It can install anti-spoofing filters and ARP entries based on the DHCP lease state. Static subscriber hosts are not required to have any SLA or subscriber profile associations and are not required to have a subscriber identification string defined.
- Enhanced mode — When enabled on a SAP, the system expands the information it stores per subscriber host, allowing SLA enforcement and accounting features on a per subscriber basis.

The operator can create a subscriber identification policy that will include a URL to a user-space script that assists with the subscriber host identification process.

→ A subscriber host is identified by a subscriber identification string instead of the limited Option 82 values (although, the identification string is normally derived from string manipulation of the Option 82 fields). A subscriber identification policy is used to process the dynamic host DHCP events to manage the lease state information stored per subscriber host. The static subscriber hosts also must have subscriber identification strings associations to allow static and dynamic hosts to be grouped into subscriber contexts.

→ Further processing by the subscriber identification policy derives the appropriate subscriber and SLA profiles used to define the hierarchical virtual schedulers for each subscriber and the unique queuing and filtering required for the hosts associated with each subscriber

→ The SLA profile information is used to identify which QoS policies and which queues will be used for each subscriber host (dynamic or static).

→ The system performs SLA enforcement functions on a per subscriber SLA profile instance basis. SLA enforcement functions include QoS (classification, filtering and queuing), security (filtering), and accounting.

When the enhanced mode is enabled on a SAP (see Subscriber SAPs on page 871), first, the router ensures that existing configurations on the SAP do not prevent proper enhanced mode operation. If any one of the following requirements is not met, enhanced mode operation is not allowed on the SAP:

- Anti-spoofing filters must be enabled and configured as IP+MAC matching.
- Any existing static subscriber hosts must have:
  - → An assigned subscriber identification string.
  - → An assigned subscriber profile name.
  - → An assigned SLA profile name.
- The system must have sufficient resources to create the required SLA profile instances and schedulers.

When the router successfully enables the enhanced mode, the current dynamic subscriber hosts are not touched until a DHCP message event occurs that allows re-population of the dynamic host information. Thus, over time, the dynamic subscriber host entries are moved from SAP-based queuing and SAP-based filtering to subscriber-based queuing and filtering. In the event that a dynamic host event cannot be processed due to insufficient resources, the DHCP ACK message is discarded and the previous host lease information is retained in the system.

## Subscriber Management Definitions

### Subscriber

A subscriber is typically defined by a unique subscriber identifier to which an assortment of polices (or subscriber profile) can be applied. A subscriber typically (but not always) maps into a VLAN, a VPI/VCI pair, an “ifentry” (a logical interface such as a SAP), a (source) MAC or IP address or a physical port, which uniquely identify a billable entity for the service provider.

### Subscriber Management

The management of all services, policies, AAA functions and configurations that relate to the concept of a subscriber. Subscriber management can be configured in a variety of ways, but it is critical that subscriber management integrates seamlessly with element and service management across the broadband infrastructure, via for instance, the Alcatel-Lucent 5750 Subscriber Services Controller (SSC). Subscriber management can also be implemented through CLI or scripted commands at the platform level, whereby a network administrator would manually configure the set of QoS, security, AAA or anti-spoofing functions that relate to a particular billable entity or subscriber. Subscriber management is typically centralized and highly integrated with the element, services and middleware management functions for streamlined management, flowthrough provisioning, and accelerated service activation, with minimized operating expenditures.

### Subscriber Policy Enforcement

Is the set of actual enforcement functions that are implemented relative to a given subscriber, possibly at multiple enforcement points in the infrastructure and as a result of a match between the subscriber profile which was defined by the subscriber management suite (Alcatel-Lucent’s 5750 SSC) and actual traffic patterns. Examples include for instance, the shaping, policing or rate limiting of traffic or the traffic of a given subscriber being dropped because it matched or violated any specific rule (packet with a mismatch between MAC and IP address suggesting an address spoof for instance)

### Subscriber SAPs

A subscriber SAP is a service access point (SAP) where enhanced subscriber management is active. Enhanced subscriber management must be explicitly enabled on a per-SAP basis with the CLI **sub-sla-mgmt** command.

A subscriber SAP can be used by a single subscriber or support multiple subscribers simultaneously. Each subscriber can be represented by one or multiple subscriber hosts on the subscriber SAP. If enhanced subscriber management is enabled on a SAP, any configured QoS and

IP filter policies defined on the SAP are ignored. A subscriber SAP must refer to an existing subscriber identification policy.

## Hosts and Subscribers

A host is a device identified by a unique combination of IP address and MAC address. Typically, the term “subscriber host” is used instead of the “host”.

A host can be an end-user device, such as a PC, VoIP phone or a set top box, or it can be the user’s Residential Gateway (RGW) if the RGW is using Network Address Translation (NAT).

Each subscriber host must be either statically provisioned or dynamically learned by the system. The host’s IP address + MAC address are populated in the subscriber host table on the appropriate SAP to allow packets matching the IP address and MAC address access to the provider’s network.

- A dynamic subscriber host is dynamically learned by the system through the DHCP snooping or relay process. Each subscriber SAP created on the system is configured (using the lease-populate command) to monitor DHCP activity between DHCP clients reached through the SAP and DHCP servers. DHCP ACKs from the DHCP server are used to determine that a certain IP address is in use by a specific DHCP client. This client IP address association is treated by the system as a dynamic subscriber host.
- When it is not possible to dynamically learn a subscriber host through DHCP, a static subscriber host can be created directly on a subscriber SAP. Since a subscriber identification policy is not applicable to static subscriber hosts, the subscriber identification string, subscriber profile and SLA profile must be explicitly defined with the hosts IP address and MAC address.

A subscriber (in the context of the router) is a collection of hosts getting common (overall) treatment. It is expected that this group of hosts originate from the same site and all hosts of a subscriber are reached by the same physical path (such as a DSL port).

Once a subscriber host is known by the system, it is associated with a subscriber identifier and an SLA profile instance. Subscriber hosts with a common subscriber identifier are considered to be owned by the same subscriber.

Depending on the network model, hosts associated with a single subscriber can be associated with a single subscriber SAP or spread across multiple subscriber SAPs on the same port.

## Subscriber Identification Policy

The subscriber identification policy contains the URL definitions for the Programmable Subscriber Configuration Policy (PSCP) scripts used for DHCP ACK message processing. Up to three URLs can be defined per subscriber identification policy. These are designated as primary, secondary and tertiary. Each URL can be individually enabled or disabled. Only one script (the URL with the highest priority active script) is used at any one time to process DHCP ACK messages. If the system detects an error with a specified script, the URL is placed in an operationally down state. If the script is shutdown, it is placed in an administratively down state. A script that is operationally or administratively down is considered inactive. The system automatically reverts to the highest priority active script. If a script becomes operationally down, it must be cycled through the administratively down then administratively up states for the system to attempt to reactivate the script.

Multiple subscriber identification policies are provided for the event that access nodes (such as DSLAMs) from different vendors are attached to the same router. Each policy's active script can be explicitly defined to process the various DHCP message formats or idiosyncrasies of each vendor.

If a script is changed, it must be reloaded by disabling and re-enabling any URL which refers to the changed script (a **shutdown** command followed by a **no shutdown** command).

Each subscriber identification policy can also contain a subscriber profile map and/or an SLA profile map. The subscriber profile map creates a mapping between the sub-profile-strings returned from the active script with an existing subscriber profile name. The SLA profile map is used to create a mapping between the sla-profile-strings returned from the active script with an existing SLA profile name.

The subscriber identification policy is designed to accept a DHCP ACK message destined for a subscriber host and return up to three string values to the system;

- The subscriber identification string (mandatory)
- The subscriber profile string (optional)
- The SLA profile string (optional).

These strings are used to derive the subscriber profile and the SLA profile to be used for this host

See Using Scripts for Dynamic Recognition of Subscribers on page 905.

## Subscriber Identification String

Subscribers are managed by the router through the use of subscriber identification strings. A subscriber identification string uniquely identifies a subscriber.

The subscriber identification string is the index key to any entry in the active subscriber table, and thus must always be available. It is derived as follows:

- For dynamic hosts, the subscriber identification string is derived from the DHCP ACK message sent to the subscriber host.
  - → The DHCP ACK message is processed by a subscriber identification script which has the capability to parse the message into an alternative ASCII string value.
  - → If enhanced subscriber management is disabled, the default value for the string is the content of the Option 82 circuit-id and remote-id fields interpreted as an octet string.
- For static hosts, the subscriber identification string must be explicitly defined with each static subscriber host.

When multiple hosts are associated with the same subscriber identification string, they are considered to be host members of the same subscriber. Hosts from multiple SAPs can be members of the same subscriber, but for proper virtual scheduling to be performed all hosts of a subscriber must be active on the same IOM.

When the first host (either dynamic or static) is created with a certain subscriber identification string, an entry is created in the active subscriber table. The entries are grouped by their subscriber identification string.

---

## Subscriber Profile

The subscriber profile is a template which contains those hierarchical QoS (HQoS) and accounting settings which are applicable to all hosts belonging to the same subscriber. These include:

- Ingress and egress scheduler policy HQoS
- Accounting policy
- RADIUS accounting policy

Subscribers are either explicitly mapped to a subscriber profile template or are dynamically associated with a subscriber profile.

Attempting to delete any subscriber profile (including the profile named ‘default’) while in use by an existing active subscriber will fail.

### SLA Profile

For the purpose of supporting multiple service types (such as high speed Internet (HSI), voice over IP (VoIP), video on demand (VoD) and Broadcast TV) for a single subscriber, the hosts associated with a subscriber can be subdivided into multiple SLA profiles.

The SLA profile contains those QoS and security settings which are applicable to individual hosts. An SLA profile acts like a template and can be used by many subscribers at one time. Settings in the SLA profile include:

- Egress and ingress QoS settings
- Egress and ingress IP filters
- Host limit

If the SLA profile does not explicitly define an ingress or egress QoS policy, the default SAP ingress or default SAP egress QoS policy is used.

Refer to Determining the SLA Profile on page 910 for information on how the SLA profile is determined for dynamic hosts.

---

### Explicit Subscriber Profile Mapping

An explicit mapping of a subscriber identification string to a specific subscriber profile can be configured.

An explicit mapping overrides all default subscriber profile definitions while processing a DHCP ACK. In an environment where dynamic and static hosts coexist in the context of a single subscriber, care will be taken to not define a subscriber profile in the explicit subscriber map that conflicts with the subscriber profile provisioned for the static host(s). If such a conflict occurs, the DHCP ACKs will be dropped.

An explicit mapping of a subscriber identification string to the subscriber profile name ‘default’ is not allowed. However, it is possible for the subscriber identification string to be entered in the mapping table without a defined subscriber profile which can result in the explicitly defined subscriber to be associated with the subscriber profile named ‘default’.

Attempting to delete a subscriber profile that is currently defined in an explicit subscriber identification string mapping will fail.

The explicit mapping entries can be removed at any time.

# ESM for IPv6

ESM for IPv6 is supported on 7750 chassis with at least IOM3-XP cards or equivalent or in 7450 chassis operating in Mixed Mode (containing one or more IOM3-XP cards that have the 7750 SR feature set enabled.) ESM for IPv6 is supported with RADIUS as the backend authentication and authorization mechanism.

## Models

- PPPoE Host
- PPPoE RG
- IPoE Host/RG

### PPPoE Host

For PPPoE, the ESR suggests the IPv6CP protocol to the client during the session setup phase if the appropriate attributes have been returned by the RADIUS server on authentication. The RADIUS attribute that indicates the setup of a PPPoE host is Framed-IPv6-Prefix, which should contain a /64 prefix for the client.

When a PPPoE host has successfully completed the IPv6CP negotiation, the ESR will transmit a Router Advertisement to the PPPoE host containing the suggested prefix and any other options that are configured. The client may use this information to pick one or more addresses from the suggested prefix; all addresses within the prefix are forwarded towards the client.

Alternatively, the Recursive DNS Server (RDNSS) Option as defined in RFC-6106 can be included in IPv6 Router Advertisements for DNS name resolution of IPv6 SLAAC hosts. Following CLI command includes the DNS info in IPv6 Router Advertisements for SLAAC hosts and sets the RDNSS liftetime:

```
config>service>ies>sub-if>grp-if>ipv6>rtr-adv
config>service>vprn>sub-if>grp-if>ipv6>rtr-adv

[no] dns-options

   [no] include-dns    - Set/reset inclusion of the RDNSS server
                         option 25 on this group-interface
   [no] rdnss-lifetime - Maximum time the RDNSS address is valid
                         in this group-interface
```

The source for DNS information to be included in Router Advertisements for IPv6 SLAAC hosts, can be (listed in priority order):

1. Local User Database IPv6 options:

```
configure subscriber-mgmt local-user-db <ludb-name> dhcp|ppp host <host-name>
options6 dns-server <ip-address> [<ip-address>...(up to 4 max)]
```

2. RADIUS attributes [26-6527-105] Alc-Ipv6-Primary-Dns and [26-6527-106] Alc-Ipv6-Secondary-Dns
3. Default IPv6 DNS Server configured at the group interface:

```
configure service ies|vprn <svc-id> subscriber-interface <sub-int-name> ipv6 default-
dns <ipv6-address> [secondary <secondary-ipv6-address>]
```

**Note:** A default IPv6 Server configuration at the group interface is a last resort IPv6 DNS info that can be used for IPoEv6 hosts (IA_NA, IA_PD and SLAAC) and PPPoEv6 hosts (IA_NA, IA_PD and SLAAC).

## PPPoE RG

Initially, a PPPoE RG follows the same procedure as a PPPoE host: the ESR receives a prefix from RADIUS (in this case through a Delegated-IPv6-Prefix attribute), which is used as a trigger to suggest the IPv6CP protocol to the client. The prefix that is suggested to the client should have the same prefix length as configured under the subscriber interface ipv6 node (delegated-prefix-length). This length should be between 48 and 64 bits, inclusive.

After the IPv6CP protocol has completed, however, the client should run the DHCPv6 protocol over its PPPoE tunnel to receive a Delegated Prefix (IA_PD) and optionally IPv6 DNS server information. This Delegated Prefix can then be subdivided by the client and distributed over its downstream interfaces. During DHCPv6, no extra RADIUS request will be made; the information is stored during the initial (PPPoE or PPP) authentication until the client starts DHCPv6.

Only after DHCPv6 has completed, the IPv6 subscriber host will be instantiated and the ESR will start sending Router Advertisements (if configured.) The router advertisements will not contain any prefix information, which has already been provided by DHCPv6, but it is used as an indication to the client that its default gateway should be the ESR.

## IPoE Host/RG

Similar to an IPv4 DHCP client, a DHCPv6 client is authenticated at its Solicit message, where it can request one or more addresses or prefixes. The address and prefix types supported are IA_NA (Non-Temporary Address) through the Alc-IPv6-Address RADIUS attribute and IA_PD (Delegated Prefix) through the Delegated-IPv6-Prefix attribute. Contrary to the IPv4 case, the ESR will always reply to a DHCPv6 request because the client may request more than one address or prefix simultaneously and not all of the requests may be honored.

The DHCPv6 protocol handling and Router Advertisement behavior are similar to the PPPoE RG case above, with the exception that for an IA_NA address, the entire /64 prefix containing the address is allocated to the client.

## Setup

IPv6 ESM hosts are only supported in the Routed CO model (both VPRN and IES).

At the ipv6 node under the subscriber interface level, the length of the prefixes that are offered is defined through the delegated-prefix-length option. This setting is fixed for the subscriber interface and can not be changed once subscriber prefixes are defined.

Subscriber prefixes define the ranges of addresses that are offered on this subscriber interface. By default only these subscriber prefixes are exported to the routing protocols to keep the routing tables small. There are three types of subscriber interfaces:

- wan-host — A range of prefixes that are assigned to PPPoE hosts and as DHCPv6 IA_NA addresses. These prefixes are always /64.
- pd — A range of prefixes that are assigned as DHCPv6 IA_PD prefixes for DHCPv6 IPoE clients and for PPPoE RGs. The length of these prefixes is defined by the delegated-prefix-length.
- both — When both 'wan-host' and 'pd' are defined, the subscriber prefix is a range that can be used for both previous types. However, the delegated-prefix-length is restricted to /64 in this case.

The IPv6 node under the group interface contains the DHCPv6 proxy configuration and the router advertisement configuration.

# Behavior



## Dual Stack

Clients may support both IPv4 and IPv6 simultaneously (dual stack hosts.) In this case one subscriber host entry will be created for the IPv4 address family and one for the IPv6 instance. The scaling limits apply for all entries, regardless of address type.

For DHCP, these subscriber hosts are fully independent (as they are set up through different protocols), but for PPPoE hosts or RGs, the ESM information in both subscriber host entries is linked together through the PPPoE session.

## Router Advertisements (RA)

RA messages are started immediately after the subscriber host is instantiated and unsolicited messages are sent in the interval defined in the configuration. Apart from unsolicited RAs, the client may also send a router solicitation (RS) to explicitly request the information. RAs are throttled so that they are not sent more often than once every three seconds.

## CoA and Disconnect-Request

For IPv6 subscriber hosts, RADIUS-triggered mid-session changes and session terminations may identify the subscriber host to be changed by the same address or prefix that was originally returned from RADIUS. Only one address attribute (framed-IP-address, framed-IPv6-prefix, delegated-IPv6-prefix or Alc-IPv6-address) may be given in a single request.

For PPPoE clients, changing either the IPv4 or IPv6 information will result in both the v4 and v6 subscriber host being modified (if they are contained within the same PPPoE session.)

The only CoA action that is allowed for IPv6 hosts is a change of ESM strings; creation of new hosts and forcing a DHCPv6 RENEW is not supported.

## Delegated-Prefix-Length

The delegated prefix length (DPL) is applicable to subscriber-hosts with IPv6 Prefix (IA-PD) assigned via DHCPv6 Server. IPv6 Prefix is more akin to a route then it is to an IP address. The length of the prefix plays crucial role in forwarding decisions, antispoofing, and prefix assignment via DHCPv6 pools in the local DHCPv6 Server.

The structure of an IPv6 prefix is shown in Figure 58.

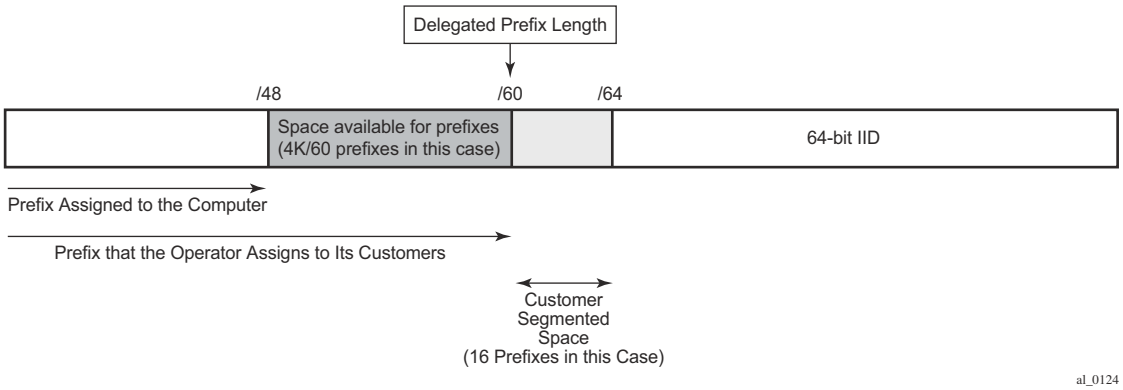



**Figure 58: IPv6 Prefix**

For example, a DHCPv6 server prefix pool contains an aggregated (configured) IPv6 prefix from which the delegated prefixes will be carved out. In Figure 58 this aggregated IPv6 prefix has length of /48. In addition, the DHCPv6 server needs to know the length of the delegated prefix (in the above case /60). These two values are marking the boundary within which a unique delegated prefix will be selected. This is represented by the purple area in Figure 58.

The delegated prefix length can be obtained via:

- RADIUS
  - Delegated-IPv6-Prefix attribute that contains the prefix and the length (Delegated-IPv6-Prefix = AAAA:BBBB::/56). The DPL in this case is /56.
  - Alc-Delegated-IPv6-Prefix-Length VSA (to be used in conjunction with the DHCPv6 pool name - Alc-Delegated-IPv6-Pool VSA)
- LUDB – configured via LUDB per IPoEv6/PPPoEv6 host:

```
configure
subscriber-mgmt
    local-user-db <name>
        dhcp | ppp
            host <name>
                ipv6-delegated-prefix-length [48..64]
```

This is to be used along with the DHCPv6 pool name (ipv6-delegated-prefix-pool) defined under the same CLI hierarchy.

Alternatively, the entire prefix, including the DPL can be returned via LUDB.

```
configure
    subscriber-mgmt
        local-user-db <name>
            dhcp | ppp
                host <name>
                    ipv6-delegated-prefix <ipv6-prefix/prefix-length>
```

- DHCPv6 server – each DHCPv6 pool can optionally be configured with a DPL:

```
configure
    service/router
        dhcp6
            local-dhcp-server <name>
                pool <pool-name>
                    delegated-prefix-length [48..64]
```

- Configured statically under the ipv6 CLI node of subscriber-interface. In this case the DPL is fixed for all subscriber-hosts under the subscriber-interface.

```
configure
    service ies/vprn
        subscriber-interface <name>
            ipv6
                delegated-prefix-length [48..64] | variable
```

---

## Order of Preference for DPL

In case that the DPL is statically provisioned under the subscriber-interface>ipv6> hierarchy, all hosts under this subscriber-interface will inherit this fixed DPL. In case that the DPL is provided via LUDB or RADIUS in addition to static configuration under the subscriber-interface then the LUDB or the RADIUS one MUST match the DPL that is statically provisioned under the subscriber-interface. Otherwise, the prefix instantiation in 7x50 will fail.

Note that the “no delegated-prefix-length” command under the **subscriber-interface>ipv6>** hierarchy means that the DPL is set to a default-value of 64.

When the delegated-prefix-length commands under the **subscriber-interface>ipv6>** hierarchy is set to variable, prefixes under such subscriber-interface can have different lengths and the DPL can be configured via one of the following means:

- LUDB
- RADIUS

- DHCP Server

## DHCP Server Address Utilization and Delegated Prefix Length

In case that the delegated prefix length is variable, for each consecutive address allocation request for the given delegated prefix, the DHCPv6 server will allocate the prefix at the end of the last delegated lease with the same delegated prefix length. This will minimize the address space fragmentation within the configured prefix.

## DHCPv6 Relay Agent

A DHCPv6 Relay Agent can support a 7x50 DHCPv6 local server (same or remote chassis) and a third party DHCPv6 external server.

An incoming DHCPv6 client message is relayed within the Relay-Forward message specified in RFC 3315. If the server responds with a valid address/prefix, the ESM process attempts to install it. If it fails, the DHCPv6 Relay Agent sends an explicit RELEASE to the server. There is no retransmission of DHCPv6 Relay-Forwards in the case of failure – it requires the client to re-start or re-send the original DHCPv6 message.

A Lightweight DHCPv6 Relay Agent may insert Relay Agent Information including the Interface ID option between the DHCPv6 client and the DHCPv6 Relay Agent.

Additional Relay Agents (non-LDRA) between the DHCPv6 client and the DHCPv6 Relay Agent are not supported.

---

### Configuring a DHCPv6 Relay Agent

A DHCPv6 Relay Agent is configured in the IPv6 DHCP6 context of a group-interface:

```
config>service>vprn>sub-if>grp-if>ipv6>dhcp6# relay ?
config>service>ies>sub-if>grp-if>ipv6>dhcp6# relay ?
  - no relay
  - relay

 [no] client-applications - Configure the set of DHCP6 relay server client
                            applications
 [no] description     - Description for DHCPv6 relay
 [no] link-address    - Configure the link address of the DHCPv6 relay messages
 [no] option          + Configure the DHCPv6 Relay information options
 [no] server          - Configure the DHCPv6 server IPv6 address
 [no] shutdown        - Administratively enable/disable DHCPv6 relay on this interface
 [no] source-address  - Configure the source IPv6 address of the DHCPv6 relay messages
```

Up to eight DHCPv6 servers can be provisioned to be served by a DHCPv6 Relay Agent. A Relay-Forward is send to all servers and the Relay-Replies from all servers are sent to the client.

The “client-applications” parameter specifies if the Relay Agent can be used for IPoE (dhcp) or PPP (ppp) hosts.

Optional configuration parameters:

- description: a free configurable description string.
- link-address: the link address field in the DHCPv6 Relay-Forward message header.

  The link address can be configured to enable link-address based pool selection in a 7x50 DHCPv6 local server. The address must be one of the IPv6 prefixes configured at the ipv6 subscriber-prefixes context for a subscriber interface. If not configured, the system selects one of the prefixes.

- option: allows to configure following options to be inserted in the Relay-Forward message:
  - → Interface-Id [18] – the interface ID option identifies the interface on which the DHCPv6 client message is received. The format options are the following:
    - – ascii-tuple: *host-name*|*service-id*|*group-interface-name*|*sap-id*
    - – ifindex: Interface index for the group-interface
    - – sap-id: SAP identifier (port and vlans)
    - – string <string>: a free configurable string (max. 80 chars)
  - → Remote-Id [37] – Relay Agent Remote Id option contains the DHCPv6 client DHCP Unique Identifier (DUID).
- source-address: the source-address of the Relay-Forward messages.

  If not configured, the outgoing interface IPv6 address is used.

  The source-address configuration is mandatory for a DHCP Relay Agent in a VPRN service when the DHCPv6 server is reachable via a tunnelled next-hop (MPLS).

---

## DHCPv6 Relay to Third Party DHCPv6 External Server

When the DHCPv6 Relay Agent is relaying to a third party DHCPv6 external server, following conditions should be met:

- The third party DHCPv6 server must return a unique IA_PD IPv6 delegated prefix (/64 or lower) for each allocation. The length of the IA_PD IPv6 delegated prefix must match the delegated-prefix-len configured on the subscriber interface on the 7750 DHCP L3 relay. This length is also included in the Relay-Forward message as PFX_LEN option (3) in a Vendor-Specific-Information-Option (17)
- For IPv6oE routed CPE's, the 3rd party DHCPv6 server must return a unique IA_NA IPv6 address (/128) from a different /64 subnet for each allocation.
- For IPv6oE hosts behind bridged CPE's,
  - → the third party DHCPv6 server must return a unique IA_NA IPv6 address (/128) from a different /64 subnet for each allocation (host) that belongs to a different CPE.
  - → the third party DHCPv6 server may return a unique IA_NA IPv6 address (/128) from the same /64 subnet for allocations (hosts) that belong to the same CPE and that are attached to the same vlan (SAP) on the BNG.

Following information is available to the third party DHCPv6 server in a Vendor-Specific-Information-Option (17) included in the Relay-Forward message:

- WAN_POOL option (1): contains the pool name from which the IA_NA IPv6 address should be allocated.
- PFX_POOL option (2): contains the pool name from which the IA_PD IPv6 delegated prefix should be allocated.

- PFX_LEN option (3): contains the IA_PD IPv6 delegated prefix length that should be allocated.

## DHCPv6 Local Server

A local DHCPv6 pool server for both addresses (IA_NA) and prefixed (IA_PD) manages the address and prefixes sent to either routing gateways or hosts.

Because IPv6 home networks lack NAT, the IPv6 addresses delegated to a routing gateway are in turn assigned to hosts in the home. These addresses are assigned with reasonably long (but configurable) lifetimes such that the loss of the WAN connection will not result in the IPv6 hosts in the LAN losing their IPv6 addresses. One consequence of these long lifetimes is that the IPv6 hosts will retain any IPv6 address provided the valid-lifetime is greater than zero. Should an operator delegate a prefix and then at a later time delegate a second IPv6 prefix, a host may end up with two or more valid prefixes. This situation plays havoc with IPv6 source address selection and may result in impaired service.

To overcome the problems of multiple IPv6 prefixes in the home, the operator must ensure that the individual subscriber has the same IPv6 prefix even across modem reboots (that is, if a subscriber session is destroyed and later re-created, an attempt should be made to use the previously delegated prefix). In release 8.0, the operator used RADIUS for all address and prefix assignment, but in release 9.0, with the introduction of the local DHCPv6 server, it requires the 7750 to process and maintain some state even after a session disconnects.

For the DHCPv6 local server to function, a DHCPv6 relay or proxy function must also operate alongside ESM. For the purposes of this document, to relay means to implement a DHCPv6 Relay as indicated in RFC 3315 : a relay encapsulates the client DHCP message within a DHCP Relay-Forward message and unicasts it to a specified destination.

A proxy is an internal concept. Unlike a DHCPv6 relay, the DHCPv6 proxy does NOT encapsulate the client message in a Relay-Forward, nor does it send packets towards the Local DHCPv6 Server. The DHCPv6 proxy is exclusively used as an interface between the RADIUS Access-Accept or local user database lookup and the DHCPv6 client in the consumer device.

The use of the DHCPv6 relay or proxy function depends on the attributes returned from authentication phase (RADIUS or LUDB).

1. DHCPv6 Proxy:
   - → If only IPv6 address/prefix information provided (Framed-IPv6-Prefix, Alc-IPv6-Address or Delegated-IPv6-Prefix).
2. DHCPv6 Relay:
   - → If no IPv6 address/prefix (Framed-IPv6-Prefix, Alc-IPv6-Address or Delegated-IPv6-Prefix) and no IPv6 pool (Framed-Pool, Delegated-Pool) information provided.
   - → If no IPv6 address/prefix (Framed-IPv6-Prefix, Alc-IPv6-Address or Delegated-IPv6-Prefix) and IPv6 pool (Framed-Pool, Delegated-Pool) information provided.
3. If both IPv6 address/prefix (Framed-IPv6-Prefix, Alc-IPv6-Address or Delegated-IPv6-Prefix) and IPv6 pool (Framed-Pool, Delegated-Pool) information are present, the DHCP packet is DROPPED.

**Note:** If IPv6 DNS parameters are returned in RADIUS AND a pool is specified then the DNS parameters are ignored. It is the DHCPv6 server that will need to reply with appropriate DNS servers.

# Dynamic Subscriber Host Processing

## Dynamic Tables

To support all processing for Enhanced Subscriber Management, several tables are maintained in the router (Figure 59).

- Active Subscriber Table on page 891
- SLA Profile Instance Table on page 891
- Subscriber Host Table on page 891
- DHCP Lease State Table on page 893

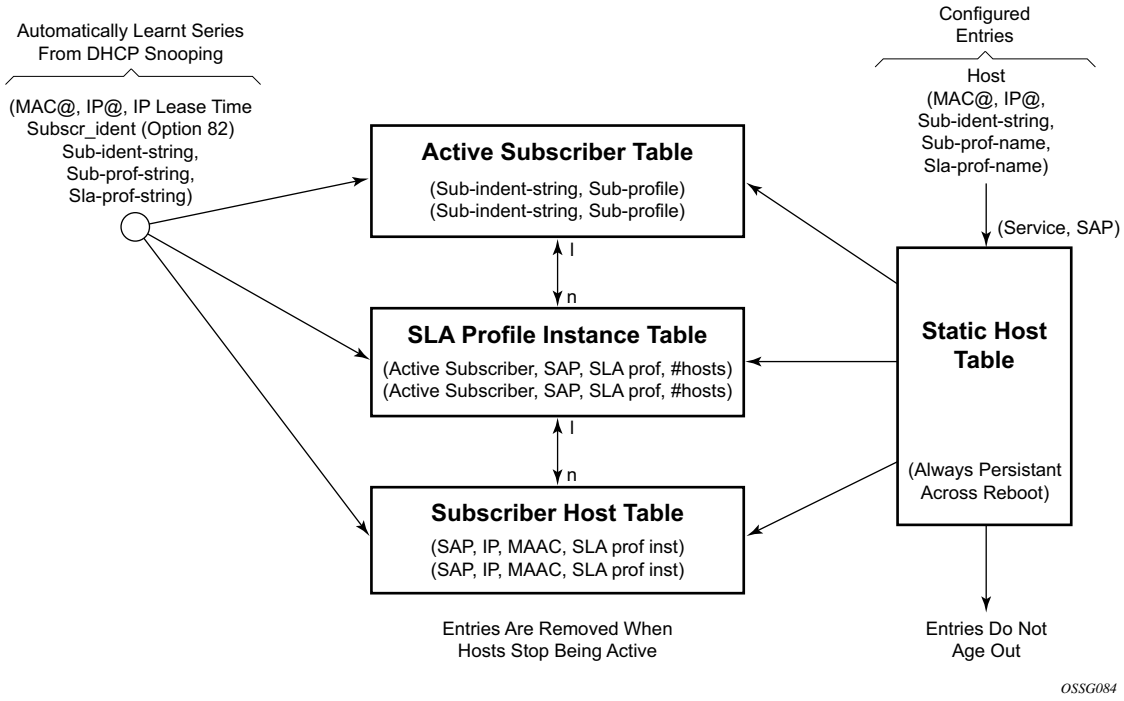



**Figure 59: Enhanced Subscriber Management Dynamic Tables**

### Active Subscriber Table

An entry is created in the active subscriber table when the first host (either dynamic or static) is created with a certain subscriber identification string. The entries are grouped by their subscriber identification string.

Fields for each entry in the active subscriber table include:

- The subscriber identification string (see Subscriber Identification String on page 875).
- In use subscriber profiles (see Subscriber Profile on page 875).

---

### SLA Profile Instance Table

An entry is created in the SLA profile instance table when the first subscriber host on a certain SAP is created that uses a certain SLA profile. All subsequent hosts of the same subscriber on the same SAP that use the same SLA profile will be associated with this entry. When the last host on this SAP, using this SLA profile disappears, the SLA profile instance is deleted from the table and the associated queues are removed.

SLA profile instances can not span multiple subscriber SAPs. If subscriber hosts from the same subscriber exist on multiple SAPs and are associated with the same SLA profile template, a separate SLA profile instance is created for each SAP.

Fields for each entry in the SLA profile instance table include:

- Active subscriber
- SAP
- SLA profile
- Number of active subscriber hosts that share this instance

---

### Subscriber Host Table

An entry is created in the subscriber host table if anti-spoofing is enabled as well as:

- The first host (dynamic or static) with a specific IP and MAC combination is created. If the anti-spoof is IP only, the MAC address is masked to all 0's. If anti-spoof is MAC, only the IP address is 0.0.0.0. All dynamic hosts and static hosts with the same IP and MAC combination will be associated with the same subscriber host entry. If the anti-spoof type includes IP (IP-only or IP/MAC), there can be at most two hosts associated with the entry: one dynamic and one static. If the anti-spoof type is MAC-only, there can be a combination of several dynamic and static hosts associated with the entry.
- The non-prof-traffic is provisioned. Both IP and MAC address are all 0's.

Fields for each entry in the subscriber host table include:

- SAP

- IP address
- MAC address
- SLA profile instance (enhanced mode only)

## DHCP Lease State Table

An entry in the DHCP lease state table is created for each dynamic host. Fields for each entry in the lease state table include:

- Assigned IP address
- Assigned MAC address
- Persistence key

# Enhanced Subscriber Management Entities

Figure 60 illustrates the relationship between the main entities in Enhanced Subscriber Management:

- A subscriber is associated with only one subscriber profile.
- A subscriber can be associated with one or more SLA profile (a VPLS service with 2 different SAPs can have different SLA profiles for the same subscriber).
- A maximum of one SLA profile instance is generated (including ingress and egress queues) per SAP per SLA profile.
- One or more hosts can be assigned to each SLA profile instance (these will share the same queues).

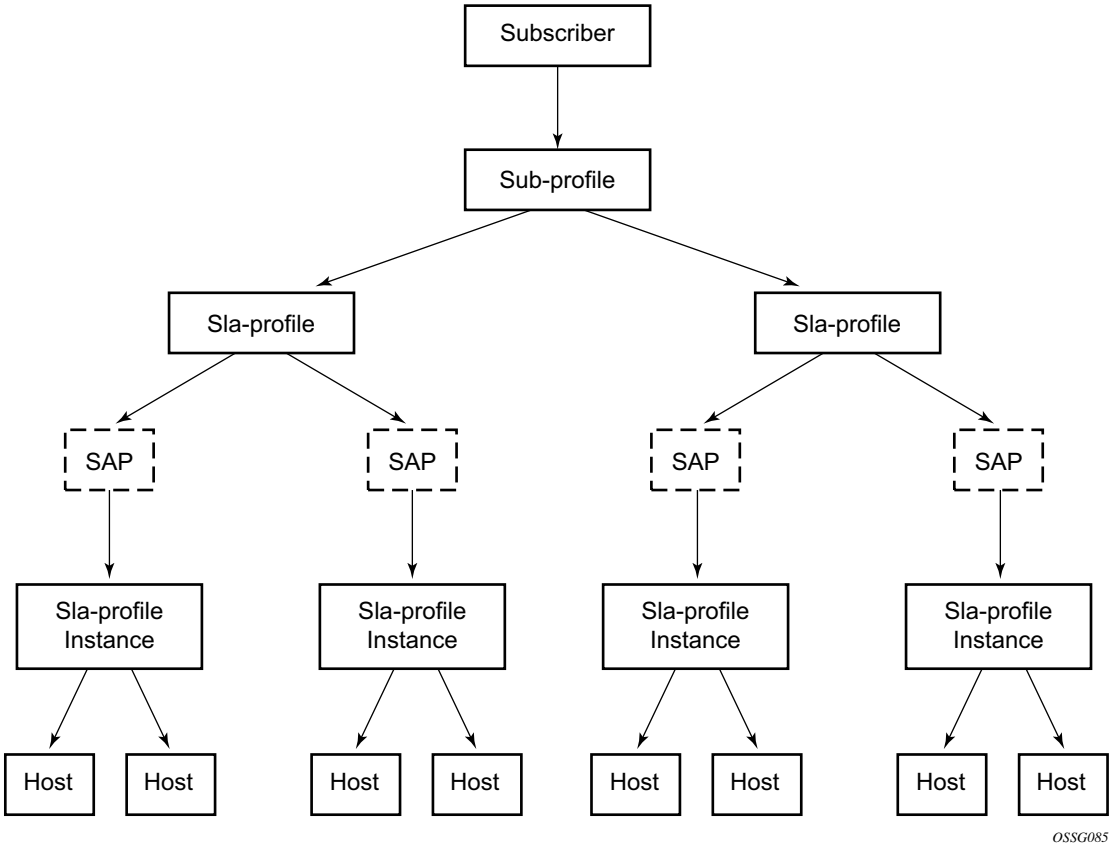



**Figure 60: Relationship Between Enhanced Subscriber Management Entities**

## Instantiating a New Host

When a DHCP ACK is received for a new subscriber host on a particular SAP:

- The ACK message is parsed using the appropriate script.
- An entry is generated in the subscriber host table with indices:
  - → The SAP on which the host resides
  - → The assigned IP address
  - → The assigned MAC address and as lookup parameters:
    - – The subscriber profile and
    - – The SLA profile to be used (derived from using the script).

If this is the first host of a subscriber, an HQoS scheduler is instantiated using the ingress and egress scheduler policies referred to in the subscriber profile. Otherwise, if the subscriber profile of the new host equals the subscriber profile of the existing subscriber, the new host is linked to the existing scheduler. If the subscriber profile is different from the subscriber profile of the existing subscriber, a new scheduler is created and all the hosts belonging to that subscriber are linked to this new scheduler. Notice that the new subscriber profile will not conflict with the subscriber profile provisioned for a static host or non-sub-traffic under the same SAP.

If this is the first host of a subscriber on a particular SAP using a particular SLA profile, an SLA profile instance is generated and added to the SLA profile instance table. This includes instantiating a number of queues, according to the ingress and egress QoS profiles referred to in SLA profile, optionally with some specific overrides defined in the SLA profile. Otherwise the host is linked to the existing SLA profile instance for this subscriber on this SAP.

Notes:

- Any QoS and IP filter policies defined on the SAP are still processed even if Enhanced Subscriber Management is enabled on the SAP. For IPv4 traffic that is dropped due to anti-spoofing, counters, logging, and mirroring can be used. All other Layer 2 traffic that is never blocked by anti-spoofing can be processed by applying a QoS policy on the SAP and can still be classified differently, by the dot1p value.
- If insufficient hardware resources (queues) or software resources (profile instances) are available to support the new host, the DHCP ACK is dropped and an event is generated.

## Packet Processing for an Existing Host

Whenever an IP packet arrives on a subscriber-facing SAP on which Enhanced Subscriber Management is enabled, a lookup is done in the subscriber host table using as the index the SAP, source IP address, and source MAC address.

- If there is no entry, this means that the host is not using his assigned IP address, so the packet is dropped;
- If there is an entry, this will refer to the subscriber profile and SLA profile to be used.

# ESM Host Lockout

This feature is applicable to the 7750 SR and the 7450 ESS.

This feature increasingly penalizes hosts that fail repeated login attempts within a configurable time interval. This is done by holding off on creation attempts for these hosts for a configured but adaptable time period. A transient failure, due to a mis-configuration, is quickly corrected and does not prevent the host from logging in within a reasonable amount of time. At the same time, a malicious client or a constantly mis-configured client is locked-out and will not take up resources impacting other clients.

A lockout time per host supports exponential back-off with each retry and failure cycle, starting with a configured minimum value and increasing up to a configured maximum. The lockout time can be reset to the configured minimum value if there is no failed retry within a configured time threshold. The configurable values include:

**lockout-reset-time** *seconds*
**lockout-time** [**min** *seconds*] [**max** *seconds*]
**max-lockout-hosts** *hosts*

If multiple retries/failure cycles occur within the lockout time, then lockout period is exponentially increased starting from configured minimum value up to the configured maximum value. The lockout is reset to the minimum value if there is no failed retry till this lockout time.

This mechanism is supported for both single and dual-stack PPPoE and IPoE (DHCP) hosts over 1:1 or N:1 static or managed SAPs. The hold-off timer maintenance is on a per host basis (as follows):

- For 1:1 VLAN (PPPoE or IPoE hosts) per <VLAN, MAC address>
- For N:1 VLAN (PPPoE or IPv4oE hosts) per <VLAN, agent-circuit-id, agent-remote-id, MAC@>
- For 1:1 VLAN (IPv6oE hosts) per <VLAN, DUID>

A show lockout state for hosts is supported, given one or more of <SAP, MAC@, agent-circuit-id, agent-remote-id>.

A clear lockout state is supported for hosts given one or more of <SAP, MAC@, agent-circuit-id, agent-remote-id>.

Any changes in configured lockout values will not apply to hosts currently under lockout and will only apply once these hosts are out of lockout.

## Functionality

ESM lockout is supported for dual-stack PPPoE hosts, dual-stack IPoE hosts and ARP hosts. ESM Lockout will track the following:

- PPPoE PADI and PADR
- DHCPv4 discover, DHCPv4 request, DHCPv6 solicit, DHCPv6 request
- ARP Request

During lockout, authentication and ESM host creation is suppressed. A lockout context will be created when a client first enters lockout. The context maintains state and timeout parameters for the lockout. If a lockout policy is configured for the underlying SAP for a host that has failed authentication or host creation, the host enters lockout for the configured minimum time (1 — 86400 seconds). When the lockout time expires, normal authentication and ESM host creation will be resumed on relevant PPP or DHCP messages. In case of another failure, the host will again enter the lockout state. The lockout time for the host on each failure will be exponentially increased up to the configured maximum time (1 — 86400 seconds). The lockout time for a client will be reset to the configured minimum value, and the corresponding lockout context will be deleted, if there is no authentication (and host creation) failure within a configured amount of time that needs to elapse after the client initially enters lockout. This time is called the **lockout-reset-time**.

The host identification for lockout includes <SAP, MAC@, circuit ID, remote ID>.

# ANCP and GSMP



## ANCP

Access Node Control Protocol Management (ANCP) can provide the following information to the router:

- ANCP can communicate the current access line rate to the router. This allows the router to adjust the H-QoS subscriber scheduler with the correct rate or potentially change alarm when the rate goes below a set threshold. This allows a policy manager to change the entire policy when the rate drops below a minimal threshold value. The ANCP actual upstream synchronization rate is mapped to the ingress while ANCP actual downstream synchronization rate is mapped to the egress.
- The router can send DSL line OAM commands to complete an OAM test from a centralized point or when operational boundaries prevent direct access to the DSLAM.

When ANCP is used with Enhanced Subscriber Management (ESM), a new string ancp-string can be returned from the Python script or from RADIUS. If not returned it defaults to the subscriber ID.

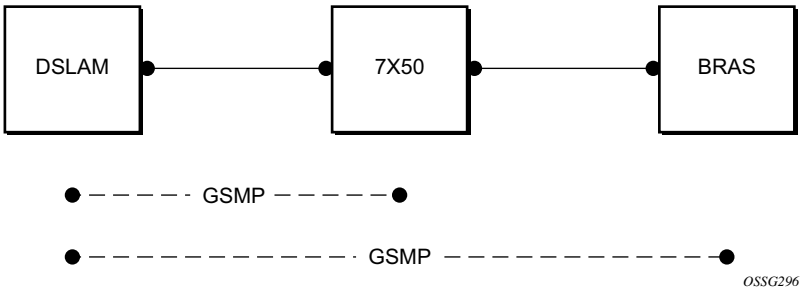
ANCP version 0x31 and 0x32 are both supported and will be auto detected at the start of each ANCP session. Within version 0x32, partitioning is also supported.

Multiple partitions from the same Access Node are also supported. If partitions are used, they are automatically detected during the start of an ANCP session.

### Static ANCP Management

As depicted in Figure 61, a DSLAM is connected to an aggregation network that is connecting the DSLAM to a BRAS. ANCP is used to provide SAP level rate management. The DSLAM in this application maintains multiple ANCP connections. The primary connection is to the BRAS, providing rate and OAM capabilities while the secondary is to the router to provide rate management.

7750 SR and 7450 ESS:




7710 SR:

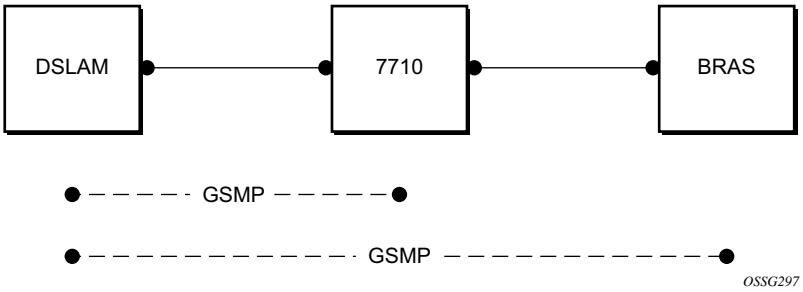



**Figure 61: Static ANCP Management Example**

## Enhanced Subscriber Management (ESM) Dynamic ANCP

In this application ANCP is used between the DSLAM and the router to provide line control. There are multiple attributes defined as described below. Figure 62 depicts the connectivity model.

This application is used to communicate the following from the DSLAM to the router (the policy control point):

- Subscriber rate
- OAM

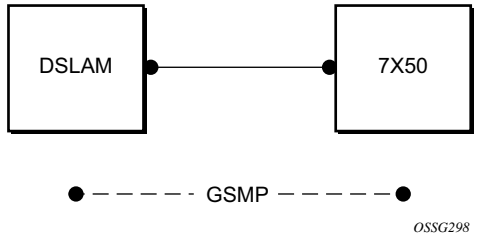



**Figure 62: ESM Dynamic ANCP Example**

## ANCP String

To support node communication with the access device the line rate, OAM commands, etc. the node can use an “ANCP string” that serves as a key in the out-of-band channel with the access node. The string can be either provisioned in the static case, retrieved from RADIUS or from the Python script.

## ANCP Persistency Support

Persistency is available for subscriber's ANCP attributes and is stored on the on-board compact flash card. ANCP data will stay persistence during an ISSU as well as nodal reboots. During recovery, ANCP attributes are first restored fully from the persistence file and incoming ANCP sessions are temporarily on hold. Afterwards new ANCP data can overwrite any existing values. This new data is then stored into the compact flash in preparation for the next event.

## General Switch Management Protocol Version 3 (GSMPv3)

General Switch Management Protocol version 3 (GSMPv3) is a generic protocol that allows a switch controller node to establish and maintain connections with one or more nodes to exchange operational information. Several extensions to GSMPv3 exist in the context of broadband aggregation. These extensions were proposed to allow GSMPv3 to be used in a broadband environment as additional information is needed to synchronize the control plane between access nodes (such as DSLAMs) and broadband network gateways (such as BRAS).

In the TPSDA framework, nodes fulfill some BRAS functionality, where per subscriber QoS enforcement is one of the most important aspects. To provide accurate per-subscriber QoS enforcement, the network element not only knows about the subscriber profile and its service level agreement but it is aware of the dynamic characteristics of the subscriber access circuit.

The most important parameters in this context are the subscriber-line capacity (DSL sync-rate) and the subscriber's channel viewership status (the actual number of BTV channels received by the given subscriber in any point in time). This information can be then used to adjust parameters of aggregate scheduling policy.

Besides, the above-mentioned information, GSMPv3 can convey OAM information between a switch controller and access switch. The node can operate in two roles:

- As the intermediate controller — The router terminates a connection from the DSLAM.
- As the terminating controller— The router fulfils full the roll of BRAS.

The DSL forum working documents recommends that a dedicated Layer 2 path (such as, a VLAN in an Ethernet aggregation network) is used for this communication to provide a certain level of security. The actual connection between DSLAM and BRAS is established at TCP level, and then individual messages are transported.

## DHCP Release Messages

The node supports DHCP release messages. A DHCP release message removes state from the DHCP server when the node rejects ACKs or removes hosts.

### DHCP Release

DHCP release messages will be controlled by the node and sent to the DHCP server to clear stale state. There are two examples:

1. If the node drops a DHCP ACK (because of resources, duplicate host or other reasons) the servers state must be cleared and the node will send a DHCP release.
2. When a host state is removed, based on SHCV, ANCP, user clear, etc., the node will send a DHCP release to the server and the MAC will be flushed from SDPs. A new flag will allow the user to elect not to send the release message If when using a clear lease command the host was removed by the user (using a clear command) a new flag will allow the user to elect not to send the release message.

## DHCP Client Mobility

Client mobility allows the node to use host monitoring (SHCV, ANCP, split DHCP) to remove network and server state when a host is removed locally. This allows for MAC addressed learned and pinned to move based on policy parameters.

Subscriber Host Connectivity Verification (SHCV) configuration is mandatory. This allows clients to move from one SAP to another SAP in the same service. This is only applicable in a VPLS service and group interfaces.

The first DHCP message on the new SAP with same MAC address (and IP address for group-interfaces) will trigger SHCV and will always be discarded.

SHCV will check that the host is no longer present on the SAP where the lease is currently populated to prevent spoofing. When SHCV detects that the host is not present on the original SAP, the lease-state will be removed. The next DHCP message on the new SAP can initiate the host.

## DHCP Lease Control

DHCP lease control allows the node to be configured to present a different lease to the client. This can be used to monitor the health of the client.

# Using Scripts for Dynamic Recognition of Subscribers

Whenever a host belonging to a subscriber is activated (when a PC or set-top box (STB) is turned on), the host will typically request an IP address from the network using DHCP. Refer to DHCP Management on page 335 for an explanation of DHCP and DHCP snooping in the router.

The DHCP ACK response from the DHCP server can be parsed and the contents of the message can be used to identify the "class" to which this host belongs, and thus, the QoS and security settings to apply.

The information necessary to select these settings can be codified in, the IP address by the DHCP server and/or the Option 82 string inserted by the DSLAM or other access node.

---

## Python Language and Programmable Subscriber Configuration Policy (PSCP)

PSCP is an identification mechanism using the Python scripting language. The PSCP references a Python script that can use regular expressions to derive the sub-ident-string, sub-profile-string and sla-profile-string from the DHCP response. A tutorial of regular expressions is beyond the scope of this guide, and can be found on the Internet (refer to http://www.amk.ca/python/howto/regex/).

A tutorial of Python is beyond the scope of this guide but can be found on the Internet (refer to http://www.python.org/).

Example scripts, using some regular expressions, can be found in Sample Python Scripts on page 1727. Additional information about the service manager scripting language, see Service Manager Scripting Language on page 1717.

One or more scripts can be written by the operator and stored centrally on a server (in a location accessible by the router). They are loaded into each router at bootup.

Note that if a centrally stored script is changed, it is not automatically re-loaded onto the router. The reload must be forced by executing the **shutdown** / **no shutdown** commands on the affected URL(s).

## Determining the Subscriber Profile and SLA Profile of a Host

Figure 63 describes the data flow while determining which subscriber profile and SLA profile to use for a certain subscriber host based on a snooped/relayed DHCP ACK for that subscriber host.

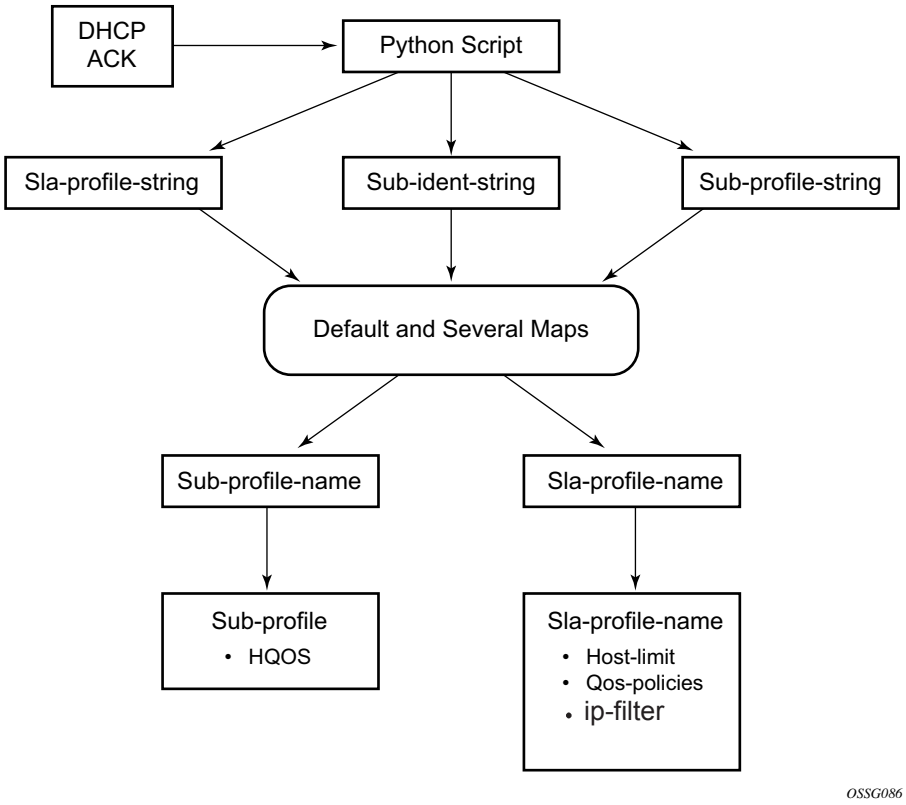



**Figure 63: Data Flow in Determining Subscriber Profile and SLA Profile**

An incoming DHCP ACK (relayed or snooped) is processed by the script provisioned in the sub-ident-policy defined in the SAP on which the message arrived. This script outputs one or more of the following strings:

- sub-ident — Identifies the subscriber (always needed).
- sub-profile — Identifies the subscriber class (optional).
- sla-profile — Identifies the SLA Profile for this subscriber host (optional).

These strings are used for a lookup in one or more maps to find the names of the sub-profile and sla-profile to use. If none of the maps contained an entry for these strings, the names will be determined based on a set of defaults.

Only when the names for both the sub-profile and sla-profile are known, the subscriber host can be instantiated. If even no default is found for either profile, the DHCP ACK is dropped and the host will not gain network access.

## Determining the Subscriber Profile

All hosts (devices) belonging to the same subscriber will be subject to the same HQoS processing. The HQoS processing is defined in the sub-profile. A sub-profile refers to an existing scheduler policy and offers the possibility to overrule the rate of individual schedulers within this policy.

Because all subscriber hosts of one subscriber use the same scheduler policy instance, they must all reside on the same I/O module.

The figure below shows how the sub-profile is derived, based on the sub-ident string, the sub-profile string and/or the provisioned data structures. The numbers associated with the arrows pointing toward the subscriber profiles indicate the precedence of the checks.

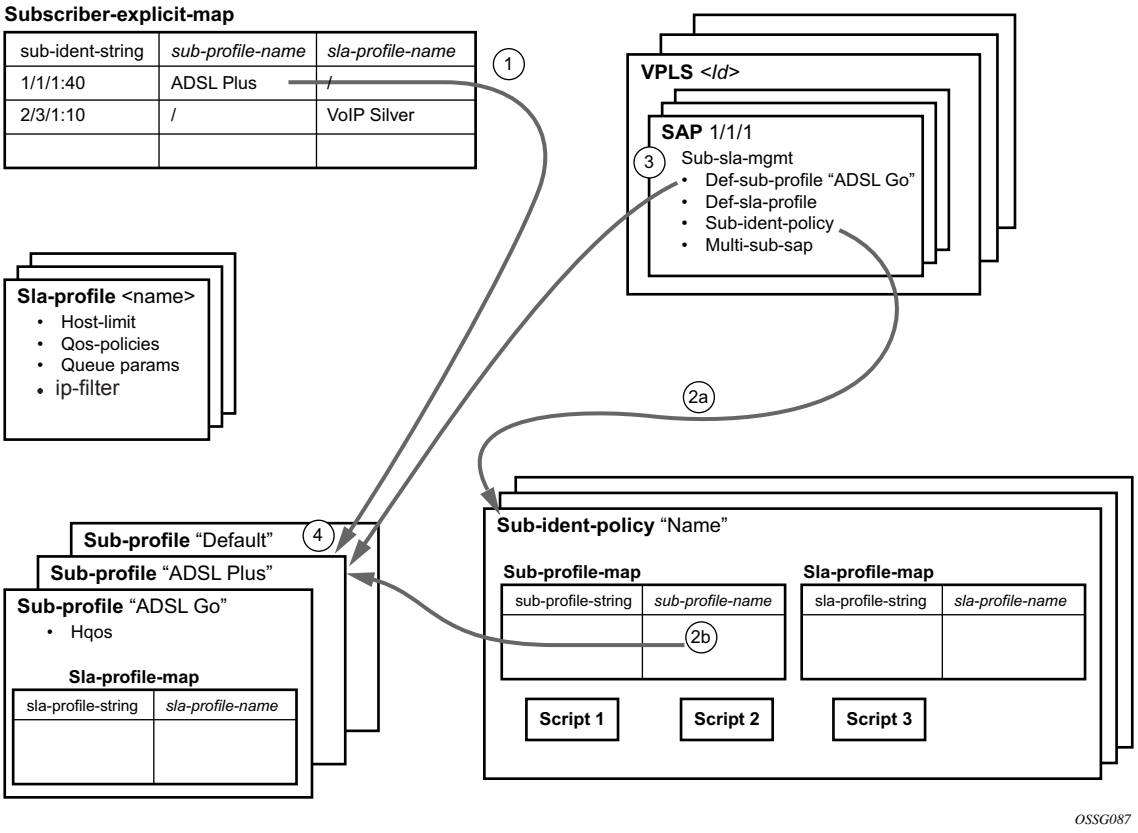



**Figure 64: 7750 SR Determining the Subscriber Profile**

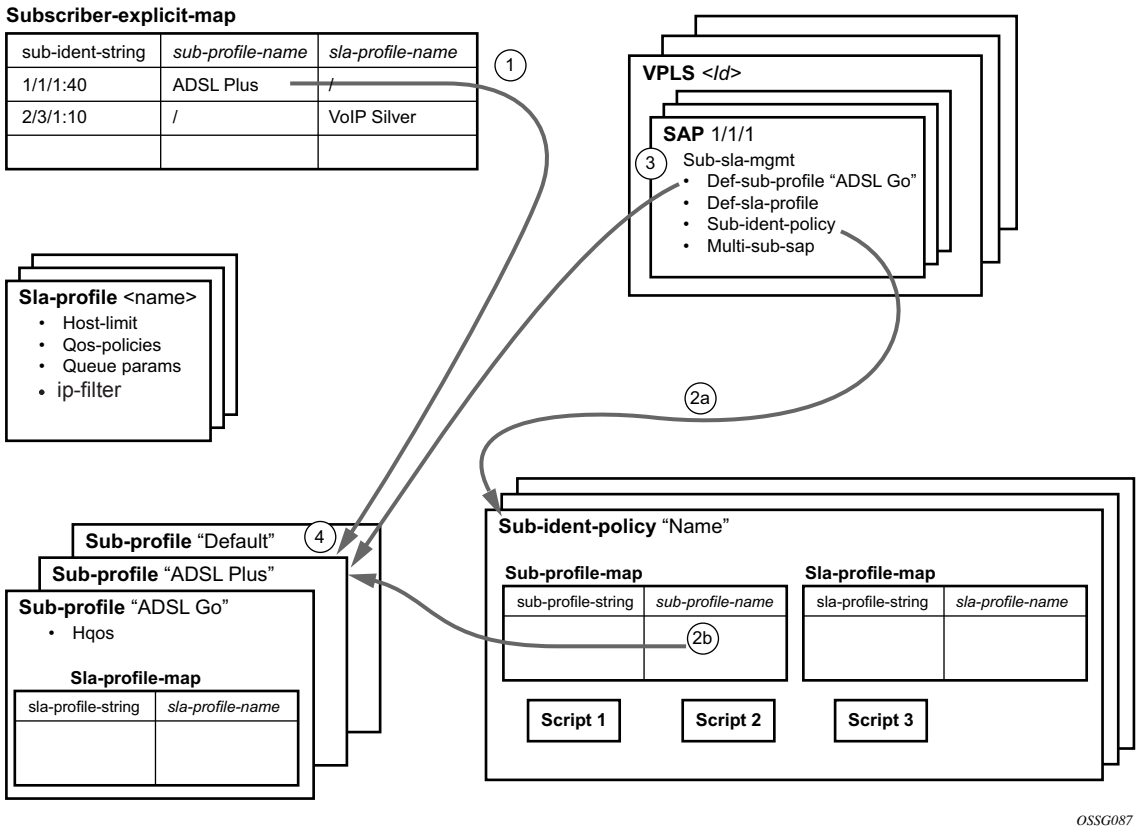



**Figure 65: 7450 ESS Determining the Subscriber Profile**

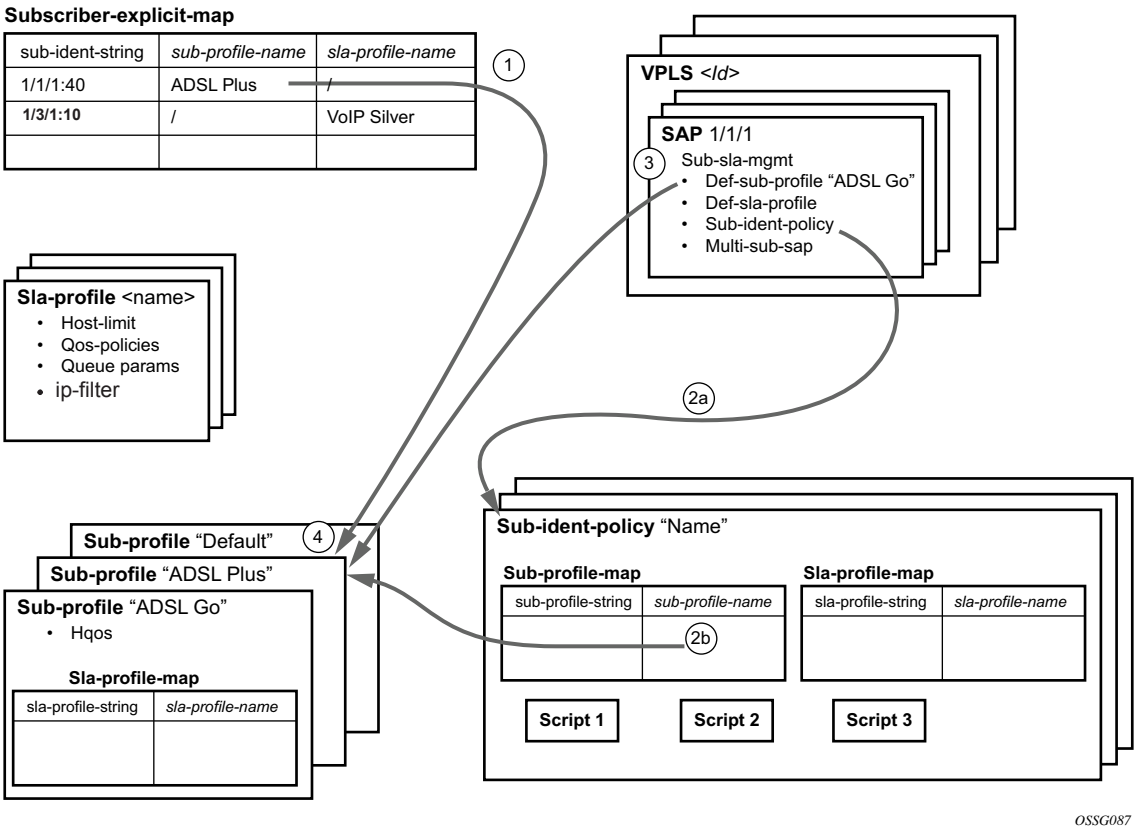



**Figure 66: 7710 SR Determining the Subscriber Profile**

1. A lookup in the **explicit-subscriber-map** is done with the sub-ident string returned by the script. If a matching entry is found, the sub-profile-name (if defined) is taken. Otherwise:
2. If a **sub-ident-policy** is defined on the SAP, a lookup is done on its **sub-profile-map** with the sub-profile string from the script. The sub-profile-name is taken from the entry. If no entry was found, then:
3. If provisioned, the sub-profile-name is taken from the **def-sub-profile** attribute on the SAP. If not provisioned, then:
4. The **sub-profile** with the name "default" is selected (if provisioned). If this is not provisioned, there are no other alternatives, the ACK is dropped, and the host will not gain access.

## Determining the SLA Profile

For each host that comes on-line, the router also needs to determine which SLA profile to use. The SLA profile will determine for this host:

- The QoS-policies to use:
  - → classification
  - → queues
  - → queue mapping
- The IP filter to use.

The SLA profile also has a host-limit attribute which limits the total number of hosts (belonging to the same subscriber) on a certain SAP that can be using this SLA profile.

The classification and the queue mapping are shared by all the hosts on the same forwarding complex that use the same QoS policy (by their SLA profile).

The queues are shared by all the hosts (of the same subscriber) on the same SAP that are using the same SLA profile. In other words, queues are instantiated when, on a given SAP, a host of a subscriber is the first to use a certain SLA profile. This instantiation is referred to as an SLA profile instance.

The figure below shows a graphical description of how the SLA profile is derived based on the subscriber identification string, the SLA profile string and the provisioned data structures. The numbers on the arrows towards the SLA profile indicate the “priority” of the provisioning (the lower number means the higher priority).

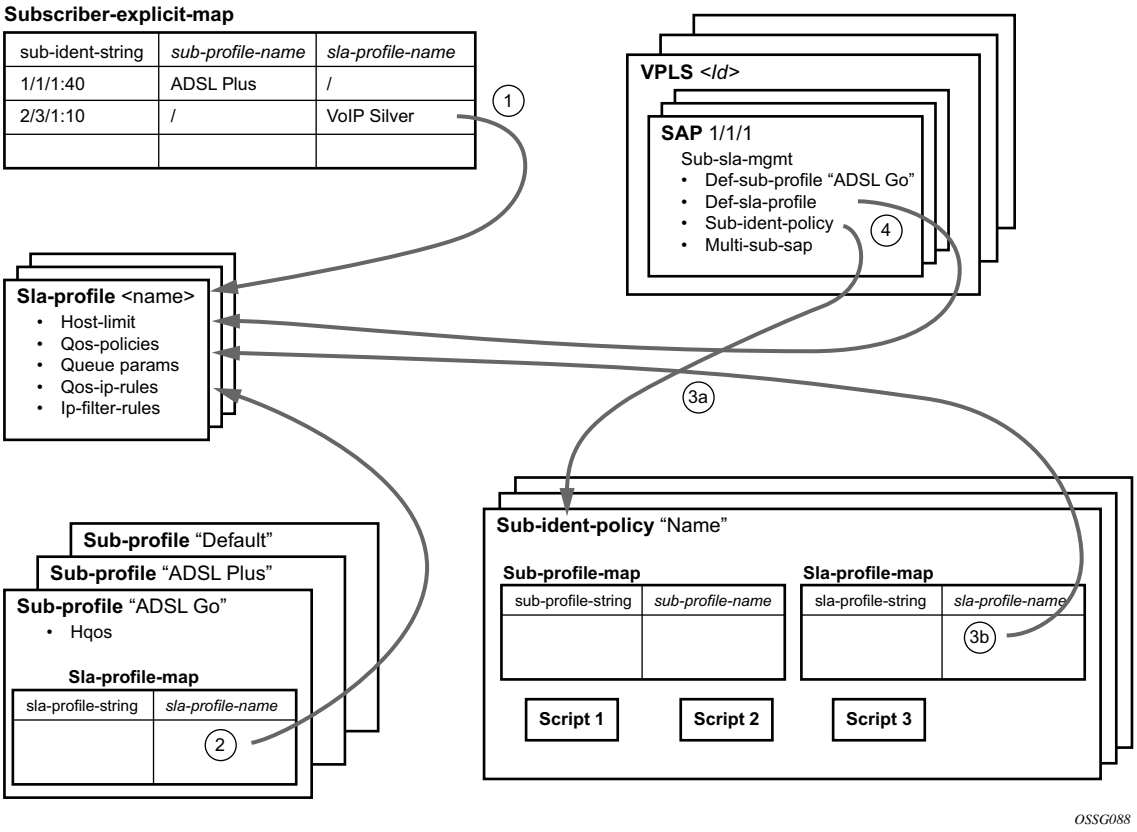



**Figure 67: 7750 SR Determining the SLA Profile**

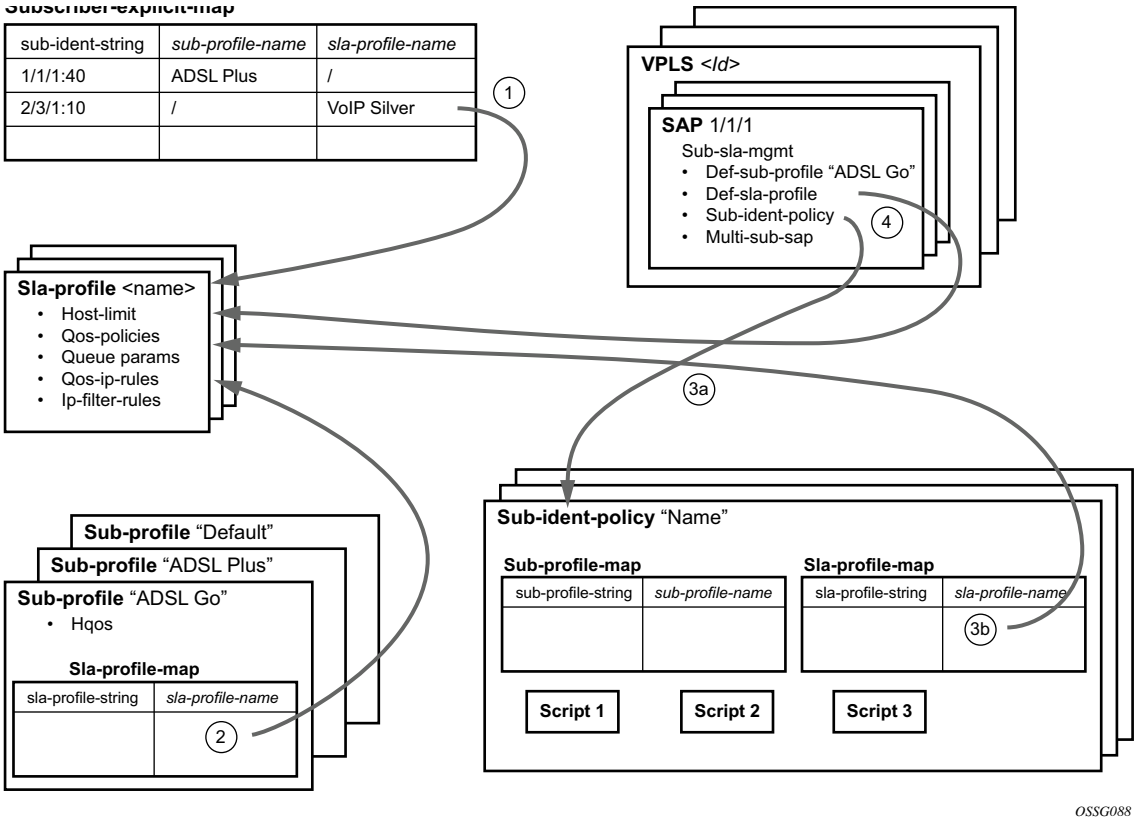



**Figure 68: 7450 ESS Determining the SLA Profile**

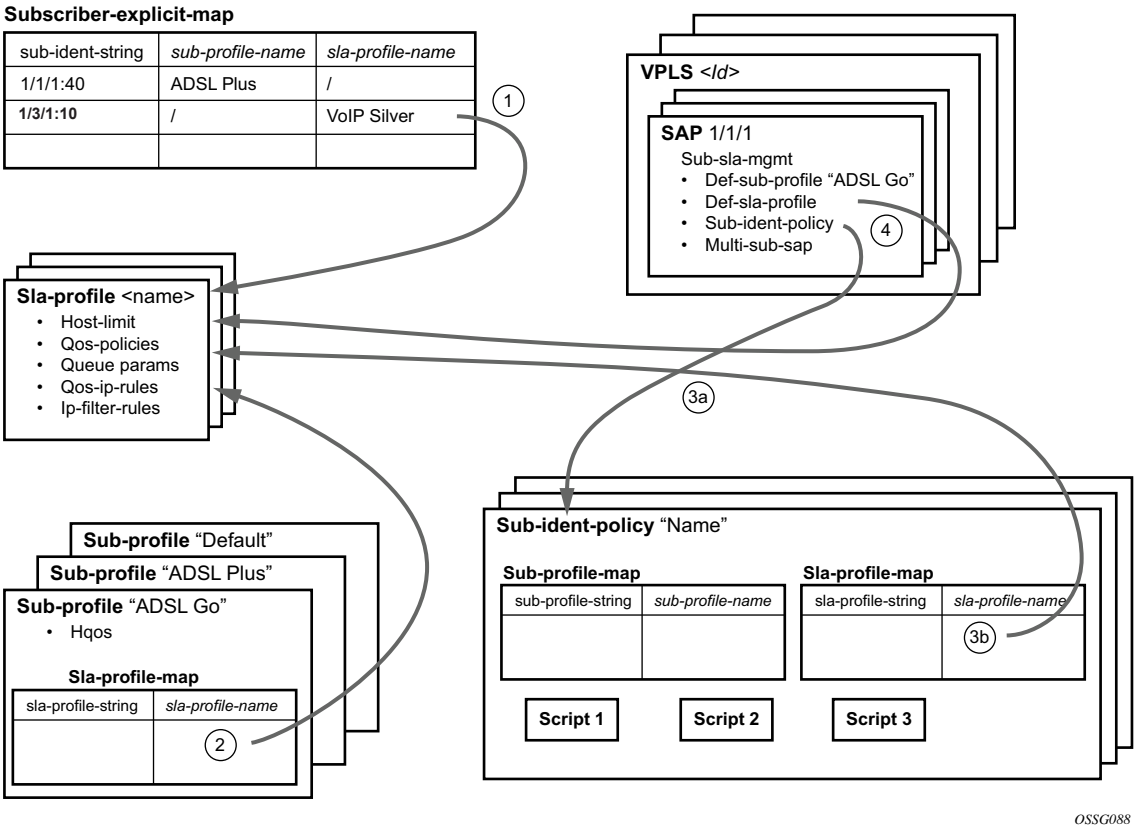



**Figure 69: 7710 SR Determining the SLA Profile**

1. A lookup is done with the sub-ident string returned by the script in the **explicit-subscriber-map**. If a matching entry is found, the sla-profile-name is taken from it – if defined. Otherwise:
2. A lookup with the sla-profile string from the script is done in the **sla-profile-map** of the sub-profile found earlier. The sla-profile-name from the found entry is taken. If no entry was found, then:
3. A lookup is done with the sla-profile string in the **sla-profile-map** of the **sub-ident-policy** configured on the SAP. The sla-profile-name from the found entry is taken. If no **sub-ident-policy** was configured on the SAP or no entry was found, then:
4. If provisioned, the sla-profile-name is taken from the **def-sla-profile** attribute on the SAP. If not provisioned, there are no more alternatives, the ACK is dropped, and the host will not gain access.

## SLA-Based Egress QoS Marking

The egress QoS marking for subscriber-host traffic is derived from SAP-egress QoS policy associated with a corresponding SAP, rather than from the SLA profile associated with the corresponding subscriber-host. As a consequence, no egress QoS marking (or Dot1p marking is set to 0, the dscp/prec field is kept unchanged) is performed for traffic transmitted on a managed-SAP because per default, sap-egress policy 1 is attached to every managed-SAP.

The default value of the "qos-marking-from-sap" flag is enabled. This means that the qos-marking defined in the SAP egress QoS policy associated with the SAP will be used. The default setting of this flag in a combination with managed-SAP will result in the same behavior as in the current system (dot1p=0, dscp/prec is unchanged).

If "no qos-marking-from-sap" is executed, then both the Dot1p marking (all IOMs) and DSCP marking (IOM2/3 only) are derived from the sla-profile.

Changing the flag setting in the SLA profile being used by any subscriber-hosts (this includes subscriber-hosts on managed-SAPs as well) will be allowed.

The following MC traffic characteristics apply:

- On Layer 3 subscriber-interfaces, MC is not supported so it is impossible to enable it at the SAP level or at the sla-instance level.
- On Layer 2 SAPs IGMP snooping is supported while it is not supported on the sla-instance level. Therefore, any MC traffic transmitted at egress belongs to a SAP (meaning it will use SAP queues), rather than to sla-instance.
- The special case are SAPs with a profiled-traffic-only flag enabled. Although it is possible to define an sla-profile applicable to a Layer 2-host, this will not be taken as reference for marking mc-traffic, but rather SAP settings will be used.

# Auto-Sub ID

The subscriber ID name (sub-id) is a mandatory object that binds all hosts of a given subscriber together. Briefly, the sub-id name represents a residential household. Many management/ troubleshooting and even billing operations rely on the sub-id name entity. The sub-id name is required for the host creation process, and it can be supplied by RADIUS or LUDB. It is derived from the sap-id or is statically provisioned in the form of a string.

In many ESM deployments with RADIUS, it is desirable that the sub-id is auto-generated within the 7x50 rather than burdening the OSS and the RADIUS server with this function. A typical application for auto sub-id is as follows:

- RADIUS server provides the sla-profile string and the sub-profile string but not the sub-id string.
- The sub-id name is auto-generated and formatted based on the configured options.

The following are the properties of auto sub-id generation:

- The auto-generation of the sub-id name can be based on any combination of the following fields:
  - → MAC address
  - → sap-id
  - → circuit-id
  - → remote-id
  - → session-id

There can be only a single set of subscriber identification fields defined per host type (IPoE or PPPoE) per chassis. If the combination of the fields must be modified, the existing subscribers with an auto-generated sub-id must be manually terminated. Considering that remote termination of the IPoE subscribers by a DHCP server is not supported by all DHCP client vendors through the FORCERENEW DHCP message (RFC 3203, *DHCP reconfigure extension*), changing the subscriber fields while subscribers with auto generated sub-id are active should be avoided.

The sub-id name generation will take place at the end of the host initiation process (as after the authentication phase is completed) and only in case whereby the sub-id had not been already provided by any other more specific means (RADIUS, LUDB). This means that if the sub-id is supplied by other means (RADIUS, LUDB), then the sub-id name will not be auto-generated.

The format of the sub-id name can be either a random 10 characters encoded string or a user-friendly string based on the subscriber identification fields. Note that the maximum length of the sub-id name is 32 characters.

The sub-id name will not be passed in the Access-Request to the RADIUS server since it is generated after the authentication phase.

The sub-id name can be auto-generated regardless of how the sla/sub-profile strings are obtained (RADIUS, LUDB or static).

The subscriber identification fields used in auto-generation of the sub-id name are enabled on the global level.

**CLI Syntax:**
```
configure
        subscriber-mgmt
           auto-sub-id-key
              ppp-sub-id-key [mac] [sap-id] [circuit-id] [remote-
              id] [session-id]
              ipoe-sub-id-key [mac] [sap-id] [circuit-id] [remote-
              id]
```

If no sub-id-key per host type is configured, then the defaults are:

PPPoE host type: <mac, sap-id, session-id>

IPoE host type : <mac, sap-id>.

The order in which the fields are configured is important because the sub-id name will potentially become a concatenated string of the subscriber host identifiers in the order in which they are provisioned. Note that the sub-id cannot be longer than 32 characters.

- In case that the length of the concatenated fields for the sub-id name is larger than 32 characters, the host creation will fail.
- In case that the circuit-id/remote-id is in the key and they contain non-printable characters, their place in sub-id name will be formatted in hex instead of ASCII. ASCII printable characters contain byte values 0x20..0x7E. All other values are ASCII non-printable and thus are formatted in hex characters.

The following would generate a sub-id name: xx:xx:xx:xx:xx:xx|1/1/3:23|44. The length of such sub-id name would be 29B.

- mac: xx:xx:xx:xx:xx:xx
- sap: 1/1/3:23
- session-id: 44 (16bits length)

In case that the key contains the circuit-id as: 0x610163 (3 bytes), then the sub-id name will be formatted as ‘610161’ (hex) since ‘01’ hex is non printable in ASCII. In this case the sub-id name will be of length 6B.

However, if the circuit-id is 0x616263 (3 bytes), then the string will be formatted as ASCII string ‘abc’ (3 characters). The sub-id name is 3B long.

The assignment of the sub-id to dynamic hosts is as follows:

- From RADIUS (sub-ident-policy including use-direct-map-as-default)
- From LUDB (sub-ident-policy including use-direct-map-as-default)

- Configured (explicit) defaults:
  - → use-sap-id: sap-id
  - → auto-id: combination of sub-id identifiers specified in auto-sub-id-key. The sub-id name will be in a human friendly format, i.e. concatenation of the fields in the pppoe|ipoe-sub-id-key command separated by a "|" character.
  - → string: custom string
- Non-configured (implicit) defaults:

PPPoE host types: random 10 character string based on fields defined in the

- → **ppp-sub-id-key** command. If no such fields are explicitly defined, the default ones will be assumed: <mac, sap-id, session-id>.
- → IPoE host types: random 10 character string based on the **ipoe-sub-id-key** command. If no such fields are explicitly defined, the defaults will be assumed: <mac, sap-id >.

The way in which the default sub-id is generated is configured under the SAP level in the following manner:

**CLI Syntax:**
```
configure
service ies/vprn
   subscriber-interface <sub-if-name>
        group-interface <grp-if-name>
           sap <sap-id>
              sub-sla-mgmt
                 def-sub-id use-sap-id|use-auto-id|string
```

Under the msap-policy:

**CLI Syntax:**
```
configure
subscriber-mgmt
   msap-policy <name> (msap-policy referenced in msap-de-
      faults under the capture sap)
      sub-sla-mgmt
         def-sub-id <use-sap-id|use-auto-id|string <sub-
         id>
```

The **use-auto-id** keyword parameter of the def-sub-id string consists of concatenated auto-sub-id-keys separated by a '|' character. In the absence of the **use-auto-id** keyword, the sub-id name will be a random 10 characters encoded string based on the ipoe|ppp-sub-id-keys. This random encoded 10 character string is unique per chassis as well as in dual-homed environment.

This command will have no effect if it is configured directly under the capture SAPs in VPLS (in the **config>service>vpls>sap>sub-sla-mgmt** context). Managed SAPs in ESM are instantiated by a capture SAP and the msap-policy in this case is mandatory. An auto-id keyword in case of managed SAP will be looked only under the msap-policy.

Static subscribers are required to have the sub-id manually configured.

## Sub-id Identifiers

The sub-id can be based on any combination of the following identifiers:

- The sap-id, in combination with any other allowable identifier, will be used as the search key. This assumes a 1:1 (subscriber per SAP) deployment model.
- The circuit-id, in combination with any other allowable identifier, will be used to identify subscribers. This can be used in 1:1 deployment model, or in service per SAP deployment model. Circuit-id is applicable to IPoE v4 type hosts (option 82), to IPoE v6 type hosts (option 18 – interface-id) and PPPoE hosts (remote agent option signaled by PPPoE tags). The format of circuit-id is identical for IPv4 and IPv6 hosts.
- The remote-id, in combination with any other allowable identifier, will be used to identify subscribers. This can be used in 1:1 deployment model, or in service per SAP deployment model. The remote-id is applicable to IPoE v4 type hosts (option 82), to IPoE v6 type hosts (option 37) and PPPoE hosts (remote agent option signaled via PPPoE tags).
- The mac address (in combination with any other allowable identifier will be used to identify subscribers. This assumes a 1:1 deployment model.
- The PPPoE session id, in combination with any other allowable identifier, is applicable only to PPPoE hosts. The session-id used will be of the first host that is instantiated for the subscriber.

## Dual Stack Hosts

Autogeneration of sub-id names for subscribers with a single dual stack hosts (IPoE and PPPoE) is enabled by default by not explicitly provisioning anything for the def-sub-id. The sub-id name would be semi-randomly generated based on the <mac, sap-id, session-id> for PPPoE hosts and the <mac, sap-id> combination for IPoE host.

## Mixing Hosts with Auto-Generated IDs and non Auto-Generated IDs

Hosts with different sub-id names but identical auto-sub-id keys are not linked into the same subscriber. Such scenarios can arise with hosts with the same auto-sub-id keys but different methods for obtaining the sub-id name. For example, one host relying on auto-generated sub-id name while the other is using explicit configuration methods (sap-id, string, RADIUS or LUDB). If the auto-generated sub-id name and explicit sub-id name are the same, the host will be tied into the same subscriber.

For example:

The default auto-sub-id for the following two hosts are <mac, sap-id>.

Host X on SAP 1/1/1:1 with MAC 00:00:00:00:00:01 obtains sub-id through RADIUS.

Host Y on SAP 1/1/1:1 with MAC 00:00:00:00:00:01 has sub-id auto-generated.

Regardless of which host comes up first, those two hosts at the end will belong two different subscribers as long as their sub-ids are different.

## PPPoA/PPPoEoA Considerations

PPPoA/PPPoEoA hosts will adhere to the same rules as PPPoE. Fields that are supported in PPPoE but not PPPoA/PPPoEoA will be simply ignored.

Fields that are not supported in PPPoA are:

- Remote-id
- Circuit-id
- MAC address

Fields that are most likely not applicable to PPPoEoA are:

- Remote-id
- Circuit-id

## Deployment Considerations

The following is a possible deployment example scenario.

**CLI Syntax:**
```
configure
subscriber-mgmt
   auto-sub-id-key
      ppp-sub-id-key sap-id
      ipoe-sub-id-key mac circuit-id
```

**CLI Syntax:**
```
configure
service vprn 10
   subscriber-interface <sub-if-name>
   authentication-policy <auth-pol-name>
   group-interface <grp-if-name>
      sap 1
         sub-sla-mgmt
            def-sub-id use-sap-id
            sub-ident-policy <ident-pol-name>
      sap 2
         sub-sla-mgmt
```

```
            def-sub-id auto-id
            sub-ident-policy <ident-pol-name>
      sap 3
         sub-sla-mgmt
            def-sub-id "sub3"
            sub-ident-policy <ident-pol-name>
      sap 4
         sub-sla-mgmt
               sub-ident-policy <ident-pol-name>
```

Assume the following cases:

1. RADIUS returns the sub-id on all four SAPs.
2. RADIUS does not return the sub-id string on any of the SAPs.

In the first case where RADIUS returns the sub-id string, the following will occur:

- On all 4 SAPs the sub-id string will be assigned by the RADIUS server. Defaults have no effect, and neither do identifiers specified under the auto-sub-id-key node.

In the second case, the effects are the following:

- On SAP1 the sub-id name will be the <sap-id> (1/1/1:3)
- On SAP 2 the sub-id name will be <sap-id> for PPPoE hosts and <mac>-<circuit-id> concatenation for IPoE type hosts.

Example:

1/1/1:100 for PPPoE

AC:AB:AA:AD:AE:AE-AN-id eth 1/1/1/1:2 for IPoE

(circuit-ID format is: Access-Node-Identifier atm slot/port:vpi.vci or Access-Node-Identifier eth slot/port:[vlan-id]).

Note that the circuit-ID can itself be 63B in length whereas the length of the sub-id name is limited to 32 Bytes. So in the above case, the sub-id name length would be 38 Bytes (>32B) and the host instantiation would fail.

- On SAP3 the sub-id name will be the literal ‘sub3’ for PPPoE and IPoE hosts.
- On SAP4 the sub-id name will be a semi-random value based on <sap-id> for PPPoE hosts and the <mac, circuit-id> combination for IPoE hosts.

## Caveats

Only a single combination of the subscriber fields used to auto generate sub-id is allowed per host type (IPoE or PPPoE) and per chassis. In case that the combination of the fields needs to be changed, the existing subscribers with an auto-generated sub-id must be manually terminated. Considering that remote termination of the IPoE subscribers by DHCP server is not supported by all DHCP client vendors through FORCERENEW DHCP message (RFC 3203), changing the subscriber fields while subscribers with auto generated sub-id are active should be avoided.

## Limiting Subscribers and Hosts on a SAP

A number of configuration parameters are available to control the maximum amount of subscribers and/or hosts that can be simultaneously active on a SAP:

- multi-sub-sap — Limits the number of subscribers (dynamic + static) on a SAP
- lease-populate — Limits the number of dynamic hosts on a SAP
- host-limit — Limits the number of hosts (dynamic + static) per SLA profile instance.

If any of these limits are reached, a new host will be denied access and the DHCP ACK will be dropped. The only exception is when **host-limit** command is configured with the keyword **remove-oldest** specified, then the oldest active host is dropped and the new host is granted access. The dynamic host with the least remaining lease time will be considered the oldest host.

---

## Static Subscriber Hosts

While it is typically preferred to have all hosts provisioned dynamically through DHCP snooping, it may be needed to provide static access for specific hosts (those that do not support DHCP).

Since a subscriber identification policy is not applicable to static subscriber hosts, the subscriber identification string, subscriber profile and SLA profile must be explicitly defined with the host's IP address and MAC address (if Enhanced Subscriber Management is enabled).

If an SLA profile instance associated with the named SLA profile already exists on the SAP for the subscriber, the static subscriber host is placed into that SLA profile instance. If an SLA profile instance does not yet exist, one will be created if possible. If the SLA profile cannot be created, or the host cannot be placed in the existing SLA profile instance (the **host-limit** was exceeded), the static host definition will fail.

# QoS for Subscribers and Hosts

## QoS Parameters in Different Profiles

QoS aspects for subscribers and hosts can be defined statically on a SAP or dynamically using

Enhanced Subscriber Management. For example, in a VLAN-per-service model, different services belonging to a single subscriber are split over different SAPs, and thus the overall QoS (such as a scheduler policy) of this subscriber must be assigned using Enhanced Subscriber Management.

QoS parameters are shared among the subscriber profile and SLA profile as follows:

- The subscriber profile refers to HQoS ingress and egress scheduler policies which define the overall treatment for hosts of this subscriber.
- The SLA profile refers to specific queue/policer settings for each host (BTV, VoIP, PC).
- The subscriber profile also refers to CFHP ingress and egress policer-control-policies which define the overall treatment for hosts of this subscriber.

The primary use of the subscriber profile is to define the ingress and egress scheduler policies/ policer-control-policies used to govern the aggregate SLA for all hosts associated with a subscriber. To be effective, the queues/policers defined in the SLA profile's QoS policies will reference a scheduler/arbiter from the scheduler policy/policer-control-policy respectively as their parent.

## QoS Policy Overrides

Generic QoS queue/policer parameters could be specified for the SAP in a QoS policy and overridden for some customers by queue/policer parameters defined in the SLA profile. This allows for a single SAP ingress and SAP egress QoS policy to be used for many subscribers, while providing individual subscriber parameters for queue/policer operation.

## ATM/Ethernet Last-Mile Aware QoS for Broadband Network Gateway

This feature allows the user to perform hierarchical scheduling of subscriber host packets such that the packet encapsulation overhead and ATM bandwidth expansion (when applicable) due to the last mile for each type of broadband session, that is, PPPoEoA LLC/SNAP and VC-Mux, IPoE, IPoEoA LLC/SNAP and VC-Mux, etc., is accounted for by the 7x50 acting as the Broadband Network Gateway (BNG).

The intent is that the BNG distributes bandwidth among the subscriber host sessions fairly by accounting for the encapsulation overhead and bandwidth expansion of the last mile such that packets are less likely to be dropped downstream in the DSLAM DSL port.

The last mile encapsulation type can be configured by the user or signaled using the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoE Tags or DHCP Relay Options as per RFC 4679.

Furthermore, this feature allows the BNG to shape the aggregate rate of each subscriber and the aggregate rate of all subscribers destined to a given DSLAM to prevent congestion of the DSLAM. The subscriber aggregate rate is adjusted for the last mile overhead. The shaping to the aggregate rate of all subscribers of a given destination DSLAM is achieved via a new scheduling object, referred to as Virtual Port or vport in CLI, which represents the DSLAM aggregation node in the BNG scheduling hierarchy

### Broadband Network Gateway Application

An application of this feature in a BNG is shown in Figure 70.

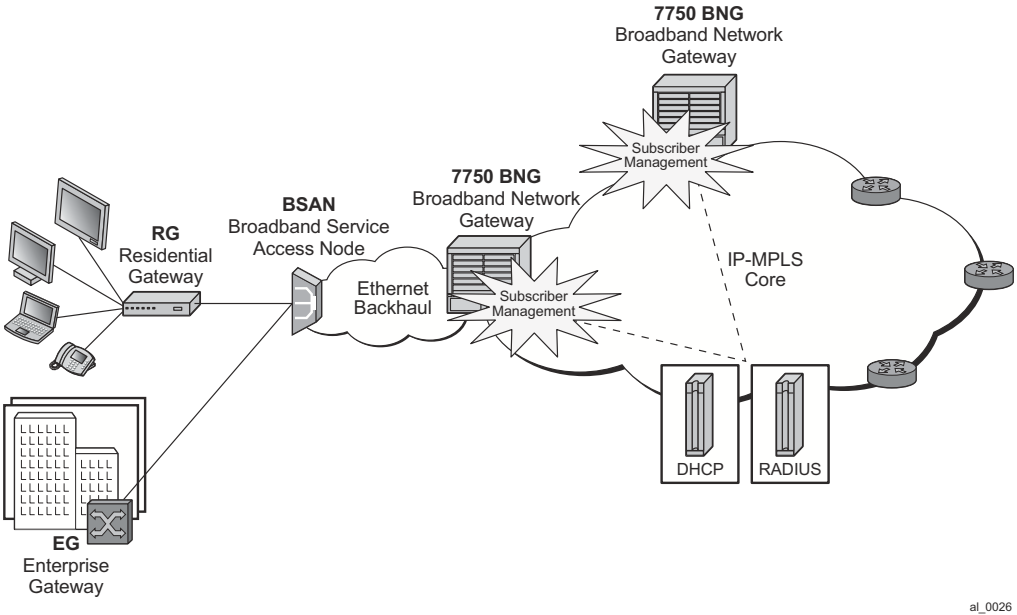



**Figure 70: BNG Application**

Residential and business subscribers use PPPoEoA, PPoA, IPoA, or IPoEoA based session over ATM/DSL lines. Each subscriber host can use a different type of session. Although Figure 1 illustrates ATM/DSL as the subscriber last mile, this feature supports both ATM and Ethernet in the last mile.

A subscriber SAP is auto-configured via DHCP or RADIUS authentication process, or is statically configured, and uses a Q-in-Q SAP with the inner C-VLAN identifying the subscriber while the outer S-VLAN identifies the Broadband Service Access Node (BSAN) which services the subscriber, i.e., the DSLAM. The SAP configuration is triggered by the first successfully validated subscriber host requesting a session. Within each subscriber SAP, there can be one or more hosts using any of the above session types. The subscriber SAP terminates on an IES or VPRN service on the BNG. It can also terminate on a VPLS instance.

When the 7750 BNG forwards IP packets from the IP-MPLS core network downstream towards the Residential Gateway (RG) or the Enterprise Gateway (EG), it adds the required PPP and Ethernet headers, including the SAP encapsulation with C-VLAN/S-VLAN. When the BSAN node receives the packet, it strips the S-VLAN tag, strips or overwrites the C-VLAN tag, and adds padding to minimum Ethernet size if required. It also adds the LLC/SNAP or VC-mux headers plus the fixed AAL5 trailer and variable AAL5 padding (to next multiple of 48 bytes) and then segments the resulting PDU into ATM cells when the last mile is ATM/DSL. Thus the packet size will undergo a fixed offset due to the encapsulation change and a variable expansion due to the AAL5 padding when applicable. Each type of subscriber host session will require a different amount of fixed offset and may require a per packet variable expansion depending of the encapsulation used by the session. The BNG node learns the encapsulation type of each subscriber host session by inspecting the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoE Tags as specified in RFC 4679. The BNG node must account for this overhead when shaping packets destined to subscriber.

## Queue Determination and Scheduling

Figure 71 illustrates the queuing and scheduling model for a BNG using the Ethernet/ATM last-mile aware QoS feature.

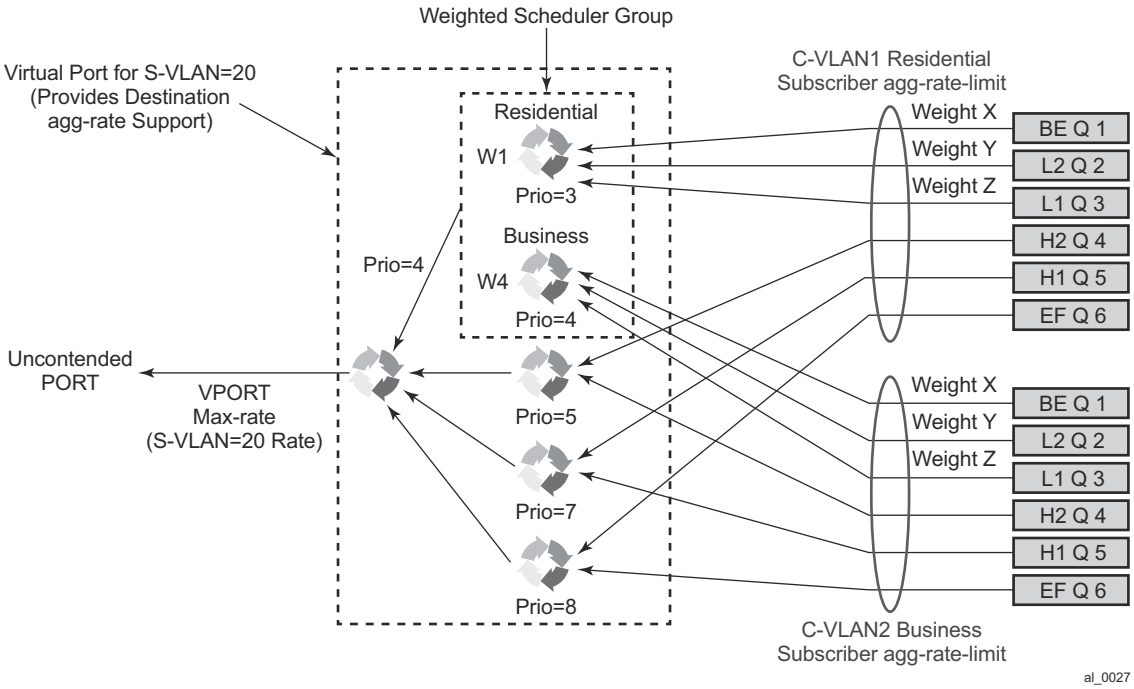



**Figure 71: BNG Queuing and Scheduling Model**

**CLI Syntax:** A set of per FC queues are applied to each subscriber host context to enforce the packet rate within each FC in the host session as specified in the subscriber's host SLA profile. A packet is stored in the queue corresponding the packet's FC as per the mapping of forwarding class to queue-id defined in the sap-egress QoS policy used by the host SLA profile. In the BNG application however, the host per FC queue packet rate is overridden by the rate provided in the RADIUS access-accept message. This rate represents the ATM rate that will be seen on the last mile, that is, it includes the encapsulation offset and the per packet expansion due to ATM segmentation into cells at the BSAN.

In order to enforce the aggregate rate of each destination BSAN, a scheduling node, referred to as virtual port, and vport is in the CLI. The vport operates exactly like a port scheduler with the difference that multiple vport objects can be configured on the egress context of an Ethernet port. The user adds a vport to an Ethernet port using the following command:

**CLI Syntax: configure>port>ethernet>access>egress>vport** *vport-name* **create**

The vport is always configured at the port level even when a port is a member of a LAG. The vport name is local to the port it is applied to but must be the same for all member ports of a LAG. It however does not need to be unique globally on a chassis.

**CLI Syntax: configure>port>ethernet>access>egress>vport** *vport-name* **create**

The vport is always configured at the port level even when a port is a member of a LAG. The vport name is local to the port it is applied to but must be the same for all member ports of a LAG. It however does not need to be unique globally on a chassis.

The user applies a port scheduler policy to a vport using the following command:

**CLI Syntax: configure>port>ethernet>acess>egress>vport>port-scheduler-policy** *port-scheduler-policy-name*

A vport cannot be parented to the port scheduler when it is using a port scheduler policy itself. It is thus important the user ensures that the sum of the **max-rate** parameter value in the port scheduler policies of all vport instances on a given egress Ethernet port does not oversubscribe the port's hardware rate. If it does, the scheduling behavior degenerates to that of the H/W scheduler on that port. A vport which uses an agg-rate-limit can be parented to a port scheduler. This is explained in Section Applying Aggregate Rate Limit to a VPORT.

Each subscriber host queue is port parented to the vport which corresponds to the destination BSAN using the existing port-parent command:

**CLI Syntax: configure>qos>sap-egress>queue>port-parent** [**weight** *weight*] [**level** *level*] [**cir-weight** *cir-weight*] [**cir-level** *cir-level*]

This command can parent the queue to either a port or to a vport. These operations are mutually exclusive in CLI as explained above. When parenting to a vport, the parent vport for a subscriber host queue is not explicitly indicated in the above command. It is determined indirectly. The determination of the parent vport for a given subscriber host queue is described in VPORT Determination and Evaluation on page 936.

Furthermore, the weight (cir-weight) of a queue is normalized to the sum of the weights (cir-weights) of all active subscriber host queues port-parented at the same priority level of the vport or the port scheduler policy. Since packets of ESM subscriber host queues are sprayed among the link of a LAG port based on the subscriber-id, it is required that all subscribers host queues mapping to the same vport, i.e., having the same destination BSAN, be on the same LAG link so that the aggregate rate towards the BSAN is enforced. The only way of achieving this is to operate the LAG port in active/standby mode with a single active link and a single standby link.

The aggregate rate of each subscriber must also be enforced. The user achieves this by applying the existing agg-rate-limit command to the egress context of the subscriber profile:

**CLI Syntax: configure>subscriber-mgmt>sub-profile>egress>agg-rate-limit** *agg-rate*

In the BNG application however, this rate is overridden by the rate provided in the RADIUS access-accept message. This rate represents the ATM rate that will be seen on the last mile, that is, it includes the encapsulation offset and the per packet expansion due to ATM segmentation into cells at the BSAN.

## Weighted Scheduler Group

The existing port scheduler policy defines a set of eight priority levels with no ability of grouping levels within a single priority. In order to allow for the application of a scheduling weight to groups of subscriber host queues competing at the same priority level of the port scheduler policy applied to the vport, or to the Ethernet port, a new group object is defined under the port scheduler policy:

**CLI Syntax: configure>qos>port-scheduler-policy>group group-name rate pir-rate** [**cir** *cir-rate*]

Up to eight groups can be defined within each port scheduler policy. One or more levels can map to the same group. A group has a rate and optionally a cir-rate and inherits the highest scheduling priority of its member levels. For example, the scheduler group shown in the vport in Figure 2 consists of level priority 3 and level priority 4. It thus inherits priority 4 when competing for bandwidth with the standalone priority levels 8, 7, and 5.

In essence, a group receives bandwidth from the port or from the vport and distributes it within the member levels of the group according to the weight of each level within the group. Each priority level will compete for bandwidth within the group based on its weight under congestion situation. If there is no congestion, a priority level can achieve up to its rate (cir-rate) worth of bandwidth.

The mapping of a level to a group is performed as follows:

**CLI Syntax: configure>qos>port-scheduler-policy>level priority-level rate pir-rate** [**cir** *cir-rate*] **group** *group-name* [**weight** *weight-in-group*]

Note that CLI will enforce that mapping of levels to a group are contiguous. In other words, a user would not be able to add priority level to group unless the resulting set of priority levels is contiguous.

When a level is not explicitly mapped to any group, it maps directly to the root of the port scheduler at its own priority like in existing behavior.

## Queue and Subscriber Aggregate Rate Configuration and Adjustment

### Software-Based Implementation (8.0R4)

The subscriber aggregate rate is adjusted and it will be based on an average frame size.

The user enables the use of this adjustment method by configuring the following option in the egress context of the subscriber profile:

**CLI Syntax: configure>subscriber-management>sub-profile>egress>encap-offset [type** *type***]**

This command allows the user to configure a default value to be used by all hosts of the subscriber in the absence of a valid signaled value. The following is a list of the configurable values:

**Values** pppoa-llc, pppoa-null, pppoeoa-llc, pppoeoa-llc-fcs, pppoeoa-llc-tagged, pppoeoa-llc-tagged-fcs, pppoeoa-null, pppoeoa-null-fcs, pppoeoa-null-tagged, pppoeoa-null-tagged-fcs
ipoa-llc, ipoa-null, ipoeoa-llc, ipoeoa-llc-fcs, ipoeoa-llc-tagged, ipoeoa-llc-tagged-fcs, ipoeoa-null, ipoeoa-null-fcs, ipoeoa-null-tagged, ipoeoa-null-tagged-fcs,
pppoe, pppoe-tagged,
ipoe, ipoe-tagged

Otherwise, the fixed packet offset will be derived from the encapsulation type value signaled in the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoE Tags as explained in Section Signaling of Last Mile Encapsulation Type. Only signaling using PPPoE Tags is supported in the software based implementation. The last signaled valid value is then applied to all active hosts of this subscriber. If no value is signaled in the subscriber host session or the value in the fields of the Access-loop-encapsulation sub-TLV are invalid, then the offset applied to the aggregate rate of this subscriber will use the last valid value signaled by a host of this subscriber if it exists, or the user entered default type value if configured, or no offset is applied.

The user also configures the average frame size value to be used for this adjustment:

**CLI Syntax: configure>subscriber-management>sub-profile>egress>avg-frame-size** *bytes*

The value entered by the user must include the FCS but not the Inter-Frame Gap (IFG) or the preamble. If the user does not explicitly configure a value for the **avg-frame-size** parameter, then it will also be assumed the offset is zero regardless of the signaled or user-configured value.

The computation of the subscriber aggregate rate consists of taking the average frame size, adding the encapsulation fixed offset including the AAL5 trailer, and then adding the variable offset consisting of the AAL5 padding to next multiple of 48 bytes. The AverageFrameExpansionRatio is then derived as follows:

AverageFrameExpansionRatio = (53/48 x (AverageFrameSize + FixedEncapOffset + AAL5Padding)) / (AverageFrameSize + IFG + Preamble).

When the last mile is Ethernet, the formula simplifies to:

AverageFrameExpansionRatio = (AverageFrameSize + FixedEncapOffset + IFG + Preamble) / (AverageFrameSize + IFG + Preamble).

The following are the frame size and rate applied to the subscriber queue and scheduler:

Subscriber Host Queue (no change):

Size = ImmediateEgressEncap + Data

Rate = ImmediateEgressEncap + Data

Subscriber Aggregate Rate Scheduler:

Size = ImmediateEgressEncap + Data

Rate = sub-agg-rate / AverageFrameExpansionRatio

Note that the CPM applies the AverageFrameExpansionRatio adjustment to the various components used in the determination of the net subscriber operational aggregate rate. It then pushes these adjusted components to IOM which then makes the calculation of the net subscriber operational aggregate rate.

The formula used by the IOM for this determination is:

sub-oper-agg-rate = min(sub-policy-agg-rate/*AverageFrameExpansionRatio*, ancp_rate/*AverageFrameExpansionRatio*) + (igmp_rate_delta/*AverageFrameExpansionRatio*),

where *sub-policy-agg-rate* is either the value configured in the **agg-rate-limit** parameter in the subscriber profile or the resulting RADIUS override value. In both cases, the CPM uses an internal override to download the adjusted value to IOM.

The value of *sub-oper-agg-rate* is stored in the IOM's subscriber table.

The following are the procedures for handling signaling changes or configuration changes affecting the subscriber profile:

1. If a new RADIUS update comes in for the aggregate subscriber rate, then a new subscriber aggregate ATM adjusted rate is computed by CPM using the last configured **avg-frame-size** and then programmed to IOM.
2. If the user changes the value of the **avg-frame-size** parameter, enables/disables the **encap-offset** option, or changes the parameter value of the **encap-offset** option, the CPM will immediately trigger a re-evaluation of subscribers using the corresponding subscriber profile and an update the IOM with the new subscriber aggregate rate.
3. If the user changes the value of the **agg-rate-limit** parameter in a subscriber profile which has the **avg-frame-size** configured, this will immediately trigger a re-evaluation of subscribers using the corresponding subscriber profile. An update to the subscriber aggregate rate is performed for those subscribers which rate has not been previously overridden by RADIUS.
4. If the user changes the **type** value of the encap-offset command, this will immediately trigger a re-evaluation of subscribers using the corresponding subscriber profile. An update to the subscriber aggregate rate is performed for those subscribers which are currently using the default value.

5. If two hosts of the same subscriber signal two different encapsulation types, the last one signaled gets used at the next opportunity to re-evaluate the subscriber profile.
6. If a subscriber has a DHCP host, a static host or an ARP host, the subscriber aggregate rate will continue to use the user-configured default encapsulation type value or the last valid encapsulation value signaled in the PPPoE tags by other hosts of the same subscriber. If none was signaled or configured, then no rate adjustment is applied.

## Hardware-Based Implementation

The data path will compute the adjusted frame size real-time for each serviced packet from a queue by adding the actual packet size to the fixed offset provided by CPM for this queue and variable AAL5 padding.

Like in the software based implementation, the user enables the use of the fixed offset and per packet variable expansion by configuring the following option in the egress context of the subscriber profile:

**CLI Syntax: configure>subscriber-management>sub-profile>egress>encap-offset [type** *type***]**

When this command is enabled, the fixed packet offset will be derived from the encapsulation type value signaled in the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoE Tags or DHCP Relay Options as explained in Section Signaling of Last Mile Encapsulation Type.

If the user specifies an encapsulation type with the command, this value will be used as the default value for all hosts of this subscriber until a host session signaled a valid value. The signaled value is applied to this host only and the remaining hosts of this subscriber continue to use the user entered default type value if configured, or no offset is applied. Note however that hosts of the same subscriber using the same SLA profile and which are on the same SAP will share the same instance of FC queues. In this case, the last valid encapsulation value signaled by a host of that same instance of the SAP egress QoS policy will override any previous signaled or configured value.

The procedures for handling signaling changes or configuration changes affecting the subscriber profile are the same as in the Software based implementation with except for the following:

1. The **avg-frame-size** parameter in the subscriber profile is ignored.
2. If the user specifies an encapsulation type with the command, this value will be used as the default value for all hosts of this subscriber until a host session signaled a valid value. The signaled value is applied to this host and other hosts of the same subscriber sharing the same SLA profile and which are on the same SAP. The remaining hosts of this subscriber continue to use the user entered default type value if configured, or no offset is applied.
3. If the user enables/disables the **encap-offset** option, or changes the parameter value of the **encap-offset** option, the CPM will immediately trigger a re-evaluation of subscribers hosts using the corresponding subscriber profile and an update the IOM with the new fixed offset value.
4. If subscriber host session signals an encapsulation type at the session establishment time and subsequently sends a DHCP renewal message using a Layer 2 DHCP relay which does not insert option82 in a unicast message, the encapsulation type for this host will not change. Note that TR-101 states that option82 is mandatory for DHCP broadcast messages).
5. If a subscriber has a static host or an ARP host, the subscriber host will continue to use the user-configured default encapsulation type value or the last valid encapsulation value signaled in the PPPoE tags or DHCP relay options by other hosts of the same subscriber which

use the same SLA profile instance. If none was signaled or configured, then no rate adjustment is applied.

6. The encapsulation type value signaled in DHCP relay options or PPPoE tags are not cross-checked against the host type. So, a host signaling PPPoA/LLC encapsulation type via DHCP relay options will be handled as if the packet included a PPPoE header when forwarded over the local Ethernet port. This results in applying an encap-offset in the data path which assumes the PPPoE header is added to forwarded packets over the local Ethernet port.

The **encap-offset** option forces all the rates to be either last-mile frame over the wire or local port frame over the wire, referred to as **LM-FoW** and **FoW** respectively. The system maintains a running average frame expansion ratio for each queue to convert queue rates between these two formats as explained in Frame Size, Rates, and Running Average Frame Expansion Ratio on page 935. Here are the details of the queue and scheduler operation:

1. When the **encap-offset** option is configured in the subscriber profile, the subscriber host queue rates, that is, CLI and operational PIR and CIR as well as queue bucket updates, the queue statistics, that is, forwarded, dropped, and HQoS offered counters use the **LM-FoW** format. The scheduler policy CLI and operational rates also use **LM-FoW** format. The port scheduler **max-rate** and the priority level rates and weights, if a Weighted Scheduler Group is used, are always entered in CLI and interpreted as **FoW** rates. The same is true for an **agg-rate-limit** applied to a vport. Finally the subscriber **agg-rate-limit** is entered in CLI as **LM-FoW** rate. When converting between **LM-FoW** and **FoW** rates, the queue running average frame expansion ratio value is used.
   - If the user enabled **frame-based-accounting** in a scheduler policy or **queue-frame-based-accounting** with subscriber **agg-rate-limit** and a port scheduler policy, the queue operational rate will be capped to a user configured **FoW** rate. The scheduler policy operational rates will also be in the **FoW** format. Note that a user configured queue **avg-frame-overhead** value is ignored since the running average frame expansion ratio is what is used when the **encap-offset** option is enabled.
   - If the user configured queue **packet-byte-offset** value, it is ignored and is not accounted for in the net packet offset calculation.
2. When **no encap-offset** is configured in the subscriber profile, that is, default and pre-R9.0 behavior, queue CLI and operational PIR and CIR rates, as well as queue bucket updates, the queue statistics, use data format. The scheduler policy CLI and operational rates also use data format. The port scheduler **max-rate** and the priority level rates and weights, if a Weighted Scheduler Group is used, and the subscriber **agg-rate-limit** are entered in CLI and interpreted as **FoW** rates. When converting between **FoW** and data rates, the queue a**vg-frame-overhead** value is used and since this an Ethernet port, it is not user-configurable but constant and is equal to +20 bytes (IFG and preamble).
   - If the user enabled **frame-based-accounting** in a scheduler policy or **queue-frame-based-accounting** with subscriber **agg-rate-limit** and a port scheduler policy, the queue operational rate will be capped to a user configured FoW rate in

CLI which is then converted into a data rate using the queue **avg-frame-overhead** constant value of +20 bytes. The scheduler policy operational rates will also be in the **FoW** format.

– If the user configured queue **packet-byte-offset** value, it adjusts the immediate packet size. This means that the queue rates, that is, operational PIR and CIR, and queue bucket updates use the adjusted packet size. In addition, the queue statistics will also reflect the adjusted packet size. Scheduler policy rates, which are data rates, will use the adjusted packet size. The port scheduler **max-rate** and the priority level rates and weights, if a Weighted Scheduler Group is used, as well as the subscriber **agg-rate-limit** are always **FoW** rates and thus use the actual frame size

## Frame Size, Rates, and Running Average Frame Expansion Ratio

The following are the details of the rates and frame sizes applied to the subscriber host queues, the subscriber aggregate rate, and the vport root scheduler for the scheduling model and when the **encap-offset** option is enabled in the subscriber profile.

Subscriber Host Queue:

Size = LastMileFrameOverWireEncap + Data

Rate = (48/53)* x (LastMileFrameOverWireEncap + Data)

*Applicable to ATM last-mile only.

Subscriber Aggregate Rate:

Size = LastMileFrameOverWireEncap + Data

Rate = (48/53)* x (LastMileFrameOverWireEncap + Data)

*Applicable to ATM last-mile only.

Vport/Port Port Scheduler and Weighted Scheduler Group

Size = FrameOverWireEncap + Data

Rate = FrameOverWireEncap + Data

When a frame arrives at the queue, it size will be *ImmediateEgressEncap+Data*. This size is stored as the *OfferedFrameSize* so that the queue offered stats used in HQoS calculations are correct. Let us refer to the HQoS offered statistics as Offered.

This size is then adjusted by removing the *ImmediateEgressEncap* and adding the *LastMileFrameOverWireEncap*. This new adjusted frame size, let us refer to it as *LastMileOfferedFrameSize*, is then used for checking compliance of the frame against the queue PIR and CIR bucket sizes and for updating the queue forwarded and dropped stats.

The *LastMileOfferedFrameSize* value is computed dynamically for each packet serviced by the queue.

A new HQoS stat counter *OfferedLastMileAdjusted* is maintained for the purpose of calculating the running average frame expansion ratio, which is the ratio of the accumulated *OfferedLastMileAdjusted* and Offered of each queue:

RunningAverageFrameExpansionRatio = OfferedLastMileAdjusted / Offered

The vport/port port scheduler will hand out its **FoW** bandwidth in terms of Fair Information Rate (FIR) bandwidth to each subscriber queue. This queue FIR must be converted into **LM-FoW** format to cap it by the queue PIR (*adminPIR*) and to make sure the sum of *FIRs* of all queues of the same subscriber does not exceed the subscriber **agg-rate-limit** which is also expressed in **LM-FoW** format. The conversion between these two rates makes use of the cumulative *RunningAverageFrameExpansionRatio* value.

Note that a queue **LM-foW** AdminPIR value will always be capped to the value of the local port **FoW** rate even if the conversion based on the current *RunningAverageFrameExpansionRatio* value indicates that a higher AdminPIR may be able to fill in the full line rate of the local port.

## VPORT Determination and Evaluation

In the BNG application, host queues of all subscribers destined to the same downstream BSAN, for example, all SAPs on the egress port matching the same S-VLAN tag value, are parented to the same vport which matches the destination ID of the BSAN.

The BNG determines the parent vport of a subscriber host queue, which has the **port-parent** option enabled, by matching the destination string **dest** string associated with the subscriber with the string defined under a vport on the port associated with the subscriber.

The user configures the dest string match under the egress context of the Ethernet port associated with the subscriber:

**CLI Syntax:** configure>port>ethernet>acess>egress>vport>host-match dest *string* create

If a given subscriber host queue does not have the **port-parent** option enabled, it will be foster-parented to the vport used by this subscriber and which is based on matching the **dest** string. If the subscriber could not be matched with a vport on the egress port, the host queue will not be bandwidth controlled and will compete for bandwidth directly based on its own PIR and CIR parameters.

By default, a subscriber host queue with the **port-parent** option enabled is scheduled within the context of the port's port scheduler policy. In order to indicate the option to schedule the queue in the context of a port scheduler policy associated with a vport, the user enters the following command in SLA profile used by the subscriber host:

**CLI Syntax: configure>subscriber-mgmt>sla-profile>egress>qos** *sap-egress-qos-policy-id* **vport-schedule**r

This command is persistent meaning that the user can re-enter the **qos** node without specifying the **vport-scheduler** argument each time and the system will remember it. The user can revert to the default setting without deleting the association of the SLA profile with the SAP egress QoS policy by explicitly re-entering the command with the following new argument:

**CLI Syntax:** configure>subscriber-mgmt>sla-profile>egress>qos *sap-egress-qos-policy-id* port-scheduler

---

## Applying Aggregate Rate Limit to a VPORT

The user can apply an aggregate rate limit to the vport and apply a port scheduler policy to the port.

This model allows the user to oversubscribe the Ethernet port. The application of the **agg-rate-limit** option is mutually exclusive with the application of a port scheduler policy to a vport.

When using this model, a subscriber host queue with the **port-parent** option enabled is scheduled within the context of the port's port scheduler policy. However, the user must still indicate to the

system that the queues are managed by the aggregate rate limit instance of a vport by enabling the **vport-scheduler** option in the subscriber host SLA profile:

**CLI Syntax: configure>subscriber-mgmt>sla-profile>egress>qos** *sap-egress-qos-policy-id* **vport-scheduler**

A subscriber host-queue which is port-parented will be parented to the port scheduler policy of the port used by the subscriber and aggregate rate limited within the instance of the vport used by this subscriber and which is based on matching the **dest** string and **org** string. If the vport exists but the port does not have a port scheduler policy applied, then the host queue will be orphaned and no aggregate rate limit can be enforced.

If a given subscriber host queue does not have the port-parent option enabled, it will be foster-parented to the port used by this subscriber and aggregate rate limited within the instance of the vport used by this subscriber. If the vport exists but the port does not have a port scheduler policy applied, then the host queue will be orphaned and no aggregate rate limit can be enforced.

## Signaling of Last Mile Encapsulation Type

A subscriber host session can signal one of many encapsulation types each with a different fixed offset in the last mile. These encapsulation types are described in RFC 4679 and are illustrated in Figure 72 and Figure 73. The BNG node learns the encapsulation type of each subscriber host session by inspecting the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoE Tags as specified in RFC 4679.When Ethernet is the last mile, the encapsulation type will result in a fixed offset for all packet sizes. When ATM/DSL is the last mile, there will be an additional expansion due to AAL5 padding to next multiple of 48 bytes and which varies depending on the packet size.

Both ATM and Ethernet access using PPP encapsulation options are supported in the software and hardware based implementations. Thus both provide support for the Access-loop-encapsulation sub-TLV in the Vendor-Specific PPPoEv4/PPPoEv6 Tags with the ATM encapsulation values and Ethernet encapsulation values.

ATM and Ethernet access using IP encapsulation are only supported using default encapsulation offset configuration in the subscriber profile in the software based implementation. Support for signaling the Access-loop-encapsulation sub-TLV in the DHCPv4/DHCPv6 Relay Options is included in the hardware based implementation. There is no support for DHCPv6 relay options.

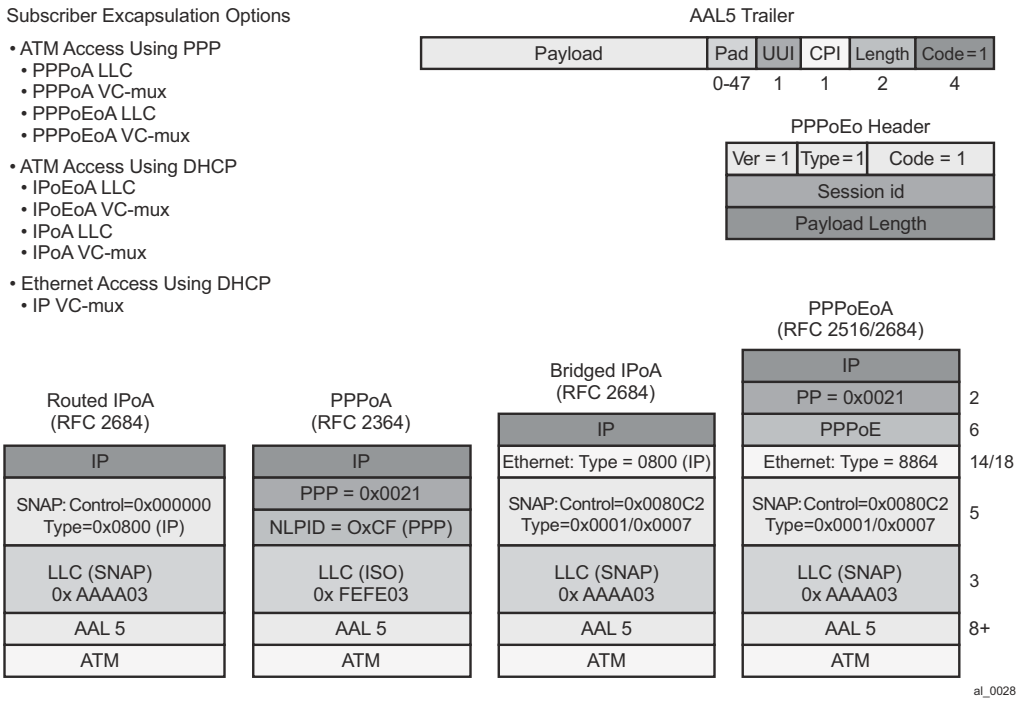



**Figure 72: Subscriber Host Session Encapsulation Types**

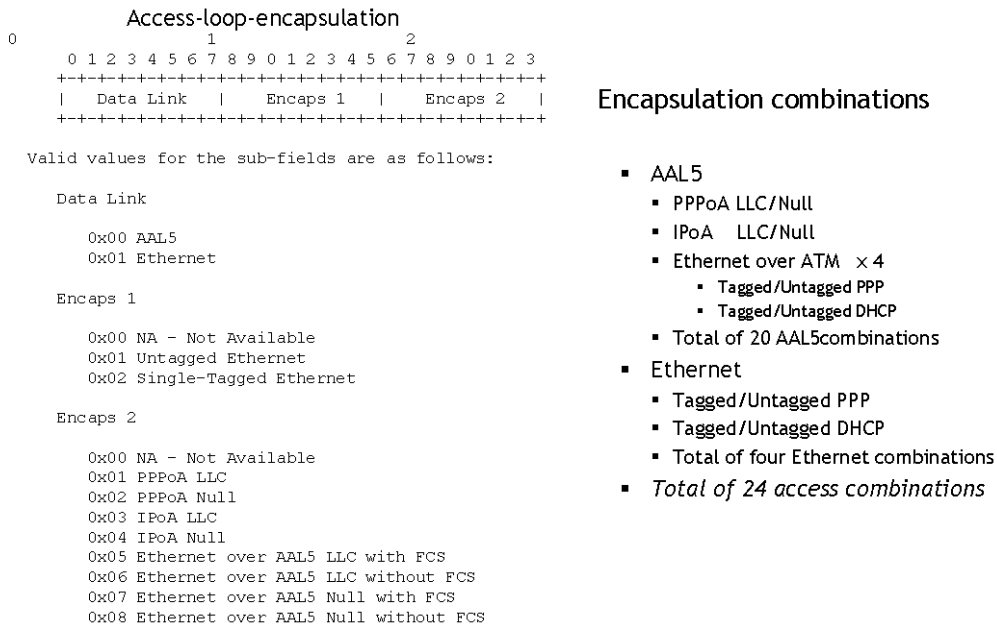



**Figure 73: Access-Loop-Encapsulation Sub-TLV**

The operational last-mile values for hosts on the same sap, having the same SLA profile are displayed in following the show-command:

**CLI Syntax: show>service active-subscribers>ale-adjust**

The data-link can have values: atm, other and unknown. If no offset is supplied it will be set to unknown. Other is used when the data-link is non-atm, otherwise it will state atm.

Operational per-queue values can also be found in the show-command:

**CLI Syntax: show>qos>scheduler-hierarchy**

Here, one can see whether the queue is operating in last-mile mode. Note that this command is not available on HSMDAv2,

**Last Mile ATM**

```
*A:Dut-C# /show service active-subscribers ale-adjust
===============================================================================
Active Subscriber Access Loop Encapsulation adjustment
===============================================================================
Subscriber
   SAP                                          SLA profile
   Data-link Offset(bytes)
```

```
-------------------------------------------------------------------------------
hpolSub81
   1/1/11:2000.1                                hpolSlaProf1
   atm       -10
-------------------------------------------------------------------------------
No. of Access Loop Encapsulation adjustments: 1
===============================================================================

*A:Dut-C# show qos scheduler-hierarchy subscriber "hpolSub81"
===============================================================================
Scheduler Hierarchy - Subscriber hpolSub81
===============================================================================
Ingress Scheduler Policy:
Egress Scheduler Policy :
-------------------------------------------------------------------------------
Root (Ing)
|
No Active Members Found on slot 1


Root (Egr)
| slot(1)
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->8->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->7->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->6->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->5->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->4->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->3->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->2->ATM  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->1->ATM  (Port 1/1/11)
|
```

### Last Mile Ethernet

```
*A:Dut-C# show service active-subscribers ale-adjust
===============================================================================
Active Subscriber Access Loop Encapsulation adjustment
===============================================================================
Subscriber
   SAP                                          SLA profile
   Data-link Offset(bytes)
-------------------------------------------------------------------------------
hpolSub81
   1/1/11:2000.1                                hpolSlaProf1
   other     +12
-------------------------------------------------------------------------------
No. of Access Loop Encapsulation adjustments: 1
===============================================================================

*A:Dut-C# show qos scheduler-hierarchy subscriber "hpolSub81"
```

```
===============================================================================
Scheduler Hierarchy - Subscriber hpolSub81
===============================================================================
Ingress Scheduler Policy:
Egress Scheduler Policy :
-------------------------------------------------------------------------------
Root (Ing)
|
No Active Members Found on slot 1


Root (Egr)
| slot(1)
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->8->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->7->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->6->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->5->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->4->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->3->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->2->Eth  (Port 1/1/11)
|
|--(Q) : Sub=hpolSub81:hpolSlaProf1 2000->1/1/11:2000.1->1->Eth  (Port 1/1/11)
|
```

### Next Mile ATM

```
*A:Dut-C# show service active-subscribers ale-adjust subscriber "hpolSub321"
===============================================================================
Active Subscriber Access Loop Encapsulation adjustment
===============================================================================
Subscriber
   SAP                                          SLA profile
   Data-link Offset(bytes)
-------------------------------------------------------------------------------
hpolSub321
   5/1/1:100/1                                  hpolSlaProf1
   atm       +8
-------------------------------------------------------------------------------
No. of Access Loop Encapsulation adjustments: 1
===============================================================================

*A:Dut-C# show qos scheduler-hierarchy subscriber "hpolSub321"
===============================================================================
Scheduler Hierarchy - Subscriber hpolSub321
===============================================================================
Ingress Scheduler Policy:
Egress Scheduler Policy :
-------------------------------------------------------------------------------
Root (Ing)
|
No Active Members Found on slot 5
```

```
Root (Egr)
| slot(5)
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->8:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->7:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->6:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->5:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->4:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->3:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->2:ATM  (Port 5/1/1)
|
|--(Q) : Sub=hpolSub321:hpolSlaProf1 2000->5/1/1:100/1->1:ATM  (Port 5/1/1)
|
===============================================================================
```

## Configuration Example

The following CLI configuration achieves the specific use case shown in Figure 71.

```
config
    qos
       port-scheduler-policy "dslam-vport-scheduler"
         group res-bus-be create
              rate 1000
         level 3 rate 1000 group res-bus-be weight w1
         level 4 rate 1000 group res-bus-be weight w4
         level 5 rate 1000 cir-rate 100
         level 7 rate 5000 cir-rate 5000
         level 8 rate 500 cir-rate 500
         max-rate 5000

       sap-egress 100                // residential policy
            queue 1                 // be-res
            port-parent weight x level 3
            queue 2                 // l2-res
            port-parent weight y level 3
            queue 3                 // l1-res
            port-parent weight z level 3
            queue 4                 // h2-res
            port-parent level 5
            queue 5                 // h1-res
            port-parent level 7
            queue 6                 // ef-res
            port-parent level 8
            fc be queue 1
            fc l2 queue 2
            fc l1 queue 3
            fc h2 queue 4
            fc h1 queue 5
            fc ef queue 6
        exit
        sap-egress 200               // business policy
            queue 1                 // be-bus
              port-parent weight x level 4
            queue 2                 // l2-bus
              port-parent weight y level 4
            queue 3                 // l1-bus
              port-parent weight z level 4
            queue 4                 // h2-bus
              port-parent level 5
            queue 5                 // h1-bus
              port-parent level 7
            queue 6                 // ef-bus
              port-parent level 8
            fc be queue 1
            fc l2 queue 2
            fc l1 queue 3
            fc h2 queue 4
            fc h1 queue 5
            fc ef queue 6
        exit
    exit
```

```
config
    sub-mgmt
        sla-profile "residential"
            egress
                qos 100 vport-scheduler
            exit
        exit
        sla-profile "business"
            egress
                qos 200 vport-scheduler
            exit
        exit
        sub-profile "residential"
            egress
               encap-offset
               avg-frame-size 1500
               agg-rate-limit 100
               exit
            exit
        exit
        sub-profile "business"
            egress
               encap-offset type pppoeoa-llc-tagged-fcs
               avg-frame-size 500
               agg-rate-limit 200
               exit
            exit
        exit
    exit

config
    port 1/1/1
        ethernet
            access
                egress
                    vport "dslam-1" create
                       port-scheduler-policy "dslam-vport-scheduler"
                        host-match dest “20” create
                        exit
                    exit
                exit
            exit
        exit
    exit
exit
```

# Configuring IP and IPv6 Filter Policies for Subscriber Hosts

This section applies to the 7750 SR and 7450 ESS.

Access Control Lists (ACLs) for subscriber traffic are defined as IP and IPv6 filter policies and are configured in the SLA-profile associated with the subscriber. For information about IP and IPv6 filter policy configurations, refer to the 7750 SR-OS Router Configuration Guide.

**CLI Syntax:**
```
config>subscr-mgmt>sla-prof
    sla-profile sla-profile-1 create
        ingress
            ip-filter 100
            ipv6-filter 300
        exit
        egress
            ip-filter 200
            ipv6-filter 400
        exit
    exit
```

Traffic from different hosts of a single subscriber and associated with the same sla-profile instance, is subject to the filter policies defined in the SLA profile.

The IP or IPv6 filter policy configuration of subscriber hosts can be dynamically updated using different mechanisms:

1. Assign a new SLA profile.

   This can be done dynamically by, for example, a RADIUS CoA message. As the SLA profile also defines the QoS configuration for the subscriber hosts, this change may result in a discontinuity in accounting.

   Note: changing the ip-filter policy in an SLA profile in use by an active subscriber is allowed in the CLI, but not recommended. Changing the IPv6 filter policy in an SLA profile in use by an active subscriber is prevented in the CLI.
2. Override the IP and IPv6 filter policies

   Alternatively, it is also possible to dynamically override the IP and IPv6 filter policies per subscriber-host through a RADIUS Access-Accept or CoA message. Following VSA should be included in the RADIUS message:
3. Insert subscriber host specific filter entries

   A subscriber host specific entry is dynamically created from a RADIUS access accept or CoA message using the NAS-Filter-Rule or Alc-Ascend-Data-Filter-Host-Spec attribute (see also IP Filter Attribute Format Details on page 952 for a detailed description of the attribute format):

| Attribute ID | Attribute Name | Type | Limits | SR OS Format |
|---|---|---|---|---|
| 92 | NAS-Filter-Rule | String | max. 10 attributes per message or max. 10 filter entries per message | The format of a NAS-Filter-Rule is defined in rfc-3588 section-4.3. A single filter rule is a string of format “<action> <direction> <protocol> from <source> to <destination> <options>”. Multiple rules should be separated by a NUL (0x00). A NAS-Filter-Rule attribute may contain a partial rule, one rule, or more than one rule. Filter rules may be continued across attribute boundaries.<br>A RADIUS message with NAS-Filter-Rule attribute value equal to 0x00 or “ “ (a space) removes all host specific filter entries for that host.<br>See also IP Filter Attribute Format Details on page 952.<br>For example:<br>Nas-Filter-Rule = "permit in ip from any to 10.1.1.1/32" |
| 26-6527-159 | Alc-Ascend-Data-Filter-Host-Spec | octets | max. 10 attributes per message or max. 10 filter entries per message. min. length 22 bytes (IPv4), 46 bytes (IPv6) max. length: 110 bytes (IPv4), 140 bytes (IPv6) | a string of octets with fixed field length (type (ipv4/ipv6), direction (ingress/egress), src-ip, dst-ip, ...). Each attribute represents a single filter entry. See IP Filter Attribute Format Details on page 952 for a description of the format.<br>For example:<br># "permit in ip from any to 10.1.1.1/32"<br>Alc-Ascend-Data-Filter-Host-Spec = 0x01010100000000000a010101002000000000000000000 |

The Alc-Subscriber-Filter VSA is a comma separated list of strings:

| Field | Use |
|---|---|
| Ingr-v4:<number> | Ingress ipv4 filter |

| Field | Use |
|---|---|
| Egr-v4:<number> | Egress ipv4 filter |
| Ingr-v6:<number> | Ingress ipv6 filter |
| Egr-v6:<number> | Egress ipv6 filter |

The filter number can have following values:

| <number> | Result |
|---|---|
| 1..65535 | Ignore filter from sla-profile configuration and assign corresponding pre-configured filter |
| 0 | Ignore filter from sla-profile configuration and do not assign a new filter (only allowed if no dynamic subscriber host specific rules are present) |
| -1 | No change in filter configuration |
| -2 | Restore filter from sla-profile configuration |

Notes:

- Not relevant fields (IPv4 filters for an IPv6 host) will be ignored.
- RADIUS CoA message: if the ingress or egress field is missing in the VSA, there will be no change for that direction.
- RADIUS Access-Accept message: if the ingress or egress field is missing in the VSA, then the IP filters as specified in the SLA profile will be active for that direction.

An SLA profile IP filter override is applicable to all dynamic host types, including L2TP LNS but excluding L2TP LAC.

4. Insert subscriber host specific filter entries

   A subscriber host specific entry is a filter entry where the match criteria is automatically extended with the subscriber host IP or IPv6 address as source (ingress) or destination (egress) IP. They represent a per host customization of a generic filter policy: only traffic to/from the subscriber host will match against these entries.

   A subscriber host specific entry is dynamically created from a RADIUS access accept or CoA message using the NAS-Filter-Rule attribute:

The format used to specify host specific filter entries (NAS-Filter-Rule format or Alc-Ascend-Data-Filter-Host-Spec format) cannot change during the lifetime of the subscriber host. A RADIUS message can only contain a single format for host specific filter entries.

Up to 10 host specific filter rules can be specified in a single RADIUS message. Each new RADIUS CoA message containing host specific filter attributes overwrites the previous subscriber host-specific filter entries for that host provided that there are enough free entries in the reserved range.

Subscriber host specific filter entries can be removed with a RADIUS CoA message with NAS-Filter-Rule attribute value equal to 0x00 or " " (a space).

When the subscriber host session terminates or is disconnected the corresponding subscriber host specific filter entries are also deleted.

Note that subscriber host-specific filter entries are moved if the subscriber host filter policy is changed (new SLA profile or ip filter policy override) and the new filter policy contains enough free reserved entries (sub-insert-RADIUS).

A range of entries must be reserved for subscriber host specific entries in a filter policy:

**CLI Syntax: config>filter**
ip-filter 100 create
sub-insert-radius start-entry 1000 count 100

High and low watermarks can be configured to raise an event when the thresholds of free entries in the reserved range are reached:

**CLI Syntax: config>filter>ip-filter# sub-insert-wmark ?**
- no sub-insert-wmark
- sub-insert-wmark low <low-watermark> high <high-watermark>

<low-watermark> : [0..100]
<high-watermark> : [0..100]

Use following show commands to check filter policy details and the filter configuration for a subscriber host:

```
# show filter ip <ip-filter-id> type <entry-type>
# show filter ipv6 <ipv6-filter-id> type <entry-type>
     <entry-type>        : fixed|radius-insert|credit-control-insert

# show service active-subscribers filter [subscriber <sub-ident-string>] [origin <origin>]
     <sub-ident-string>   : [32 chars max]
     <origin>             : radius|credit-control
```

5. Insert shared filter entries

   The target application for RADIUS shared filter entries is operators that have a predefined limited number of different filter lists that each are shared with multiple subscriber hosts and that are to be managed and activated from RADIUS at authentication.

   A local configured ip or ipv6 filter associated with a host (sla-profile or host filter override) can be enhanced with dynamic filter entries that can be shared with multiple subscriber hosts. The shared dynamic filter entries are inserted with a set of RADIUS attributes "[242]

Ascend-Data-Filter" or "[26-6527-158] Alc-Nas-Filter-Rule-Shared" received in a RADIUS Access-Accept or CoA message. A CoA message containing a set of one of those attributes overrides the previous set of shared filter entries active for that subscriber host.

For each unique set of dynamic filter entries received per type (ipv4/ipv6) and direction (ingress/egress), a copy is made of the local filter with the dynamic entries included at a preconfigured insert point. If the same set of dynamic filter entries is sent to subscriber hosts that have the same associated local filter, then they will share the same filter copy. When there are no more subscriber hosts associated with a filter copy, then the filter copy is deleted. A filter copy is identified as *local filter id*:*number*. For example: show filter ip 10:2

Shared filter entries are moved if the subscriber host filter policy is changed (new SLA profile or ip filter policy override) and if the new filter policy contains enough free reserved entries.

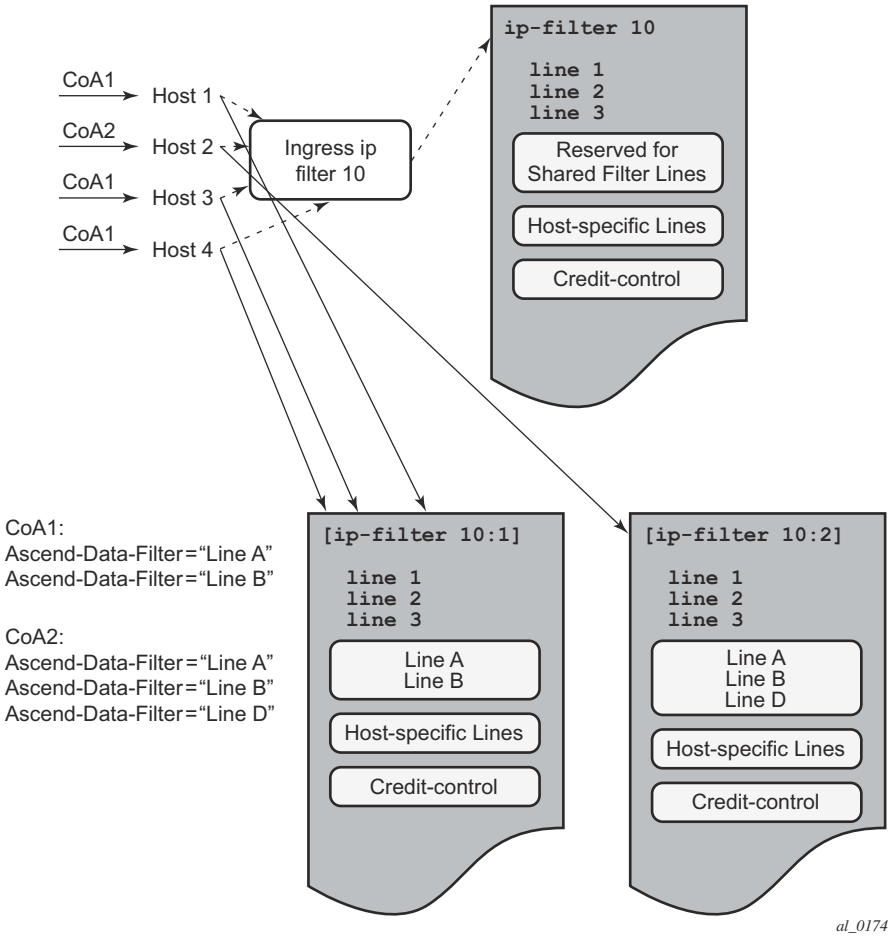



**Figure 74: Insert Shared Filters**

A range of entries must be reserved for shared entries in a filter policy:

**CLI Syntax:** **config>filter**
ip-filter 10 create
sub-insert-shared-radius start-entry 100 count 10

High and low watermarks can be configured to raise an event when the thresholds of dynamic filter copies are reached:

**CLI Syntax:** **config>filter>ip-filter# shared-radius-filter-wmark** ?
- no shared-radius-filter-wmark
- shared-radius-filter-wmark low *low-watermark* high *high-watermark*
low-watermark : [0..8000]
high-watermark : [0..8000]

The format used to specify shared filter entries (Alc-Nas-Filter-Rule-Shared format or Ascend-Data-Filter format) cannot change during the lifetime of the subscriber host. A RADIUS message can only contain a single format for shared filter entries.

Shared filter entries can be removed with a RADIUS CoA message with Alc-Nas-Filter-Rule-Shared attribute value equal to 0x00 or " " (a space).

| Attribute ID | Attribute Name | Type | Limits | SR OS Format |
|---|---|---|---|---|
| 242 | Ascend-Data-Filter | Octets | multiple attributes per RADIUS message allowed. min. length 22 bytes (IPv4), 46 bytes (IPv6) max. length: 110 bytes (IPv4), 140 bytes (IPv6) | a string of octets with fixed field length (type (ipv4/ipv6), direction (ingress/egress), src-ip, dst-ip, ...). Each attribute represents a single filter entry. See IP Filter Attribute Format Details on page 952 for a description of the format.<br>For example:<br># "permit in ip from any to 10.1.1.1/32"<br>Ascend-Data-Filter = 0x01010100000000000a010101002000000000000000000 |
| 26-6527-158 | Alc-Nas-Filter-Rule-Shared | string | Multiple attributes per RADIUS message allowed. | The format is identical to [92] NAS-Filter-Rule and is defined in rfc-3588 section-4.3. A single filter rule is a string of format “<action> <direction> <protocol> from <source> to <destination> <options>”. Multiple rules should be separated by a NUL (0x00). An Alc-Nas-Filter-Rule-Shared attribute may contain a partial rule, one rule, or more than one rule. Filter rules may be continued across attribute boundaries.<br>A RADIUS message with Alc-Nas-Filter-Rule-Shared attribute value equal to 0x00 or “ “ (a space) removes the shared filter entries for that host.<br>See also IP Filter Attribute Format Details on page 952.<br>For example:<br>Alc-Nas-Filter-Rule-Shared = "permit in ip from any to 10.1.1.1/32" |

## IP Filter Attribute Format Details

The format for [92] Nas-Filter-Rule and [26-6527-158] Alc-Nas-Filter-Rule-Shared is a string formatted as: "*action direction protocol* from *source* to *destination options*". Refer to the table below for details on the respective fields.

| Action or Classifier | Value | | Corresponding SR-OS Filter Function |
|---|---|---|---|
| <direction> | in | | ingress |
| | out | | egress |
| <protocol> | ip | | |
| | any number [0..255] | | |
| | ip | | |
| | any number [1..42] | | |
| | any number [45..49] | | |
| | any number [52..59] | | |
| | any number [61..255] | | |
| | any number 43\|44\|50\|51\|60 | | |
| from <source> | any | 100 | ingress: src-ip = host-ip-address; src-port eq 100<br>egress: src-ip = 0.0.0.0/0 \| ::/0; src-port eq 100 |
| | | 200-65535 | ingress: src-ip = host-ip-address; src-port range 200 65535<br>egress: src-ip = 0.0.0.0/0 \| ::/0; src-port range 200 65535 |
| | ip-prefix/ length | 100 | ingress: src-ip = host-ip-address; src-port eq 100<br>egress: src-ip = ip-prefix/length; src-port eq 100 |
| | | 200-65535 | ingress: src-ip = host-ip-address; src-port range 200 65535<br>egress: src-ip = ip-prefix/length; src-port range 200 65535 |
| | any | 100 | ingress: dst-ip = 0.0.0.0/0 \| ::/0; dst-port eq 100<br>egress: dst-ip = host-ip-address; dst-port eq 100 |
| | | 200-65535 | ingress: dst-ip = 0.0.0.0/0 \| ::/0; dst-port range 200 65535<br>egress: dst-ip = host-ip-address; dst-port range 200 65535 |

| Action or Classifier | Value | | Corresponding SR-OS Filter Function |
|---|---|---|---|
| | ip-prefix/ length | 100 | ingress: dst-ip = ip-prefix/length; dst-port eq 100<br>egress: dst-ip = host-ip-address; dst-port eq 100 |
| | | 200-65535 | ingress: dst-ip = ip-prefix/length; dst-port range 200 65535<br>egress: dst-ip = host-ip-address; dst-port range 200 65535 |
| <options: frag> | frag | | fragment true (ipv4 only) |
| <options: ipoptions> | ssrr | | ip-option 9 / ip-mask 255 |
| | lsrr | | ip-option 3/ ip-mask 255 |
| | rr | | ip-option 7/ ip-mask 255 |
| | ts | | ip-option 4/ ip-mask 255 |
| | !ssrr | | not supported |
| | !lsrr | | not supported |
| | !rr | | not supported |
| | !ts | | not supported |
| | ssrr,lsrr,rr, ts | | not supported |
| <options: tcpoptions> | mss | | not supported |
| | window | | not supported |
| | sack | | not supported |
| | ts | | not supported |
| | !mss | | not supported |
| | !window | | not supported |
| | !sack | | not supported |
| | !ts | | not supported |
| | mss,window,sack,ts | | not supported |
| <options: established> | established | | not supported |
| | | | not supported |
| | | | not supported |

| Action or Classifier | Value | | Corresponding SR-OS Filter Function |
|---|---|---|---|
| <options: setup> | setup | | tcp-syn true |
| | | | tcp-ack false |
| | | | protocol tcp |
| <options: tcpflags> | syn | | tcp-syn true |
| | !syn | | tcp-syn false |
| | ack | | tcp-ack true |
| | !ack | | tcp-ack false |
| | fin | | not supported |
| | rst | | not supported |
| | psh | | not supported |
| | urg | | not supported |
| <options: icmptypesv4> | echo reply | | protocol 1 / icmp-type 0 |
| | destination unreachable | | protocol 1 / icmp-type 3 |
| | source quench | | protocol 1 / icmp-type 4 |
| | redirect | | protocol 1 / icmp-type 5 |
| | echo request | | protocol 1 / icmp-type 8 |
| | router advertisement | | protocol 1 / icmp-type 9 |
| | router solicitation | | protocol 1 / icmp-type 10 |
| | time-to-live exceeded | | protocol 1 / icmp-type 11 |
| | IP header bad | | protocol 1 / icmp-type 12 |
| | timestamp request | | protocol 1 / icmp-type 13 |
| | timestamp reply | | protocol 1 / icmp-type 14 |
| | information request | | protocol 1 / icmp-type 15 |
| | information reply | | protocol 1 / icmp-type 16 |
| | address mask request | | protocol 1 / icmp-type 17 |
| | address mask reply | | protocol 1 / icmp-type 18 |

| Action or Classifier | Value | Corresponding SR-OS Filter Function |
|---|---|---|
| | - | protocol 1 / icmp-type [0..255] |
| | 3-9 ( range) | not supported |
| | 3,5,8,9 ( comma seper-ated) | not supported |
| <options: icmptypesv6> | destination unreachable | icmp-type 1 |
| | time-to-live exceeded | icmp-type 3 |
| | IP header bad | icmp-type 4 |
| | echo request | icmp-type 128 |
| | echo reply | icmp-type129 |
| | router solicitation | icmp-type 133 |
| | router advertisement | icmp-type 134 |
| | redirect | icmp-type 137 |

The format for [242] Ascend-Data-Filter and [26-6527-159] Alc-Ascend-Data-Filter-Host-Spec is an octet string with fixed length fields. Refer to the table below for details on the respective fields.

| Field | Length | Value |
|---|---|---|
| Type | byte | 1 = IPv4 |
| | | 3 = IPv6 |
| Filter or forward | 1 byte | 0 = drop |
| | | 1 = accept |
| Indirection | 1 byte | 0 = egress |
| | | 1 = ingress |
| Spare | 1 byte | ignored |
| Source IP address | IPv4 = 4 bytes | IP address of the source interface |
| | IPv6 = 16 bytes | |
| Destination IP address | IPv4 = 4 bytes | IP address of the destination interface |
| | IPv6 = 16 bytes | |
| Source IP prefix | 1 byte | Number of bits in the network portion |

| Field | Length | Value |
|---|---|---|
| Destination IP prefix | 1 byte | Number of bits in the network portion |
| Protocol | 1 byte | Protocol number.<br>Note: match the inner most header only for IPv6 |
| Established | 1 byte | ignored (not implemented) |
| Source port | 2 bytes | Port number of the source port |
| Destination port | 2 bytes | Port number of the destination port |
| Source port qualifier | 1 byte | 0 = no compare |
| | | 1 = less than |
| | | 2 = equal to |
| | | 3 = greater than |
| | | 4 = not equal to (not supported) |
| destination port qualifier | 1 byte | 0 = no compare |
| | | 1 = less than |
| | | 2 = equal to |
| | | 3 = greater than |
| | | 4 = not equal to (not supported) |
| Reserved | 2 bytes | ignored |

## Checking Filter Policy Details

Use following show commands to check filter policy details and the filter configuration for a subscriber host:

**CLI Syntax:**

show filter ip *ip-filter-id* detail
show filter ipv6 *ip-filter-id* detail
show filter ip *ip-filter-id* type entry-type
show filter ipv6 *ipv6-filter-id* type *entry-type*
*entry-type* : fixed | radius-insert | credit-control-insert | radius-shared
show service active-subscribers filter [subscriber *sub-ident-string*] [origin *origin*]
*sub-ident-string* : [32 chars max]
*origin* : radius | credit-control”

# ESM PPPoA/PPPoEoA

This section applies to the 7750 SR and 7450 ESS.

The main goal of PPP in the subscriber context is to provide authentication, to negotiate link layer parameters (such as MTU) and to negotiate IP parameters (IP address, WINS, DNS, Default Gateway, etc.).

Each PPP session is carried over a single ATM VC over to the BNG. In PPPoA environment, PPP session is directly encapsulated over ATM transport on a DSL Customer Premise Equipment (CPE).

In the PPPoEoA environment, an additional layer is added to accommodate Ethernet medium. In this fashion, a PPP session can be directly terminated on any host within the customer Ethernet network and then transported over an ATM network to the BNG. Multiple PPPoE sessions can be carried over a single ATM PVC.

However, the majority of current implementations in ATM transport networks have PPPoE session terminated at the DSL CPE and not on a host within the customer network. This PPPoE session is then transported over ATM to the BNG. Although it seems unnecessary to add the overhead associated with Ethernet on an ATM-equiped DSL CPE, mainly for historical reasons PPPoE became engrained in the home network and as such has moved into DSL CPE.

## PPPoA

In a PPPoA environment, services are offered to residential and business customers. DSL CPE and BNG are the originating and terminating points of a PPP connection. In a residential example, a DSL CPE is a Home Gateway (HG) as shown in Figure 75.

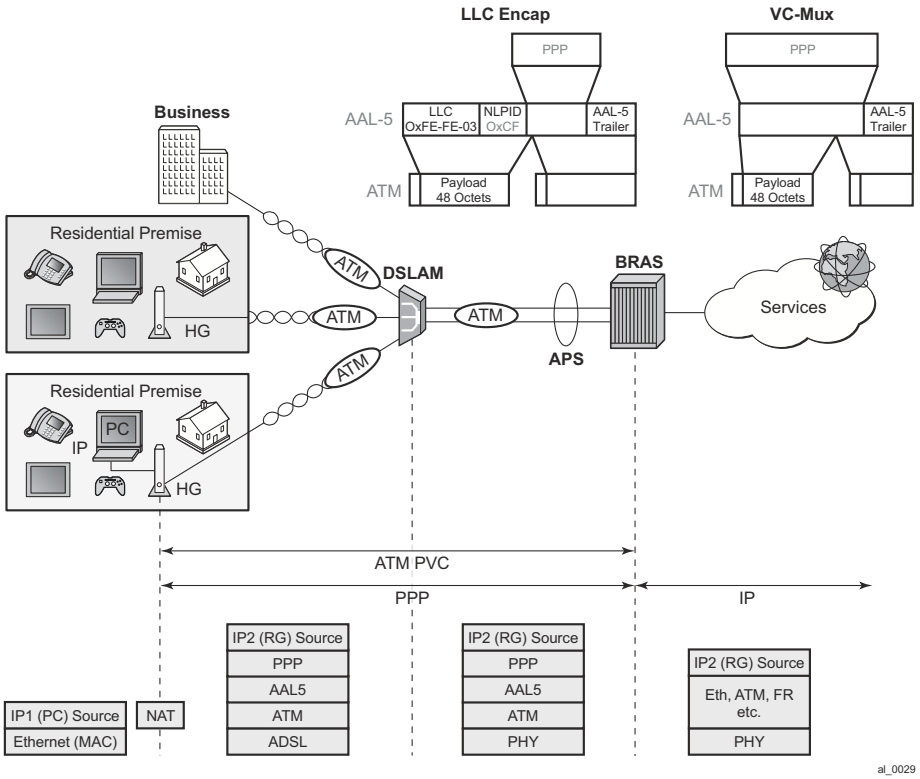



**Figure 75: PPPoA Architecture and Packet Encapsulation**

An ATM VC must exist between a CPE and a BNG before a PPP session can be established. From a QoS perspective, this VC is one of the following service categories: CBR, rt-VBR, nrt-VBR or UBR(+). CPE starts negotiating PPP link/session parameters over the VC. Once the IP address is obtained from the Service Provider (SP) side, the customer is ready for data transfer. There is a 1:1 mapping between a PPPoA session and a VC.

There is no need for the customer to run PPP within its own network. CPE, as a default gateway, accepts Ethernet IP packets, strips off the Ethernet header, performs a NAT function, and encapsulates the IP packets into PPPoA before it sends them on to the BNG.

In most residential cases, this PPPoA session is used for pure data transport (Internet access). As such, the BNG side would require a single service queue per VC.

In certain cases, customers use PPPoA for VoIP. In this case, they use an IAD (Integrated Access Device) to gain access to the ATM network.

It must be noted here that nothing in this architecture precludes customers from using multiple services over a single PPPoA session. Services would be differentiated through DSCP bits and each service in this case would require a separate queue. This is most likely a scenario with business customers where a customer runs multiple services over a PPPoA session. DSL CPE

would play a role here in differentiating and appropriately marking the services, either through a separate physical port per service or through some other means.

## PPPoEoA

In the PPPoE model the originating PPP point could extends beyond the DSL CPE and into the customer Ethernet network where any host can originate a PPP session. In such case, ATM is generally used as a transport to carry PPPoE sessions that are originated by hosts within the customer network (beyond DSL CPE). The DSL CPE would operate in a bridged mode. This is shown in Figure 76.

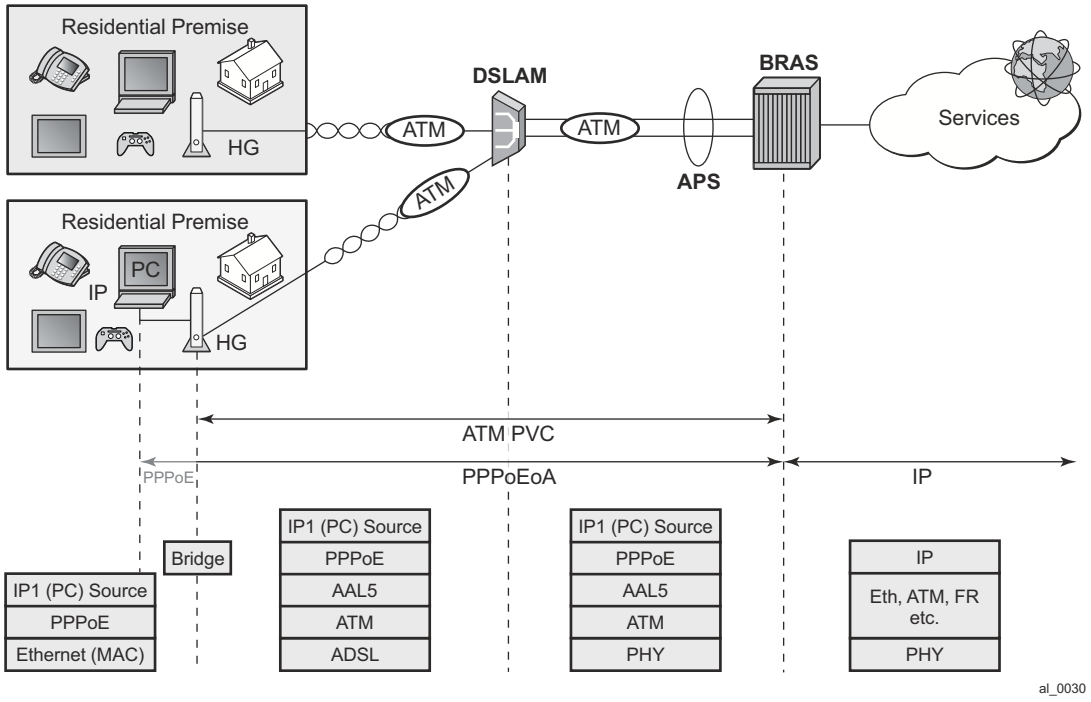



**Figure 76: PPoEoA Host Terminated Session**

However, the majority of deployments have PPPoE sessions terminated in DSL CPE. Although it is inefficient to add an extra Ethernet encapsulation layer over ATM-equiped DSL CPE, the evolution of PPP in broadband is the chief reason for this deployment scenario as shown in Figure 77.

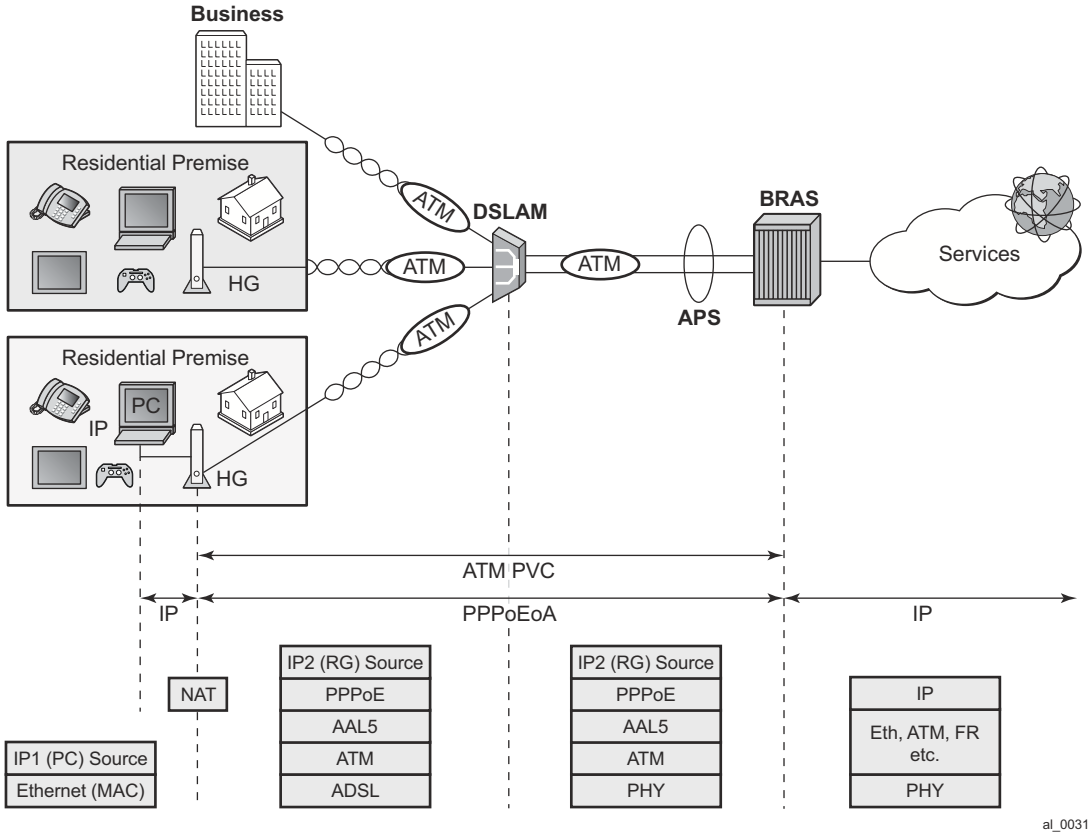



**Figure 77: PPPoEoA DSL CPE Terminated Session**

PPPoEoA implementation must allow multiple PPPoE sessions of the same subscriber to be carried over a single ATM PVC.

## Hardware Support

This feature is supported on ATM MDA on:

- 7750SR (SR-12, SR-7) platforms
- 7750 SR-c4/12 platforms
- 7450 platforms in Mixed Mode.

This feature is implemented only on IOM3 based hardware.

ATM MDAs:

- 16 port ATM OC-3/STM-1 (single rate) –
- 4 port ATM OC-3/12c/STM-1/4c (dual rate on a per port basis; oc3/stm1 or oc12/stm4; port speed can only be changed in groups of four ports).

Chassis modes B,C and D are supported.

PPPoEoA/PPPoA will NOT be supported on the following modules:

ASAP MDAs:

- 4 port channelized OC-3/STM-1 (IOM2-20G and IOM3-XP)
- 1 port channelized OC-12/STM-4
- 12 port channelized DS3/E3 (coax)
- 4 port channelized DS3/E3 (coax)

ATM CMA:

- 8 port T1/E1 ATM (RJ-48)

The 7750-c4/12 currently supports only the four port ATM MDA.

## Termination Points within 7x50

PPPoA/PPPoEoA sessions are terminated on the access ATM SAP in IES and IP-VPN service context through subscriber/group interfaces.

However, for the wholesale/retail deployment scenarios, the ATM VCs are terminated on the LAC while the PPP(oE) sessions will be terminated on the LNS.

PPPoA/PPPoEoA is not supported in the wholesale/retail VRF model (wholesale VRF + retailer VRF). However it is supported in wholesale/retail MSAP model (capture SAP in VPLS that is mapped into a VRF).

## PPPoA Encapsulation

PPP frames are transported over ATM using ATM AAL5 framing mechanism. This is defined in RFC 2364. In short, each PPP packet is appended by an 8-octet AAL5 trailer with some control information (16-bit length field and a 32-bit CRC being the most important). This new frame is not self-identifying, in other words, kind of payload it carries (PPP, IPv4, IPv6, ARP, MPLS, etc.) can not be identified. This means that if you want to send it as such, it can carry only a single protocol type which must be agreed upon in advance by configuration at each side of a PVC connection. Only then will both ends of the connection be able to recognize the payload type inside of it.

To add more flexibility to the payload type and allow multiple protocol types to be multiplexed within a single VC session, an additional header must be defined in the AAL-5 packet. This header which allows protocol multiplexing over ATM VC is called Link Layer Control (LLC) header. PPP transport over ATM using LLC is defined in RFC 2364, *PPP over AAL5*. A more extensive version for multiplexing protocols over VC is defined in RFC 2684 “Multiprotocol Encapsulation over ATM AAL5” (LLC/SNAP encap).

7x50 supports these two types of PPPoA encapsulation:

1. LLC (RFC 2364)

   LLC header is extended with a NLPID (Network Layer Protocol Identifier) which identifies the protocol type inside of the AAL5 frame. NLPID for PPP in LLC is 0xCF.

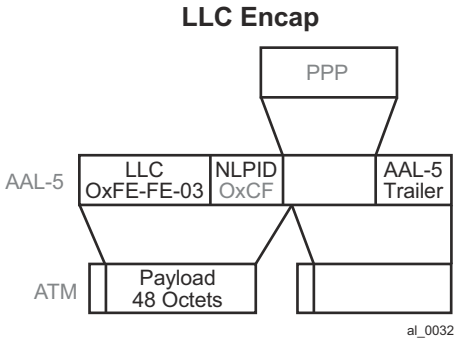



**Figure 78: PPPoA LLC Encapsulation**

Possible CLI syntax:

```
encapsulation aal5nlpid-ppp
```

The **ppp** keyword in the **aal5nlpid-ppp** command is used to indicate that ppp is the only encapsulation that currently supported in NLPID.

SNAP is an extension to LLC for protocols that are not defined in LLC NLPID. PPP protocol identifier is defined within NLPID and therefore it does not need additional SNAP header (this reduces the overall AAL5 header overhead).

LLC/SNAP encapsulation is defined in RFC 2684, *Multiprotocol Encapsulation over ATM Adaptation Layer 5*, and it is needed for PPPoEoA encapsulation.

2. VC-MUX

In VC-MUX mode there is no additional (LLC/SNAP) header used for protocol multiplexing. Instead, VC endpoints must agree before hand on the payload type that they will transport. For PVCs this is done during the provisioning phase on each side of the connection. For example:

encapsulation aal5mux-ppp

aal5mux-ppp will tell each end of the VC in advance that the payload inside of the AAL5 frame is PPP.

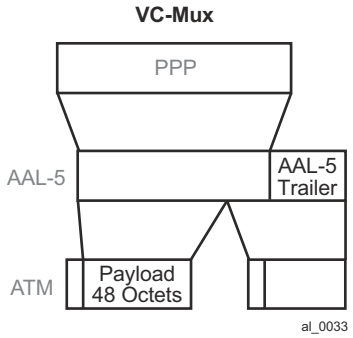



**Figure 79: PPPoA AAL5MUX Encapsulation**

## PPPoEoA Encapsulation

Similar to PPPoA, two types of encapsulation are defined for PPPoEoA:

- Protocol multiplexing (Layer 2 bridging defined by additional headers - LLC/SNAP).
- VC-Multiplexing where payload type within a VC is agreed upon in advance by static configuration.

Both types are supported in our implementation:

1. "Multiprotocol Encapsulation over ATM AAL5" (LLC/SNAP) defined in RFC 2684.

   This encapsulation adds support for protocols that are currently not defined in the LLC header. There are two basic types of encapsulations defined under this RFC:

   – Routed

   – Bridged

   PPPoEoA encapsulation is defined as 'Bridged' where Layer 2 information is preserved through transitioning between two different Layer 2 network types (Ethernet -> ATM) which is shown in Figure 80.

   0xAA-AA-03 in the LLC header indicates the presence of the SNAP header.

   0x00-80-C2 in the OUI indicates that a bridged PDU is encapsulated.

   0x00-01 or 0x00-07 in PID indicates that the encapsulated Layer 2 network type is 802.3 Ethernet with or without preserved FCS.

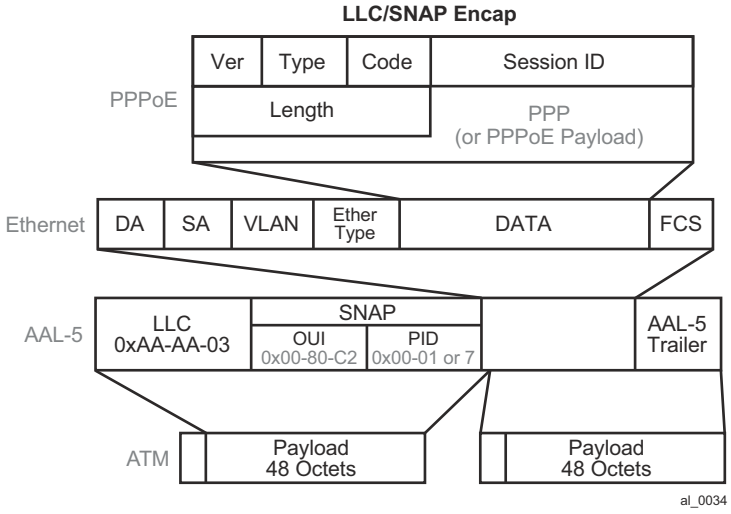



**Figure 80: PPPoEoA Bridged LLC/SNAP Encapsulation**

The CLI syntax is:

```
encapsulation aal5snap-bridged
```

2. VC-MUX

   In the VC-MUX mode there is no additional (LLC/SNAP) header used for protocol multiplexing. Instead, the VC endpoints must agree before hand on the payload type that they will transport. For PVCs this is done during the provisioning phase on each side of the connection. For example:

   ```
   encapsulation aal5mux-bridged-eth-nofcs
   ```

   `aal5mux-bridged-eth-nofcs` tells each end of the VC in advance that the payload inside of the AAL5 frame is an Ethernet frame. In this case, it accepts the frame and treat it as an Ethernet frame inside AAL5. The EtherType within the frame must be set to 0x8863 (PPPoE Discovey Phase) or 0x8864 (PPPoE Session Phase).

   The '-nofcs' portion indicates that the FCS is not supported on those Ethernet frames.

   PPPoEoA encapsulation is shown in Figure 81.

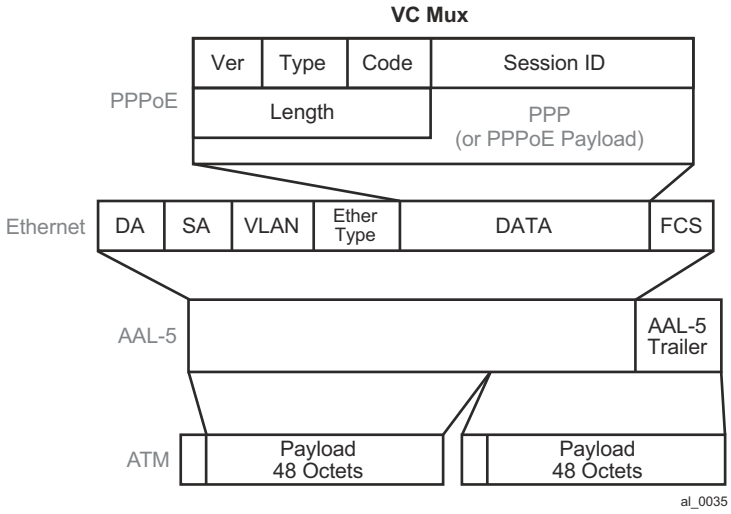



**Figure 81: PPPoA AAL5MUX Encapsulation**

## Encapsulation Summary

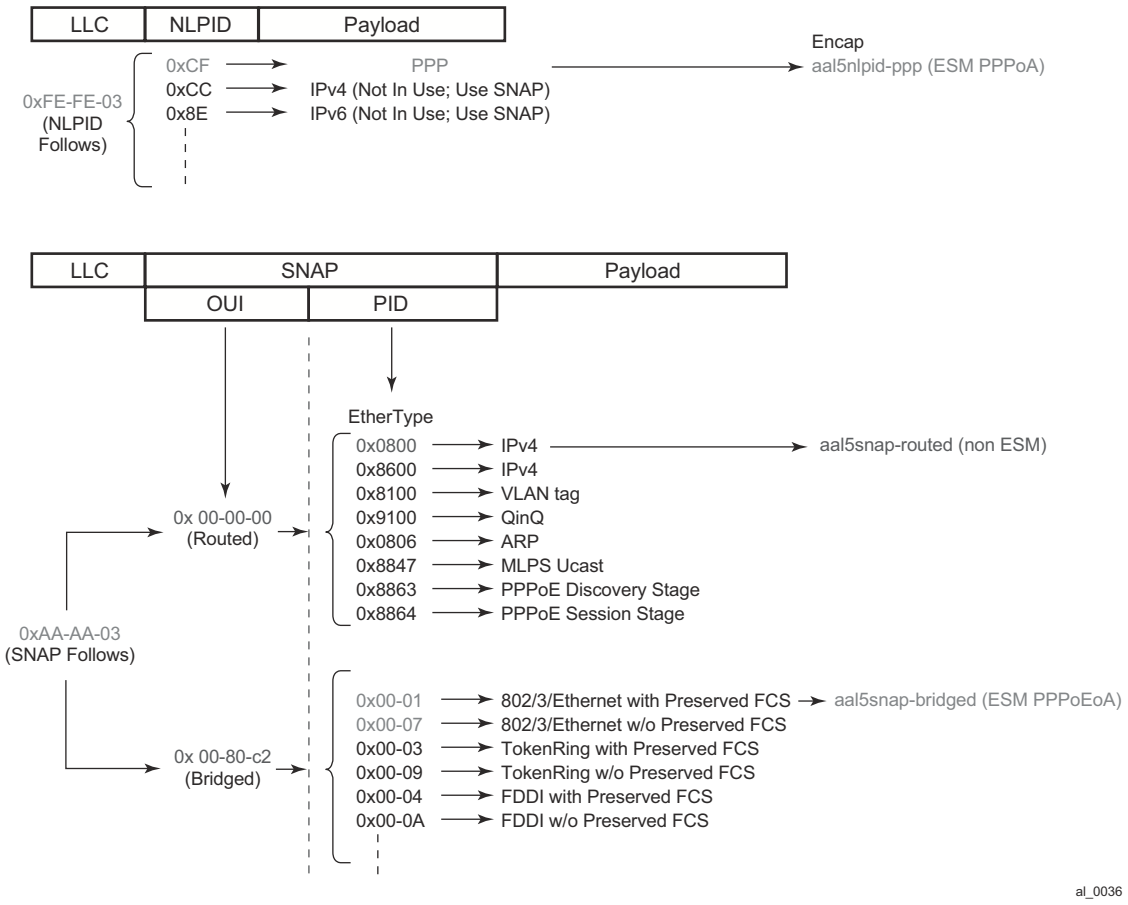



**Figure 82: LLC/SNAP Encapsulation**

There is a 0-2Byte additional padding (optional) in the SNAP Bridged encapsulation header that is not shown in Figure 82. This padding, according to RFC 2684, is necessary to align the info field (payload of the Layer 2 encapsulated frame) on a 4-Byte boundary.

At the end the following types of ATM encapsulation are supported on 7x50 in ESM:

`aal5snap-bridged`

This encapsulation type is used for PPPoEoA encapsulation with FCS or without FCS on ingress. On egress only frames without MAC FCS can be sent. PPPoE session type will be determined based on the EtherType in the Ethernet frame.

`aal5mux-bridged-eth-nofcs`

This encapsulation type is used for mux PPPoEoA sessions without MAC FCS.

`aal5nlpid-ppp`

This encapsulation type is used for LLC/NLPID PPPoA encapsulated packets.

`aal5mux-ppp`

This type encapsulation is used for PPPoA traffic without LLC/SNAP header (VC-MUX).

`aal5auto` (new command)

This encapsulation type is supported in ESM and it is used for auto detecting the encapsulation type. This is also called autosensing.

All hosts for the same subscriber will use the same encapsulation type.

## Concurrent Support for Different Service Types on the Same Port

An ATM port on 7x50 can concurrently support various service types. One service type can be mapped only to one VC. This would normally be the case if there is an aggregation network in front of 7x50. A Service Provider could run a variety of services over this aggregation network (ATM switches) connected on the same physical port within the 7x50.

For example, the following services can be run on the same physical port:

- PPPoA/PPPoEoA sessions connecting residential/business customers through a DSLAM to 7x50
- Providing access for xPIPE services
- Plain aal5snap-ip or vc-mux-ip PVC for Internet access, etc.

## Restrictions in Scaled ATM MDA Mode

**Note:** ATM concatenation mode for Apipe is not supported in the 16K VC mode.

In the concatenated mode, cells are delayed so that they can be concatenated and delivered in a single packet over to pseudowires to the other side. Without concatenation, each cell is transported individually. The implication is that each cell is individually encapsulated into Eth/MPLS which results in wasted bandwidth on the link.

Support for the concatenated ATM pseudowires is not removed from the CLI in the 16K mode of operation:

```
configure
    service apipe <id> [vc-type <cell-type>]
         spoke-sdp <sdp:pw> cell-concatenation
              [no] aal5-frame-aware
              [no] clp-change
              [no] max-cells
              [no] max-delay
```

Instead:

- Adding an ATM port or a connection profile on an MDA in 16-VC mode to an APIPE is disabled if the vc-type is set to atm-cell AND cell-concatenation is enabled.

  ```
  A:BNG>config>service>apipe# sap x/y/z:cp.w create
  MINOR: SVCMGR #2603 Cell-concatenation is not allowed on 16k VC-mode ATM MDAs
  ```

- cell-concatenation on an APIPE of the vc-type atm-cell is disabled if it already contains an ATM port or a connection profile.

  ```
  A:BNG>config>service>apipe>spoke-sdp>cell-concat# max-delay X
  MINOR: SVCMGR #2603 Cell-concatenation is not allowed on 16k VC-mode ATM MDAs
  ```

Note that regular VPI/VCI SAPs (sap:vpi/vci) are not allowed to be configured on an Apipe of vc-type atm-cell.

Cell-concatenation is supported on Apipe services with a VC on an ATM MDA in the 8-VC mode.

AAL5 SDU mode is continued to be supported.

---

## QoS Implementation

In addition to our system QoS that is provided in the “Q” chip on IOM, the ATM MDA offers QoS capabilities at the ATM cell level. In the context of this document, the system QoS is referred to as ‘IOM QoS’ and to ATM MDA provided QoS simply as ‘ATM QoS’.

Both, IOM QoS as well as ATM QoS defined by Traffic Descriptors working at the cell level, play a role in the overall QoS for the SAP (or virtual circuit). ATM QoS defines rates of each VC stream and defines the behavior under port congestion. IOM QoS defines bandwidth allotment and the scheduling scheme for each service within a VC stream.

In general, both MDAs, ATM and ASAP, support four traffic categories:

- CBR
- Rt-VBR
- Nrt-VBR
- UBR

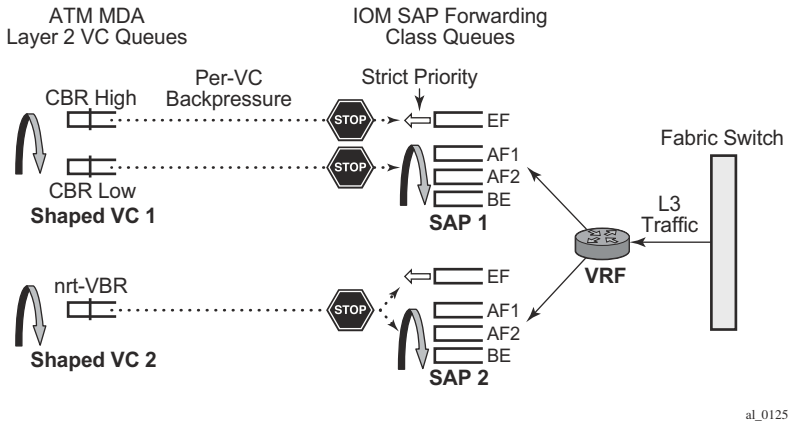



**Figure 83: Scheduling on ATM MDA**

Each ATM traffic category is defined by a set of parameters, such as PIR, SIR, MIR, MBS and CDTV.

Currently, policing is supported only on ingress for CBR and VBR traffic classes. CBR traffic class police at PIR, while rt/nrt-VBR police at SIR.

Shaping is supported only on egress for CBR and rt/nrt-VBR traffic classes. CBR shapes traffic at PIR, while rt/nrt-VBR shape traffic at SIR. Egress shaping can be disabled only for nrt-VBR traffic class.

Scheduling at the ATM layer is shown in Figure 83. Shaped CBR and VBR traffic classes have two queues, an HP and an LP queue. Packets from IOM are marked according to scheduling priority of the Forwarding Class (expedited or best-effort) from which they were sent and are accepted into the ATM VC queue (HP | LP) accordingly. For example, at the IOM level packets from an expedited FC (queue) are marked as HP, and the packets from a best-effort FC (queue) are marked as LP. When these packets arrive to MDA, they will be admitted into appropriate queues, pending VC (or port) buffer availability.

Non-shaped VBR and UBR have only one queue at the MDA level.

In terms of the ATM QoS scheduling, CBR has the highest scheduling priority (strict priority) followed by rt-VBR (also strict priority). The remaining traffic classes (non-shaped VBR and UBR) are serviced in WRR fashion, the weight being their configured SIR (non shaped nrt-VBR) or MIR (UBR) rate.

There are several areas of QoS that are addressed in relation to integration of IOM QoS and ATM QoS:

- Association between the subscriber and ATM VC traffic descriptor.
- Rate adjustments between Layer 3 and ATM QoS rates due to difference in frame/cell sizes on which they operate.
- Per VP shaping.

## Association Between the Subscriber and ATM VC Traffic Descriptor (QoS)

Each PPPoA PVC is of a certain traffic class – CBR, rt/nrt-VBR or UBR/UBR+MIR which along with other parameters is defined in the ATM traffic descriptor profile. For ATM service categories and traffic descriptors. There should be a uniform mapping between service offered and L3 and ATM QoS.

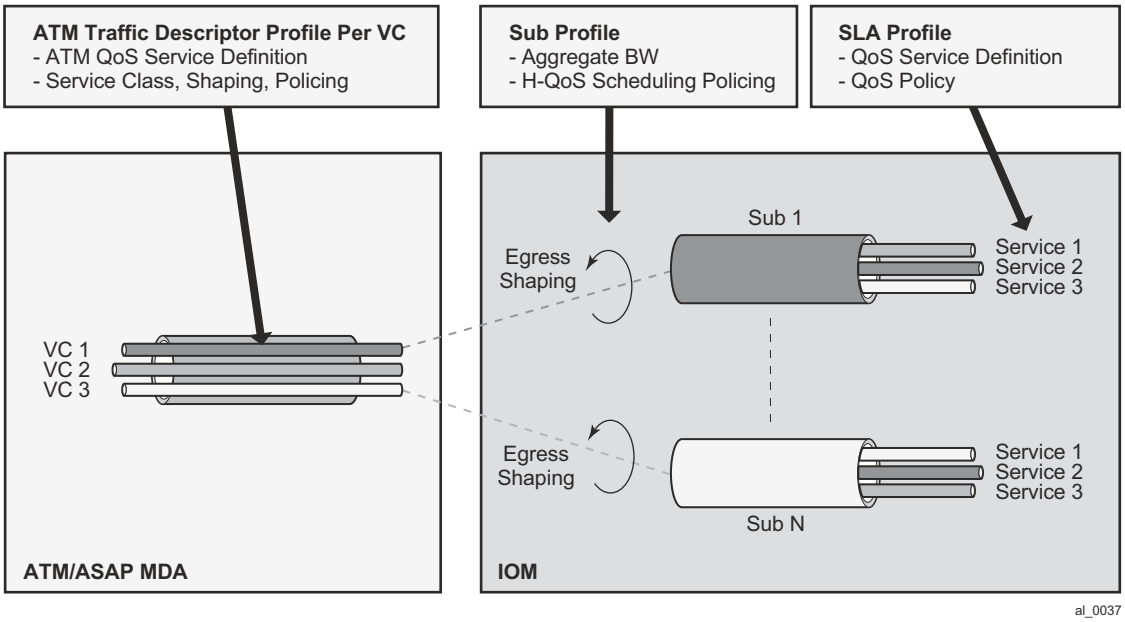



**Figure 84: ATM Traffic Descriptor Association with Subscriber**

The ATM traffic descriptor (atm-td) is applied to a VC under the SAP CLI hierarchy. In this fashion, MDA related QoS (ATM QoS) is referenced outside of the subscriber context (SUB/SLA-profiles).

The following describes the operability:

atm-td for the VC is applied under the SAP:

```
configure
         services ies/vprn
              subscriber-interface <sub-if-name>
                   group-interface <grp-if-name>
                        sap <sap-id>
                             atm
                                  ingress
                                       traffic-desc <id>
                                  egress
                                       traffic-desc <id>
```

and for MSAP:

```
con  figure
         subscriber-management
              msap-policy <name>
                   atm
                        ingress
                             traffic-desc <id>
                        egress
                             traffic-desc <id>
```

msap-policy is then invoked via LUDB, RADIUS or the default-msap-policy under the capture SAP.

This allows 7x50 to have preconfigured MSAP policies, each corresponding to a specific VC type with its own traffic-class parameters (CBR/rt-nrt-VBR/UBR.). Each subscriber with corresponding hosts is then associated with a msap-policy that determines the VC type.

A more flexible way to go about this would be to allow the subscriber to reference the atm-td directly by the sub-host at the host creation time, independently of the msap-policy. This is supported in the following manner:

- atm-td is referenced via a RADIUS VSA in the Access-Accept message.
- There are two VSAs, one for ingress and one for egress atm-td:
  - alc-ingress-atm-td
    (ATTRIBUTE Alc-ATM-Ingress-TD-Profile 128 integer)
  - alc-egress-atm-td
    (ATTRIBUTE Alc-ATM-Egress-TD-Profile 129 integer)
- Only the first host of the subscriber can overwrite the atm-td that is defined under the msap-policy or under the static SAP.
- Consecutive hosts (second, third, etc.) of the same subscriber will not have any effect on the atm-td of the VC. For example, if the second host tries to overwrite the existing atm-td with a different atm-td, it will fail. Once the atm-td is set via the Access-Accept message for the first host, it cannot be changed as long as the subscriber is active in the system. This implies that all hosts of the same subscriber will have the same atm-td.

- If another hosts of the same subscriber initiates a session with a different atm-td name in the Access-Accept message, the host will be created but the atm-td for the VC will not be changed. A trap/syslog will be generated with a 'VC-parameter-mismatch:VSA ignored' info message.
- atm-td change will not be supported through CoA.

The following guidelines should be followed in configuring ATM traffic descriptors:

- ATM policing is the only function that can discard ATM cells on an ATM MDA
- Ingress:
  - ATM policing can only be enabled on ingress.
  - Policing is disabled by default.
  - Policing can only be enabled for the following traffic categories: CBT, rt-VBR and nrt-VBR.
  - CBR police at PIR. CBR is CLP transparent (it does not look at the CLP bit – aggregate traffic policing).
  - rt-VBR and nrt-VBR police at SIR and can operate either in CLP Significant (CLP marking) or CLP transparent mode (only relevant for Apipes and not for ESM).
  - In CLP significant mode with policing enabled, traffic is policed at the PIR value, but it can be marked with CLP 1 if it exceeds the SIR rate (only relevant for Apipes and not for ESM).
  - Ingress classification: the service category and CLP bits only affect Apipe traffic. If ATM traffic is terminated in the BNG, rely on the existing QoS classification implemented in the IOM.
- Egress:
  - Shaping is supported only on egress. Traffic shaping is supported for CBR, rt/nrt-VBR service categories.
  - ATM MDA is cell "lossless" – It sends backpressure to IOM which drops packets.

## Per VP Shaping

VCs traversing the same DSLAMs will typically use the same VPI. To prevent overrunning DSLAM capacity (intermediate destinations) in case an aggregation network is in place (between BNG and a DSLAM), per VP shaping will be implemented.

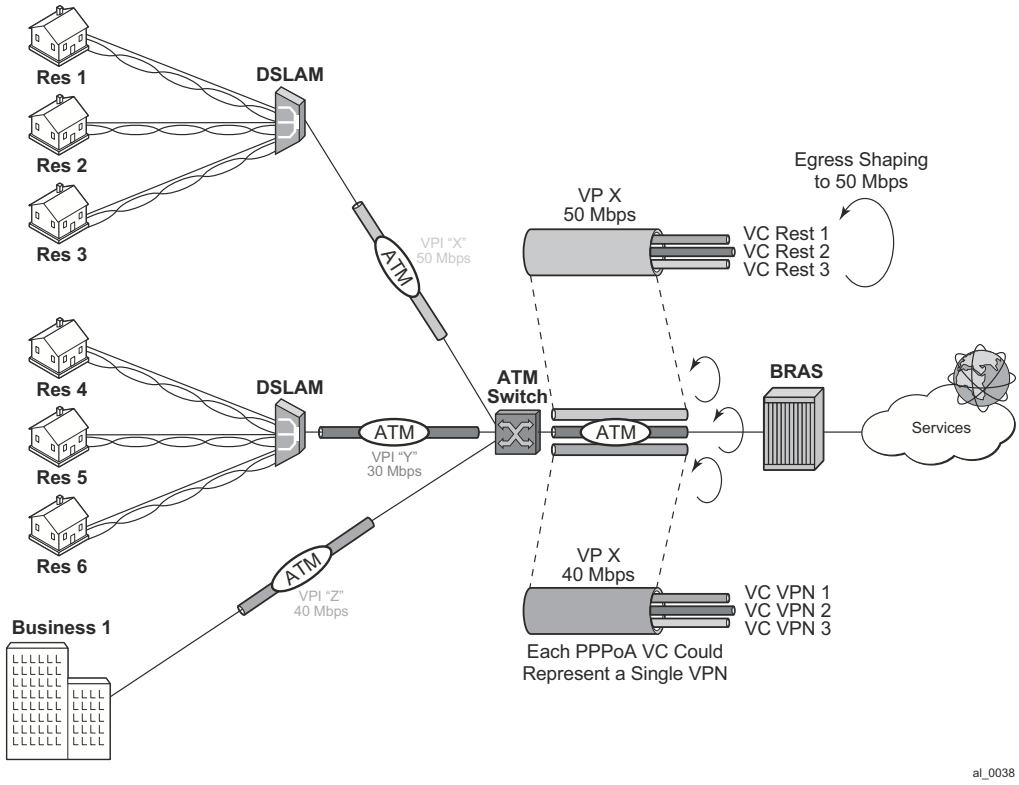



**Figure 85: VP Shaper**

The shaping rate per VP and the VP service type is provisioned manually via a traffic descriptor on a per port level. The cli syntax is:

```
config
    port <port-id>
        sonet-sdh
            path
              atm
                vp <vpi> egress-traffic-desc <atm-td-profile-id>
```

Where the vpi is a VPI identifier and the egress-traffic-desc is the traffic descriptor id. Only traffic descriptors with service-category of cbr, rt-vbr and nrt-vbr can be used in VP Shapers. However, only the rate in the traffic descriptor can be changed on-the-fly, and not the service category (nrtVBR, rtVBR, CBR). A VP shaper can be added to or removed from active VCs.

CBR VP Shaper shapes cells at the exact PCR rate. There is no burst concept in CBR shaping. Excessive traffic is back-pressured towards the IOM (Q-chip).

The IOM will never send more frames to the ATM MDA than the MDA cannot buffer. This is implemented through a combination of software and hardware backpressure mechanisms. This backpressure mechanism utilizes a combination of hardware and software. Software backpressure aims to have around 100ms of traffic queued against a VC based on its configured shaping/ scheduling rate, but being a software mechanism that is only a guideline. As ATM MDA detects more traffic that it can accept, a hardware backpressure is exerted.

A rt/nrt-VBR type VP Shaper has three parameters associated with it: PCR (peak cell rate), SCR (sustained cell rate) and MBS (maximum burst size at a peak rate). As long as there is enough MBR credit, traffic will be shaped at the PCR rate. Once all MBR credit (burst) is exhausted, traffic will be shaped at the SCR rate. Bursting above the SCR is configurable via the MBS parameter.

In both cases cells will be spaced at 1/PCR or 1/SCR as perfectly as possible with minimum jitter.

VP shaping is supported only when the ATM MDA is in max16k-vc mode. The maximum number of VP shapers per MDA to 128.

The maximum number of VCs that can feed into a single VP shaper is 16K. This includes the sum of all VC-ranges on the VP plus any statically configured VC on that VP.

VCs within the VP tunnel is serviced by a single scheduler assigned to each VP tunnel. The ATM VP shaper will condition the aggregate traffic for all ATM VCs within the VP tunnel. VCs within the shaped VP tunnel are degraded from the originally assigned service category to a common UBR service category (default traffic descriptor). If the VP shaper is removed from the VCs, the VCs will be reverted to their original service category. Scheduling between VCs will be WRR based with a weight parameter that is explicitly configured. The weights assigned to VCs within the VP tunnel are in range 1-255. By default, VCs are assigned a priority based on the originally assigned service category:

- VC degraded from CBR = wight 10
- VC degraded from rt-VBR = weight 7
- VC degraded from nrt-VBR = weight 5
- VC degraded from UBR+ = weight 2
- VC degraded from UBR = weight 1

The weight parameter is user configurable under the traffic descriptor hierarchy.

```
configure
     qos
          atm-td-profile <td-profile-id> [create]
               weight <weight>

<weight> : 1-255.
```

If weight is not specifically configured, the defaults are taken as described above.

The explicitly configured weight parameter is honored only on ATM MDA in the max16k-vc mode. On all other ATM capable MDAs (ASAP or ATM MDA in max8K-VC mode), the weight parameter is ignored.

Note that in the current ATM implementation there is already a WRR scheme in place based on internally calculated weights. This WRR scheme is used to service traffic from the VC queues of equal priority (where there are two queues per VC - a HiPrio and a LowPrio queue). Weights are assigned to VCs automatically based on the rate of the VC,

## ATM/IOM QoS Integration

There are major differences between the QoS mode of operation at the ATM MDA level and the IOM level.

ATM QoS operates on fixed size cells that contain additional transport overhead. In addition, ATM shaping is very accurate so that traffic is paced into the ATM network with nodes that are sensitive to bursts (ingress policing). Buffering and service differentiation (number of queues) at the ATM layer is not as flexible as it is on the IOM level.

On the other hand, HQoS in the IOM is less accurate and less responsive to sudden traffic fluctuations (bursts). It operates on Layer 2 frame lengths. Bursts of traffic are usually let into the network more freely than the ATM network would like to accept.

To combine the extensibility of IOM HQoS with the stringent ATM QoS requirements, the two modes of operation are integrated.

The congestion in the ATM network is treated by using extensive IOM HQoS.

When the VC ATM queue becomes congested, it exerts backpressure to the subscriber queues in the IOM on the corresponding VC.

To avoid the condition where a VC becomes overly oversubscribed and excessive in exserting the backpressure, our HQoS in IOM has to be proactive and has to be able to detect congestion on the IOM level based on the ATM VC configured rate or the VP configured rate. IOM HQoS must deal with this congestion before it actually becomes the problem in the ATM layer. Occasional and short-lived backpressure from VCs, but persistent QoS backpressure and congestion at the ATM layer is avoided. A prerequisite for treating ATM congestion at the IOM level is to adjust the frame size in HQoS calculations (on all levels - including queues and aggregate rate limits) so that it reflects the ATM overhead associated with ATM transport (AAL5 encap, cellification).

### Intermediate Node Rate Limit/Shaper

In many cases it is desired or even necessary to shape traffic per DSLAM. In the ATM network, a DSLAM corresponds to a VP (Virtual Path) which is at the ATM cell level identified by a VPI (Virtual Path Identifier). The VPI/VCI pair represents the addressing mechanism in the ATM network.

Another level of hierarchy is introduced into our HQoS model - an aggregate rate limit in the IOM that will correspond to the ATM VP shaper. The purpose of this new construct in the IOM is to help detect congestion at the IOM level before it becomes a severe problem at the ATM layer.

As a final solution, our HQoS hierarchy in IOM is a 4 tier hierarchy consisting of:

- subscriber queues that are parented to the port scheduler
- egress agg-rate-limit that represents the egress aggregate subscriber rate
- tier-2 aggregate-rate-limit that represents the ATM VP rate limiter (related to DSLAM). This is a pure rate limiter and not a scheduler (it does not receive any scheduling opportunity).
- port-scheduler to which are subscriber queues parented. The port-scheduler delegates bandwidth to its children based on the priority levels associated with the children queues.

The four tiered hierarchy is needed to deal with:

- port congestion. The port bandwidth can be overbooked the desired service levels can still be ensured.
- ATM VP congestion. An ATM VP can be effectively protected from prolonged congestion that would result into significant backpressure to the IOM queues.
- subscriber congestion. Within the subscriber, bandwidth is managed within the subscriber bandwidth limit.

The four-tiered hierarchy looks like Figure 86.

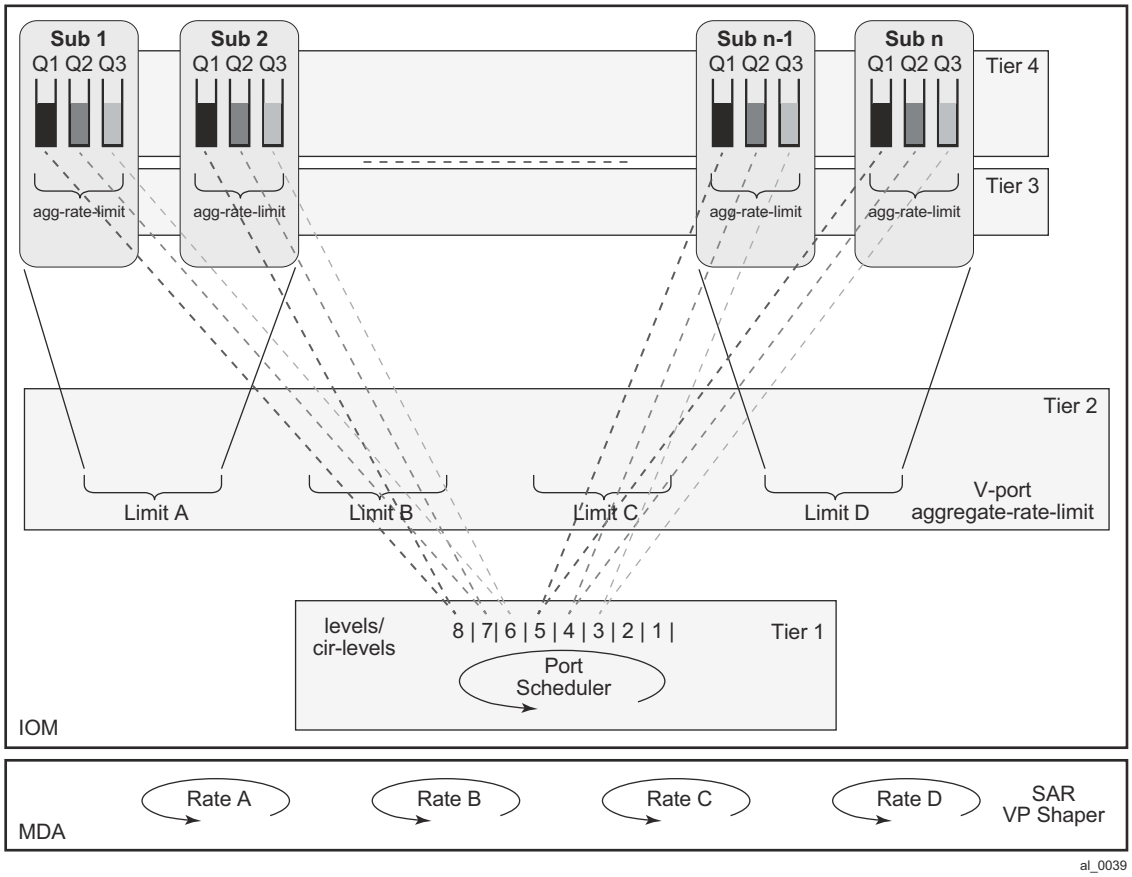



**Figure 86: Tier HQoS**

The key point in such HQoS model is that the port-scheduler delegates its available bandwidth to the subscriber queues directly according to the queue priority on a configured level. Higher priority queue are served over all subscribers before any lower priority queues, up to the limits imposed by the tier 1 and tier 2 aggregate rate limits.

## Provisioning Aspects

The tier 2 aggregate rate limit that corresponds to the VP shaper on the IOM level is provisioned in the context of a Vport. The Vport is a container that is configured directly under the port and it can contain either an aggregate rate limit or another port scheduling policy (port scheduler). These two constructs (Vport aggregate and V-port port scheduling policy) are mutually exclusive. In addition, if a V-port port scheduling policy is configured instead of the V-port aggregate, then the port scheduler cannot be used (currently two port schedulers applied under the same port cannot be used in a parent/children relationship).

The V-port aggregate have to have the agg-rate-limit defined explicitly. In other words, the rate can NOT be implicitly inherited by configuration from the VP shaper traffic descriptor.

Example:

```
config
    port <port-id>
        sonet-sdh
            path
                egress-scheduler-policy <port-scheduler-policy-name>
                access
                    egress
                        vport <vport-name>
                            description <description-string>
                            host-match dest <dest-string>
                            agg-rate-limit <agg-rate>
                            port-scheduler-policy <port-scheduler-policy-name>

                atm
                    vp <vpi> egress-traffic-desc <atm-td-profile-id>
```

The association between the vport aggregate and the subscriber host is done in three ways:

- via a RADIUS attribute - a dest-string VSA that is matched against the string defined under the corresponding Vport. This dest string VSA is returned via RADIUS at the subhost instantiation time.
- via LUDB – similar to the RADIUS method. The dest-string comes from the LUDB instead from RADIUS.
- via the VP identifier (VPI) - The VPI is known to the CPM from the beginning of the session initiation process – raw (unknown encapsulation) packets and passed to the CPM. These raw packets contain VPI,VCI identifiers. The vport container for the subscriber is referenced implicitly via the VPI in the following fashion (Figure 87):
  - CPM determines the VPI from the first packet for the session.
  - ATM VP Shaper mapping: The VPI is used to make the subscriber association with the VP Shaper which is defined under the VP Shaping node: **port>sonet-sdh>path>atm>vp**. The VP Shaping node name must be the VPI number.
  - Vport mapping: The vport-name in the **port>sonet-sdh>path>access>egress>vport <vport-name>** hierarchy is matched against the VPI number.

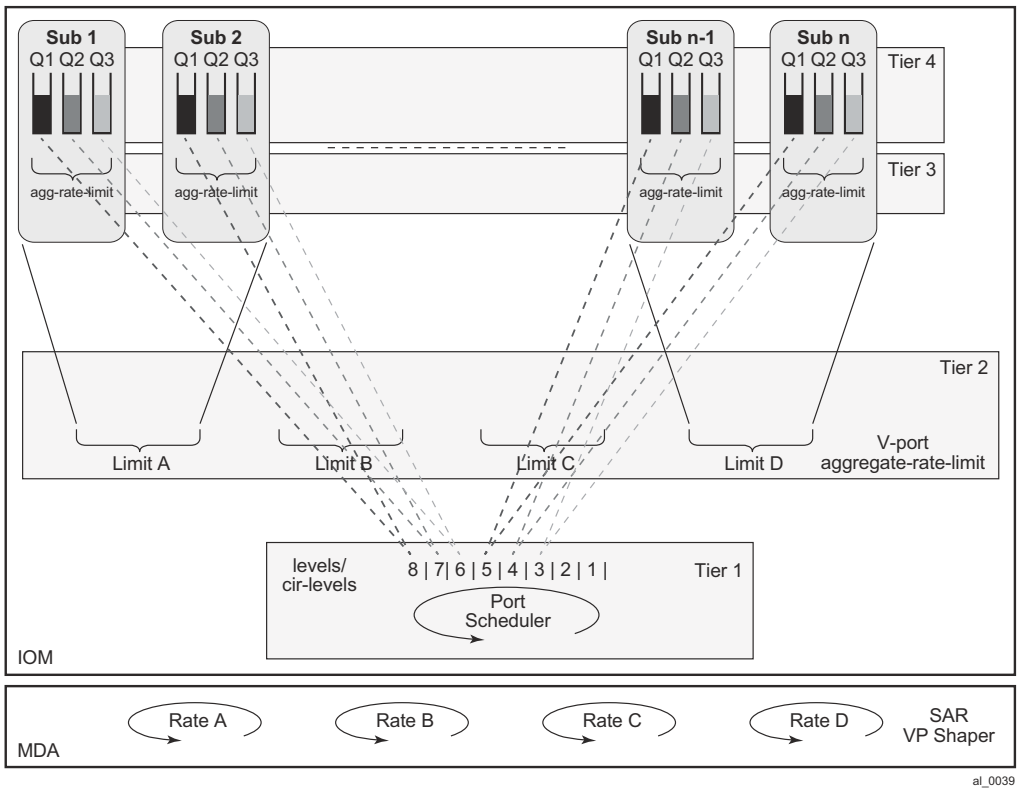



**Figure 87: VPI Based V-Port <-> Subscriber Association**

The association method (automatic via VPI or based on RADIUS/LUDB) between the subscriber host and the vport is defined under the SAP (or MSAP) where the subscriber resides. In most cases subscriber management is used with MSAPs.

This is the syntax:

```
configure
subscriber-mgmt
msap-policy <name>
sub-sla-mgmt
def-inter-dest-id string <inter-dest-string>
def-inter-dest-id {use-top-q  | use-vpi}

configure
     services ies/vprn
          subscriber-interface <sub-if-name>
               group-interface <grp-if-name>
                    sap <sap-id>
                         sub-sla-mgmt
                              def-inter-dest-id string <inter-dest-string>
def-inter-dest-id {use-top-q | use-vpi}
```

The **def-inter-dest-id** stand for a ‘default inter-destination identifier’.

If the use-vpi method is used and the VPI derived from the incoming traffic points to a non-existing VP container, the association between the subhost and the V-port container will fail and a message/trap is logged. This however will not prevent the creation of the subhost that can be parented to the port-scheduler.

If the use-vpi is used on an Ethernet port where this parameter is not applicable, the parameter is ignored and defaulted to ‘string’. A message/trap is logged.

Note that if the vport contains an aggregate-rate-limit, then there is no need for the indication of a vport construct in the sub-profile or sla-profile of the subscriber. On the contrary, in case that the vport contains a port-scheduling-policy, the sla-profile template must contain the indication that the subscriber is tied to a vport (**configure>subscriber-mgmt>sla-profile>egress>qos sap-egress-qos-policy-id vport-scheduler**.)

## HQoS Combinations

These are the possible HQoS combinations that are supported:

- port-scheduler assigned to path, agg-rate-limit assigned to vport, agg-rate-limit assigned to subscriber.
- port-scheduler assigned to path, no vport, agg-rate-limit or scheduler-policy assigned to subscriber.
- no port-scheduler assigned to path, port-scheduler assigned to V-port, agg-rate-limit or scheduler-policy assigned to subscriber.
- no port-scheduler assigned to path, no vport, scheduler-policy assigned to subscriber.

In case that the vport contains the agg-rate-limit, any subscriber host queue that is parented to a virtual scheduler will not be rate-limited by the vport aggregate rate. The queue will compete for bandwidth directly on the port's port scheduler, at the priority level and weighted scheduler group the virtual scheduler is port-parented to. If the virtual scheduler is not port-parented or if there is no port scheduler policy on the port, the host queue will be orphaned and will compete for bandwidth directly based on its own PIR and CIR parameters.

## ATM Rate Adjustment

The difference in cell/frame overhead on the ATM level and IOM Level (Layer 2) lead to inconsistent behavior in integrated QoS. For example, IOM based QoS operates on Layer 2 frames (PPP header is included in rate calculations). On the other hand, ATM QoS operates on ATM 53byte cells. Each PPP frame with AAL5 overhead is segmented into 48-byte chunk cells prepended by a 5 byte ATM header. Assuming that ATM QoS operates on 53-octet cells, a significant rate discrepancy might arise from this difference.

For example, consider two flows:

Flow 1, 1000pps, PPP+IP packet size 190Bytes => L2 (IOM) QoS rate is 1520 kbps

Flow 2, 1000pps, PPP+IP Packet size 232Bytes => L2 (IOM) QoS rate is 1856 kbps

Assuming that VC MUX encap is used, and that AAL5 trailer (8 octets + padding) is used:

Flow 1 = AAL5 length is 190+8+48B_boundary_padding= 240Bytes => 5 ATM cells

Flow 2 = AAL5 length is 232+8+48B_boundary_padding= 240Bytes=>5 ATM cells

Each cell has a 5 byte header which translates into a 2120 kbps rate for each flow.

As a result, more bandwidth is required at the ATM layer than the QoS on the IOM level has budgeted. Moreover, the two flows appear to be of the same rate at the ATM layer, even though the IP rate of flow 1 is ~20% less than the flow 2.

From the example above, Flow 1 consumes 40% more bandwidth configured at the ATM layer than it is given in IOM QoS. The rate adjustment depends on the packet size which additionally complicates conversion. This issue is more pronounced with smaller packet sizes.

An example is shown in Figure 88.

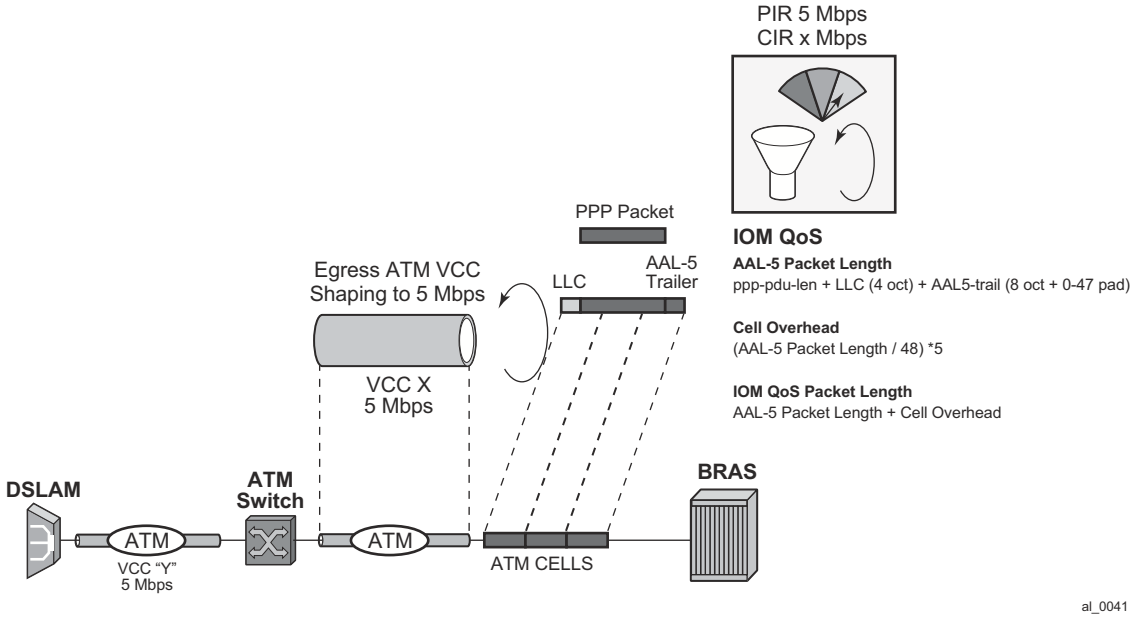



**Figure 88: QoS Adjustment**

Depending on the session encapsulation (VC-MUX, LLC/SNAP) the packet length on which IOM QoS operates is adjusted.

The rates are adjusted on a per queue level, per subscriber egress agg-rate-limit level, per V-port aggregate level and on per port-scheduler level. Because the ATM termination points are on the

BNG, there is a direct view of the encapsulation. The encapsulation information is supplied to the forwarding plane via the control plane.

The rate adjustment are examined in the following commands:

- **frame-based-accounting** under the qos>scheduler-policy hierarchy
- **queue-frame-based-accounting** under the sub-profile>agg-rate-limit hierarchy
- **avg-frame-overhead** under the queue hierarchy
- **encap-offset** under the sub-profile hierarchy

In PPPoA/PPPoEoA scenario there is no last mile rate adjustment (due to the difference in encapsulation) performed. It is assumed that the encapsulation in the last mile and in the intermediate mile is the same. In other words, it is assumed that the DSLAM does not add/change any encapsulation but instead it only acts as a VPI/VCI cross-connect. If the last mile encapsulation is PPPoA, then the intermediate encapsulation is considered to be PPPoA as well. Similar is valid for PPPoEoA encapsulation.

There are two configuration scenarios possible:

1. The **encap-offset** command in the sub-profile is configured. This command overwrites any other command related to rate conversion that might be configured (**frame-based-accounting**, **queue-frame-based-accounting**, **avg-frame-overhead**, or **avg-frame-size**).
   The *encap-offset* command forces dynamic wire rates calculation in the intermediate mile (directly connected ports) on all levels in the QoS hierarchy. The wire overhead in the intermediate mile takes into account the length of the fixed ATM encapsulation, the variable length of AAL5 encapsulation (including AAL5 48bit boundary padding) and the ATM cellification overhead. The queue stats are also wire based. All calculations are performed in the data plane using the actual packet size. In other words, this command will ensure that the rates on the queue level and the subscriber aggregate level (either through virtual schedulers or egress aggregate-rate limits) are wire based. Port-scheduler and V-port rates are already by default wire based rates and this cannot be changed.
2. The **encap-offset** command is not configured by default. In this case the other rate conversion related commands are in effect (**frame-based-accounting**, **queue-frame-based-accounting**, **avg-frame-overhead**, or **avg-frame-size**). The behavior is the following:
   - Wire rates in the intermediate mile (directly connected ports) are based on the **avg-frame-overhead** command which is provisioned via CLI. If avg-frame-overhead is not provisioned via CLI, by default it is assumed to be 0[%] and the wire rates effectively become data rates (IP payload + IP header + PPP(oE) header + fixed ATM encapsulation).
   - Queue stats (used in accounting) are always ‘data’ stats. This includes the IP Payload + IP header + PPP(oE) header + fixed ATM encapsulation.
   - *agg-rate-limit* (subscriber or vport) rates are always wire rates (as defined in the first bullet – based on the avg-frame-overhead).
   - Rates in the *port-scheduler-policy* (vport or physical port) are always wire based (rates (as defined in the first bullet – based on the avg-frame-overhead).
   - The **frame-based-accounting** command under the scheduler-policy will affect rate calculation for virtual schedulers and queues:

– If this option is configured, the virtual scheduler and queue rates will be wire based.

– if this option is NOT configured, the virtual scheduler and queue rates will be data rates

→ *queue-frame-based-accounting* configuration option under the subscriber agg-rate-limit command (in sub-profile) will affect rate calculations for queues. If this command is configured, the queue rates will be wire rates, otherwise they will be data rates

Avg-frame-size command in PPPoA/PPPoEoA is ignored.

Currently queue rates and subscriber virtual scheduler rates are allowed to be either data rates (one in Figure 89) or *on-the-wire-rates* (three in Figure 89). *Port-scheduler* rates, vport rates and the subscriber *agg-rate-limit* (in sub-profile) are always on *on-the-wire* rates.

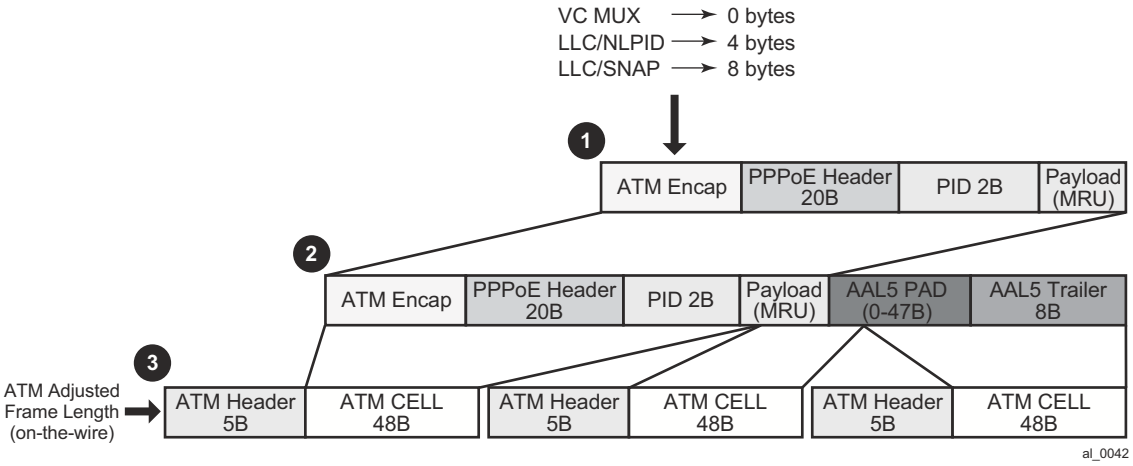



**Figure 89: ATM Wire Overhead**

## Subscriber Instantiation Use Cases

Note that multiple subscribers per single VC are not supported. However, multiple VCs per subscriber are supported.

The following displays examples of how subscriber hosts could be instantiated in PPPoA/PPPoEoA environment.

**Case 1**

- One host per subscriber.
- One ATM VC per subscriber.
- Authentication is done via RADIUS or LUDB.
- Authentication methods: PAP/CHAP or PADI for PPPoE.
- If there is no RADIUS/LUDB authentication, the subscriber-id should be by default the SAP (with VPI:VCI) with default sub-strings. For MSAP this would be configured under the msap-policy – all subs would have the same service in this case (SLA/SUB profiles).

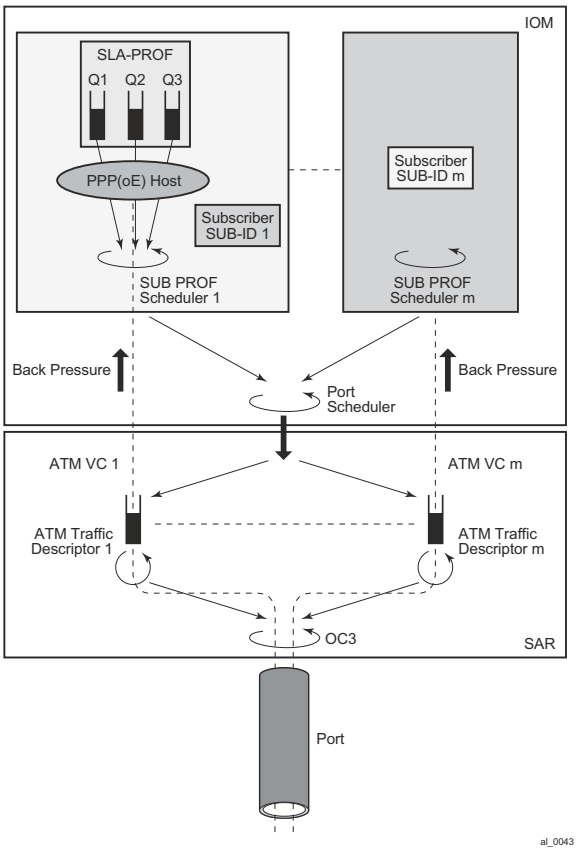



**Figure 90: Subhost per VC**

**Case 2**

- Multiple hosts per subscriber
- Single VC per host
- Multiple VCs per subscriber
- PPPoA or PPPoEoA
- Subhosts for the same subscriber can authenticate using a unique username or the same username

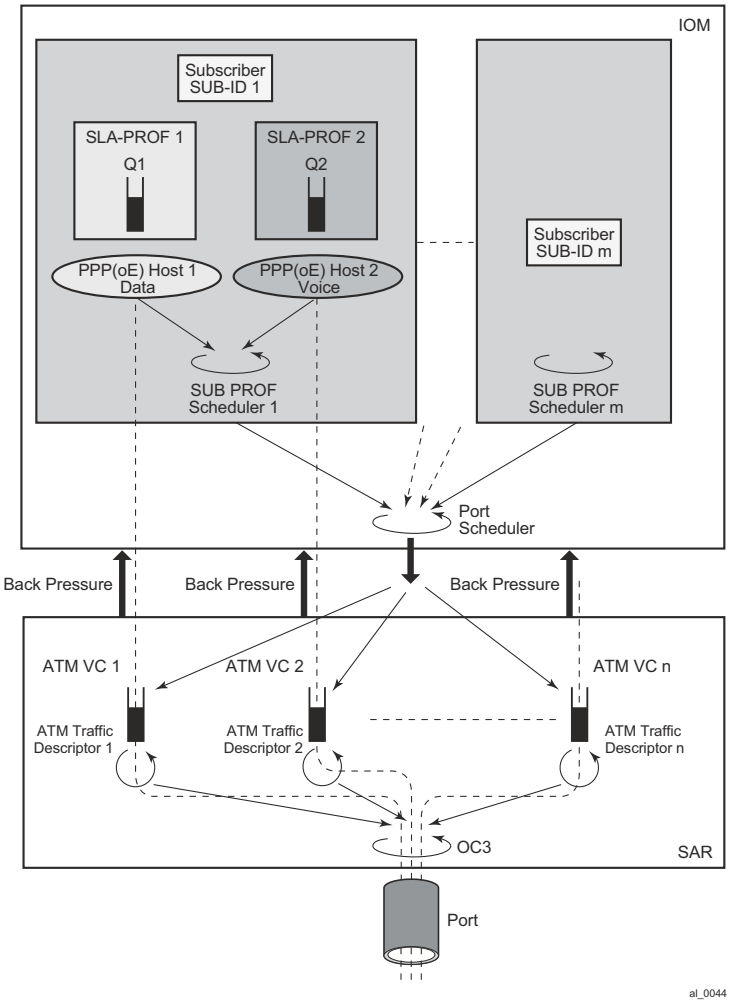



**Figure 91: Multiple VCs per Subscriber**

**Case 3**

- Multiple hosts per subscriber
- Singe VC
- PPPoEoA only
- Per host SLA-PROFILE instantiation

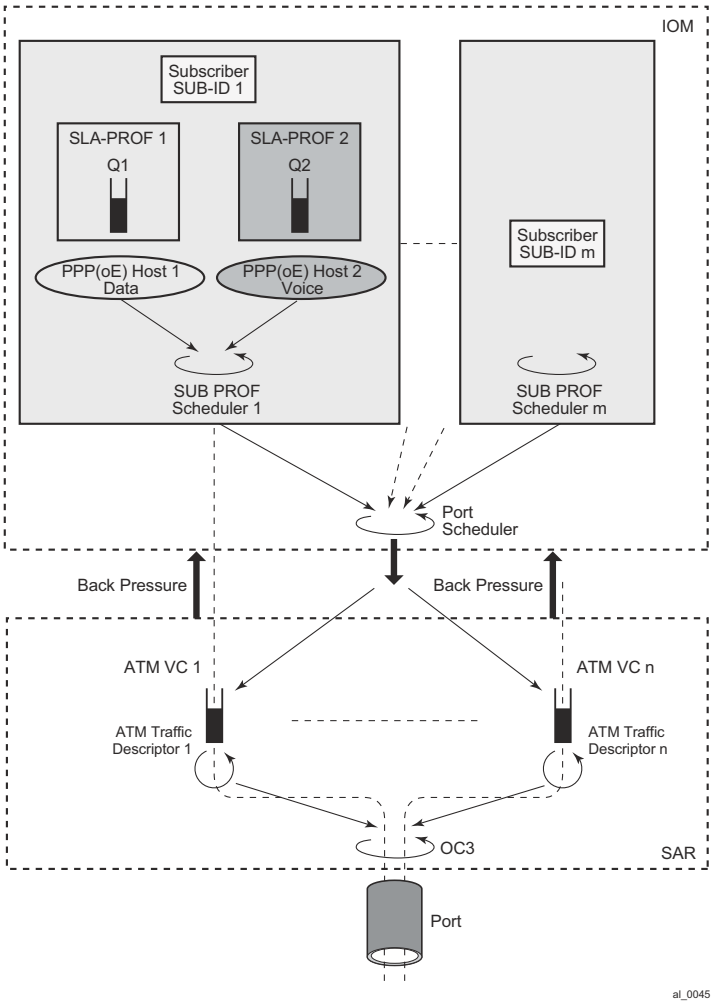



**Figure 92: Multiple Hosts per Subscriber, Single VC**

**Case 4**

- VP Shaping

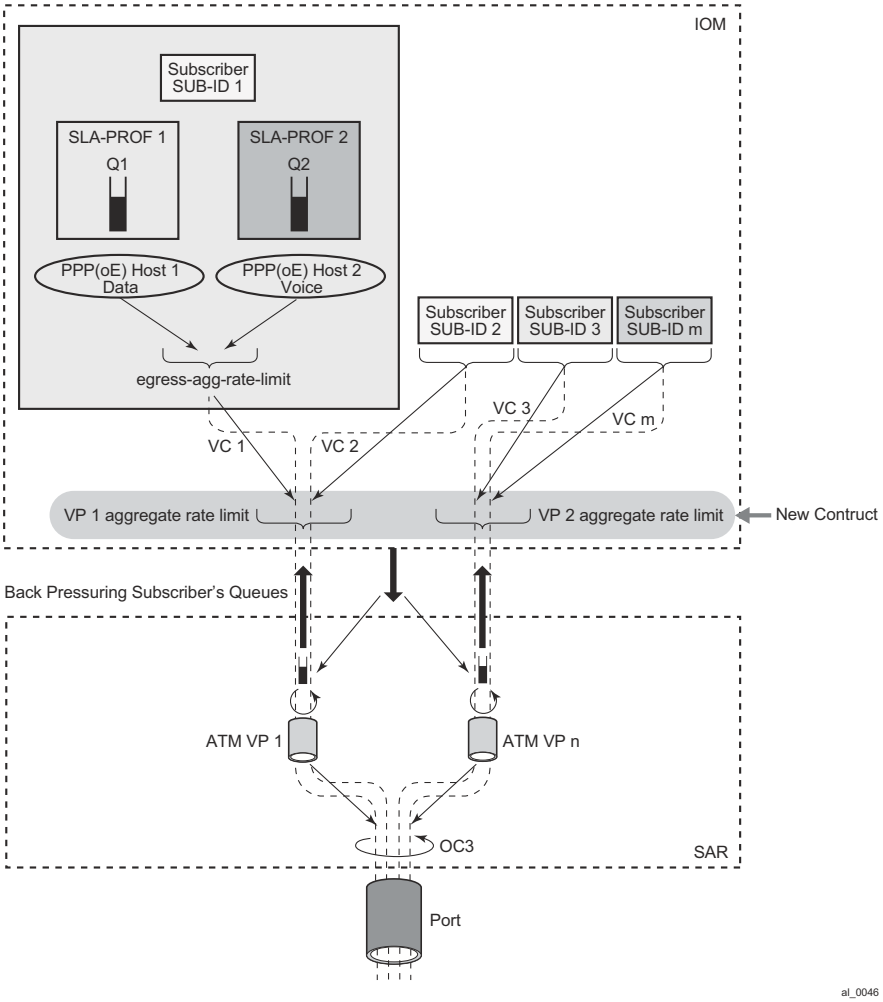



**Figure 93: VP Shaping**

## Authentication

Authentication for PPPoEoA is the same as in PPPoE:

- based on PADI
- based on PAP/CHAP
- No authentication

Authentication in PPPoA is based on:

- PAP/CHAP
- No Authentication — There is a small percentage of cases where this might be required.

An example of no authentication configuration for PPPoEoA would be:

```
configure
     service ies/vprn
          subscriber-interface <sub-if-name>
               group-interface <grp-if-name>
                    dhcp
                         server <ip-address>
                         client-application pppoe  => make PPPoE session the client of DHCP
server
                         local-dhcp-server <dhcp-server-name>
                              user-db <ludb-name>       => allow DHCP server to query LUDB
     subscriber-management
          local-user-db <ludb-name>
               pppoe
                    host <host-name>
                         host-identification [circuit-id | mac | remote-id | service-name |
username]
                    options
                    address
                    identification-strings
```

In the case that there is no PAP/CHAP authentication, a PPPoE host can be identified by a MAC address, circuit/remote-id (inserted by the BNG) or service-name in PPPoE.

## LUDB Access via Capture SAP

Access to LUDB via a capture SAP is enabled for this feature:

```
configure
     service vpls <id>
          sap <sap-id> capture-sap
               pppoe-user-db <ludb-name>
               ppp-user-db <ludb-name>
```

Note that if the authentication-policy (RADIUS authentication) is specified under the capture SAP, it (the RADIUS authentication) will take precedence over LUDB.

## Encapsulation Autosensing

As previously discussed, these four static encapsulation types will be supported:

1. aal5nlpid-ppp (LLC/SNAP for PPPoA)
2. aal5snap-bridged (LLC/SNAP for PPPoEoA)
3. aal5mux-ppp (mux PPPoA)
4. aal5-mux-bridged-eth-nofcs (mux PPPoEoA)

These four types of encapsulation are supported for ATM MSAPs as well as for fixed configuration ATM SAPs.

In addition, the LLC/SNAP encapsulation type can be autosensed. This is called autosensing. The keyword for autosensing is **aal5auto**. An option is given to provision encapsulation statically, if needed.

```
sap x/y/z:* capture-sap
atm
encapsulation aal5auto | aal5mux-ppp | aal5nlpid-ppp | aal5mux-pppoe | aal5snap-bridged-
eth-nofcs

sap x/y/z:w/z
atm
encapsulation aal5auto | aal5mux-ppp | aal5nlpid-ppp | aal5mux-pppoe | aal5snap-bridged-
eth-nofcs
```

The aforementioned encapsulation options (including autosensing) is visible only under the SAP hierarchy on group-interfaces and on capture SAPs in VPLS. This is allowed only in 16K-VP mode ATM MDA.

---

## SAP Autoprovisioning

In order to simplify the provisioning of the subscriber access ports in the ESM context, a concept similar to managed SAP(MSAP) on Ethernet is introduced. In Ethernet MSAP, an ingress access SAP is automatically created upon receipt of the first (VLAN) tagged packet from the customer side (pending the authentication process).

In our ESM over PPPoA/PPPoEoA case, a number of PVCs is pre-provisioned that will initially be only in a provisioned (or listening, passive) state. This is sometimes referred as a bulk configuration of VC ranges. Once the initial ESM processing in the CPM is completed (for example, the user is authenticated), the ATM SAP is created in the appropriate context (VRF or GRT) and the ATM VC is activated. ATM VC activation means that the ATM VC is associated with a SAP.

The ratio between the maximum number of provisioned VCs vs the maximum number of active VCs on an MDA supporting PPPoA/PPPoEoA is 2:1. This amounts to 32K ATM VCs in the listening state. Out of the total 32K ATM VCs, 16K ATM VCs can be active simultaneously.

Obviously, there are some differences between the Ethernet MSAP processing and the ATM MSAP processing. On Ethernet there is not need to pre-configure VCs, while on ATM this is necessary due to complexity of ATM layer comparing to Ethernet.

Our current Ethernet based MSAP is configured under VPLS. The same approach is adopted for ATM capture SAP.

Since a combination of ESM over PPPoA/PPPoEoA VCs and other non PPPoA/PPPoEoA VCs on a single physical port is supported, ranges of VCs that will be supported in autoconfiguration are defined. ATM VCs outside this range are available for manual creation. Multiple ranges are necessary in order to address non-contiguous sets of VPI/VCIs.

The following is the CLI syntax:

```
configure
    service vpls <id> customer <customer-id> [create]
        sap x/y/x:*/* capture-sap [create]
            atm
                vc-range <num> vpi-range <vpi-range> vci-range <vci-range>
```

Note that the capture SAP must be configured as '*/*' in place of the VPI/VCI identifiers. Any other combination for the VPI/VCI identifiers is not allowed.

Up to 5 ranges are allowed per capture SAP. VCs configured in this way can carry a single PPP session per VC or multiple PPPoE sessions per VC. Ranges are not allowed unless the ATM MDA is in the 16K-VCs mode.

The total number of VCs that can be fed into a VP is 16K. This includes all VC ranges associated with the VP plus any statically configured VC on the VP.

## PPP Nodes and ppp-policy

Differences in operation between PPP and PPPoE warrant creation of a new ppp node under the *group-interface* in the CLI that will cover PPP aspects of operation. The existing pppoe node under the *group-interface* is preserved in the CLI. This allows referencing different *user-dbs* for authentication purposes and different session parameters defined in the *ppp-policy* for each session type (pppoe or plain ppp).

PPP node under the *group-interface* is used to cover PPPoA operation while PPPoE node is used to cover PPPoE and PPPoEoA operation. ATM in PPPoEoA is just a transport and as such does not carry any information relevant to PPP operation (like PADx does in PPPoE).

For dynamic SAPs (managed SAPs), the same PPP(oE) related structures are referenced under the capture SAP hierarchy. Under the capture SAP there are no ppp/pppoe nodes (like they are under the group-interface hierarchy). In order to differentiate between *ppp* and *pppoe* clients, a new ppp-policy command is introduced in addition to the *pppoe-policy* command. The *ppp-policy* under the

capture SAP is needed for the definition of session parameters before the *group-interface* (where normally session parameters are referenced) is determined.

Two commands for LUDB access are available under the same *capture-sap* hierarchy (*pppoe-user-db* and *ppp-user-db*).

The *ppp-policy* under the *subscr-mgmt* hierarchy contains PPP and PPPoE session parameters. PPP parameters are applicable to both session types (PPP and PPPoE) while PPPoE parameters are applicable only to PPPoE session type. The PPPoE parameters are ignored for PPP sessions.

## MTU Considerations

MRU configuration option negotiated during the LCP phase in PPPoA is based on the following command:

```
configure>subscr-mgmt>ppp-policy#
     ppp-mtu
```

By default, this command is disabled and consequently will negotiate the default MRU of 1500B, as long as the ATM port's MTU can accommodate at least 1500B:

```
configure>port>sonnet-sdh>path#
     mtu
```

The MRU option in PPPoA refers to the PPP packet length (PID+Information+Padding) that is ATM encapsulated.

## PPP(oE) Session Antispoofing

Antispoofing filters need to be in place in order to prevent the hijacking of a PPPoA/PPPoEoA session. For successful anti-spoofing, the following fields are accessible:

- Source IP address (PPPoA and PPPoEoA)
- VPI/VCI pair which is equivalent to the SAP (PPPoA/PPPoEoA)
- Source MAC Address (PPPoEoA only)
- session-ID (PPPoEoA only)

1. For locally terminated PPPoEoA subscriber hosts, access to all fields are required (a, b, c and d). Antispoofing is defined under the following hierarchy:

```
configure
service <svc-name>
subscriber-interface <sub-if-name>
group-interface <grp-if-name>
         pppoe
anti-spoof [mac-sid | mac-sid-ip]
```

   This behavior matches our current PPPoE behavior. The default antispoofing option is set to mac-sid, which means that the incoming traffic is checked against the source MAC address and the session-ID. Antispoofing under this hierarchy cannot be disabled.

   The source IP address can be added to the source MAC and session-ID (mac-sid-ip).

   Note that the group-interface is the lowest granularity at which options can be anabled. In other words, these two options (mac+sid or mac+sid+ip) cannot be changed at the SAP level.

   The VPI/VCI pair is always be checked against the incoming traffic, regardless of the configuration option as the VPI/VCI pair is an intrinsic part of the SAP (similar to VLAN tags on Ethernet).

   In case that a subscriber host is a routed host (managed routes), the nh-mac antispoofing option must be enabled under the managed sap (msap-policy) or group-interface->sap level. Otherwise, managed routes for the host would NOT be installed. The nh-mac option forces a lookup of the incoming packet based on the mac+sid of the originally created host (CPE device). Note that only the group-interface->sap>anti-spoof and msap-policy hierarchies contain the nh-mac option, and NOT the pppoe->anti-spoof hierarchy.

   For locally terminated PPPoA subscriber hosts, only access to source IP address is available (src MAC and session-ID fields are non-existent in PPP).

   The default antispoofing is based on the VPI/VCI pair + IP. This cannot be changed by configuration, except when the session has managed routes or is a LAC session. In the case of managed routes (routed host) a lookup will be done based on the SAP (VPI/VCI) only. The antispoofing for PPPoA should be set to antispoof nh-mac, under msap-policy or grp-if>sap hierarchy even though PPPoA has no MAC address.

For LAC, the behavior is the following:

- For PPPoE traffic, antispoofing is always done based only on the SAP+mac+session-id.
- For PPP traffic, antispoofing is always based on VPI/VCI pair (SAP only). The IOM automatically determines whether IP based antispoofing can be done (e.g. no IP based antispoofing for LAC and managed routes).

There are two other nodes on which antispoofing can be configured:

```
configure
     service ies/vprn
          subscriber-interface <sub-if-name>
               group-interface <grp-if-name>
                    sap <sap-id>
                         anti-spoofing [ip|ip-mac|nh-mac]
```

A few points in regards to the above hierarchy:

- **ip** - This option can be only used in BSM mode. This is used only for IPoE.
- **ip-mac** – Lookup is performed based on the combination of the IP address and MAC address (incipient host). This is used for IPoE. For PPPoE, it will be overwritten by the configuration option under the group-interface-> ppp node (mac+sid or mac-sid+ip).
- **nh-mac** – Lookup is based on the MAC only for IPoE or mac+sid for PPPoE.

```
configure
     subscriber-managemnt
          msap-policy <msap-pol-name>
               ies-vprn-only-sap-parameters
                    anti-spoof [ip-mac|nh-mac]
```

This is similar to the previous case (under the group-interface>sap hierarchy) with the exception that the pure IP option is not supported under the MSAP. The reason is that the IP option can only be used in BSM (under the SAP node) whereas MSAP can only be used in ESM (PPPoX).

Note that for IPoE, the entire control over antispoofing under the (M)SAP node, while for PPPoE, the anti-spoofing control is distributed between the (M)SAP node and the *group-interface->ppp* node.

# Multi-Chassis Synchronization

Figure 94 shows the configuration under which synchronization of subscriber management information is performed. As depicted, a single access node aggregating several subscriber lines is dual- homed to redundant-pair of nodes.

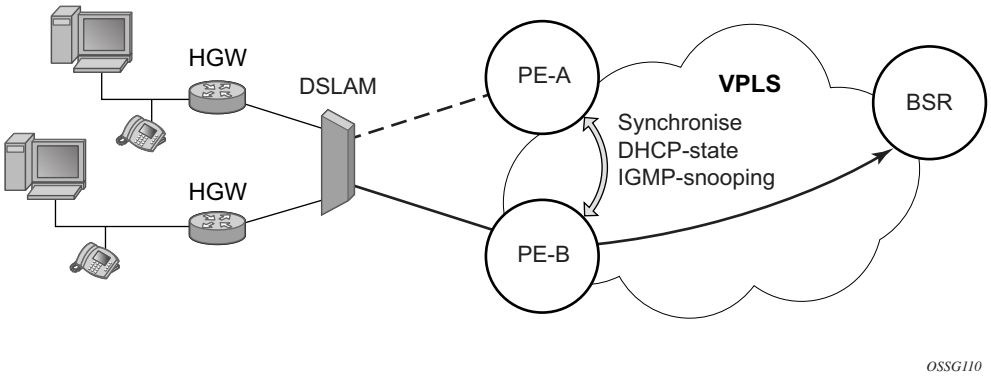



**Figure 94: Dual-Homing Configuration**

Enabling subscriber management features (whether basic subscriber-management (BSM) or enhanced subscriber management (ESM)) causes the node to create and maintain state information related to a given subscriber-host. This information is synchronized between redundant-pair nodes to secure non-stop service delivery in case of the switchover.

## Overview

The synchronization process provides the means to manage distributed database (the Multi-Chassis Synchronization (MCS) database), which contains the dynamic state information created on any of the nodes by any application using its services. The individual entries in the MCS database are always paired by peering-relation, sync-tag and application-id. At any time the given entry is related two the single redundant-pair objects (two SAPs on two different nodes) and hence stored in a local MCS database of the respective nodes.

Internally, peering-relation and sync-tag are translated into a port and encapsulation value identifying the object (SAP) that the given entry is associated with. The application-id then identifies the application which created the entry on one of the nodes. There are three basic operations that the application can perform on MCS database. The MCS database will always synchronize these operations with its respective peer for the given entry.

The following principles apply:

- add-operation — Any dynamic-state created in the application is pushed to the MCS database. MCS then creates and synchronizes with the corresponding peer provided (if configured). The application in the peer node is then notified as soon as the entry has been created. Similarly, the application in the local node (the node where the state has been created) is notified that entry has been synchronized (MCS is "in-sync" state). This operation will be also used to modify existing MCS database entry.
- local-delete — The MCS database entry is marked as no longer in use locally and this information is sent to the peer node. If the information is no longer used by applications on both nodes (the application in remote-node has already issued local-delete before), it is removed from database.
- global-delete — The MCS database entry is removed from both nodes and from the application in the remote node.

The choice of the operation in corresponding situation is driven by the application. The following general guidelines are observed:

- An event which leads to a dynamic-state deletion on a standby chassis will be handled as "local-delete".
- An event which leads to a dynamic-state deletion on an active chassis will be handled as "global-delete".
- An exception to above the rules is an explicit "clear" command which will be handled as "global-delete" regardless of where the command was executed.

As previously stated, the MCS process automatically synchronizes any database operation with the corresponding peer. During this time, the MCS process maintains state per peer indicating to the applications (and network operator) the current status, such as in-sync, synchronizing or sync_down. These states are indicated by corresponding traps.

## Loss of Synchronization and Reconciliation

Each time the connection between the redundant pair nodes is (re)established the MCS database will be re-synchronized. There are several levels of connectivity loss which may have different effect on amount of data being lost. To prevent massive retransmissions when the synchronization connection experiences loss or excessive delay, the MCS process implementation will take provisions to ensure following:

- In the case of a reboot of one or both nodes or establishing the peering for the first time, the full MCS database will be reconciled.
- In the case that the MCS communication is lost and then re-established but neither node rebooted during the connection loss, only the information not synchronized during this time will be reconciled (using sequence numbers helps identify information which was not synchronized).
- In the case that MCS communication is lost because of excessive delay in ACK messages but no information has been effectively lost, the MCS process indicates a loss of synchronization but no reconciliation is performed.

# Subscriber Routed Redundancy Protocol (SRRP)

Subscriber Routed Redundancy Protocol (SRRP) is closely tied to the multi-chassis synchronization 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 hosts 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 will be mapped to the redundant SAP on the other node.

Once SRRP is active on the group IP interface, the SRRP instance will attempt 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 will enter a master state while the other router will enter 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 15 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 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 15 lists the SRRP's state effect on subscriber hosts associated with group IP interfaces.

**Table 15: 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 | • Will respond to ARP for all owned subscriber subnet IP addresses.<br>• Will not respond to ARP for subscriber subnet SRRP gateway IP addresses.<br>• All ARP responses will contain the native MAC of the group IP interface (not the SRRP gateway MAC). | • Will respond to ARP for all subscriber hosts on the subscriber subnet. | • Will respond to ARP for all reachable remote IP hosts. | • All routing out the group IP interface will use the native group IP interface MAC address.<br>• The group IP interface redundant IP interface will not be used.<br>• 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) | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• Will respond to ARP for subscriber subnet SRRP gateway IP addresses (hardware address = SRRP gateway IP address, source MAC = group IP interface native MAC). | • Will respond to ARP for all subscriber hosts on the subscriber subnet (hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC). | • Will respond 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 native group IP interface MAC address.<br>• Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface.<br>• Will not accept packets destined to the SRRP gateway MAC received on the group IP interface. |

**Table 15: 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 | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• Will respond to ARP for subscriber subnet SRRP gateway IP addresses (hardware address = SRRP gateway IP address, source MAC = group IP interface native MAC). | • Will respond to ARP for all subscriber hosts on the subscriber subnet (hardware address = SRRP gateway MAC address, source MAC = group IP interface native MAC). | • Will respond 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.<br>• Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface.<br>• Will accept packets destined to the SRRP gateway MAC received on the group IP interface. |
| Becoming Backup (redundant IP interface operational) | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• 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.<br>• Subscriber hosts mapped to the group IP interface are remapped to the redundant IP interface.<br>• Will accept packets destined to the SRRP gateway MAC received on the group IP interface. |

**Table 15: 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) | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• 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.<br>• Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface.<br>• Will accept packets destined to the SRRP gateway MAC received on the group IP interface |
| Backup (redundant IP interface operational) | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• 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.<br>• Subscriber hosts mapped to the group IP interface are remapped to the redundant IP interface.<br>• Will not accept packets destined to the SRRP gateway MAC received on the group IP interface. |

**Table 15: 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) | • Will respond to ARP for all owned subscriber subnet IP addresses (hardware address and source MAC = group IP interface native MAC).<br>• 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.<br>• Subscriber hosts mapped to the redundant IP interface are remapped to the group IP interface.<br>• 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 has been expanded to support 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 master's 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 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.

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 / 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 advertisements stating that the SAP is unavailable from the remote node, an SRRP Remote Receive advertisement SAP Unavailable event will be generated. This details 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, the owning node must advertise the IP addresses as /32 host routes into the core. This is important since the subscriber subnet is advertised into the core by multiple routers and the network will follow the shortest path to the closest available router which may not own the IP address if the /32 is not advertised within the IGP.

## 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 will respond 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.

# PPPoE MC Redundancy

This feature minimizes the downtime for PPPoE clients in an ESM environment when a single node fails.

But it is not necessary that an entire BNG fails before it triggers the corrective action. The solution outlined in this document will natively include protection against interfaces and line card failures within the BNG. The redundant (protective) entity, however, does not reside within the same BNG on which the failure occurs but instead it is on a separate BNG node.

The PPPoE MC Redundancy is based on SRRP and MC-LAG because SRRP is already established in ESM providing IPoE MC Redundancy. With some modifications, SRRP approach is adopted to PPPoE deployments.

## Hardware Support

This feature is supported on the following platforms:

- 7750 SR-7/12
- 7750-c4/12
- 7450 in mixed chassis mode.

MCS across different platform types (7750 SR-7/12, 7750-c4/12, 7450) is not supported. For example, MCS between 7750 SR-7/12 and 7750-c12 is not supported.

This feature is supported in the following chassis modes:  B, C, D and SR Mixed Mode (IPv6 on IOM3 while IOM1 cards can be present in the same system).

IPv4 functionality is supported on IOM2 cards and IPv4/IPv6 on IOM3 cards.

**Note:** ESM v6 is supported only on IOM3 cards. IPv6 forwarding in ESM between IOM3 and IOM2 cards is not supported – for example, if the access side is IOM3 and the network side is IOM2. However, plain routing (non-ESM related) is supported between these two cards.

## SRRP Considerations for PPPoE

SRRP is based on VRRP whose purpose is to provide a default gateway redundancy for clients sharing the transport medium such as Ethernet. IPoE would be a typical example of this where IPoE clients use a virtual IP and MAC address that is shared between two default gateway nodes in the Master/Backup configuration. SRRP supports only two nodes in a cluster but VRRP allows multiple nodes to be configured in a cluster with a priority that will determine which node will assume Mastership. Although it is mandatory for the proper operation of IPoE clients that the same SRRP IP address is shared between the two BNG nodes providing redundancy, having the same SRRP IP address is not necessary for the operation of SRRP itself. In other words, SRRP itself (Master/Backup states) will work with different SRRP IP addresses on each node. Same is valid for MAC addressing. It is possible by configuration that the redundant BNG nodes use different IP/MAC addresses on a pair of SRRP instances.

Upon a switchover, a gratuitous ARP is sent from a newly selected active node so that each IPoE client can update the ARP table, if the MAC address has indeed changed (it does not have to). More importantly, if an Layer 2 aggregation network is in place between the BNG and the IPoE client, all intermediate Layer 2 devices will have to update their port-to-mac mappings (Layer 2 FIB). The above described process will ensure proper packet addressing on the IPoE client side as well as the proper forwarding path through Layer 2 aggregation network to the newly activated BNG.

When considering PPPoE in conjunction with SRRP, keep in mind that PPP protocol (point-to-point protocol) is adopted for the Ethernet (shared medium) by enabling an extra Ethernet related layer in PPP that allows sharing of point-to-point sessions over Ethernet (shared medium). The result is a PPPoE protocol designed to ‘tunnel’ each PPP session over Ethernet.

PPPoE is not aware of ARP (Address Resolution Protocol) and it will not react to gratuitous ARP packets sent by a newly active BNG. The destination MAC address that PPPoE clients will use when sending traffic is determined not by ARP but by the PPPoE Discovery phase at the beginning of the session establishment. This originally discovered destination MAC is used throughout the lifetime of the session. This has a couple of consequences:

1. If SRRP is used for PPPoE then the ‘SRRP’ MAC address between the redundant BNG nodes must be shared. It is not allowed to use a unique ‘SRRP’ MAC address per BNG in the redundant pair of BNG nodes (as it is allowed today for IPoE). Every PADx conversation is based on the SRRP shared MAC address, that is, the PADO reply must have the shared SRRP MAC address as the source MAC. This has a significant impact on the operation of MSAP in conjunction with this feature.
2. Since PPPoE sessions are not ARP aware, the only purpose of the gratuitous ARP would be to update the Layer 2 FIB in the aggregation network (and not the PPPoE client destination MAC address). For IPoE, the gratuitous ARP is sent for ALL subnet gateway IP addresses found under the subscriber interface over either all SAPs (default) or top-tags only. For PPPoE, the gratuitous ARP is sent only for the system IP address. The purpose of the gratuitous ARP in PPPoE scenario is only to update Layer 2 network path which is otherwise IP unaware. It is not necessary to send the gratuitous ARP for every default-gateway address found under the subscriber-interface. Since this feature is only applicable to PPPoE deploy-

ments, therefore, only PPPoE is present under the group interface. This is indicated by the following command under the SRRP node:

```
group-interface <name>
srrp <id>
     one-garp-per-sap
```

### SRRP Fact-Checks

1. Once Multi-chassis Synchronization (MCS) for subscriber management and SRRP is enabled, both BNG nodes, Master and Backup will in general forward packets (for subscribers) in both directions.
2. Traffic flows through an SRRP enabled node according to the entries in the SRRP sync database and the SRRP state of the node:
   - Backup SRRP directs downstream traffic over the redundant-interface towards the Master SRRP node. If the redundant interface is unavailable, traffic is sent directly to the subscriber.
   - Master SRRP always directly forwards the downstream traffic towards the subscriber.
   - In the upstream direction, the active SRRP node accepts subscriber traffic addressed either to the MAC address of the SRRP active group OR the native interface MAC address.
   - The standby node accepts in the upstream direction only packets addressed to its native interface MAC address.
3. If both SRRP nodes become Masters then both forward traffic to/from subscribers unaware of the link failure somewhere in the Layer 2 network. As a result, downstream traffic can be blackholed. Whether downstream traffic will be lost depends on the native routing on the network side, which is unaware of the failures in the aggregation network.

## State Synchronization

PPPoE sessions are synchronized between the redundant BNG nodes. The subscriber synchronization is achieved through Multichasiss Synchronization (MCS) protocol in a similar way it is performed for IPoE.

```
multi-chassis
            peer <IP@>create
                sync
                    local-dhcp-server
                    SRRP
                    sub-mgmt [ipoe | pppoe]
                 :
                                                :
                    no shutdown
                exit
                no shutdown
            exit
```

A two keywords, **ipoe** and **pppoe** enable a more granular control over which type of subscribers the MCS should be enabled.

Subscriber synchronization is important for following reasons:

1. Forwarding of downstream traffic between the redundant BNG nodes through a redundant interface is an artifact of how natural routing steers traffic through the network.
2. Subscriber instantiation on the node which did not originally create subscriber session. This drastically reduces downtime during the SRRP switchover.
3. Monitors operational aspects of the subscriber management through show commands.

---

## PPPoE Multi-chassis Synchronization (MCS) Model

PPPoE MCS model is based on SRRP synchronization and can be used in a centralized or distributed environment with or without Layer 2 aggregation network in-between DSLAMs and BNG nodes. The failure detection speed is dependent on SRRP timers. Traffic load can be balanced per SRRP group over the two links.  In this model (Figure 95), PPPoE states are synchronized between the redundant BNG nodes. If one BNG fails, the newly activated BNG sends out a 'MAC update' (gratuitous ARP) message prompting the intermediate Layer 2 nodes to update their forwarding tables so that forwarding can resume. The SRRP timers can be configured in the sub-second range. In reality, the limiting factor for timer values is the scale of the deployment, in particular the number of SRRP groups per node.

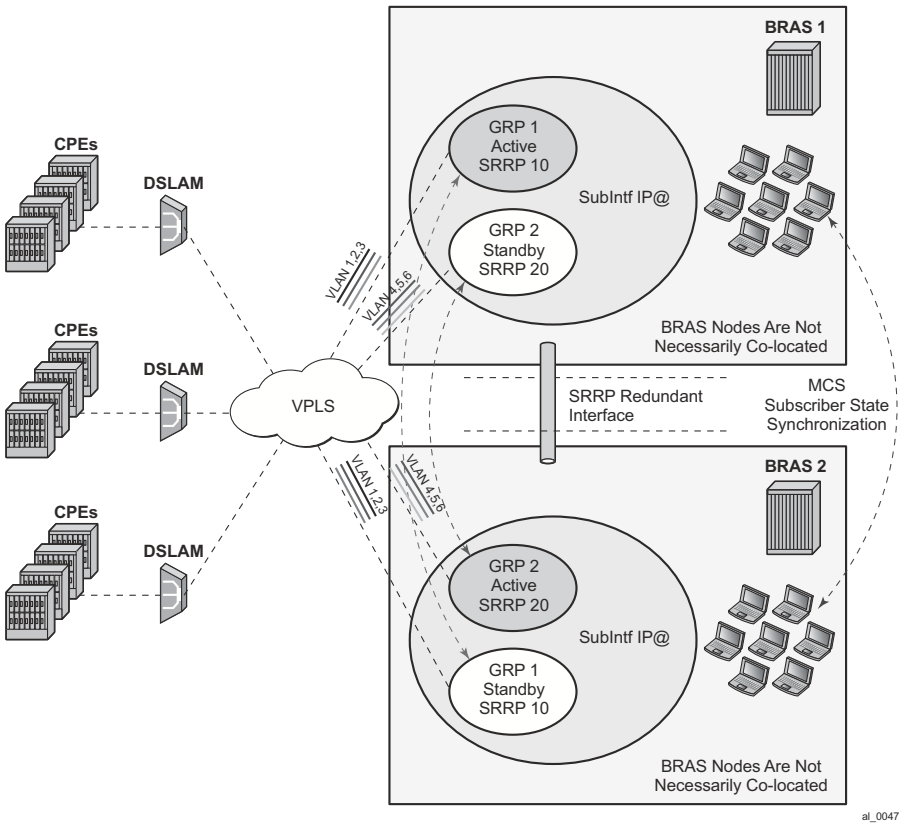



**Figure 95: Fully Redundant "Statefull 1:1" Model**

## Traffic Control and Redundant Interface

To preserve QoS and Accounting, subscriber's traffic must flow in both directions through the Master BNG node.

In the upstream direction, this is always true as traffic is steered to the master SRRP node just by the virtue of SRRP operation.

In the downstream direction which represents bulk of traffic, SRRP can not be relied up on to steer traffic through the Master node. This poses a problem in a very common environment where IP subnets are shared over multiple group-interfaces with SRRP enabled. A particular subnet will be advertised to the network side from both BNG nodes, Master and Backup. Natural routing on the network side will determine which BNG node will receive subscriber's traffic in the downstream direction. If the Backup SRRP node receives the traffic, it cannot simply send the traffic directly to the access network where the subscriber resides by just inserting the source MAC address of the SRRP instance in the outgoing packet. This would break the operation of SRRP. Instead, the Backup BNG node must send the traffic to the Master BNG node via a redundant-interface. The Master SRRP node would then forward traffic directly to the subscriber. Source MAC address of this traffic would then be the MAC address of SRRP instance. This traffic shunting over the redundant interface can result in a substantial load on the link between the two BNG nodes.

The increase in shunted traffic can quickly become an issue if the redundant BNG nodes that are not collocated. To minimize the shunt traffic, more granular routing information must be presented to the network core. This would lead to more optimal routing where downstream subscriber traffic would be directed towards the Master BNG node, without the need to cross the redundant interface. The downside of this approach is that this would further fragment the IP address space within the network core. In the extreme case where /32 (subscriber) IP addresses are advertised, the churn that /32s can cause in the core routing would most likely be unsustainable. In this case, routing updates in the core would be triggered by subscribers coming on/off-line.

Optimal operation would call for the shunt traffic to be eliminated and at the same time, a high IP route aggregation on the network side is achieved. The existence of the shunt traffic stems from the fact that routing protocols advertise subscriber subnets into the network with no awareness of the SRRP activity state (Master/Standby). To address this problem along with better aggregation of advertized subnets, two SRRP enhancements are introduced:

- SRRP fate-sharing
- SRRP aware routing

Both of this concepts are described under the 'SRRP Enhancements" section.

Traffic destined to/from the subscriber is forwarded under the condition that the subscriber-interface is operationally UP. This applies also to shunting of downstream subscriber traffic from the STANDBY to MASTER node. It is always necessary to keep the subscriber-interface operationally UP by configuring a dummy *group-interface* with a command *oper-up-while-empty* under it. This is especially true for the MC-LAG which causes the messaging SAP on the STANDBY node always to be in the INIT state. In case that MSAPs are used on such group-interfaces, the group-interfaces would be also operationally DOWN, causing the subscriber-interface to be operationally DOWN.

## Subnet Assignment and Advertisement - Option ‘A’

A single IP subnet is used for all subscribers terminated within the redundant BNG nodes. The upside of the Option ‘A’ is that it offers aggregated IP addressing in the network core per pair of redundant BNG nodes. The downside is that the subscriber termination point (active BNG for the SRRP group) is hidden from the network core. Since both BNG nodes share the same IP subnet for the subscribers, the natural routing can cause downstream traffic to be sent to the standby BNG which in turn will have to shunt the traffic to the active BNG. It is likely that half of the traffic will be shunted over the redundant-interface with this approach. This scenario is shown in Figure 96.

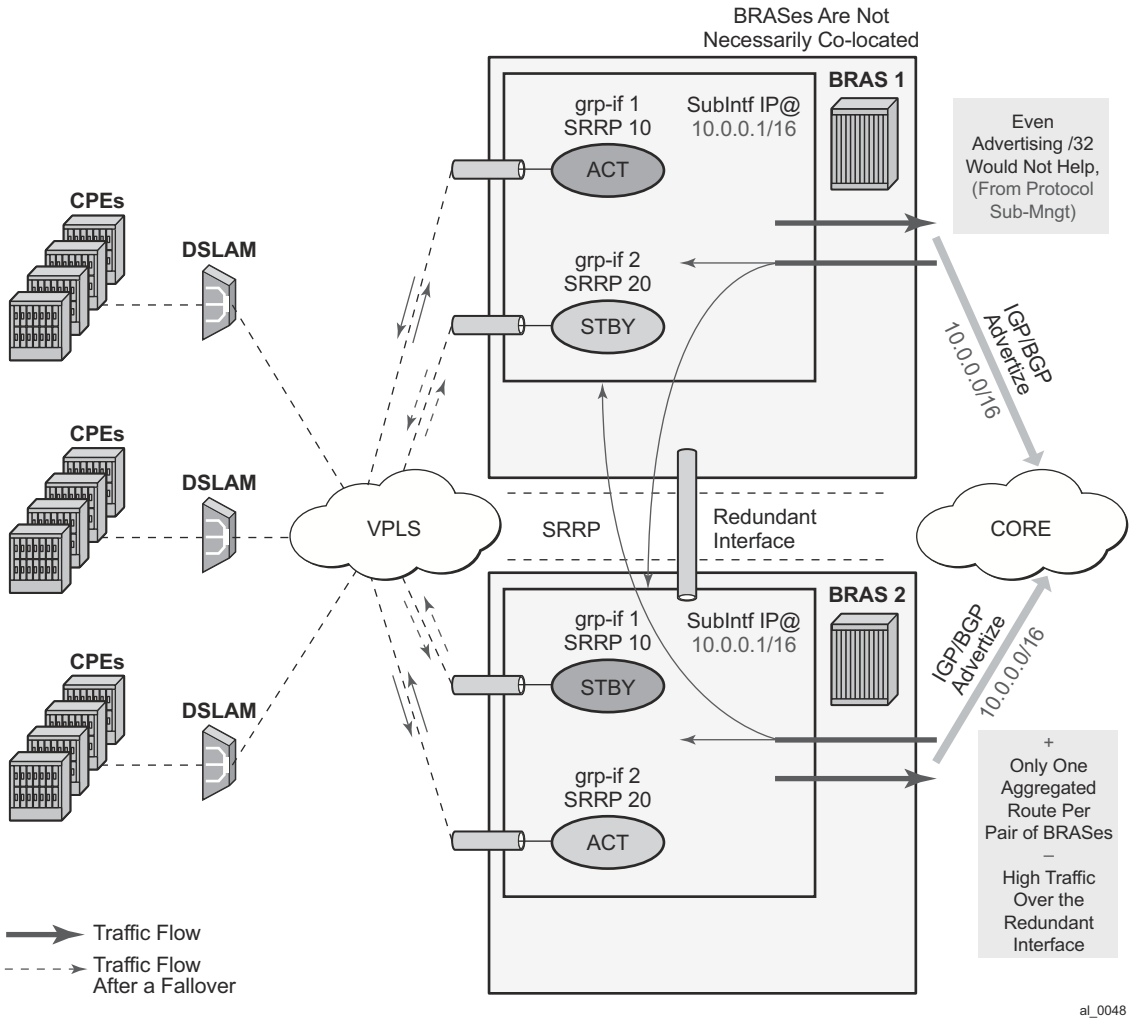



**Figure 96: Shared Subscriber IP Space**

## Subnet Assignment and Advertisement - Option ‘B’

With the option ‘B’, an IP address pool (or subnet) can be allocated per group of SRRP instances that are in the Master state. The routing decision on the network side is further influenced by the static increase of the metric of the advertised route on the BNG node hosting the active SRRP groups (Figure 97).

This approach would cause greater IP space segmentation in the network core, but at the same time, it would indirectly provide more information about the subscriber whereabouts and thus minimize or eliminate the shunt traffic during the normal operation. However, in the case of a SRRP switchover, the shunt traffic would ensue. The amount of the shunted traffic would depend on the scale of the failure. From the Figure 97, it can be concluded that:

- In the depicted scenario on Figure 97, there is no shunted traffic.
- If any of the SRRP instances transitions out of the Master state, traffic for an entire IP network associated with this failed SRRP instance would be shunted. The reason for this is that the advertised route metric is static and it does not follow changes in SRRP state.

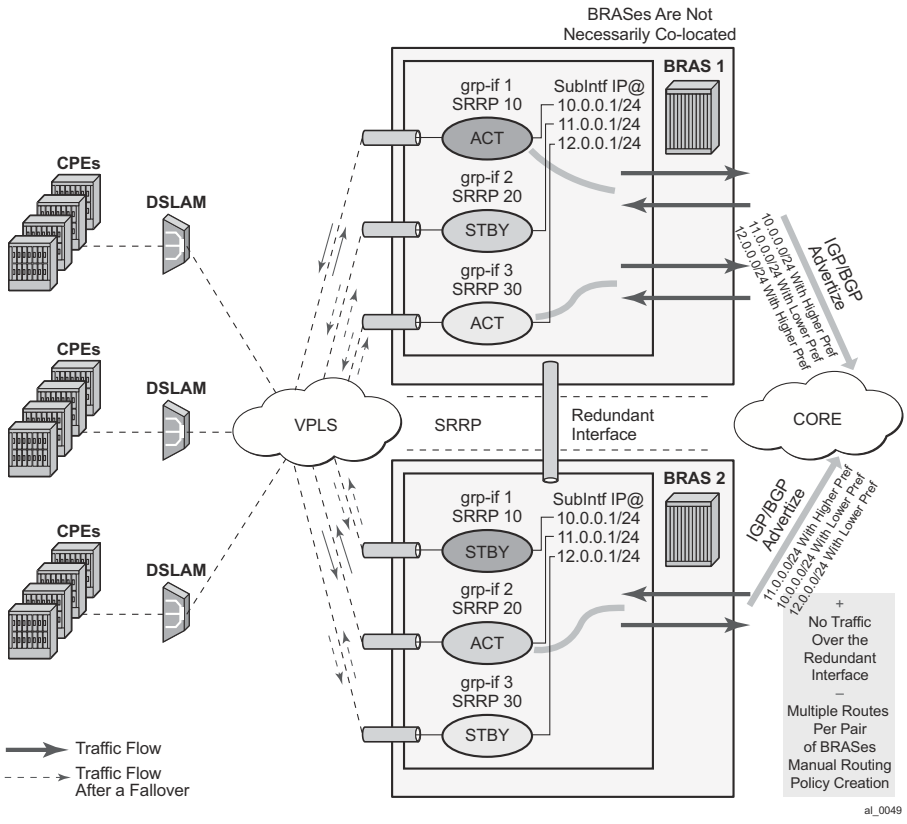



**Figure 97: Option ‘B’ – IP Subnet per Active SRRP Group**

## MSAP Considerations

As per RFC 2516 (PPPoE), this has the implications on the operation of the capture SAP. In IPoE environment, the initial DHCP traffic related to host establishment will use its native MAC of the physical port on 7x50. Once the group-interface is learnt (later in the process, via RADIUS or msap-policy), the MAC address is switched to SRRP MAC address (virtual MAC). The IPoE client will adapt easily to this change. On the contrary, for the proper operation of PPPoE with SRRP, the initial destination MAC address learned by the PPPoE client does not change during the lifetime of the session.

This is ensured by indirectly referencing the grp-if under the capture SAP:

```
configure>service>vpls
     sap 1/1/1:1.* capture-sap
track-srrp 10
     sap 1/1/1:2.* capture-sap
track-srrp 20

configure>service>vprn>
          subscriber-interface <if-name>
               group-interface <grp-if-name>
                    sap 1/1/1:1.1
                    srrp 10
                         message-path 1/1/1:1.1

               group-interface <grp-if-name>
                    sap 1/1/1:2.1
                    srrp 20
                         message-path 1/1/1:2.1
```

With this approach the grp-if is nailed during the session initiation phase by referencing the SRRP instance in track-srrp statement (srrp is a grp-if wide concept). RADIUS returned grp-if name must match the one on which referenced SRRP instance runs.

The capture SAP of the form

```
sap port-id:*.* capture-sap
     track-srrp X
```

assumes that there is only one grp-if associated with all msaps under this capture SAP.

A check is put in place to make sure that the MAC addresses associated with the SRRP instance is the same as the MAC address of the associated capture SAP. A log is raised if there is a discrepancy between the MAC addresses while the grp-if is operationally UP. If there is a MAC address change (user misconfiguration) then the existing PPPoE sessions will time out and the new sessions will fail to establish until the condition is corrected.

## Unnumbered Interface Support

For unnumbered subscriber-interface support in PPPoE, the gw IP address that is used to send gratuitous ARP is not available. For this reason, the system IP address is used to send gratuitous ARPs. Gratuitous ARP is used to update the Layer 2 network forwarding path towards the BNG node in the upstream direction.

The system IP address is used automatically if the subscriber interface is unnumbered.

## Compatibility with MC-LAG

SRRP for PPPoE works in an environment where MC-LAG is enabled. For example, the standby LAG link automatically puts the SRRP node in a Backup state and the SRRP becomes master on the active MC-LAG link. It is important that the SRRP on the standby leg of the MC-LAG is forced into a Backup state, or any new state that will force the downstream traffic to use the redundant interface.

Traffic destined to/from the subscriber is forwarded under the condition that the subscriber-interface is operationally UP. This applies also to shunting of downstream subscriber traffic from the STANDBY to MASTER node. It is always necessary to keep the subscriber-interface operationally UP by configuring a dummy group-interface with a command oper-up-while-empty under it. This is especially true for the MC-LAG which causes the messaging SAP on the STANDBY node always to be in the INIT state. In case that MSAPs are used on such group-interfaces, the group-interfaces would be also operationally DOWN, causing the subscriber-interface to be operationally DOWN.

## IPv6 Support

Prerequisite for MC IPv6 Redundancy is to synchronize PPPoEv6 and IPoEv6 subscribers between the nodes via MCS.

In PPPoE environment, SRRP is used to refresh the forwarding path (MAC addresses) in the access aggregation network (via gratuitous ARP). SRRP ensures that the upstream traffic is steered to the Master BNG node. In the downstream direction, the Backup BNG directs traffic over to the Master BNG node via redundant-interface.

The IPv6 functionality currently relies on IPv4 based SRRP and IPv4 based redundant-interface. In other words, IPv4 is required to run on the access side as well as on the redundant-interface.

The redundant-interface is used in the downstream direction. Traffic arriving on the network links on the Standby node is shunted over to the Master node over the redundant-interface. This is required to ensure consistent QoS and accounting functionality across the nodes (upstream and downstream traffic on the access links for a subscriber must traverse the same BNG node). There is no IPv6 related CLI associated with the redundant-interface.

All IPv6 subscriber traffic that arrives on the Standby node in the downstream direction is automatically shunted over the IPv4 redundant-interface to the Master node. When IPv6 traffic arrives over the redundant-interface on the Master node, it is either PPPoEv6 encapsulated or left as plain IPoEv6 before it is forwarded to the subscriber.

In the upstream direction (AN->BNG) the behavior is the following:

- PPPoEv6

  On the switchover, gratuitous ARPs is sent from the new Master on each vlan. The IP address in gARP is the IPv4 gw-ip address or the system IP in the case of unnumbered interfaces. This updates the Layer 2 network path with the proper SRRP MAC address.

- IPoEv6

  IPv4 based SRRP is used to update the Layer 2 forwarding path in the case of a switchover. A gratuitous ARP is sent in the same fashion as it is used for IPoE v4 hosts. Router Advertisements (RA) are not sent out in the case of the switchover.

  However, the two BNG nodes share the same virtual Link Local (LL) IPv6 address. This address is used by the clients as a default-gw and only the Master BNG advertises this LL address in RAs. RAs are suppressed on the Standby node. As already mentioned, RAs are not sent during the switchover. RAs are sent:

  → When the client first gets established – this is how the client learns its default-gw (in PPPoE case RA can also be used for SLAAC – stateless address configuration).

  → As a reply to Router Solicitations messages sent by the clients.

  → Periodically to each client.

  Note that RAs are unicasted to each client.

  Neighbor Advertisements (NA) used for address resolution are sent only from the Master. NA has the SRRP MAC address in the target link layer option on SRRP enabled group interfaces (on non-SRRP enabled group-interfaces, NAs contains the group interface MAC address).

  The syntax to configure the LL address on the subscriber interface is the following:

```
configure>service>ies | vprn>
          subscriber-interface <if-name>
               ipv6
[no] link-local-address <ipv6-address>

          <ipv6-address> : ipv6-address - x:x:x:x:x:x:x:x:x:x:x:x:x:x:d.d.d.d
          x    [0..FFFF]H
          d    [0..255]D
```

  The LL IPv6 address must be the same on both nodes. In addition, the gw-mac address must be the same on both nodes. The IPv6 clients will not be aware of the switchover and therefore they will not send NS to solicit the update of its neighbor cache with the possibly different gw-mac address.

Note that the current version of SRRP relies only on IPv4 routes. The connection between SRRP and IPv4 routes is done via the subnets with gw IP addresses defined under the subscriber-

interfaces in the ESM context. This connection is needed so that SRRP can send Gratuitous ARP properly.

These are the cases for PPPoEv6 MC Redundancy that are supported:

- unnumbered subscriber-interfaces (config>service>subscriber-interface hierarchy)
- numbered IPv4 subscriber-interfaces (config>service>subscriber-interface hierarchy)
- numbered IPv4 AND IPv6 subscriber-interfaces (config>service>subscriber-interface and config>service>subscriber-interface>ipv6 hierarchy)

numbered IPv6 only subscriber-interfaces (config>service>subscriber-interface>ipv6 hierarchy) is not supported

## Considerations with Local DHCP Server

When local DHCP Server redundancy/synchronization is used in conjunction with PPPoE in multi-chassis environment, both DHCP servers must be referenced under the corresponding group-interface on each node:

```
subscriber-interface <sub-if>
     group-interface <grp-if>
          dhcp
               server <local-dhcp-ip-address> <remote-dhcp-ip-address>
```

Otherwise, the PPPoE clients will not be synchronized via MCS.

Note that this is not the requirement in IPoE environment. In IPoE environment, it is enough that the DHCP server points to the IP address of the local DHCP server. If the IP lease is originally assigned by the peer DHCP server, the request for renewal is automatically forwarded to the remote DHCP server by the virtue of the IP address of the original DHCP server that is included in the renewal request.

It is necessary for the successful renewal of the IP address on the remote DHCP server, that the remote DHCP server has a valid return path back to the gi-address of the forwarder of the renewal request.

## Redundant Interface Considerations

In PPPoE dual-chassis environment without the redundant-interface in place, SRRP aware routing should always be used. Otherwise, if the downstream traffic arrives on the backup node, it will get forwarded directly to the client over the access network (assuming that the access network is operational) with the source MAC address of the group-interface (instead of gw-mac). This grp-if MAC address is different from the MAC address (gw-mac) negotiated during the initial PPPoE phase, and therefore, this traffic will be dropped by the client. It must be ensured that the downstream traffic is always attracted to the Master node in the absence of redundant.

# Routed Central Office (CO)

The Routed CO feature allows a network operator to connect a DSLAM to a router. Figure 98 shows a DSLAM connected to a router using a Layer 3 interface within an IES service. Operators that do not require an aggregation network can implement this topology. Typical DSLAM connection models include:

- One SAP for all subscribers with all services.
  - → Subscriber management will be used for subscriber separation with DSCP/Dot1p service separation.
- One SAP per service.
  - → Subscriber management will be used for subscriber separation with the SAP being the service differentiator.
- One SAP per subscriber.
  - → Model with SAP level subscriber separation with DSCP/Dot1p service separation.

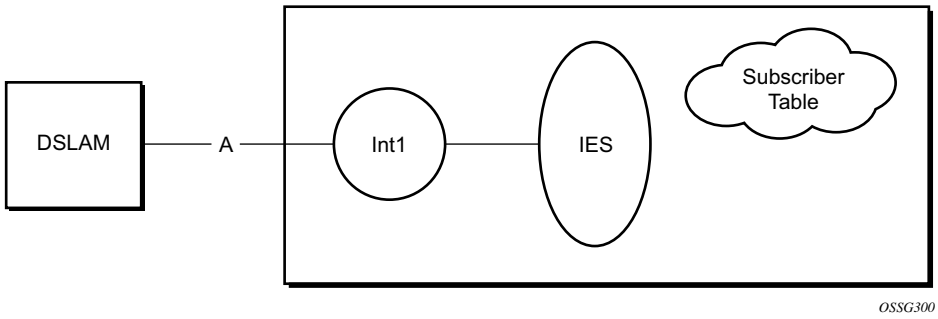



**Figure 98: DSLAM Connection**

In early releases of the OS, only one SAP could be associated with each Layer 3 interface. Now a group-interface allows multiple SAPs to be configured as part of a single interface. All SAPs in a single group-interface must be within the same port. Since broadcast is not allowed in this mode, forwarding to the subscriber is based on IP/MAC addresses information gathered by the subscriber management module and stored in the subscriber management table. These entries are based on both static and dynamic DHCP hosts. Routed CO must be used with standard subscriber management or enhanced subscriber management. DSLAMs are typically deployed with Ethernet interfaces.

This model is a combination of two key technologies, subscriber interfaces and group interfaces. While the subscriber interface defined the subscriber subnets, the group interfaces are responsible for aggregating the SAPs.

As depicted in Figure 99, an operator can create a new subscriber interface in the IES service. A subscriber interface allows for the creation of multiple group interfaces. The IP space is defined by the subnets of the subscriber interface's addresses. Figure 100 shows the details of group interface A.

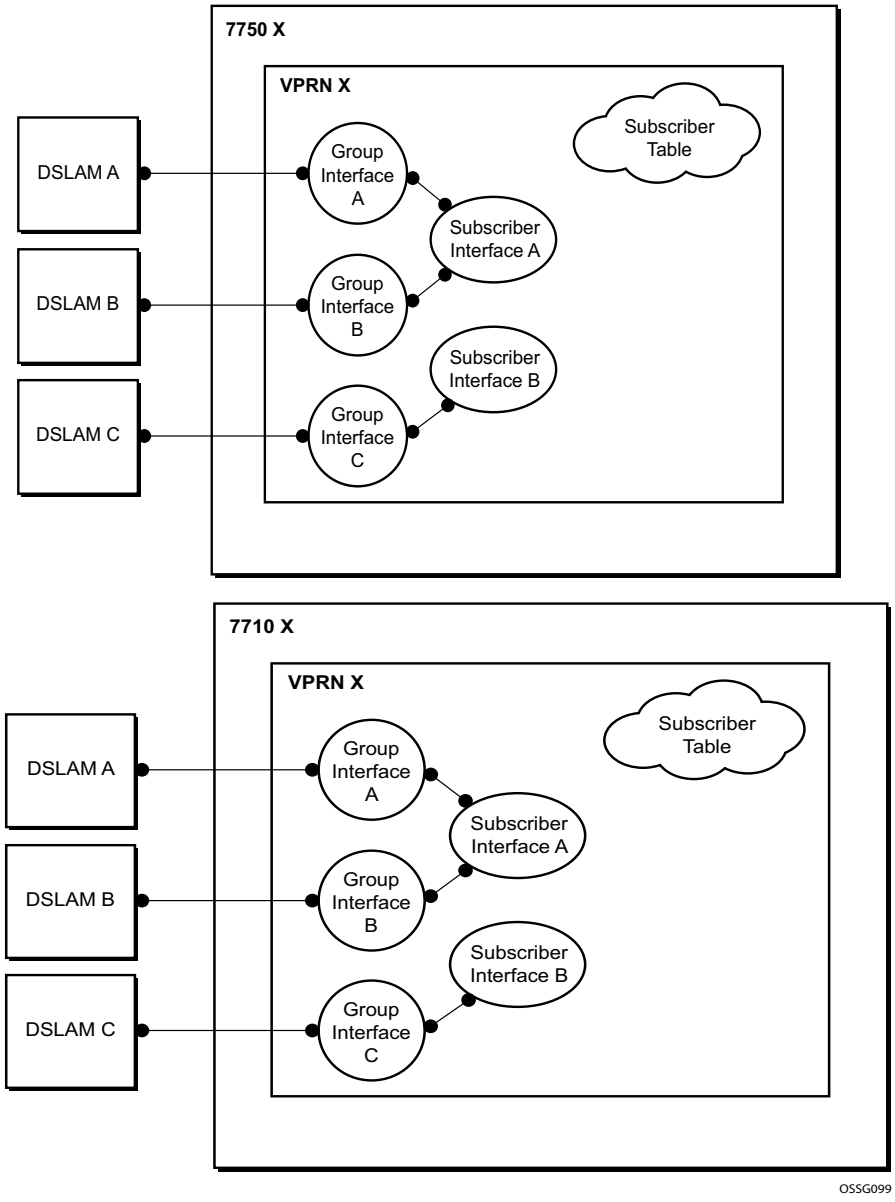



**Figure 99: Subscriber Interface in an IES Service**

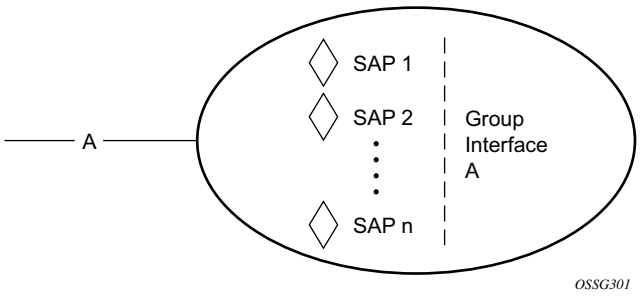



**Figure 100: Details of a Group-Interface**

## Layer 3 Subscriber Interfaces

This section describes the Alcatel-Lucent routers acting as a Broadband Service Router (BSR), with Enhanced Subscriber Management enabled.

In this model, a router is positioned directly behind a DSLAM. This design removes the need for a Layer 2 aggregation network between the router and the DSLAM, however it does involve more routing entities in the network.

Figure 101 shows a network diagram where the DSLAM are connected directly to a Broadband Service Router (BSR) providing access to an IP subnet. Subscribers from multiple DSLAMs can be part of the same subnet. Note that BSR is referred to as BBNG, Broadband Network Gateway, in the DSL Forum.

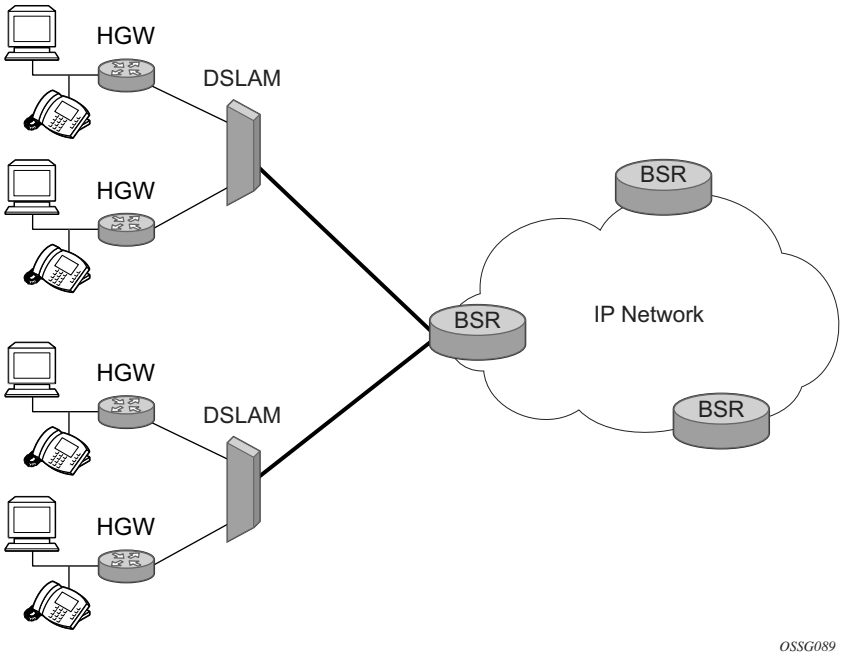



**Figure 101: Aggregation Network with Direct DSLAM-BSR Connection**

The BSR can be configured with multiple subnets, allowing subscribers to be part of a single subnet as well as providing mechanisms for re-addressing or expanding existing services without affecting existing users.

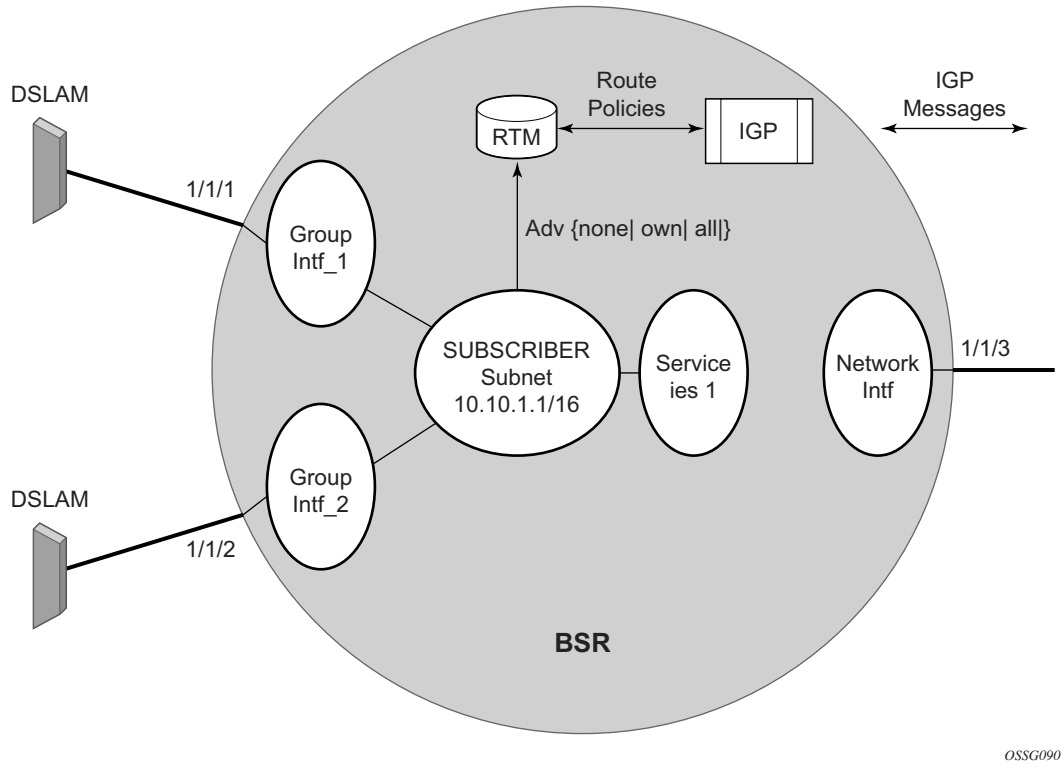



**Figure 102: Detailed View of Configurable Objects Related to Layer 3 Subscriber Interfaces**

Figure 102 shows a detailed view of a router and the configuration objects implemented to support Layer 3 subscriber interfaces.

- A subscriber service is defined by an IES Service. One or more IES services can be created.
- Each IES service concentrates a number of subscriber-interfaces. The operator can create multiple subscriber interfaces (represented as a subscriber subnet). A subscriber interface will define at least one subnet.
- A group-interface will be provisioned within the subscriber interface for each DSLAM connected. All group interfaces created under the subscriber interface will share the same subnet (or subnets). Group interfaces (shown as intf_1 and intf_2 Figure 102) are configured as unnumbered and are associated with the subscriber-interface under which they are configured.
- SAPs can be configured under the group-interface. In a VLAN-per-DSLAM model only, one SAP per group-interface is needed, while in the VLAN-per-subscriber model, a subscriber of the DSLAM will require its own SAP. All SAPs on a group-interface must be on the same physical port or LAG.

The individual features related to subscribers, such as DHCP relay, DHCP snooping and anti-spoofing filters, are enabled at group-interface level. For a Routed CO model of subscriber

management, and when enhanced subscriber management (if sub-sla-mgmt is configured). Then, hashing will be based on an internally assigned subscriber-ID. Having a unique subscriber ID configured in CLI will ensure that each subscriber is assigned a unique internal subscriber ID.

It is assumed that individual end-user devices (further referred to as subscriber hosts) get their IP address assigned through either DHCP or static configuration. The management of individual subscriber hosts (such as creation, queue allocation, etc.) is performed by Enhanced Subscriber Management (see Triple Play Enhanced Subscriber Management on page 827).

The operator can provision how the system advertises routes. While most deployments will advertise the full subnet it is possible to have the system advertise only the active, discovered or static host routes.

The distribution of this information into routing protocols will be driven by import/export route-policies configured by the operator.

## Routed Subscriber Hosts

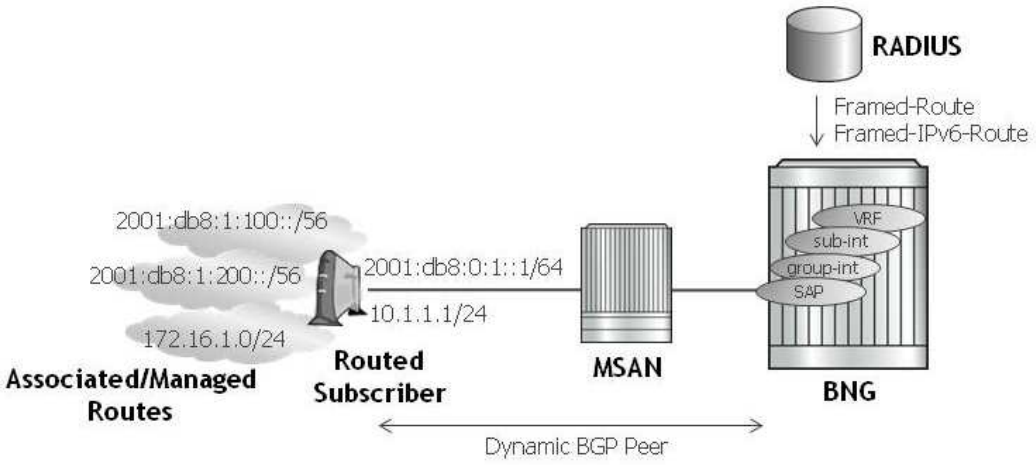



**Figure 103: Router Subscriber Hosts**

A routed subscriber host associated route is a global routable subnet/prefix behind a routed CPE or Home Gateway. The routed CPE is identified in the BNG as an ESM subscriber host: QoS, accounting and anti-spoofing is enforced per CPE. The associated routes are installed in the BNG route table with next-hop pointing to the routed subscriber host's WAN address.

Routed subscriber host associated routes are supported on IES/VPRN subscriber interfaces in a routed CO configuration. To put a SAP or MSAP in routed subscriber mode, the anti-spoof type for the SAP or MSAP must be configured to nh-mac:

```
configure
    service ies/vprn <service-id>
        subscriber-interface <ip-int-name>
            group-interface <ip-int-name>
                sap <sap-id>
                            anti-spoof nh-mac

configure
    subscriber-mgmt
        msap-policy <msap-policy-name>
            ies-vprn-only-sap-parameters
                anti-spoof nh-mac
```

There are three ways to learn about a routed subscriber host associated IPv4 route:

1. Configuration for a static host
2. A dynamic BGP peer
3. The RADIUS [22] Framed-Route attribute

A routed subscriber host associated IPv6 route can only be learned with the RADIUS [99] Framed-IPv6-Route attribute.

## Static Configured IPv4 Managed Route

The routes associated with a static host are populated in the routing table as "Remote Managed" routes. Up to sixteen managed routes can be configured for a static host.

```
config>service>ies>sub-if>grp-if>sap
config>service>vprn>sub-if>grp-if>sap

    static-host ip 10.1.1.20 create
        sla-profile "sla-profile-1"
        sub-profile "sub-profile-1"
        subscriber "static-host-1"
        managed-routes
            route 172.20.1.0/24
            . . .
            route 172.20.16.0/24
        exit
        no shutdown
           exit
```

To display the managed routes associated with a routed subscriber host, use following commands:

**show service id** *service-id* **static-host detail**

## Dynamic BGP Peering

Routed subscriber host associated IPv4 routes can be learned over a dynamic BGP IPv4 peer that is automatically set up when a subscriber host is instantiated. The parameters for the BGP peer are configured in a BGP peering policy or obtained in Radius VSA attributes. The subscriber-host IPv4 address is used as the BGP peer IP address. The BGP peering is torn down and the associated routes removed from the routing table as soon as the subscriber-host is removed.

Dynamic BGP peering is supported for routed subscriber hosts terminated in a VPRN service and is not supported for routed subscriber hosts terminated in an IES service. The BGP learned routes scaling is limited by the BGP scaling limits. The routes learned via a dynamic BGP peer are populated in the routing table as "Remote BGP" routes.

To display the BGP learned routes associated with a routed subscriber host, use the regular BGP commands. For example:

**show router** *service-id* **bgp neighbor** *ip-address* **received-routes**

A dynamic BGP group must be configured in the BGP cli context of the VPRN service where the subscriber host is started:

**config>service>vprn>bgp**

```
    group "dynamic-peer-1" dynamic-peer
    exit
```

The BGP peering policy to be used must be configured in the subscriber-mgmt CLI context:

```
config>subscr-mgmt

    bgp-peering-policy "bgp-policy-1" create
    exit
```

A dynamic BGP peer is established for a subscriber host if the RADIUS attribute [26-6527-55] “Alc-BGP-Policy” returned in the Access-Accept contains the name of a local configured bgp-peering-policy and if a dynamic peer group is configured in the VPRN BGP context.

BGP peering parameters can be specified from multiple sources:

- Use BGP peering parameters returned in Radius VSA attributes
- If not available from RADIUS, use BGP peering parameters configured in the bgp-peering-policy
- If not configured in the bgp-peering-policy, use BGP peering parameters configured for the dynamic-peer group
- If not configured in the dynamic-peer group, use the BGP peering parameters configured in the VPRN service BGP CLI context.
- If not configured in the VPRN service BGP CLI context, use the defaults

The import and export policies to be used for the dynamic bgp peer are determined in following priority order:

1. Use import/export policies returned in RADIUS VSA attributes and append policies configured in the bgp-peering-policy.
2. If not available from RADIUS AND not configured in the bgp-peering-policy, use the policies configured in the dynamic-peer group.
3. If not configured in the dynamic-peer group, use the policies configured in the VPRN service BGP CLI context.

Table 16 details the RADIUS VSA attributes that can be used to setup dynamic BGP peering.

**Table 16: RADIUS VSA Attributes to Setup Dynamic BGP Peering**

| Attribute-ID | Attribute Name | Description |
|---|---|---|
| 26-6527-55 | Alc-BGP-Policy | Mandatory attribute to setup a dynamic BGP peer. References a bgp peering policy configured in the "configure subscriber-mgmt bgp-peering-policy <policy-name>" CLI context. |
| 26-6527-56 | Alc-BGP-Auth-Keychain | Optional. References a keychain configured in the "configure system security keychain <keychain-name> " CLI context. |
| 26-6527-57 | Alc-BGP-Auth-Key | Optional. The MD5 authentication key used between BGP peers for BGP session establishment. |
| 26-6527-58 | Alc-BGP-Export-Policy | Optional. References a pre-configured BGP export routing policy. |
| 26-6527-59 | Alc-BGP-Import-Policy | Optional. References a pre-configured BGP import routing policy. |
| 26-6527-60 | Alc-BGP-PeerAS | Optional. Specifies the Autonomous System number for the remote peer |

## Radius: Framed-Route and Framed-IPv6-Route

RADIUS attribute [22] Framed-Route can be specified in a Radius Access-Accept message to associate an IPv4 route with an IPv4 routed subscriber host and Radius attribute [99] Framed-IPv6-Route can be used to associate an IPv6 route with an IPv6 routed subscriber wan host (DHCPv6 IA-NA or SLAAC). These routes are populated in the routing table as "Remote Managed" routes. Up to sixteen managed routes can be installed for a routed subscriber host; this corresponds with up to sixteen Framed-Routes and sixteen Framed-IPv6-Routes for a dual stack routed subscriber. Framed-IPv6-Routes cannot be associated with a Prefix Delegation host (DHCP IA-PD).

The Framed-Route and Framed-IPv6-Route attributes should be formatted as:

"<ip-prefix>[/<prefix-length>] <space> <gateway-address> [<space> <metric>] [<space> tag <space> <tag-value>] [<space> pref <space> <preference-value>]"

where:

<space> — is a white space or blank character.

<ip-prefix>[/prefix-length] — is the managed route to be associated with the routed subscriber host. The prefix-length is optional for an IPv4 managed route. When not specified, a class-full class A,B or C subnet is assumed. The prefix-length is mandatory for an IPv6 managed route.

<gateway-address> — must be the routed subscriber host IP address.
"0.0.0.0" is automatically interpreted as the host IPv4 address for managed IPv4 routes.

"::" and "0:0:0:0:0:0:0:0" are automatically interpreted as the wan-host IPv6 address for managed IPv6 routes.

[<metric>] — Optional. Installed in the routing table as the metric of the managed route. If not specified, metric zero is used. Value = [0.. 65535].

[tag <tag-value>] — Optional. The managed route will be tagged for use in routing policies. If not specified, or tag-value = 0, then the route is not tagged. Value = [0..4294967295].

[pref <preference-value>] — Optional. Installed in the routing table as protocol preference for this managed route. If not specified, preference zero is used. Value = [0..255].

If the optional metrics (metric, tag and/or preference) are specified in a wrong format or with out of range values, then the defaults are used for all metrics: metric=0, no tag and preference=0. No event is logged.

If the Framed-Route or Framed-IPv6-Route is invalid (for example because the gateway address specified does not match the host wan IP address or because the host bits are not zero) then the routed subscriber host is instantiated without the ill defined managed route. An event is logged in this case.

If an identical managed route is associated with different routed subscriber hosts in the context of the same IES/VPRN service only one managed route is installed in the routing table. The selection criteria are (in order of priority):

1. Lowest preference
2. Lowest metric
3. Lowest ip next-hop

Other identical managed routes are shadowed and an event is logged.

Valid Framed-Routes and Framed-IPv6-Routes are persistent (stored in the persistency file for recovery after reboot) and synchronized in a Multi-Chassis Redundancy configuration.

RADIUS-learned Framed-Route/Framed-IPv6-Route and static host associated managed routes that are installed in the routing table can be identified in routing policies for redistribution as protocol "managed".

To display the managed routes associated with a routed subscriber host, use following commands:

**show service id** *service-id* **dhcp lease-state detail**

**show service id** *service-id* **dhcp6 lease-state detail**

**show service id** *service-id* **slaac host detail**

**show service id** *service-id* **ppp session detail**

**show service id** *service-id* **pppoe session detail**

**show service id** *service-id* **arp-host detail**

Valid RADIUS-learned managed routes can be included in Radius accounting messages with following configuration:

```
configure
    subscriber-mgmt
        radius-accounting-policy <name>
            include-radius-attribute
                framed-route
                framed-ipv6-route
```

Associated managed routes for an instantiated routed subscriber host are included in RADIUS accounting messages independent of the state of the managed route (Installed, Shadowed, HostInactive, etc.)

In case of a PPP session, when a Framed-Route or Framed-IPv6-Route is available while the corresponding routed subscriber host is not yet instantiated, the managed route is in the state “notYetInstalled” and will not be included in Radius accounting messages.

## Interaction Notes



### Routed CO for VPRN Interactions

Much like the Routed CO model, the Routed CO model for VPRN depends on subscriber management to maintain the subscriber host information. To create a group-interface, the operator must first create a subscriber interface in the **config>service>vprn** context. The subscriber interface can maintain up to 256 subscriber subnets and can be configured with a host address for each subnet. The wholesaler subscriber interface can contain addresses that are used with its own end-customers. The host IP address can be installed as a result of both relaying to a DHCP server and proxy to a RADIUS server. In both cases the host IP address must be in the subnet defined by either the wholesaler's or retailer's VPRN's subscriber interface.

Basic subscriber management is allowed only in a subscriber/SAP model and can be used only in a dedicated VPRN architecture. A RADIUS service selection will require Enhanced Subscriber Management. The subscriber interface's subnets are allowed to be advertised to both IGPs and BGP within a VPRN.

If a packet is received by a wholesaler instance sourced from a subscriber host it will be forwarded based on the routing table of the retailer's VPRN. The destination IP address can be part of the same retailer subnet of another local interface (subscriber or regular interface). If the destination IP address is part of a different VPRN (the destination subnet is not available in the retailer's VPRN) the packet will be dropped. However, if the destination IP address is of a subscriber that is within the same retailer's VPRN the packet can be delivered. When a subscriber-split-horizon is enabled, the system will divide the retailer VPRN into two VRFs. The upstream VRF will contain routes other than the subscriber interface while the downstream VRF will contain all routes.

When an authentication policy is specified for a group-interface, DHCP snooping must be enabled to intercept DHCP discover and renew messages for RADIUS authentication. Subscriber management RADIUS extensions are allowed if the operator chooses to have the RADIUS server return the subscriber identification, subscriber profile and sla-profile strings using RADIUS.

In the retailer's VPRN, DHCP settings (RADIUS and DHCP server assignments) are allowed. The entire DHCP node is allowed only for subscriber interfaces that are associated with a wholesaler context. RADIUS authentication is available to the retailer's subscriber interface (when the interface is defined to be linked to another).

The node can be defined with both a DHCP relay or proxy function in each routing instance. If the user configures a DHCP relay within the instance, the local-proxy-server command will enable DHCP split leases. Note that the local DHCP proxy in this case can be different than the one defined in the wholesale context. In that configuration the node will provide the configured DHCP

lease to the client using either RADIUS or the real DHCP server as the source of the IP address to be provided.

The RADIUS server can send a Change of Authorization (CoA) message containing the DHCP forcerenew VSA which prompts the local-proxy-server to send a forcerenew message to the client. The node ACKs when the Force-Renew has been sent, regardless of whether the subscriber responds. If the client fails to respond or if a new session cannot be established due to resource management issues or otherwise the node must respond with a NACK to the RADIUS server.

If the CoA message contains an IP address that is different than the configured IP address (when RADIUS was providing IP addresses) the node must send a forcerenew message to the client and NAK the request and provide a new IP address. If the node fails to receive a request the CoA is ACK'd when the force-renew has been sent

The operational state of group and subscriber interfaces are dependent on the state of active SAPs. A group interface can become operationally up only if at least one SAP is configured and is in an operationally up state. A subscriber interface becomes operationally up if at least one group interface is operationally up or the associated wholesale forwarding interface is operationally up. This ensures that, in a failure scenario that affects all group interfaces in a given subscriber subnet, the node will stop advertising the subnet to the network. The SRRP state will affect this behavior as well and can cause the subnet to be removed if all group interfaces (and SRRP instances) are in backup state.

The system does not allow subscriber interface chaining. A retailer subscriber interface can be associated with only one wholesaler subscriber interface. The interface cannot be associated with another subscriber interface that is referencing an interface already. Further, once a retailer subscriber interface is created and linked with a wholesale subscriber interface all subscriber interfaces in that retailer's VPRN cannot be used to forward (cannot be linked) to other VPRNs.

Because the subscriber interface in the retailer context does not use a group-interface, some group-interface DHCP functions are allowed. This functionality is available to a retailer subscriber interface only (an interface that was configured to reference another).

Note that the wholesaler DHCP settings can be configured with a vendor-specific option to send the service ID to the server. The service ID that is used must be the retailer's service ID returned from RADIUS. If the option is used in the retailer's context, the retailer's service ID will always be used.

Lease-populate must be enabled for dynamic hosts in the context of the wholesaler VPRN. In that context, the limit must include all hosts for all retailer's hosts seen. A second limit (identified as lease-populate-limit) can be defined in the context of the retailer to limit the number of leases that are allowed to be associated with its subscribers.

## DHCP Interactions

The DHCP relay process has been enhanced to record incoming DHCP discover and request messages. Since forwarding to the SAPs is done by the information in the subscriber management table and multiple SAPs are allowed in one interface it was impossible to know which SAP will be used to forward the DHCP replies. The node maintains a cache of the DHCP requests. The cache can be viewed using the **tools>dump>router>dhcp>group-if-mapping** command. The cache holds an entry for 30 seconds. If an ACK/NAK packet was not received from the server within the timeout the node discards the cache entry. The node can use the Option 82 circuit-id field as part of the temporary host entry. If used, the ACK must contain the same circuit-id field in Option 82 to be found in the cache only if the **match-circuit-id** is specified at the DHCP level of the group-interface. When the **match-circuit-id** command is enabled a check is performed for option 82 circuit-id.

---

## Routed CO Interactions

The routed CO model depends on subscriber management to maintain the subscriber host information. To create a group-interface the operator must first create a subscriber interface within the service (**config>service>ies>subscriber-interface** *ip-int-name*). The subscriber interface maintains up to 256 subscriber subnets and is configured with a host address for each subnet. When a DHCP ACK is received the IP address provided to the client will be verified to be in one of the subscriber subnets associated with the egress SAP. It will be noted that when DHCP snooping is enabled for regular IES interfaces the same rule will apply.

The subscriber interface is an internal loopback interface. The operational state is driven from the child's group-interface states and the configuration of an address in the RTM.

The group interface is an unnumbered interface. The interface will be operationally up if it is in the no shutdown state and if at least one SAP has been defined and is up and the parent subscriber interface is administratively up. The first SAP defined will determine the port for the group-interface. If the user attempts to define a subsequent SAP that is on a different port will result in an error. When the subscriber-interface or the group-interface is in shutdown state no packets will be delivered/received to/from the subscriber hosts but the subscriber hosts, both dynamic and static, will be maintained based on the lease time.

In the routed CO model, the router acts as a DHCP relay agent and also serves as the subscriber-identification agent. The DHCP actions are defined in the group-interface. All SAPs in that interface inherit these definitions. The group-interface DHCP definition are a template for all SAPs.

Lease-populate is enabled by default with the number-of-entries set to 1. This enables DHCP lease state for each SAP in the group-interface.

Since the group-interface can aggregate subscribers in different subnets a GI address must be defined for the DHCP relay process. The address must be in one of the host addresses defined for the subscriber interface. The GI address can be defined at the subscriber interface level which will

cause all child group interface to inherit that route. The GI address can then be overridden at the group interface level. A GI address must be defined in order for DHCP relay to function.

Because of the nature of the group-interface, local-proxy-arp, as well as arp-populate, should be enabled. This would allow the system to respond to subscriber ARP requests if the ARP request contains an IP address which is in the same subnet as one of the subscriber interface subnets.

When an authentication policy is specified for a SAP under a group interface, DHCP will intercept DHCP discover messages for RADIUS authentication. If the system is a DHCP-relay defined in a group-interface and the GI address was not configured the operational state of DHCP will be down.

## VPRN Routed COs

VPRN Routed CO allows a provider to resell wholesaler services (from a carrier) while providing direct DSLAM connectivity. An operator can create a VPRN service for the retailer and configure the access from subscribers as well as to the retailer network. Any further action will be as if the VPRN is a standalone router running the Routed CO model. All forwarding to these servers must be done within the VPRN service. The operator can leak routes from the base interface. In this model, the operator can use RADIUS for subscriber host authentication, DHCP relay and DHCP proxy. This provides maximum flexibility to the retailer while minimizing the involvement of the wholesaler. Access cannot be shared among retailers unless a subscriber SAP is used. This requires that the wholesaler maintain a different access node (DSLAM) for each retailer that does not scale well.

### VPRN Wholesaler-Retail Routed CO

In the wholesaler/retailer model (see Figure 104), the wholesaler instance connections that are common to the access nodes are distributed to many retail instances. Upstream subscriber traffic ingresses into the wholesaler instance and after identification is then forwarded into the retail instance. The reverse will occur for traffic in the other direction. The wholesale/retail traffic flow is controlled with minimal communication with RADIUS. A RADIUS policy is defined in the wholesaler instance. The RADIUS response used during the subscriber instantiation provides the service context of the retailer VPRN. If the wholesaler has a retail business, the operator can configure a separate VPRN for their retail services.

The retailer's subscriber interface controls the DHCP configuration wherever possible instead of the wholesaler's group interface DHCP configuration. The only exception is the lease-populate value. The lease-populate value in the wholesale context controls the per-SAP limits. The lease-populate value in the retail subscriber interface controls the limits for that retailer's interface. Both limits must be satisfied before a new subscriber can be instantiated.

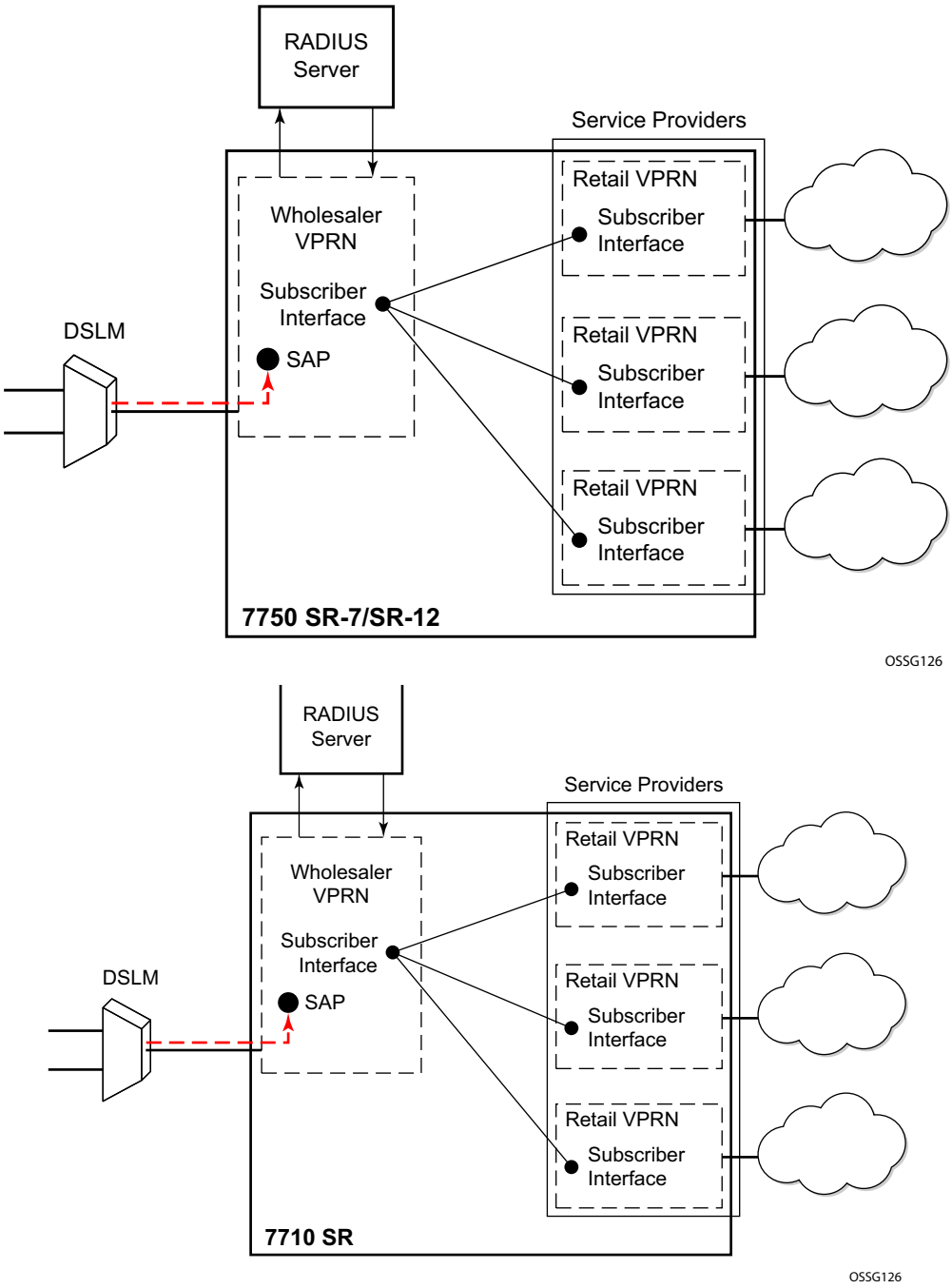



**Figure 104: Routed CO Examples**

As previously stated, in the wholesaler instance, the flow of traffic is possible with minimal communication with RADIUS. A RADIUS policy defined in the wholesaler instance provides the service context of the retailer VPRN.

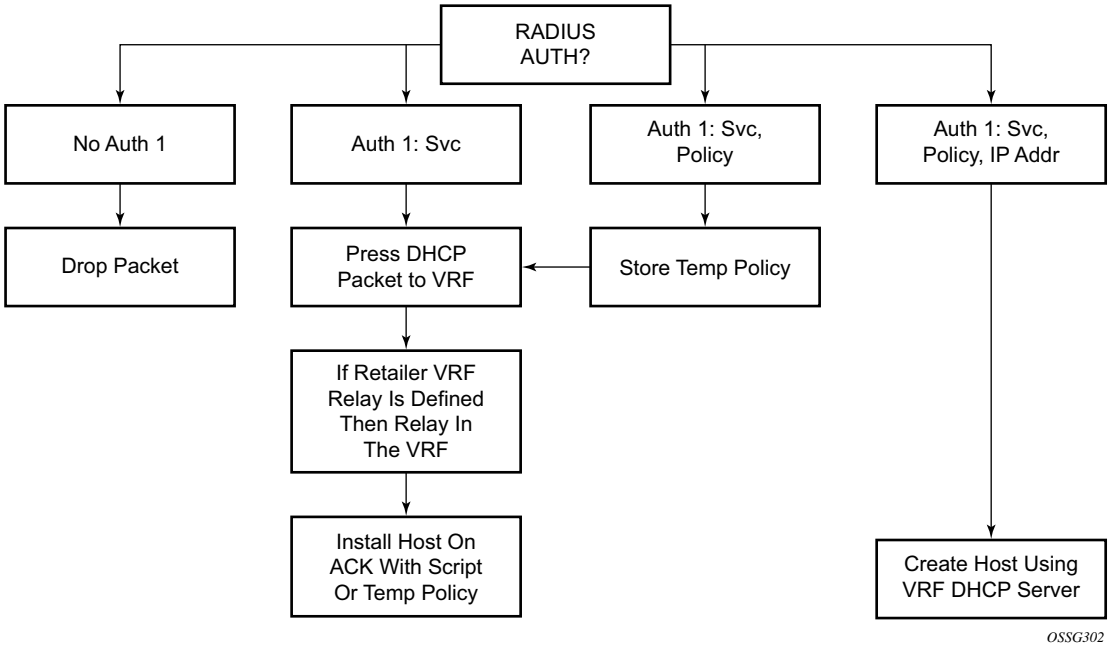



**Figure 105: Traffic Flow**

Figure 105 shows the process starting with a DHCP DISCOVER message and, optionally, with a renew message.

- The packet enters the wholesaler instance which defines the wholesaler RADIUS authentication policy.
- The node contacts RADIUS to authorize access. If RADIUS denies access, the DHCP packet is discarded.
- If access is granted, RADIUS returns the service selection VSA.
- If RADIUS does not return this VSA but authenticates the host, service is allowed only in the wholesaler router instance. In this case, processing RADIUS and DHCP continues within the routing context configured.
- The RADIUS server can return all the attributes for the host such as the service ID, policy and IP address information. The node then provides the IP address to the client using the local DHCP proxy defined in the VPRN and install the host. If the wholesaler provides the IP configuration through RADIUS, proxy DHCP on the retailer must be enabled.

- The RADIUS server can optionally return the host's policy definition. If RADIUS replies with only the VPRN, the policy definition processing will continue by the DHCP configuration of the retailer VPRN. The packet is relayed using the VPRN DHCP configuration and the host will be installed after the ACK is received.

## Linked Subscriber Interface

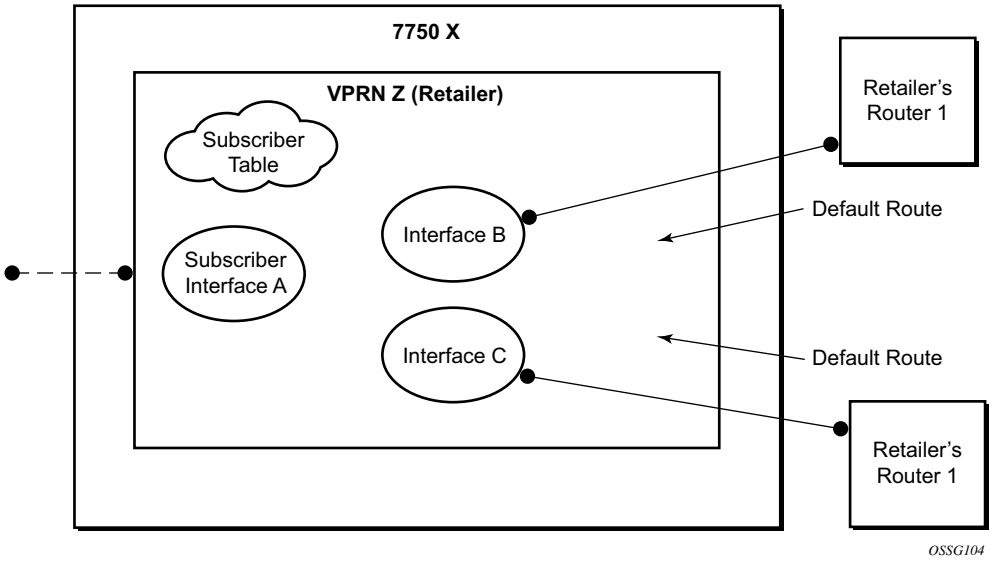
In a wholesale model that is using multiple VPRNs it is necessary to use two subscriber interfaces. The primary subscriber interface will be configured in the context of the wholesaler. This interface is defined with group interface parenting SAPs defined. A second linked subscriber interface must be defined in the retailer context. This linked subscriber interface defines the subnet and potentially the DHCP and RADIUS behavior of the subscribers in that context.

Since that linked subscriber interface does not parent any group interfaces or SAPs it cannot directly forward. It must be linked to another interface used for forwarding. A subscriber interface can be linked to a single subscriber interface and context. Multiple subscriber interfaces can be associated with a single primary subscriber interface in a 1:N (primary sub-int:linked sub-int) configuration. After a subscriber interface was linked to a context all other subscriber interfaces must be linked to the same context.

## Hub-and-Spoke Forwarding

In some cases, hub-and-spoke-type forwarding is needed for the retailer's VPRN. When the retailer expects all subscriber traffic to reach its router (for accounting, monitoring, wiretapping, etc.) normal best-hop behavior within the retailer VPRN is not desired. Any subscriber-to-subscriber traffic will be forwarded within the VPRN preventing the retailer from receiving these packets. To force all subscriber packets to the retailer network a new type of hub-and-spoke topology is defined. The new flag is "type wholesale-retail" which can be used to force all subscriber traffic (upstream) to the retailers network. The system requires that the operator will shutdown the VPRN service to enable this flag. When the flag is enabled, routes learned from MBGP, IGP through a regular interface, static routes through regular interfaces and locally attached regular interface routes will be considered HUB routes and allowed to be used for upstream traffic. Subscriber subnets cannot be used for upstream traffic. Each packet is examined against the HUB routes when the flag is set to eliminate subscriber to subscriber direct forwarding.

Figure 106 describes the forwarding model without a "type wholesale-retail" flag. Each packet is forwarded based on the VPRN (same as the VRF) Z forwarding table. Any packet that is destined to a known subscriber host will be routed directly.



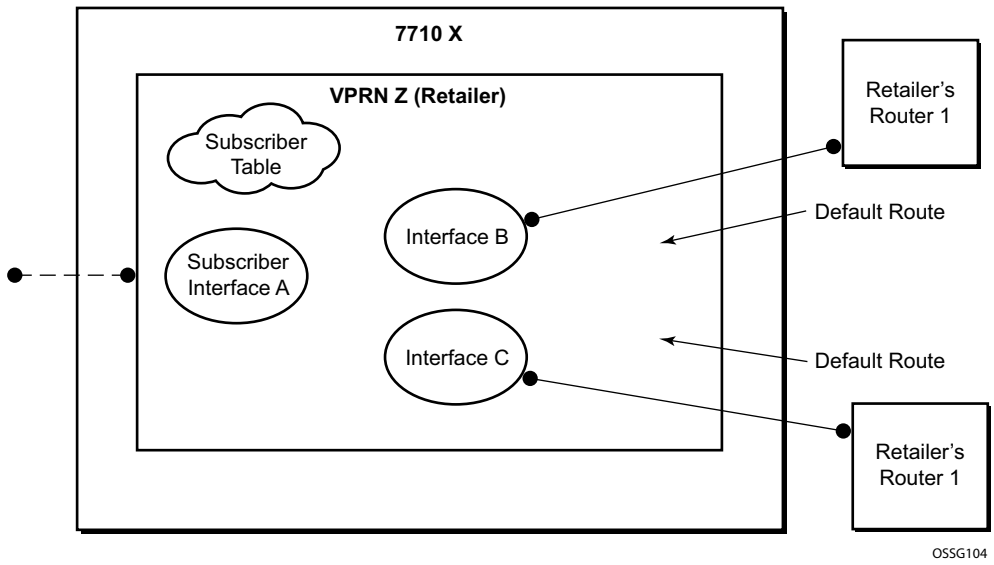
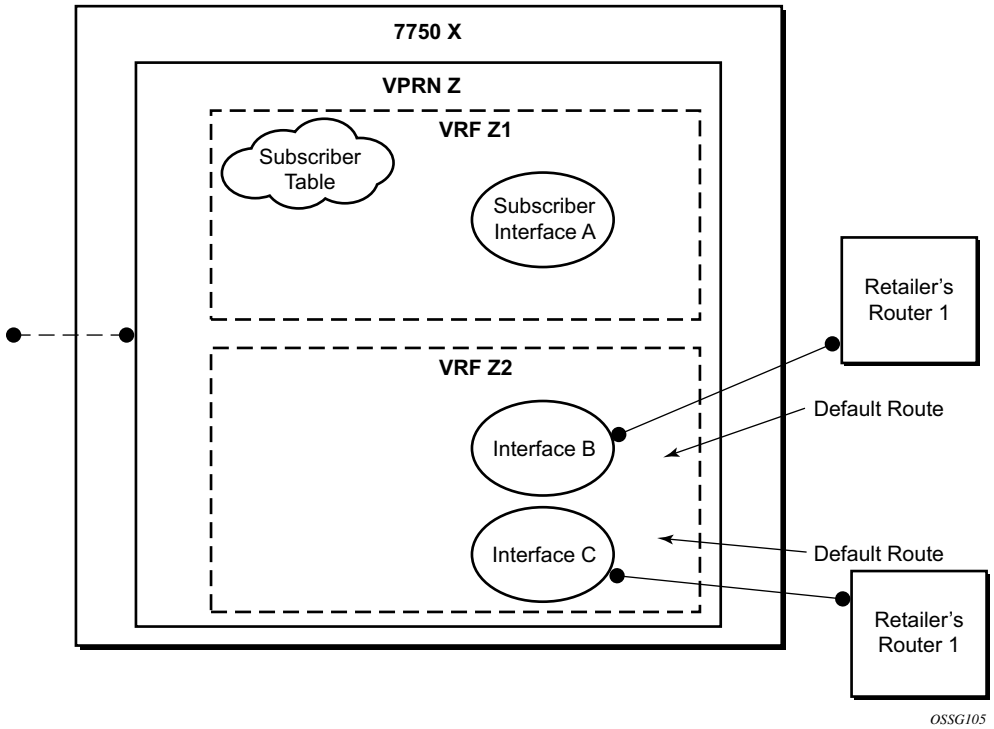



**Figure 106: Forwarding Model Without Wholesale-Retail Type Flag**

Figure 107 describes forwarding following the “type wholesale-retail” flag. Packets coming from the subscriber interface (SAP is defined in the wholesale VPRN) will be routed based on VRF Z2 routes only. Each packet coming from an LSP or Interfaces B or C will be routes based on routes from both VRF Z1 and Z2. This forwarding model is sometimes referred to as subscriber-split-horizon.



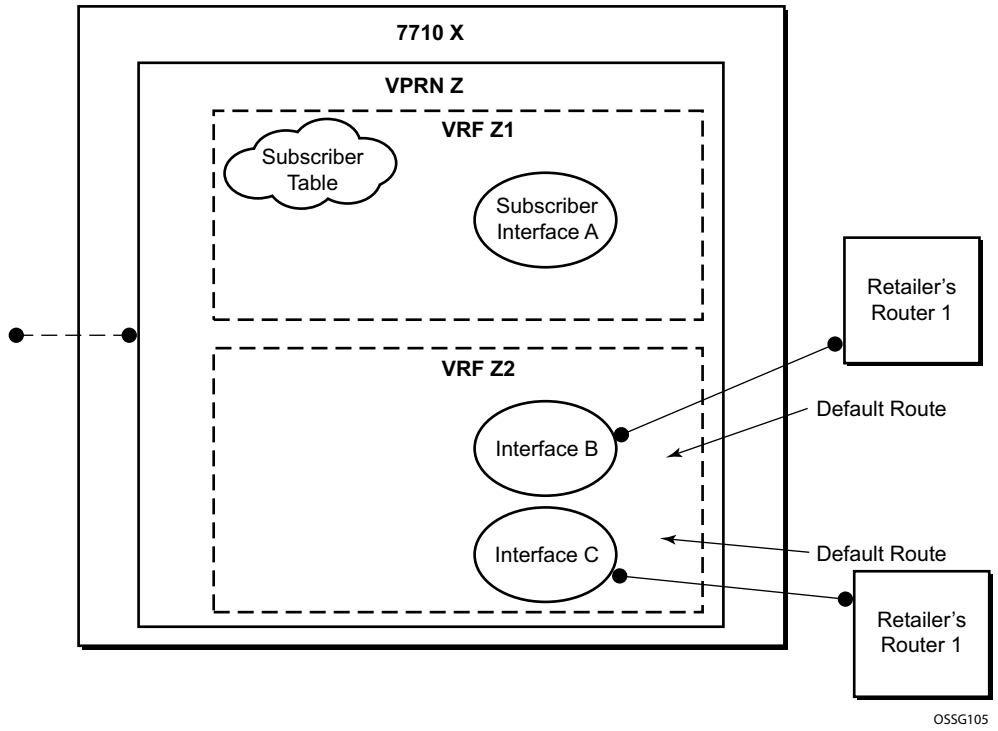



**Figure 107: Forwarding Model With Wholesale-Retail Type Flag**

# Dual Homing

All residential networks are based on two models: Layer 2 CO and Layer 3 CO. Dual homing methods for Layer2 CO include MC-LAG and MC-Ring. Dual homing for Layer 3 CO is based on SRRP and can be done in ring-topologies (l3-mc-ring or with directly attached nodes. All methods use multi-chassis synchronization protocol to sync subscriber state.

## Dual Homing to Two PEs (Redundant-Pair Nodes) in Triple Play Aggregation

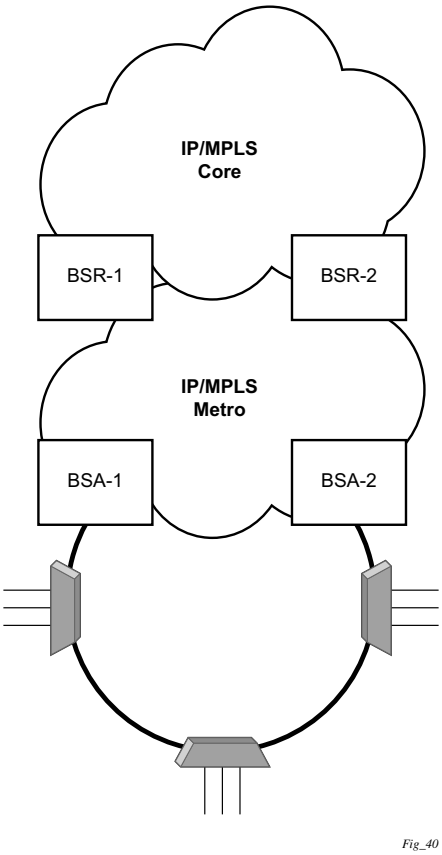



**Figure 108: Dual-Homing to Two PEs**

Figure 108 depicts dual-homing to two different PE nodes. The actual architecture can be based on a single DSLAM having two connections to two different PEs (using MC-LAG) or ring of DSLAMs dual-connected to redundant pair of PEs.

Similarly to previous configuration, both aggregation models (VLAN-per-subscriber or VLAN-per-service) are applicable.

Configurations include:

- Loop resolution and failure recovery — Can be based on MC-LAG or mVPLS.
- DHCP-lease-state persistency — Stores all required information to recover from node failure.
- DHCP-lease-state synchronization — A mechanism to synchronize the DHCP lease-state between two PE nodes in the scope of redundancy groups (a group of SAPs used for dual homing).
- IGMP snooping state synchronization — Similarly to DHCP lease-state synchronization, IGMP snooping state is synchronized to ensure fast switchover between PE nodes. In a VPLS network, a BTV stream is typically available in all PE nodes (the ring interconnecting all PEs with Mcast routers is typically used) so the switch over can be purely driven by RSTP or MC-LAG.
- ARP reply agent responses — The ARP reply agent can response to ARP requests addressing a host behind the given SAP if the SAP is in a forwarding state. This prevents the FDB table in the VPLS from being “poisoned” by ARP responses generated by the node with a SAP in a blocking state (see Figure 109).

Figure 109 shows a typical configuration of network model based on Layer 2 CO model. Individual rings of access nodes are aggregated at BSA level in one (or multiple) VPLS services. At higher aggregation levels (the BSR), individual BSAs are connected to Layer 3 interfaces (IES or VPRN) by spoke SDP termination. Every Layer 3 interface at BSR level aggregates all subscribers in one subnet.

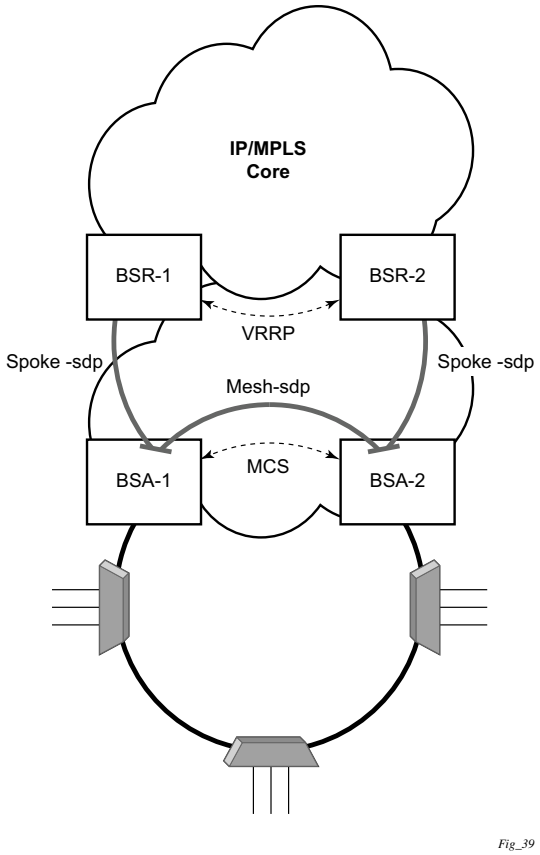



**Figure 109: Layer 2 CO Dual Homing - Network Diagram**

Typically, BTV service distribution is implemented in a separate VPLS service with a separate SAP per access-node. This extra VPLS is not explicitly indicated in Figure 109 (and subsequent figures) but the descriptions refer to its presence.

From a configuration point of view in this model, it is assumed that all subscriber management features are enabled at the BSA level and that synchronization of the information (using multi-chassis synchronization) is configured between redundant pair nodes (BSA-1 and BSA-2 shown in Figure 109 on page 1053). The multi-chassis synchronization connection is used only for synchronizing active subscriber host database and will operate independently from dual-homing connectivity control. At the BSR level, there are no subscriber management features enabled.

The operation of redundancy at the BSR level through VRRP is the same as dual homing based on MC-LAG. The operation of dual homing at BSA level is based on two mechanisms. Ring control connection between two BSAs have two components, in-band and out-of-band communication. With in-band communication, BFD session between BSA-1 and BSA-2 running through the access ring and using dedicated IES/VPRN interface configured on both nodes. This connection uses a separate VLAN throughout the ring. The access nodes provides transparent bridging for this

VLAN. The BFD session is used to continuously verify the integrity of the ring and to detect a failure somewhere in the ring.

With out-of-band communication, the communication channel is used by BSA nodes to exchange information about the reachability of individual access nodes as well as basic configurations in order to verify the consistency of the ring. The configuration information is synchronized through multi-chassis synchronization and therefore it is mandatory to enable multi-chassis synchronization between two nodes using the multi-chassis-ring concept.

In addition, the communication channel used by MC-LAG or MC-APS control protocol is used to exchange some event information. The use of this channel is transparent to the user.

Ring node connectivity check continuously checks the reachability of individual access nodes in the ring. The session carrying the connection is conducted on separate VLAN, typically common for all access nodes. SHCV causes no interoperability problems.

## Steady-State Operation of Dual-homed Ring

Figure 110 illustrates the operation of the dual-homed ring. The steady state is achieved when both nodes are configured in a consistent way and the peering relation is up. The multi-chassis ring must be provisioned consistently between two nodes.

In-Band Ring Control Connection (IB-RCC) is in operational UP state. Note that this connection is set up using a bi-directional forwarding session between IP interfaces on BSA-1 and BSA-2.

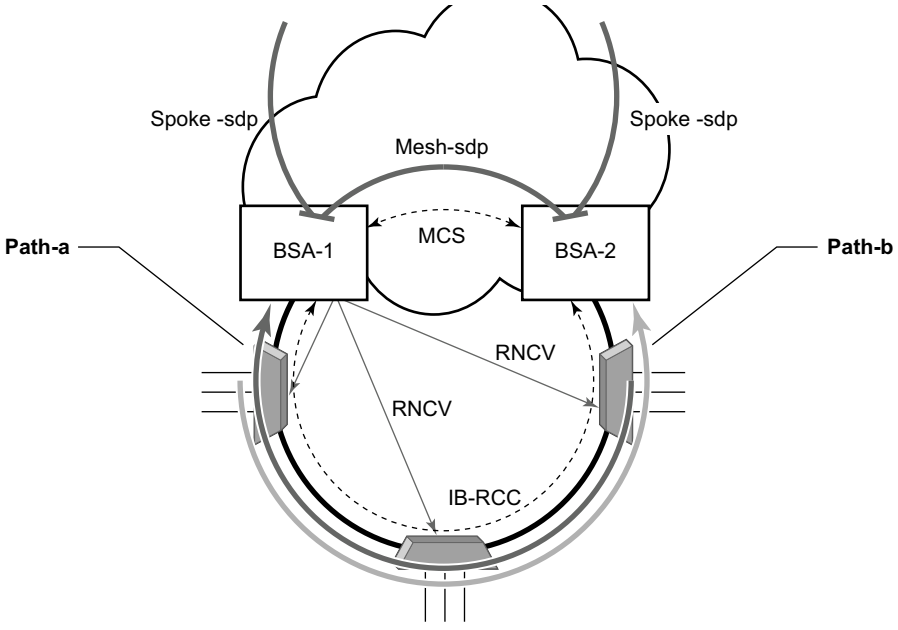



**Figure 110: Dual Homing Ring Under Steady-State Condition**

In Figure 110, the ring is fully closed and every access node has two possible paths towards the VPLS core. Figure 110 refers to these as **path-a** and **path-b**. In order to avoid the loop created by the ring, only one of the paths can be used by any given ring node for any given VLAN. The assignment of the individual VLANs to path-a or path-b, respectively, has to be provisioned on both BSAs.

The selection of the BSA master for both paths will be based on the IP address of the interface used for IB-RCC communication (bi-directional forwarding session). The BSA with the lower IP address of the interface used as IB-RCC channel will become master for ring nodes and their respective VLANs assigned to path-a. The master of path-b will be other BSA.

In this example, each path in the ring has a master and standby BSA. The functionality of both devices in steady state are as follows:

In the master BSA:

- All SAPs that belong to the path where the given BSA is a master, are operationally UP and all FDB entries of subscriber hosts associated with these SAPs point to their respective SAPs.
- The master of a path performs periodical Ring Node Connectivity Verification (RNCV) check to all ring nodes.
- In case of a RNCV failure, the respective alarm will be raised. Note that the loss of RNCV to the given ring node does not trigger any switchover action even if the other BSA appears to have the connection to that ring node. As long as the BFD session is up, the ring is considered closed and the master/standby behavior is driven solely by provisioning of the individual paths.
- The ARP reply agent replies to ARP requests addressing subscriber hosts where the BSA master.

In the standby BSA:

All SAPs that belong to a BSA's path, the standby will be operationally down and all FDB entries of subscriber hosts associated with those SAPs will be pointing towards SDP connecting to master BSA (also called a shunt SDP).

In both BSAs:

- The information on individual paths assignment is exchanged between both BSAs through multi-chassis synchronization communication channel and conflicting SAPs (being assigned to different paths on both BSA nodes) will be forced to path-a (the default behavior).
- For IGMP snooping, the corresponding multi-chassis IDs are targeting all subscriber-facing SAPs on both nodes. On the standby BSA node, the corresponding SAPs are in an operationally down state to prevent the MC traffic be injected on the ring twice.

## Broken-Ring Operation and the Transition to this State

Figure 111 illustrates the model with a broken ring (link failure or ring node failure). This state is reached in following conditions:

- Both nodes are configured similarly.
- Peering is up.
- The multi-chassis ring is provisioned similarly between two nodes
- IB-RCC is operationally down.

In this scenario, every ring node has only one access path towards the VPLS core and hence, the Path-a and Path-b notion has no meaning in this situation.

Functionally, both BSAs are now the master for the reachable ring nodes and will take action as described in Steady-State Operation of Dual-homed Ring on page 1054. For all hosts behind the unreachable ring nodes, the corresponding subscriber host FDB entries point to the shunt SDP.

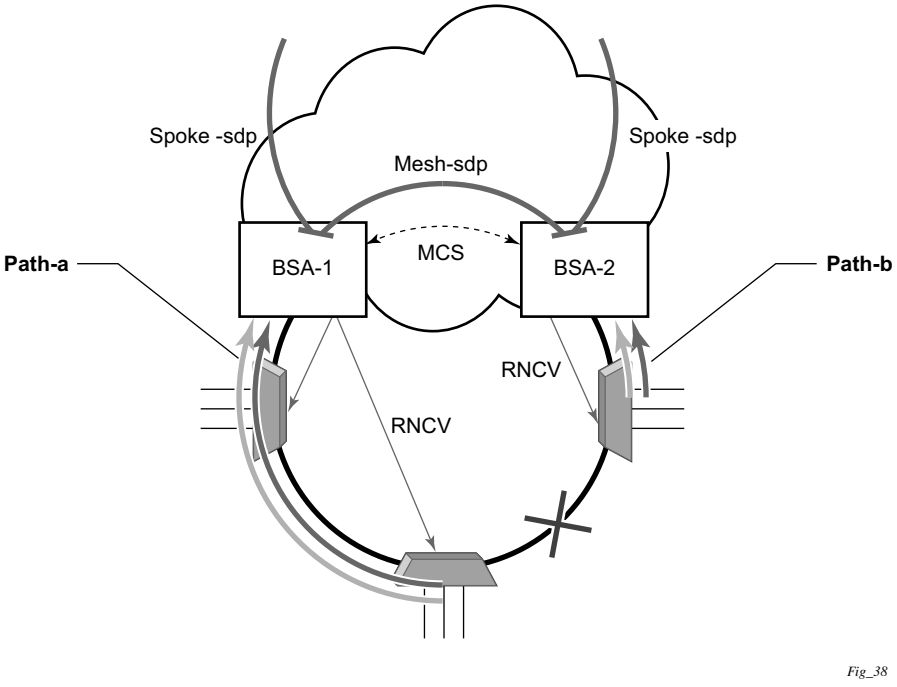



**Figure 111: Broken Ring State**

The mapping of individual subscriber hosts into the individual ring nodes is complicated, especially in the VLAN-per-service model where a single SAP can represent all nodes on the ring. In this case, a given BSA can have subscriber hosts associated with the given SAP that are behind

reachable ring nodes as well as subscriber hosts behind un-reachable ring nodes. This means that the given SAP cannot be placed in an operationally down state (as in a closed ring state), but rather, selectively re-direct unreachable subscriber states to the shunt SDP.

All SAPs remain in an operationally up state as long as the ring remains broken. This mainly applies for BTV SAPs that do not have any subscriber hosts associated with and do not belong to any particular ring node.

In order to make the mapping of the subscriber-hosts on the given ring node automatically provisioned, the ring node identity will be extracted during subscriber authentication process from RADIUS or from a Python script. The subscriber hosts which are mapped to non-existing ring node will remain attached to the SAP.

At the time both BSA detect the break in IB-RCC communication (if BFD session goes down) following actions are taken:

- Both nodes trigger a RNCV check towards all ring nodes. The node, which receives the reply first, will assume a master functionality and will inform the other BSA through an out-of-band channel. This way, the other node can immediately take actions related to the standby functionality without waiting for an RNCV timeout. Even if the other node receives an RNCV response from the given ring node later, the master functionality remains with the node the received the response first.
- After assuming the master functionality for hosts associated with the given SAP(s), the node will send out FIB population messages to ensure that new path towards the VPLS core is established. The FIB population messages are sourced from the MAC address of the default gateway used by all subscriber hosts (such as the VRRP MAC address) which is provisioned at the service level.

## Transition from Broken to Closed Ring State

By its definition, the multi-chassis ring operates in a revertive mode. This means that whenever the ring connectivity is restored, the BSA with lower IP address in the IB-RCC communication channel will become master of the path-a and vice versa for path-b.

After restoration of BFD session, the master functionality, as described in Steady-State Operation of Dual-homed Ring on page 1054, is assumed by respective BSAs. The FDB tables are updated according to the master/standby role of the given BSA and FDB population messages is sent accordingly.

## Provisioning Aspects and Error Cases

The multi-chassis ring can operate only if both nodes similarly configured. The peering relation must be configured and both nodes must be reachable at IP level. The multi-chassis ring with a corresponding sync-tag as a ring-name identifying a local port ID must be provision on both nodes. And, BFD session and corresponding interfaces needs to be configured in a consistent way.

In case the multi-chassis rings are not provisioned consistently, the ring will not become operational and the SAP managed by it will be in operationally up state on both nodes.

The assignment of individual SAPs to path-a and path-b is controlled by configuration of VLAN ranges according to the following rules:

- By default, all SAPs (and hence all VLANs on the given port) are assigned to path-a.
- An explicit statement defining the given VLAN range assigns all SAPs falling into this range to the path-b.
- An explicit statement defining the given VLAN range defines all SAPs that are excluded from the multi-chassis ring control.
- In case of a conflict in the configuration of VLAN ranges between two redundant nodes is detected, all SAPs falling into the “conflict-range” will be assigned to path-a, on both nodes regardless the local configuration.
- For QinQ-encapsulated ports the VLAN range refers to the outer VLAN.

## Dual Homing to Two BSR Nodes

Figure 112 depicts a single DSLAM dual-homed to two BSRs.

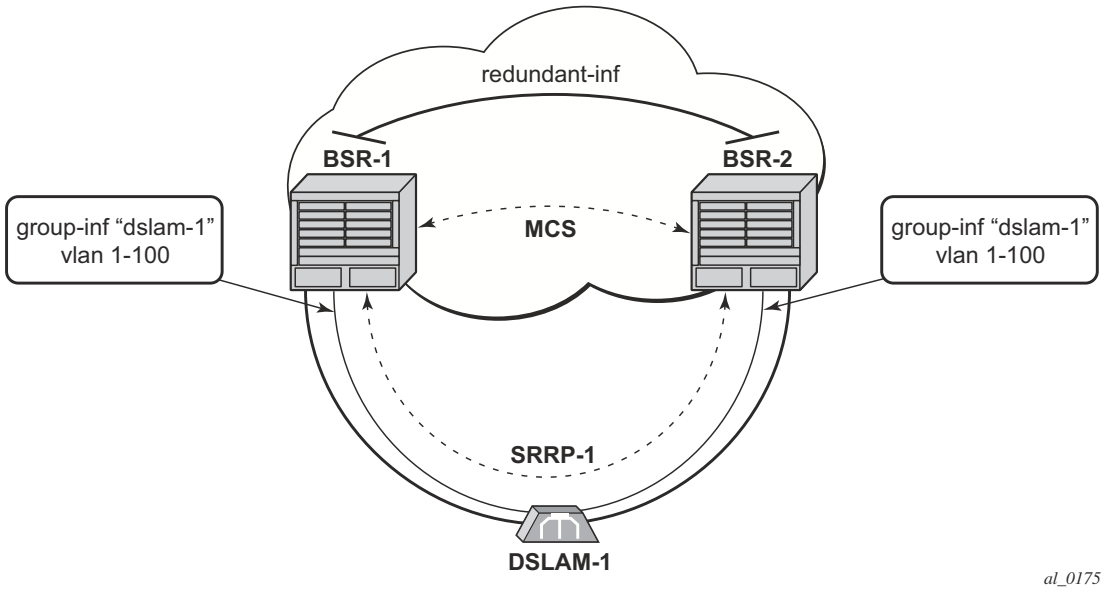



**Figure 112: Low**

In order to provide dual-homing in the context of subscriber interfaces, the following items must be configured on both BSRs:

- Group interface (dslam-1) with corresponding SAPs (vlan 1-100)
- SRRP instance controlling given group interface
- Redundant interface between BSRs to provide “shunt” connectivity
- MCS connection to provide synchronization of dynamic subscriber-host entries

During the operation, BSR-1 and BSR-2 will resolve master-backup relation and populate respective FIBs in such a way that at master side, subscriber-host entries point to corresponding group-interface while at the back-up side, subscriber-host entries point to the redundant interface. Note that the logical operation of the ring in the Layer 3 CO model is driven by SRRP. For more details on SRRP operation, refer to Subscriber Routed Redundancy Protocol (SRRP) on page 1001.

## MC Services

The typical implementation of MC services at the network level is shown in Figure 113.

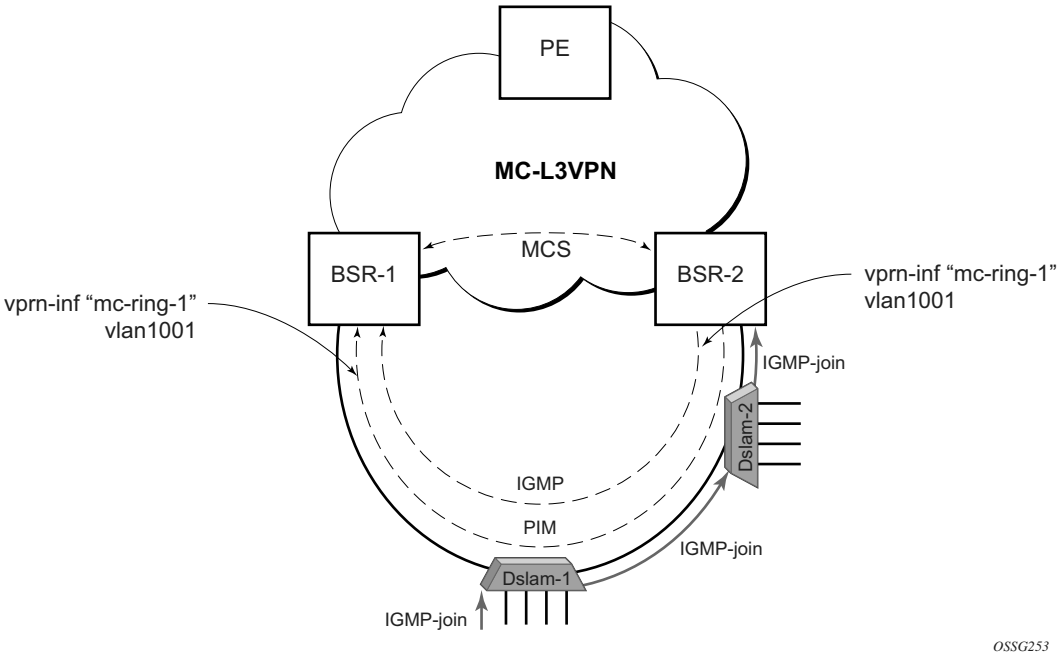



**Figure 113: MC Services in a Layer 3-Ring Topology (a)**

The IGMP is used to register joins and leaves of the user. IGMP messaging between BSRs is used to determine which router performs the querier role (BSR2 in Figure 113). PIM is used to determine which router will be the designated router and the router that sends MC streams on the ring.

The access nodes have IGMP snooping enabled and from IGMP messaging between BSR, they are aware which router is the querier. In the most generic case, IGMP snooping agents (in access nodes) send the IGMP-joins messages only to IGMP-querier. The synchronization of the IGMP entries can be then be performed through MCS. In some cases, access nodes can be configured in such a way that both ring ports are considered as m-router ports and IGMP joins are sent in both directions.

All of the above is a steady state operation which is transparent to the topology used in a Layer 2 domain.

A ring-broken state is shown in Figure 114.

In this case, IGMP and PIM messaging between BSRs is broken and both router assume role of querier and role of designated router. By the virtue of ring topology, both routers will see only

IGMP joins/leaves generated by host attached to a particular “half” of the ring. This means that both routers will have “different” view on dynamic IGMP state.

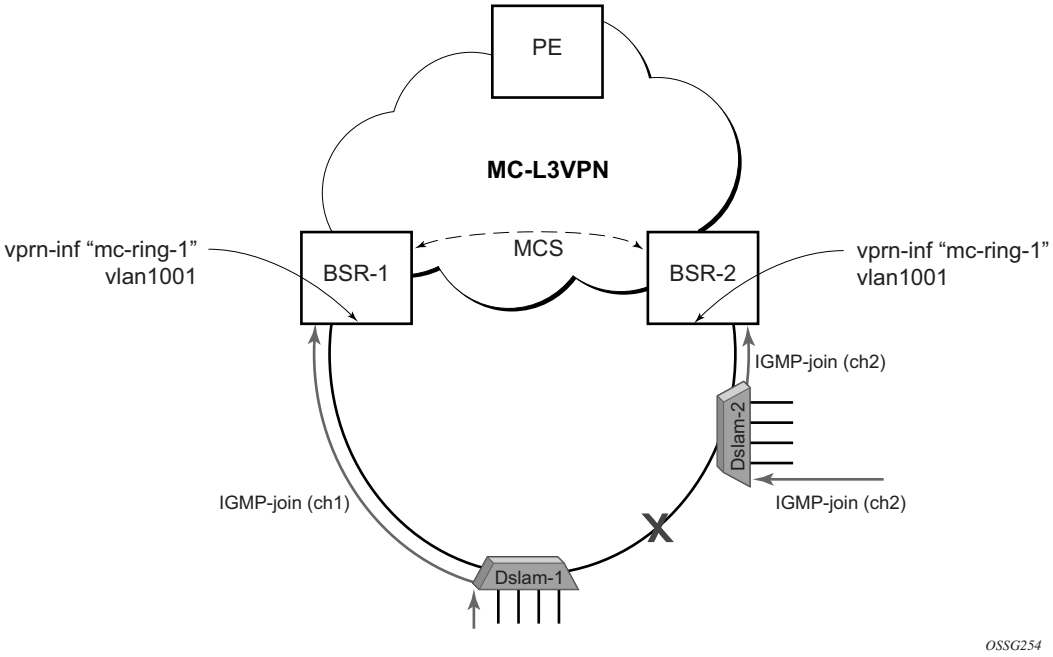



**Figure 114: MC Services on a Layer 3-Ring Topology (b)**

In principle, MCS could be used to synchronize both routers, but in case of a Layer 2 ring, the implementation sends all IGMP messages to a “ring-master: which then performs IGMP processing and consequently, MCS sync. As a result, any race conditions are avoided.

Another ring-specific aspect is related to ring healing. The ring continuity check is driven by BFD which then drives SRRP and PIM messaging. BFD is optimized for fast detection of ring-down events while ring-up events are announced more slowly. There is a time window when routers are not aware that the ring is recovered. In the case of MC, this means traffic will be duplicated on the ring.

To avoid this, the implementation of BFD provides a “raw mode” which provides visibility on “ring-up” events. The protocols, such as SRRP and PIM, use this raw mode rather than the BFD API.

## Routed CO Dual Homing

Routed CO dual homing is a solution that allows seamless failover between nodes for all models of routed CO. In the dual homed environment, only one node will forward downstream traffic to a given subscriber at a time. Dual homing involves several components:

- Redundant Interface — This is used to shunt traffic to the active node for a given subscriber for downstream traffic.
- SRRP — This is used to monitor the state of connectivity to the DSLAM. Refer to the SRRP section for more detail.
- MCS — This is used to exchange subscriber host and SRRP information between the dual homed nodes.

Routed CO dual homing can be configured for both wholesaling models. Dual homing is configured by creating a redundant interface that is associated with the protected group interfaces. The failure detection mechanism can be VRRP. If VRRP is used, each node monitors the VRRP state to determine the priority of its own interface.

Dual homing is used to aggregate a large number of subscribers in order to support a redundancy mechanism that will allow a seamless failover between nodes. Because of the Layer 3 nature of the model, forwarding is performed for the full subscriber subnet.

### Redundant Interfaces

In dual homing, a redundant interface must be created. A redundant interface is a Layer 3 spoke SDP-based interface that allows delivery of packets between the two nodes. The redundant interface is required to allow a node with a failed link to deliver packets destined to subscribers behind that link to the redundant node. Since subscriber subnets can span multiple ports it is not possible to stop advertising the subnet, thus, without this interface the node would black hole.

The redundant interface is associated with one or more group interfaces. An interface in backup state will use the redundant interface to send traffic to the active interface (in the active node). The SAP structure under the group interface must be the same on both nodes as the synchronization of subscriber information is enabled on a group interface basis. Traffic can be forwarded through the redundant interface during normal operation even when there are no failed paths. See Figure 115.

### SRRP in Dual Homing

Subscriber Router Redundancy Protocol (SRRP) allows two separate connections to a DSLAM to operate in an active/standby fashion similar to how VRRP interfaces operate. Since the SRRP state is associated with the group-interface, multiple group-interfaces may be created for a given port such that some of the SAPs will be active in one node and others active on the other node. While each SRRP pair is still allowed to be active/backup, the described configuration is allowed for load balancing between the nodes. Note that in a failure scenario subscriber bandwidth will be affected. For more information about SRRP, refer to Subscriber Routed Redundancy Protocol

(SRRP) on page 1001.

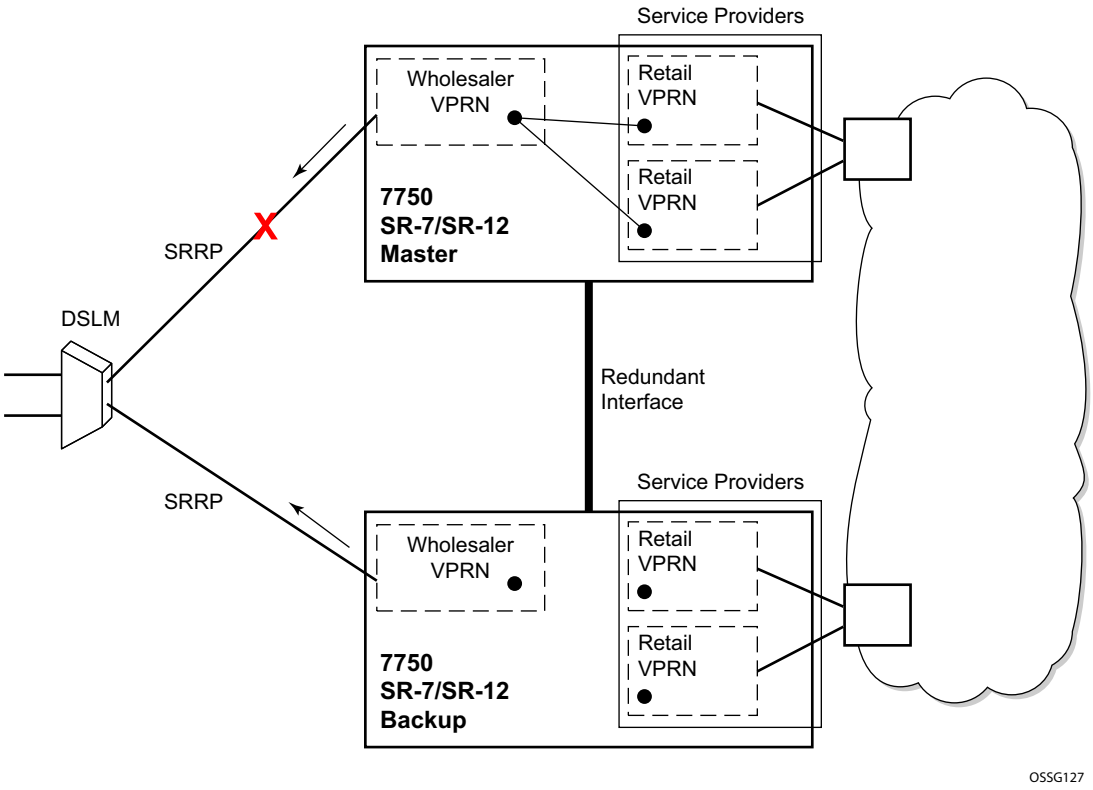
If SRRP is configured before the redundant interface is up, and in backup state the router will forward packets to the access node via the backup interface but will not use the gateway MAC address. This applies to failures in the redundant interface as well. If the redundant interface exists and up the router will send downstream packets to the redundant interface and will not use the backup group interface.

In a dual homing architecture the nodes must be configured with SRRP to support redundant paths to the access node. The nodes must also be configured to synchronize subscriber data and IGMP state. To facilitate data forwarding between the nodes in case some of the ports in a given subscriber subnet are affected a redundant interface must be created and configured with a spoke. The redundant interface is associated with one or more group interfaces.

The service IDs for both the wholesale VPRN and the retailer VPRN must be the same in both nodes.

An interface in backup state will use the redundant interface to send traffic to the active interface (in the active node). The SAP structure under the group interface must be the same on both nodes as the synchronization of subscriber information is enabled on a group interface basis.

SRRP is associated a group-interface. Multiple group-interfaces can be created for a given port such that some of the SAPs will be active in one node and others active on the other node. While every SRRP pair is still allowed to be active/backup the described configuration will allow for load balancing between the nodes. Note that in a failure scenario subscriber bandwidth will be affected.


Service Providers
Wholesaler VPRN
Retail VPRN
Retail VPRN
7750 SR-7/SR-12 Master
SRRP
X
DSLM
Redundant Interface
SRRP
Service Providers
Wholesaler VPRN
Retail VPRN
Retail VPRN
7750 SR-7/SR-12 Backup
OSSG127

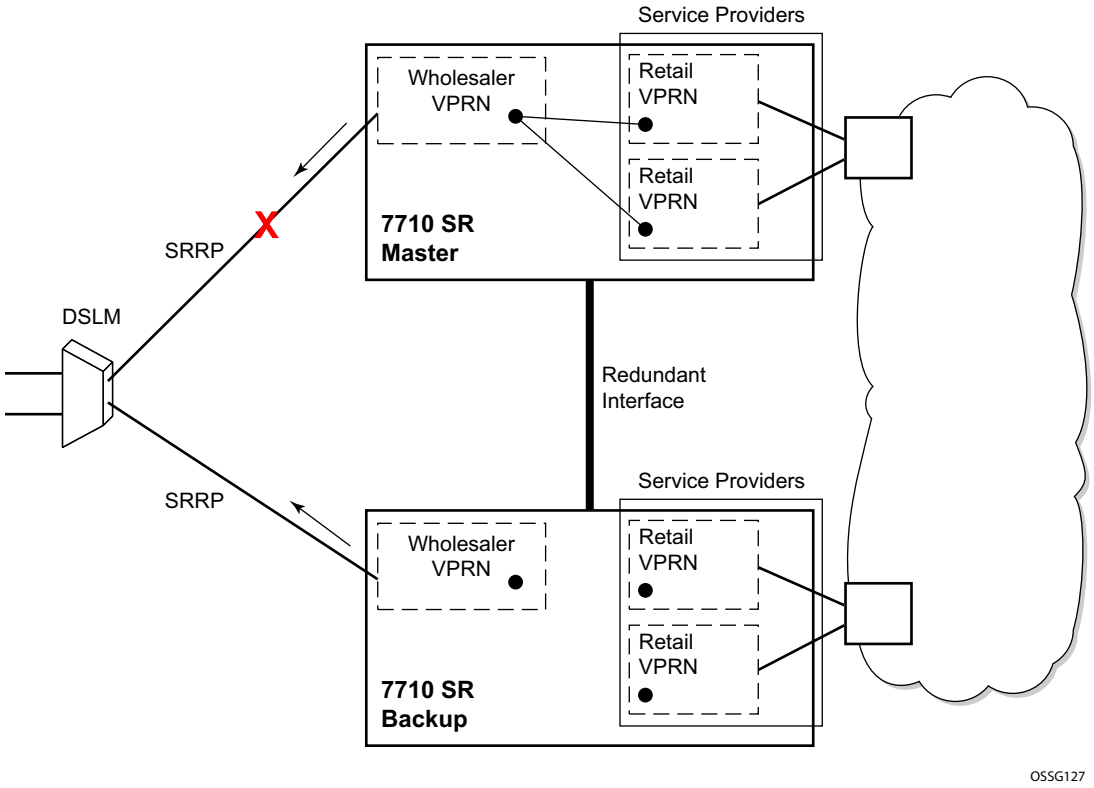



**Figure 115: Dual Homing Example**

## Synchronization

To establish subscriber state the nodes must synchronize subscriber information. Refer to the 7750 7710 SR OS Basic Configuration Guide for multi-chassis synchronization configuration information. The operator must complete the configuration and the system must have data synchronized before the backup node may deliver downstream packets to the subscriber.

If dual homing is used with regular interfaces that run IGMP the nodes must be configured to synchronize the Layer 3 IGMP state.

The service IDs for both the wholesale VPRN and the retailer VPRN must be the same in both nodes.

## SRRP and Multi-Chassis Synchronization

In order to take full advantage of SRRP resiliency and diagnostic capabilities, the SRRP instance will be tied to a MCS peering that terminates on the redundant node. 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. 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)

---

## Dual Homing and ANCP

Alcatel-Lucent provides a feature related to exchange of control information between DSLAM and BRAS (BSA is described in this model). This exchange of information is implemented by in-band control connection between DSLAM and BSA, also referred to as ANCP connection.

In case of dual homing, two separate connections will be set. As a consequence, there is no need to provide synchronization of ANCP state. Instead every node of the redundant-pair obtains this information from the DSLAM and creates corresponding an ANCP state independently.

# SRRP for IPv6 ESM

The Subscriber Routed Redundancy Protocol (SRRP) feature provides for the following:

- A rapid failover of IPv6 default-router from an ESM host perspective.
- The synchronisation of IPv6 ESM host state between two chassis
- The redirection of downstream (or peer-to-peer) traffic to an alternate chassis that has an active host.

Although an IPv6 host has built-in support for multiple upstream default routers, the failover and detection time can exceed ten's of seconds. Implementing SRRP for IPv6 ESM allows a switchover and restoration of forwarding within seconds.

To support SRRP for IPv6 ESM hosts, the 7750 SR is implemented with VRRPv3 (RFC 5798) between a pair of IPv6-enabled group-interfaces to determine master and backup roles. VRRPv3 provides a standard discovery and election process along with a path keep-alive.

The VRRPv3 RFC explicitly states that the VRRP virtual MAC address shall not be used in the creation of Interface-Id, but does not prohibit the generation of the link-local address from the virtual MAC. In the SRRPv6 implementation, this is what occurs: a virtual MAC is generated based on the SRRP instance (for example, srrp 1 results in a virtual MAC of 00-00-5E-00-02-01 and the protected/virtual link-local address would be fe80::200:5eff:fe00:201).

When SRRP is enabled on a group-interface that has IPv6 support, the VRRP virtual MAC is used to generate an Interface-Id, and in turn, link-local address is used by ESM subscriber's as the next-hop IP address (learnt through router advertisements). This link-local address is used as the IPv6 source address and the virtual MAC address used for link-layer target (when present) in any downstream Neighbour Discovery messages (including router advertisements).

Note: Per the VRRPv3 specification, router advertisements must be disabled on the standby router, and Duplicate Address Detection (DAD) should not be sent out for the link-local address generated from the virtual MAC address.

VRRPv3 provides a facility to protect a number of individual addresses. For ESM, only the link-local address exists on the subscriber-interface and is the only VRRP protected address. The subscriber subnet's configured under the subscriber-interface do not result in the creation of an IP address on the 7750 and do not need protecting.

VRRPv3 is used to elect a master router for ESM operation. Whenever a group-interface is master for IPv6, all ESMv6 hosts is associated.

SRRPv6 is configured through the addition of an ipv6 keyword under group-interface > srrp.

## SRRP Enhancement

The SRRP enhancements addressed in this section is to reduce the need for redundant-interface between the pair of redundant nodes without sacrificing the subnet aggregation on the back-end.

Redundant BNG nodes are not always collocated. This means that the logical link associated with the redundant (shunt) interfaces is taking the uplink path thus wasting valuable bandwidth (downstream traffic that arrives to the Standby node is routed via uplinks for the second time over to the Master node).

To meet the requirement to reduce the existence of shunted traffic only to the short transitioning period between SRRP switchovers while the routing on the network side is converging, the following was required (referring to Figure 116):

1. Share IP subnets over multiple SRRP instances. This is not mandatory, but it would helps to load balance traffic over the two nodes. For example, IP subnets 10 and 11 can be shared over SRRP instances 10 and 20 on node 1, and the IP subnet 12 can be associated with the SRRP instance 30 on node 2.
2. *SRRP aware routing* – this allows to dynamically increase routing metric on the IP subnets advertised from the Master SRRP node in comparison to the Standby SRRP node. It also allows to advertise/withdraw routes from a routing protocol based on the SRRP state. In this fashion, downstream traffic is routed in a predictable manner towards the Master SRRP node.
3. *SRRP Fate Sharing* for SRRP instances 10 and 11. This ensures congruency of SRRP states on the same node. This is a necessary step towards *SRRP aware routing*.

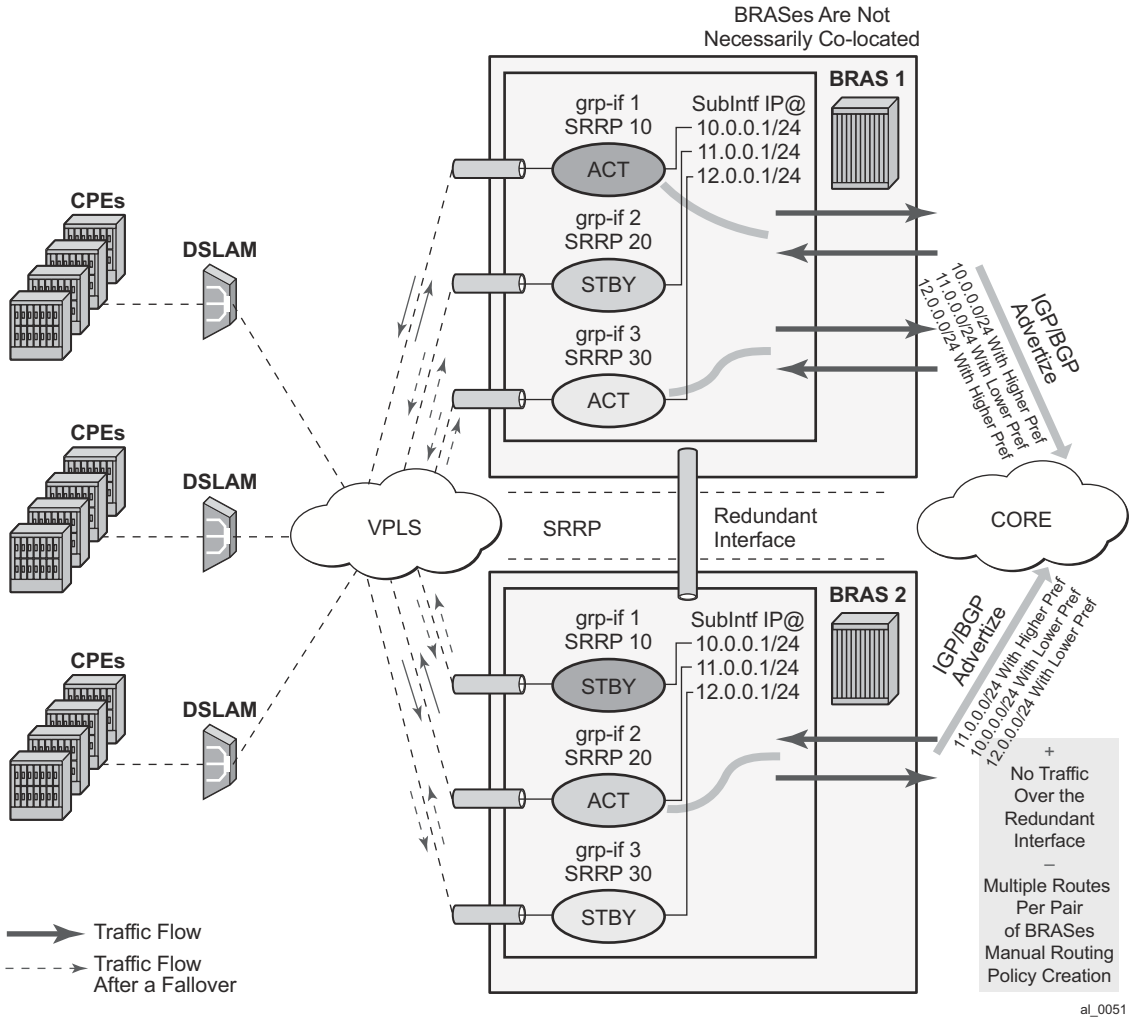



**Figure 116: IP Subnet Per SRRP Master Group**

## SRRP Fate Sharing

SRRP Fate Sharing is a concept in which a group of SRRP instances track a single operational-object comprised of SRRP messaging SAPs. The SRRP instances behave as one (in the single failure case) with regards to SRRP mastership. The group of SRRP instances that are sharing fate on a paired node are referred as a Fate Sharing Group (FSG).

Transition of a single messaging SAP within the FSG into a DOWN state forces the SRRP instance on top of it into the INIT state. Consequently, all other SRRP instances within the same FSG transitions into a Backup state. In other words, SRRP instances within the FSG all share the same fate as the failed SRRP instance as shown in Figure 117. SRRP Fate Sharing provides

optimal protection in the context of a single failure in the network.

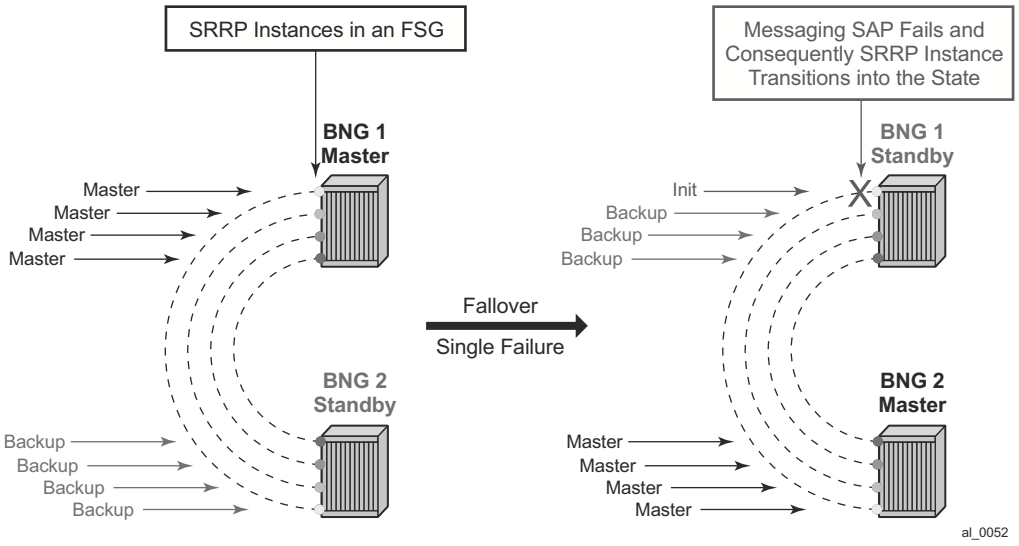



**Figure 117: FSG –— Single Network Failure**

In the case of multiple network failures, the concept of the FSG breaks as there is a possibility that a ‘FSG’ contains SRRP instances that are in any of the three possible SRRP states: Master, Backup, or Init. This Fate Sharing feature may not provide optimal protection when there are multiple network failures distributed over both redundant nodes.

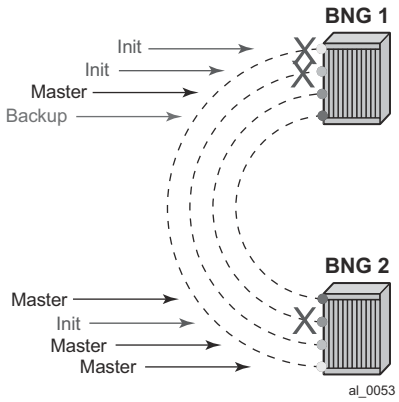



**Figure 118: Multiple Network Failures**

The whereabouts of the failure in the network path that SRRP is designed to monitor are not always clearly reflected through SRRP states. For example, if the network failure is somewhere in the aggregation network beyond the direct reach of our BNG, SRRP assumes Mastership on both BNG nodes. This is a faulty condition and the reason why solely monitoring of the SRRP states is not enough to protect against failures. On the other hand, the SRRP messaging SAP states are more indicative of the network failure since they can be tied into Eth-OAM.

Once a single network failure is detected and as a result an SRRP instance transitions into a non-Master state, the remaining SRRP instances in the FSG are forced into a Backup state. This is achieved by changing the priority of each individual SRRP instance in the FSG.

In the case of simultaneous multiple failures (multiple ports fail at the same time), it is possible that the SRRP instances within the FSG settle in any of the three possible SRRP states: Master, Backup, or Init.  In such scenario shunted traffic will ensue.

In the premise of SRRP Fate Sharing, the network failure will be reflected in the operational state of the messaging SAP over which SRRP runs. This will certainly be the case if the failure is localized to the BNG (somewhere on the directly connected link). In the case of non-localized failure (beyond the direct reach of the BNG node), Eth-OAM might be needed in to detect the remote end failure and consequently bring the SAP operationally into a DOWN state.

Once the single network failure is detected, all instance within the FSG transitions into a non-Master state.

If there are no failures in the network, all SAPs are UP and SRRP instances within the FSG are in a homogeneous and deterministic state based on their configured priorities.

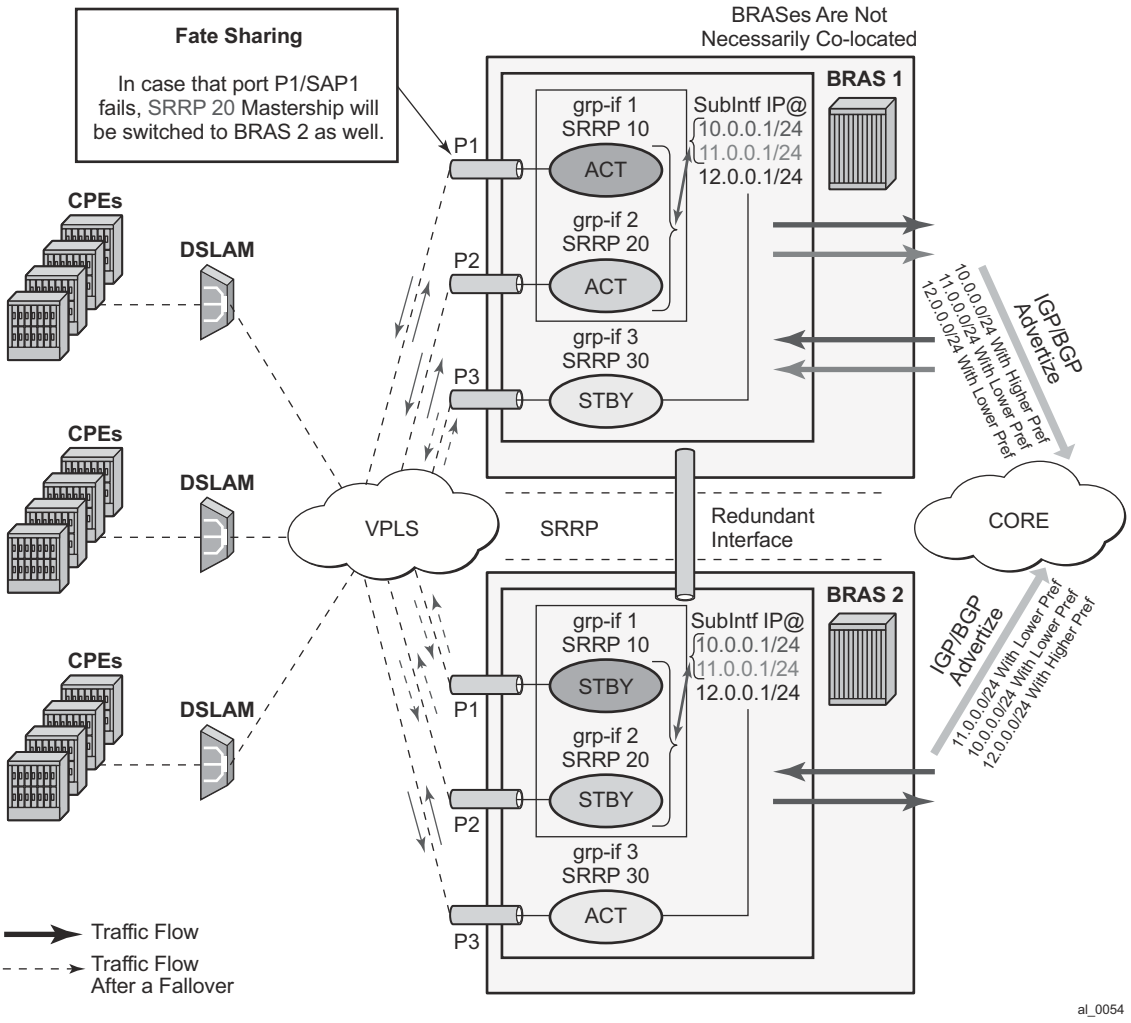



**Figure 119: SRRP Fate Sharing**

## Failure Detection in a Fate Sharing Group

1. Dual homing over directly connected ports.
   No Eth-OAM is needed, AN is directly connected to the BNG.

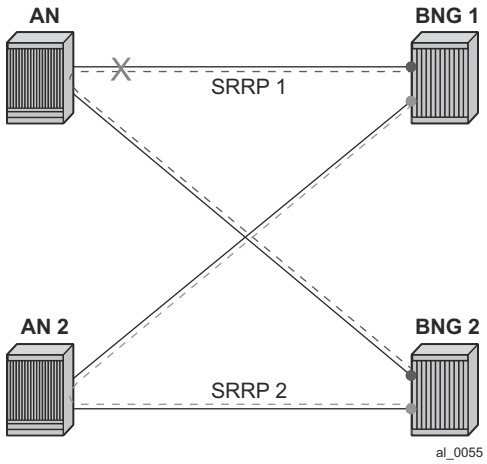



**Figure 120: Scenario 1**

2. Dual homing with aggregation network - aggregation network has no redundancy between Layer 2 switches (STP). To determine whereabouts of failure at point 1 in the figure below, Eth-OAM is needed.

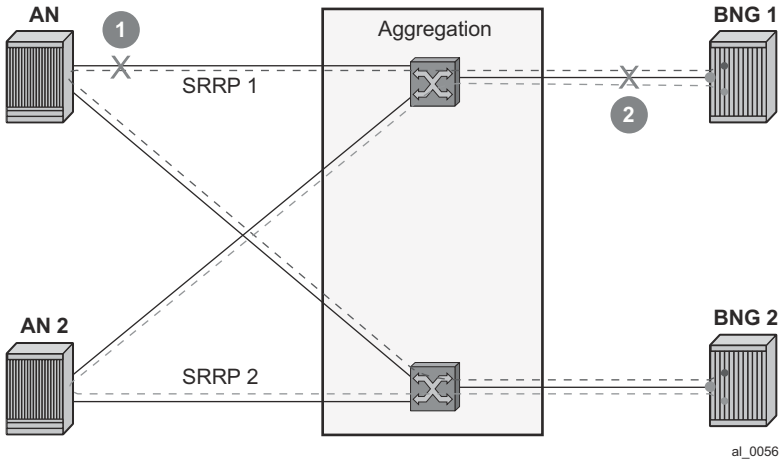



**Figure 121: Scenario 2**

3. Dual homing with aggregation network - aggregation network with redundancy between Layer 2 switches (STP).
   No Eth-OAM is needed in this case for successful operation. However, the failure detection is based on the failure of the directly attached ports.

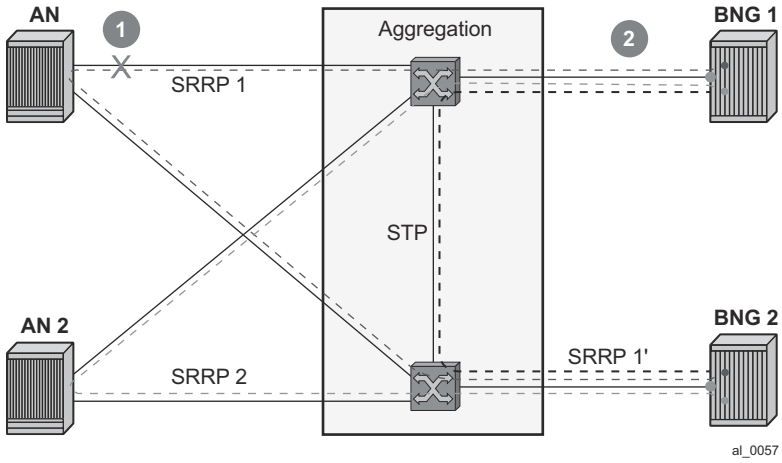



**Figure 122: Scenario 3**

4. Single homing with aggregation network.
   In this case, SRRP can protect only against direct failures. Any remote failure leaves a part of the network isolated from the subscriber point of view.

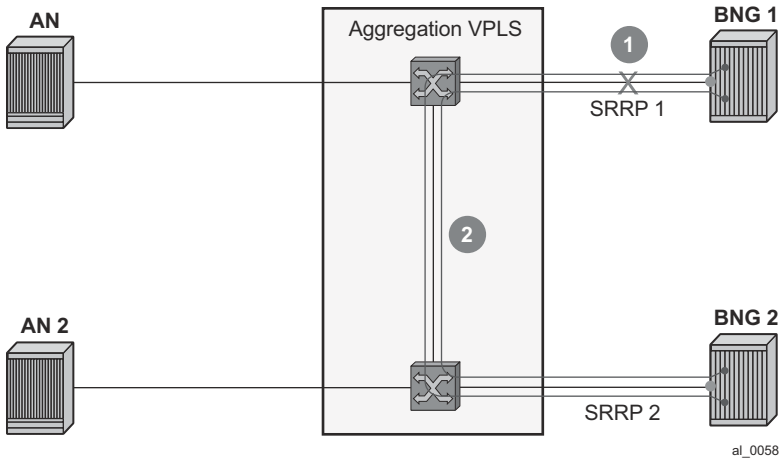



**Figure 123: Scenario 4**

## Fate Sharing Algorithm

Fate Sharing Algorithm (FSG) is relaying on tracking the state of messaging SAPs over which SRRP instances run. An SRRP instance with the messaging SAP operationally DOWN will transition into the Init state.

The transitioning of any messaging SAP in a FSG into an UP/DOWN state will trigger SRRP priority adjustment within the FSG. The SRRP priorities should be chosen carefully to achieve the desired behavior. They are modified dynamically as the SAP states change. The range in which SRRP priorities can be modified is from 1 to the SRRP priority that is initially configured under the SRRP node. Here are some general guidelines for choosing SRRP priorities in a FSG:

- Initially configured SRRP priorities for all SRRP instance within the FSG within the node should be the same.
- Initially configured SRRP priorities should be different between pairing FSGs. For example, SRRP instances in the BNG node A within an FSG will all have the same SRRP priority 'X', while corresponding SRRP instances on the paired node within corresponding FSG will all have SRRP priority 'Y'. This ensures that SRRP mastership is clearly defined between the two BNG nodes. Note that this step is not mandatory as SRRP will naturally break the Mastership tie in the case that all SRRP priorities are the same. However, following this step may provide a clearer view from an operational perspective.
- The priority-step used for dynamic SRRP priority adjustment must be greater than the difference in initially configured SRRP priorities between two BNG nodes. This ensures that a single failure event triggers the SRRP switchover. Otherwise, if the dynamically lowered SRRP priority is still greater than the one from the SRRP peer, the switchover would not be triggered. Therefore, the fate sharing concept would not function as intended.
- Initially configured SRRP priority of each SRRP instance should be greater than the (anticipated) number of SRRP instances in a FSG multiplied by the SRRP priority-step. This ensures that the dynamically priority never tries to go below 1. There is a code check that prevents SRRP priority going below 1. Nonetheless, it is recommended not to get into a situation where this needs to be enforced in the code.

**Note:** The priorities will never be less than 1 or greater than initially configured SRRP priority.

Example scenarios:

Assume 3 SRRP instances in a FSG. The SRRP instances in the FSG-1 on BNG 1 have the priority of 100, while the SRRP instances in the FSG-2 on BNG 2 have the priority of 95. The priority-step is 6. The SRRP instances and underlying messaging SAPs will be referred as SRRP 1, 2, 3 and SAP 1,2,3, respectively.

Initialization:

Scenario 1 – all SAPs are operationally UP.

BNG 1 boots up and all messaging SAPs transition into the UP state. When the first SRRP instance in FSG-1 comes up, it looks under the FSG to finds out how many messaging SAPs are operationally UP. Since all messaging SAPs are operationally UP, this first SRRP instance

assumes its initially configured priority of 100. The other two SRRP instances in the same FSG follows the same sequence of events.

BNG 2 follows the same flow of events. As a result, BNG 1 assumes mastership over BNG 1 for all SRRP instances within the corresponding FSG.

Scenario 2 – messaging SAP 1 is operationally DOWN on BNG 1, the rest of the messaging SAPS are operationally UP.

SRRP 2 and 3, during the initialization, pick up SRRP priority of 94 (100 – 1*priority-step).

On BNG 2, all messaging SAPs are UP and consequently all SRRP instances within the FSG on BNG 2 have SRRP priority of 95. The SRRP instances on BNG 2 assumes Mastership.

Scenario 3 – Continuing from scenario 2, the SAP 1 on BNG 1 transitions into the UP state. SRRP priority of each SRRP instance in FSG-1 is increased by 6, bringing it to 100, enough to assume Mastership.

---

### Adding a New Instance into an FSG

To introduce minimal network disruption, first create messaging SAPs in both BNG nodes and ensure that both SAPs are operationally UP. Then a new SRRP 4 instance should be created on both BNG nodes. The next step would be to include this new messaging SAP into a SAP monitoring group. And finally, the SRRP-4 is added into the FSG (1 and 2). This triggers the recalculation of SRRP priorities for the existing FSG-1 and FSG-2. Since all SRRP priorities are at the max (initially configured priority), nothing changes.

There are more disruptive ways of adding an SRRP instance into a FSG. One such example would be in the case where SRRP priorities are not at their maximum (initially configured) priority. If an SRRP instance is first added into an FSG that is in a ‘Backup’ state, this would increase the FSG priority and potentially cause a switchover. If the SRRP instances is then added in a FSG on the peer BNG (previously Master), the priority of this FSG would be increased again and the switchover would unnecessarily occur for the second time. The new SRRP instances, once operational, should always be added in the Master FSG first.

SRRP priority re-calculation within the FSG is triggered by the following events:

- SRRP initialization
- addition of a SAP under the monitoring group
- messaging SAP failure

This priority calculation looks into how many SAPs are in the DOWN state within the monitored SAP group. Based on this number, the priority is calculated as follows:

SRRP priority = configured-priority – priority-step * num_down_SAPs.

## SRRP Aware Routing - IPv4/IPv6 Route Advertisement Based on SRRP State

There are three cases that need to be covered, each case with its own specifics:

- Subscriber Interface Routes (IPv4/IPv6)
- Managed Routes
- Subscriber Management Routes (/32 IPv4 hosts routes and IPv6 PD wan-host routes)

Depending on the route type, the action is to either modify the route metric based on the SRRP state that the route is tracking, OR to advertise/withdraw the route based on the SRRP state that the route is tracking. The action is defined in the routing policy and it is based on the new attributes with which the routes are associated.

To achieve a better granularity of the routes that are advertized, an origin attribute is added to the subscriber management routes (/32 IPv4 routes and IPv6 PD wan-host) with three possible values:

**aaa**

IPv4

*subscriber-management* /32 host routes that are originated through RADIUS framed-ip-address VSA other than 255.255.255.254. The 255.255.255.254 returned by the RADIUS indicates that the BNG (NAS) should assign an IP address from its own pool.

IPv6

*subscriber-management* routes that are originated through framed-ipv6-prefix (SLAAC), delegated-ipv6-prefix (IA_PD) or alc-ipv6-address (IA_NA) RADIUS attributes . This is valid for IPoE and PPPoE type host.

**dhcp**

IPv4

*subscriber-management* /32 host routes that are originated via DHCP server (local or remote) and also RADIUS framed-ip-address=255.255.255.254 (RFC 2865).

IPv6

*subscriber-management* routes that are assigned via local DHCPv6 server pools whose name is obtained through Alc-Delegated-IPv6-Pool (PD pool) and Framed-IPv6-Pool (NA pool) RADIUS attributes. This is valid for IPoE and PPPoE type hosts.

In addition, for IPoEv6 only, the pool name can be also obtained via ipv6-delegated-prefix-pool (PD pool) and ipv6-wan-address-pool (NA pool) from LUDB.

**ludb**

IPv4

subscriber-management /32 host routes that are originated via LUDB. This also covers RADIUS fallback category (RADIUS falls back to system-defaults or to LUDB).

IPv6

subscriber-management routes obtained from LUDB via ipv6-address (IA_NA) or ipv6-prefix (IA_PD). This is supported only for IPoE.

Overall, the following new route attribute is added:

*state: srrp-master, srrp-non-master*

The existing origin attribute is expanded to contain the following values:

*origin: aaa, dhcp, ludb*

These two attribute types are applied in the following fashion:

The state attribute is applied to all three route types: *subscriber interface routes, managed routes* and *subscriber management routes*. Each route listens to the SRRP state.

If an attribute is defined in the routing policy as a match condition (from statement) but the route itself does not have this attribute, the route is evaluated into a non-match condition.

The origin attribute is always applied only to subscriber management routes. No additional statement is needed to explicitly apply this attribute as it may be the case for the state attribute.

Every time there is a change in the attribute associated with the route, the route is re-evaluated in the RTM by the routing policy and corresponding action is taken.

---

## Subscriber Interface Routes (IPv4 and IPv6)

Optimized routing and elimination of downstream shunt traffic during normal operation can be achieved by **statically** favoring the routes on the network side that are advertized with an increased metric by Master SRRP nodes.

The downside of this static approach is that during the port/card failure and consequently a SRRP switchover, the node with the failed port/card will continue to advertise routes with the same high metric as long as the subscriber interface is in the 'UP' state (or a single SAP under it). That is, the network side will not be aware of the switchover. It will continue to forward traffic to the standby node, and as a result, heavy shunt traffic will ensue. To effectively deal with this, the network side must be aware of the routing change that occurred in the access layer.

When failure is detected, the metric for the route is changed automatically based on the following configuration:

```
configure
     service <type> <id>
          subscriber-interface <intf-name>
address <ip-address> gw-ip-address <gw-address> track-srrp        <srrp-inst> holdup-time
<msec>
ipv6
subscriber-prefixes
     prefix <ipv6-prefix> pd track-srrp <srrp-id> holdup-time <msec>
     prefix <ipv6-prefix> wan-host track-srrp <srrp-id> holdup-time <msec>


     policy-options
          begin
          policy-statement <name>
```

```
            entry 1
                from
protocol direct
                    state 'srrp-master'
                exit
                action accept
                    metric set 100
                exit
            exit
            entry 2
                from
protocol direct
                    state 'srrp-non-master'
                exit
                action accept
                    metric subtract 10
                exit
            exit
            entry 3
                from
                    protocol direct
                exit
                action accept
            exit
        exit
```

This configuration ensures that the route metric is changed for the subscriber interface routes based on the SRRP state while the other, non-subscriber directly attached routes are unaffected by SRRP.

*Route Advertisement based on SRRP State* requirement is applicable to BGP (IPv4, IPv4-IPVPN) and IGP.

*Routing policy* also provides the flexibility to prevent route advertisement (*action reject*) instead of changing the route metric.

Although this feature is designed to minimize or eliminate the use of the redundant-interface, it is important to note that the redundant-interfaces would still be used in the case of transient conditions. An example of such condition would be:

1. Messaging SAP Fails
2. SRRP switches over
3. Stale routing in the core is still in the effect while the metric is being propagated (or the route is being advertised/withdrawn). During this time, traffic is flowing over the redundant interface.
4. Network convergence is complete
5. Traffic in the network core is redirected to the new Master SRRP node

## Managed Routes

Only the state attribute is applicable to managed routes, and only to the ones that are synchronized (static and RADIUS obtained – framed-route and framed-ipv6-route). The managed routes obtained via BGP are not synchronized and this feature is not applicable to them.

Based on the SRRP state, the managed route can be either advertized with a modified metric or be withdrawn altogether.

For example:

*Managed routes* that are tracking SRRP state are only advertised from the Master node and denied from Backup node. All other managed routes that are not tracking SRRP state are advertised regardless of the SRRP state.

```
policy-options
         begin
         policy-statement <name>
              entry 1
                   from
protocol managed
                        state 'srrp-master'
                   exit
                   action accept
                   exit
              exit
              entry 2
                   from
protocol managed
                        state 'srrp-non-master'
                   exit
                   action reject
                   exit
              exit
              entry 3
                   from
                        protocol managed
                   exit
                   action accept
              exit
         exit
```

### Subscriber Management Routes (/32 IPv4 Host Routes, IPv6 PD WAN-Host Routes)

Both attributes (state and origin) are applicable to the subscriber management routes.

For Example:

A Service Provider wants to advertise only subscriber-management routes with the origin DHCP and AAA from the Master node. Routes with the LUDB origin are not advertised. Standby node is not advertising any /32 subscriber management routes.

```
policy-options
         begin
         policy-statement <name>
              entry 1
                   from
origin dhcp
origin aaa
                        state 'srrp-master'
                   exit
                   action accept
                   exit
              exit
         exit
```

Default action is reject.

---

### Activating SRRP State Tracking

The SRRP state tracking by routes is turned on only when desired.

For subscriber-interface routes (IPv4 and IPv6), this is performed explicitly.

```
subscriber-interface <intf-name>
address <ip-address> gw-ip-address <gw-address> track-srrp        <inst-name> holdup-time
<msec>
ipv6
subscriber-prefixes
     prefix <ipv6-prefix> pd track-srrp <srrp-id> holdup-time <msec>
      prefix <ipv6-prefix> wan-host track-srrp <srrp-id> holdup-time <msec>
```

For managed and subscriber management routes, this is explicitly enabled under the group interface:

```
group-interface <name>
     srrp-enabled-routing holdup-time <msec>
```

## SRRP in Conjunction with a PW in ESM Environment – Use Case

In certain cases, subscriber traffic is terminated on the BNG via an EPIPE. In this case, the subscriber traffic can be offloaded onto a plain Ethernet port via a VSM module (a 'loop') so that it can be terminated in ESM. EPIPEs can be configured in A/S configuration and terminated on two BNG nodes in multihomed environment.

In such multi-homed environment with EPIPEs and 'loops', the ESM itself would be detached from the EPIPE, which brings the subscriber traffic to the BNG. Because of that, the ESM would not know if the PW's state is Active or Standby. As a result, in the downstream direction, traffic could end up being forwarded towards the Standby PW, effectively being black-holed.

To overcome this, SRRP can be used in conjunction with an additional mechanism to help monitor the activity of the PWs. This monitoring mechanism is very similar to Fate-sharing. The difference in this case is that the messaging SAP (instead of SRRP instance) is monitoring the activity of the PW. As a result, the SRRP messaging SAP reflects the state of the PW. For example, the PW in a Standby mode would cause the messaging SAP to be in the DOWN state while the PW Active state would cause the messaging SAP to be in the UP state. That is, the SRRP instance reflects the operational state of the messaging SAP. SRRP is indirectly tied into PW state.

Modifying the priority of SRRP instance based on PW's state as a mean of mapping the Master SRRP into the Active PW would not help here as SRRP messages are not flowing over standby PWs. This is why SRRP state must be enforced via the messaging SAP.

Fate-sharing for PW termination in conjunction with SRRP is not supported.

Metric adjustment for the subscriber routes is supported. Once the tracked SRRP instance transitions into a non-Master state, the state attribute of the route changes and the appropriate action defined in the routing-policy is taken.

---

## Group-monitor

The failure detection mechanism to trigger an action within FSG relies on the operational state of the messaging SAP. Such failure detection mechanism is referred as a group monitor.

Group monitor can also be used to detect the state change of the PW. PW state change is reflected in the messaging SAP which in turn triggers the state change of an SRRP instance.

All this is implemented through an oper-group object which is described in the 'Services Guide'. All entities that needs to be monitored (messaging SAPs and PWs) are associated with this oper-group object. Finally, an SRRP instance (in case of FSG) or a messaging SAP (in case of PW) is instructed to monitor the entities in the oper-group object. State transitions of objects in a oper-group object trigger state transitions of entities that are monitoring them (messaging SAPs and SRRP instances). State transitions of monitored objects in a oper-group will cause the following actions:

- In the case of an FSG, priorities of SRRP instances are recalculated
- In the case of PW termination on BNG, the operational state of the messaging SAP is changed.

This is an overview of the CLI syntax showing the principles of how should this work (for exact description of commands and full syntax, please see the command reference):

```
configure>service
oper-group <name>   //oper-group creation

configure>service(IES | VPRN)>sub-if>grp-if>sap
     oper-group <name>  //adds the SAP to the oper-group
     monitor-oper-group <name> // links the status of the oper-group to the SAP. In this
fashion a messaging SAP can monitor the state of a PW.

configure>service(IES | VPRN)>sub-if>grp-if>srrp x
     monitor-oper-group <name> priority-step [0-253] //with this, a state transition of the
objects in the oper-group should trigger SRRP priority recalculation. The state of the
oper-group is not important but in the state of the objects within. If an object within the
oper-group goes down, the SRRP priority is lowered by a priority-step. The SRRP priority
will be adjusted on every state transition of member objects.

configure>service>epipe>spoke-sdp
     oper-group <name>  // this will add a PW to the oper-group. A messaging SAP monitors
this PW and it assumes the state according to the state of the PW in the oper-group. A
standby or a DOWN PW state causes the messaging SAP to assume a DOWN state. Otherwise the
messaging SAP would be in the UP state. In order for the SAP to assume the DOWN state, both
RX and TX side of the PW must be shut. In other words, a PW in standby mode also must have
the local TX disabled by the virtue of the 'slave' flag (standby-signaling-slave command
under the spoke-sdp hierarchy). Without the TX disabled, the SAP monitoring the PW would
not transition in the down state.
```

Hold timer is provided within the oper-group command to suppress flapping of the monitored object (SAP or pseudowire).

Example with ESM over pseudowire through a VSM 'loop'.

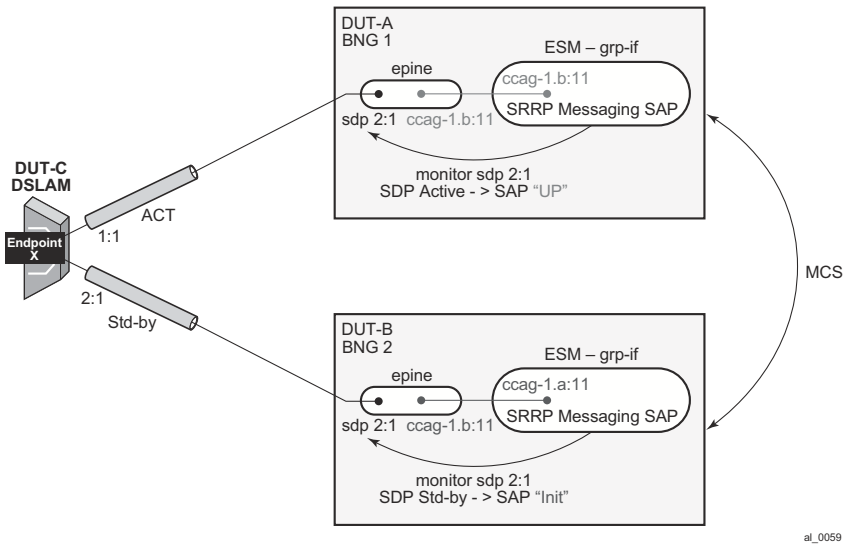



**Figure 124: Pseudowire Example**

```
*A:Dut-C>config>service>epipe# info
----------------------------------------------
            endpoint "x" create
                standby-signaling-master
            exit
            sap 1/1/7:1 create
            exit
            spoke-sdp 1:1 endpoint "x" create
                precedence primary
                no shutdown
            exit
            spoke-sdp 2:1 endpoint "x" create
                no shutdown
            exit
            no shutdown
----------------------------------------------

*A:Dut-A>config>service>epipe# info
----------------------------------------------
            sap ccag-1.b:11 create
            exit
            spoke-sdp 2:1 create
                standby-signaling-slave
                oper-group "1"
                no shutdown
            exit
            no shutdown
----------------------------------------------
*A:Dut-B>config>service>epipe# info
----------------------------------------------
            sap ccag-1.b:11 create
            exit
            spoke-sdp 2:1 create
                standby-signaling-slave
                oper-group "1"
                no shutdown
            exit
            no shutdown
----------------------------------------------
*A:Dut-A>config>service>ies# info
----------------------------------------------
            redundant-interface "redif11" create
                address 101.1.1.2/24 remote-ip 101.1.1.4
                spoke-sdp 1:1 create
                    no shutdown
                exit
            exit
            subscriber-interface "subif_1" create
                shutdown
                address 1.1.1.2/24 gw-ip-address 1.1.1.100
                group-interface "grpif_1_2" create
                    shutdown
                    redundant-interface "redif11"
                exit
            exit
            subscriber-interface "subTest" create
                address 80.1.1.2/24 gw-ip-address 80.1.1.254
                group-interface "grpTest" create
                    redundant-interface "redif11"
```

```
                    sap ccag-1.a:1 create
                    exit
                    sap ccag-1.a:11 create
                        monitor-oper-group "1"
                    exit
                    srrp 11 create
                        message-path ccag-1.a:11
                        no shutdown
                    exit
                exit
            exit
            no shutdown
-----------------------------------------------
*A:Dut-B>config>service>ies# info
-----------------------------------------------
            redundant-interface "redif11" create
                address 101.1.1.4/24 remote-ip 101.1.1.2
                spoke-sdp 1:1 create
                    no shutdown
                exit
            exit
            subscriber-interface "subif_1" create
                shutdown
                address 1.1.1.4/24 gw-ip-address 1.1.1.100
            exit
            subscriber-interface "subTest" create
                address 80.1.1.4/24 gw-ip-address 80.1.1.254
                group-interface "grpTest" create
                    redundant-interface "redif11"
                    sap ccag-1.a:1 create
                    exit
                    sap ccag-1.a:11 create
                        monitor-oper-group "1"
                    exit
                    srrp 11 create
                        message-path ccag-1.a:11
                        no shutdown
                    exit
                exit
            exit
            no shutdown
-----------------------------------------------
*A:Dut-B>config>service>ies# show srrp
===============================================================================
SRRP Table
===============================================================================
ID        Service        Group Interface                Admin     Oper
-------------------------------------------------------------------------------
11        1              grpTest                        Up        initialize
-------------------------------------------------------------------------------
No. of SRRP Entries: 1
===============================================================================
*A:Dut-A>config>service>ies# show srrp
*A:Dut-A>config>service>ies#
===============================================================================
SRRP Table
===============================================================================
ID        Service        Group Interface                Admin     Oper
-------------------------------------------------------------------------------
```

```
11        1              grpTest                         Up        master
-------------------------------------------------------------------------------
No. of SRRP Entries: 1
===============================================================================
```

## Subscriber Override

This feature provides the ability to override queue and policer parameters (CIR, PIR, CBS, MBS) as well as HQoS parameters (egress aggregate-rate and root-arbiter rate) configured at sla-profile and sub-profile level in order to provide per-subscriber-(host) customizations. The goal is to avoid an explosion of the sub-profiles and sla-profiles to cover all service level combinations. This customization of QoS related parameters can in principle occur during authentication (auth-response message) or during sub-host life time by RADIUS CoA messages.

The QoS parameter customizations are communicated by RADIUS server in form of RADIUS VSAs which can be included in RADIUS-authentication response message or in CoA message.

The Alc-Subscriber-QoS-Override VSA (126) is a string with following layout "direction:type:[key:]values" where:

- Object-type:
  - → **direction** represents single character indicating **i** for ingress and **e** for egress.
  - → **type** represents single character indicating **q** for queue, **p** for policer, **r** for aggregate-rate and **a** for arbiter.
  - → **key** is indicated the queue or policer-id. It is not used in case of aggregate-rate and root-arbiter.
  - → **values** indicates actual values preceded with keywords used in CLI (e.g., cir).

Aspects, such as parent, priority level, stats- mode are not accessible through this customization. Instead, a new policy should be created on the node.

The key identifying the subscriber-host in the RADIUS CoA message is accounting-session-id This is different in previous releases, where the *service-id* and *ip-address* are mandatory fields in RADIUS CoA message.

The operational value of the QoS objects (queues or schedulers) are derived from different inputs. As in queue/policer parameters, the following hierarchy of inputs are respected (highest priority is the first):

- On-line charging overrides
- RADIUS response/CoA overrides
- Queue overrides configured at sla-profile level
- Queue parameters set in QoS policy level

In the case of scheduler/arbiter-overrides, the following hierarchy of inputs apply:

- ANCP overrides
- RADIUS response overrides
- Scheduler/arbiter parameters as configured in scheduling/policer-control-policy.

The above rules are generic. If any given override mechanism is not yet applicable to policers (or arbiters) it will be skipped. The above rules indicate the priority if all mechanisms are supported.

The QoS overrides received in RADIUS message (in the form of a VSA) are per definition related to a given subscriber-host the given message is referring to. Internally, the overrides are applied on per SLA instance level (queue/policer-overrides) and per-subscriber-level (scheduler/arbiter overrides). The subscriber-overrides are applicable to PPPoE hosts only.

In a dual-homing environment, the adapted values are synchronized through MCS, but they do not need to be persistent.

# Dual Stack Lite

Dual Stack Lite feature is supported on the 7710 SR-Series in combination with the MS-ISA to function as a DS-Lite Address Family Transition Router (AFTR).

Dual Stack Lite is an IPv6 transition technique that allows tunnelling of IPv4 traffic across an IPv6-only network. Dual-stack IPv6 transition strategies allow service providers to offer IPv4 and IPv6 services and save on OPEX by allowing the use of a single IPv6 access network instead of running concurrent IPv6 and IPv4 access networks. Dual-Stack Lite has two components: the client in the customer network, known as the Basic Bridging BroadBand element (B4) and an Address Family Transition Router (AFTR) deployed in the service provider network.

Dual-Stack Lite leverages a network address and port translation (NAPT) function in the service-provider AFTR element to translate traffic tunneled from the private addresses in the home network into public addresses maintained by the service provider. On the 7750 SR, this is facilitated through the Carrier Grade NAT function.

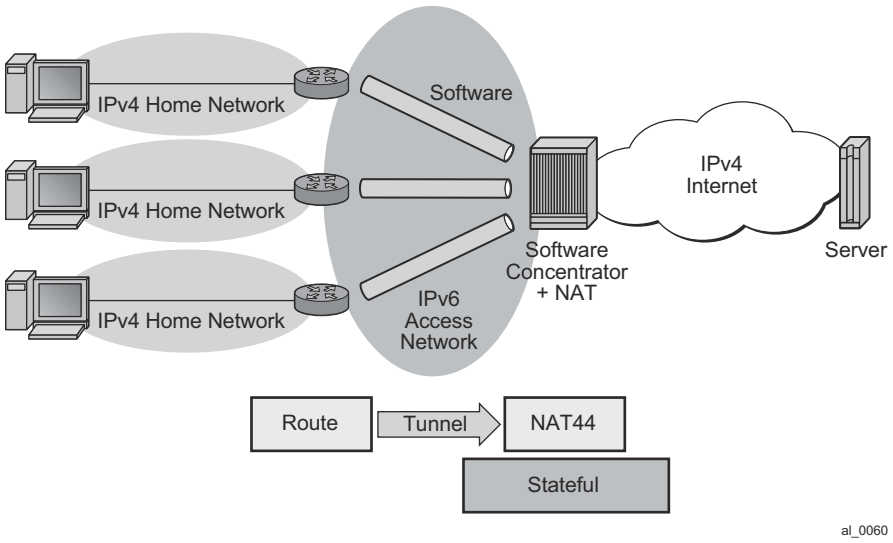



**Figure 125: Dual-Stack Lite**

As shown in Figure 125, Dual-Stack Lite has two components, a softwire initiator in the RG and a softwire concentrator, deployed in the service provider network, where control-less IP-in-IP (using protocol 4 - IPv4 in IPv6) is used for tunnelling. When using control-less protocol, packets are sent on the wire for the remote softwire endpoint without prior setup of a tunnel.

The softwire initiator in the home network is combined with a routing function, where the default route is passed to the softwire pseudo-interface. Note that there is no NAT function, therefor, the private IP addresses of the home network are encapsulated without source address modification,

and forwarded to the softwire concentrator where all NAT is performed. The softwire pseudo-interface unicasts all IPv4 traffic to the IPv6 address of the softwire concentrator, which was pre-configured.

When encapsulated traffic reaches the softwire concentrator, the device treats the source-IP of the tunnel to represent a unique subscriber. The softwire concentrator performs IPv4 network address and port translation on the embedded packet by re-using Large Scale NAT and L2-Aware NAT concepts.

## IP-in-IP

As shown in Figure 126, IP-in-IP uses IP protocol 4 (IPv4) to encapsulate IPv4 traffic from the home network across an IPv6 access network. The IPv4 traffic tunnelling is treated as best-effort with no subscriber management or policy, and does not use ESM. The scale is dependant only on the internal structures of the MS-ISA and CPM, that is, the IP-in-IP model can support more subscribers than an ESM-based approach.

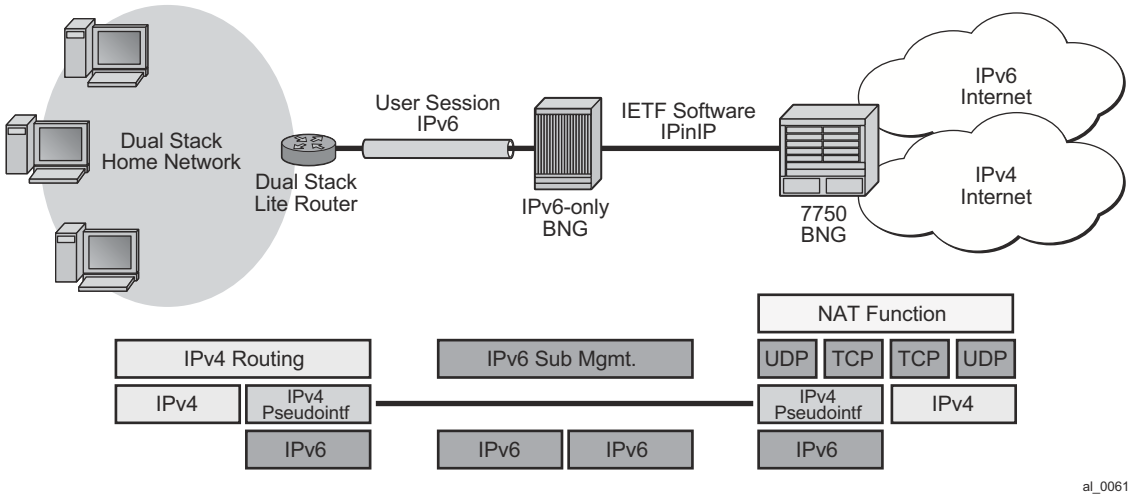



**Figure 126: IP-in-IP**

Dual-Stack Lite IP-in-IP is configured through the existing **nat** command that is inside the CLI statements that are within the base router or VPRN. A service performing large scale NAT supports Dual-Stack Lite.

Dual-Stack Lite expects a routing (non-NATing) gateway in the home, where many different IPv4 inside addresses exist for each subscriber. These inside addresses may overlap other subscriber's address, especially given the heavy use of RFC 1918 address space.

The lack of control of protocol for the IP-in-IP tunnels simplifies the functional model, since any received IPv4 packet to the ISA dual-stack-lite address can simply be:

- Checked for protocol 4 in the IPv6 header.
- Checked that the embedded IP packet is IPv4.
- Processed as if it were L2-Aware, where the source-IP of the tunnel (the source IPv6 address) is used as the subscriber identifier.

Note that the inside IP address in the NAT, tables must not be the IPv6 address of the tunnel, but the true IPv4 address of any host within the home. The subscriber-id must be the literal IPv6 address (appreciating this may be 34 characters in length).

## Configuring Dual Stack Lite

Dual Stack Lite is configured on an inside service and uses the existing Large Scale NAT nat-policies and outside pools. Dual-Stack Lite and NAT44 Large Scale NAT can operate concurrently on the same inside and outside services.

Dual Stack Lite is configured using the following CLI:

```
configure {router | service vprn service-id}
     - [no] nat
          - inside
               - [no] dual-stack-lite
                    - [no] *address ipv6-address
```

## L2TP over IPv6

In this mode, L2TP provides the transport for IPv4 that allows full ESM capabilities on the 7750 SR. From the 7750 perspective, the L2TP tunnel is no different in capability to those already supported. Only the underlying transport (IPv6 instead of IPv4) distinguishes this approach.

To support legacy IPv4 access, L2TP over IPv6 is combined with the existing L2-Aware NAT feature as shown in Figure 127.

As ESM is used, scale is limited by the number of ESM hosts supported on a chassis and any associated resources like queues.

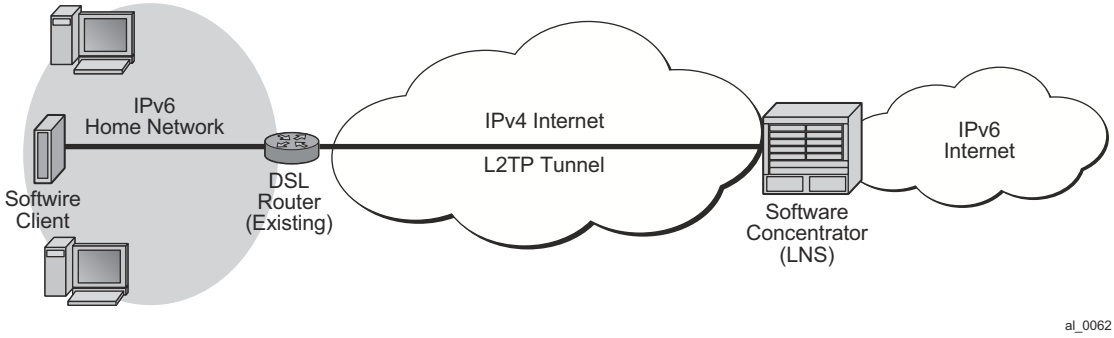



**Figure 127: L2TP over IPv6**

L2TP LNS over IPv6 is supported in both the base routing instance and VPRN that has 6VPE configured.

Like the SR-OS 8.0 LNS implementation, tunnels are terminated on any routing interface, including loopback, SAP, or network port. A single interface simultaneously supports IPv4 and IPv6 L2TP tunnel termination by having two different addresses configured.

For greater scalability, L2TP tunnel and session count per chassis are increased to allow 1 tunnel per session.

NAT capabilities are supported via existing L2-Aware NAT methods. Note that the L2TP LNS over IPv6 may be used without NAT as well and the L2TP sessions may be either IPv6-only or dual-stack.

# L2TP Tunnel RADIUS Accounting

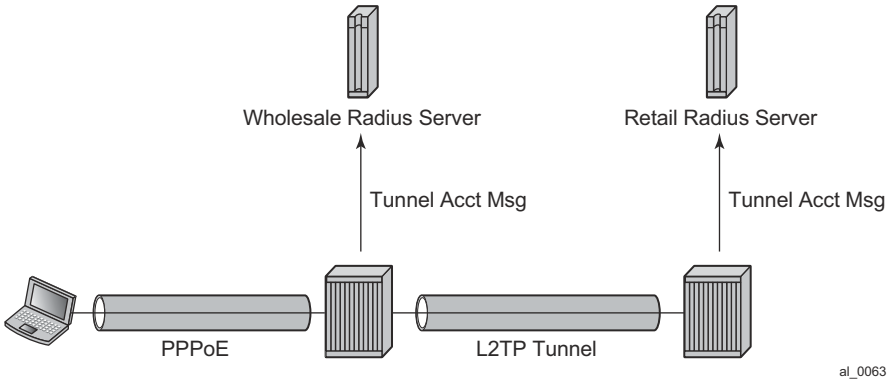



**Figure 128: L2TP Tunnel Accounting**

When L2TP tunnel accounting is enabled, except for **host** or **sla-profile**-based accounting packets and attributes, the following are additional accounting packets and attributes:

- Accounting packets: tunnel-start/stop/reject; tunnel-link-start/stop/reject — There are no interim updates for L2TP tunnel/session accounting.
- RADIUS accounting attributes:
  - → Tunnel-Assignment-Id (LAC only)
  - → Acct-Tunnel-Connection
  - → Acct-Tunnel-Packets-Lost

These attributes were added into current account-start/stop/interim-update packets (host accounting/sla-profile accounting)

Tunnel level accounting and session level accounting can be enabled or disabled independently.

New accounting packets and related RADIUS attribute list are described in Table 17.

Some considerations of RADIUS attributes are described in RADIUS Attributes Value Considerations on page 1099

# Accounting Packets List

Table 17 describes L2TP tunnel accounting behavior along with some key RADIUS attributes (apply for both LAC and LNS):

**Table 17: L2TP Tunnel Accounting Behavior**

| Act-Packet | When | Key Attributes | Remark |
|---|---|---|---|
| Tunnel-Start | A new L2TP tunnel is created | Acct-Session-ID | |
| | | Event-Timestamp | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |
| | | Tunnel-Server-Endpoint:0 | |
| | | Tunnel-Server-Auth-Id:0 | |
| Tunnel-Reject | A new L2TP tunnel creation failed | Acct-Session-Id | |
| | | Event-Timestamp | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |
| | | Tunnel-Server-Endpoint:0 | |
| | | Acct-Terminate-Cause | |
| Tunnel-Stop | An established L2TP tunnel is removed | Acct-Session-Id | |
| | | Event-Timestamp | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |

**Table 17: L2TP Tunnel Accounting Behavior (Continued)**

| Act-Packet | When | Key Attributes | Remark |
|---|---|---|---|
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |
| | | Tunnel-Server-Endpoint:0 | |
| | | Tunnel-Server-Auth-Id:0 | |
| | | Acct-Session-Time | |
| | | Acct-Input-Gigawords | |
| | | Acct-Input-Octets | |
| | | Acct-Output-Gigawords | |
| | | Acct-Output-Octets | |
| | | Acct-Input-Packets | |
| | | Acct-Output-Packets | |
| | | Acct-Terminate-Cause | |
| Tunnel-Link-Start | An L2TP session is created | User-Name | |
| | | Acct-Session-Id | This is the same as Acct-Session-id in access-request of host auth |
| | | Event-Timestamp | |
| | | Service-Type | Framed |
| | | Class | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |
| | | Tunnel-Server-Endpoint:0 | |
| | | Tunnel-Server-Auth-Id:0 | |

**Table 17: L2TP Tunnel Accounting Behavior (Continued)**

| Act-Packet | When | Key Attributes | Remark |
|---|---|---|---|
| | | Acct-Tunnel-Connection | See RADIUS Attributes Value Considerations on page 1099 |
| Tunnel-Link-Reject | A new L2TP session creation is failed | Acct-Session-Id | Should be as same as Acct-Session-id in access-request of host auth |
| | | Event-Timestamp | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |
| | | Tunnel-Server-Endpoint:0 | |
| | | Acct-Terminate-Cause | |
| | | Acct-Tunnel-Connection | |
| Tunnel-Link-Stop | A established L2TP session is removed | User-Name | |
| | | Acct-Session-Id | Should be as same as Acct-Session-id in access-request of host auth |
| | | Event-Timestamp | |
| | | Service-Type | Framed |
| | | Class | |
| | | Tunnel-Type:0 | |
| | | Tunnel-Medium-Type:0 | |
| | | Tunnel-Assignment-Id:0 | |
| | | Tunnel-Client-Endpoint:0 | |
| | | Tunnel-Client-Auth-Id:0 | |

**Table 17: L2TP Tunnel Accounting Behavior (Continued)**

| Act-Packet | When | Key Attributes | Remark |
|---|---|---|---|
| | | Tunnel-Server-Endpoint:0 | |
| | | Tunnel-Server-Auth-Id:0 | |
| | | Acct-Tunnel-Connection | |
| | | Acct-Session-Time | |
| | | Acct-Input-Gigawords | |
| | | Acct-Input-Octets | |
| | | Acct-Output-Gigawords | |
| | | Acct-Output-Octets | |
| | | Acct-Input-Packets | |
| | | Acct-Output-Packets | |
| | | Acct-Tunnel-Packets-Lost | |
| | | Acct-Terminate-Cause | |

Notes:

- Errors will occur if there are multiple hosts sharing the same sla-profile instance and then these hosts go to different tunnel.
- 7750 SRs have an internal limitation of 500 pps for accounting messages. This feature shares the same limitation

## RADIUS Attributes Value Considerations

- The value of Acct-Tunnel-Connection uniquely identify a L2TP session, and in order to match LAC and LNS accounting record, the value of Acct-Tunnel-Connection is determined by a method shared by LAC and LNS. This means for a given L2TP session, Acct-Tunnel-Connection from the LAC and LNS are the same.
- Current ESM stats are used in Tunnel-Link and tunnel level accounting. This applies for both standard attribute and the 7750's own VSA.
- Tunnel level accounting stats need to aggregate all sessions stats that belong to the tunnel. Note: there could be sessions come and go before tunnel is down, so system need to remember the stats of every session that has been created within the tunnel.

  This applies for both standard attribute and 7750's own VSA.
- The value of Acct-Tunnel-Packets-Lost is the aggregation of all discarded packets on both ingress and egress.

## Other Optional RADIUS Attributes

Table 18 lists the optional attributes that could be optionally included in tunnel accounting packet, some of them are applied for link level accounting only.

**Table 18: Optional RADIUS Attributes**

| Attribute | Tunnel/Link |
|---|---|
| nas-identifier | Both |
| nas-port | Link level only |
| nas-port-id | Link level only |
| nas-port-type | Link level only |

## RADIUS VSA to Enable L2TP Tunnel Accounting

In order to support pure RADIUS-enabled L2TP tunnel accounting on LAC side, the following RADIUS VSA are supported:

**Table 19: Supported RADIUS VSAs**

| VSA | Type | Value |
|---|---|---|
| ALC-Tunnel-Accounting-Policy | String | Policy-name; if the name is **disable** then this means L2TP tunnel accounting is disabled for this tunnel |

Note: ALC-Tunnel-Accounting-Policy takes precedence over what has been defined in CLI when Alc-Tunnel-Group is also returned.

## MLPPP on the LNS Side

With MLPPP, the counter on LNS side is only available for the bundle, not for each link, so the SR OS's behavior is:

- For each new link session system sends a tunnel-link-start.
- For each link session that is deleted system sends a tunnel-link-stop.
- For all link sessions except the last one system reports 0 for all counters.
- For the last link session, system reports the actual counters for the bundle.

# RADIUS Route Download

The RADIUS route download mechanism periodically polls a RADIUS server for routes to download. The main objective of this feature is to download, in advance, customer-assigned subnets so that they can be re-advertised to the corresponding routing protocols. In this way, subscriber bringup can potentially be done faster (as the routes are already in place and advertised) and, most importantly, reduce the routing protocol churn as subscribers connect and disconnect. The routes being learned through this mechanism could be both managed routes/delegated prefixes as well as the WAN IP assigned to the subscriber in the case PPPoE and un-numbered interfaces are being used.a

The route download process requests the routes to a configured RADIUS server by triggering an access-request message. The key identifier for this message is the username, which is a combination of the system's name (or an optionally configured value), appended by a dash ( "-") and then a monotonically increasing integer. The download process sends an access request starting with 1 (such as "hostname-1") and the RADIUS server replies with an access-accept message and a number of routes embedded within the message. The system then increases the counter and sends another access request (this time being hostname-2) and receive a reply with the next batch of routes to download. The process continues, incrementing the counter by 1 each time until the system gets an access-reject or the maximum number of routes that can be downloaded is reached.

The routes to be accepted are in the following format:

> [vrf {vprn-name | vprn-service-id}] prefix-mask {null0 | null 0 | black-hole} [metric] [tag tag-value]

The prefix-mask could be in any form as 'prefix/length', 'prefix mask' or 'prefix' (in the latter case, for IPv4 routes, the mask shall be derived from the IP class of the prefix).

The route formats are supported:

- Framed-Route (RADIUS attribute 22)

```
Framed-Route = "192.168.3.0 255.255.255.0 null0"
Framed-Route = "vrf vrfboston 192.168.1.0/24 null 0 0 tag 6"
Framed-Route = "vrf 2001 192.168.10.0/24 black-hole 0 tag 8"
```

- Cisco-AVPair (Cisco VSA 26-1)

```
cisco-avpair = "ip:route = 192.168.3.0 255.255.255.0 null0"
cisco-avpair = "ip:route = vrf vrfboston 192.168.1.0/24 null 0 0 tag
6"
```

IPv6 routes are also supported. The format is based on using the IETF-defined IPv6 Framed-IPv6-Route (attribute 99). The following text shows the supported formats.

- Framed-IPv6-Route (RADIUS attribute 99)

```
Framed-Route = "2001:100:bad:cafe::/64 null0"
Framed-Route = "vrf vrfboston 2100:5aaa:dead:beaf::/96 null 0 0 tag
6"
Framed-Route = "vrf 3000 2200:1bbbb:dead::/48 black-hole 0 tag 6"
```

All the routes downloaded will be a new protocol type "**periodic**". The downloader process re-starts the AAA requests after a given interval (a configurable value but target refresh rate is 15 minutes) and routes shall be updated according to the following process:

- When the router initiates a new download process, the routes are kept in a temporary table until the download process completes (receives an access-reject from the AAA). The temporary download table is then checked for errors and finally, any changes reflected to the actual routing table.
- Routes no longer present in the download will be removed from the routing table.
- If the AAA server responds with an access-reject for the first username (that is, an implicit empty route-download table), all routes will be removed from the routing table.
- If there are any protocol errors (at the RADIUS level), such as time-out, no response, bad record format, too many records, etc., the download process is suspended and retried after a configurable timer. The minimum retry timer is at least 1 minute and given the light load this represents control-plane-wise (concurrent downloads are not supported) the retries can continue infinitely until the next refresh period occurs, where the download restarts from the beginning. An exponential backoff algorithm with a configured minimum and maximum delay will be used to determine the retry timer.
- In any case, the routes are only purged from the routing table after a complete download process was achieved (properly terminated with an access-reject message). Under any other failure condition, the routes shall remain active. Shutting down the download process should not remove the downloaded routes. A clear command will be provided to clear the periodic routes.
- All the imported routes (blackholes) will be imported into the line-card FIBs to avoid the routing loops caused by announcing the prefixes but not installing the actual blackholes.

# Managed SAP (M-SAP)

Managed SAPs allow the use of policies and a SAP template for the creation of a SAP. Although the router supports automatic creation of subscriber hosts in a shared SAP, the most secure mode of operation and common mode is the subscriber per SAP model. In this model, each subscriber is defined with its own VLAN. This feature uses authentication mechanisms supported by the node to provide a SAP.

When enabled, receiving a triggering packet initiates RADIUS authentication that provides a service context. The authentication, together with the service context for this request, creates a managed SAP. The VLAN is the same as the triggering packet. This SAP behaves as a regular SAP but its configuration is not user editable and not maintained in the configuration file. The managed SAP remains active as long as the session is active.

The following trigger types are supported:

- DHCP discover (or requests if configured) for DHCP clients. The managed SAP lifetime is defined by the lease time.
- PPPoE PADI for the PPPoE client. The managed SAP lifetime is defined by the session time. The MSAP is installed after the IP address is provided. A short temporary state handles packets between the PADO and ACK.
- ARP defines the managed SAP lifetime. The ARP entry refresh behavior is maintained.
- DHCP6
- PPP

A “capture SAP” triggers the process. This SAP is defined in a similar way to a default SAP but does not forward traffic. A capture SAP and default SAP cannot be configured at the same time for a single port with the dot1q encapsulation or for a single port:topq combination with qinq encap. The capture SAP is used if a more specific match for the Q or Q-in-Q tags is not found by the IOM. If a capturing SAP is defined, triggering packets are sent to the CPM. Non-triggering packets captured by the capturing SAP are dropped.

An ingress VLAN ID (VID) type mac filter can be configured on a capture-sap to have additional control on the vlans that are allowed to initiate a host setup. Other filter types are not supported on a capture-sap.

Supported modes:

- Port:*: Provides a context for the trigger packet and SAP template.
- A capture SAP can be created in the format Port:Q.
- Port:Q: A specific Q-tag defined SAP for the port and already running managed SAPs.

Supported Q-in-Q modes:

- Port:*.*, or Q.*: Both q-tags will always be sent to RADIUS. The M-SAP created will bear both q-tags that arrived in the original packet if authenticated by RADIUS.

- Port:*.Q: It is an inverse capture-sap that matches on a fixed inner tag with the outer tag identifying the user. The following restrictions apply when an inverse capture-sap is configured on a port:
  - Ethernet ports only
  - It is not possible to create y.* saps when there is a *.x capture sap present on the port (y=0,1..4094,* and x=1..4094).
  - It is not possible to create a y.* network interface when there is a *.x capture SAP present on the port (y=0,1..4094,* and x=1..4094).

A set of mandatory parameters should be provisioned for M-SAP creation are as follows:

- **Service id**: service context in which the M-SAP will be created.
- **Interface id**: name of the group-interface context in which the MSAP will be created. The group-interface must exist in the provided service for the M-SAP to be installed (routed CO scenario only).
- **MSAP policy**: name of the policy that defines the M-SAP parameters. The policy must exist in the subscriber-mgmt context.

These parameters can be obtained from the following order of preference:

1. Local user database lookup.
2. RADIUS attributes.
3. Defaults configured at the capture-sap context.

For IPoEv4 hosts and PPP hosts, the MSAP parameters can be obtained from a local user database in the pre authentication phase. For this, a local user database should be configured at the capture sap and group-interface context. For example,

```
# IPoEv4 hosts
>config>service>vpls>sap# dhcp-user-db <local-user-db-name>
>config>service>ies>sub-if>grp-if>dhcp# user-db <local-user-db-name>

# PPP hosts
>config>service>vpls>sap# pppoe-user-db <local-user-db-name>
>config>service>ies>sub-if>grp-if>pppoe# user-db <local-user-db-name>
```

At the group-interface context, no authentication policy may be configured. Instead, the authentication policy is specified in the local user database. For example,

```
# IPoE hosts
>config>subscr-mgmt>loc-user-db>dhcp>host# auth-policy <policy-name>

# PPP hosts
>config>subscr-mgmt>loc-user-db>ppp>host# auth-policy <policy-name>
```

The MSAP parameters are configured at the local user database host context. For example,

```
>config>subscr-mgmt>loc-user-db>dhcp>host# msap-defaults
>config>subscr-mgmt>loc-user-db>ppp>host# msap-defaults

  - msap-defaults
```

```
[no] group-interface - Configure the group interface
[no] policy          - Configure the MSAP policy
[no] service         - Configure the service
```

The following table lists the RADIUS attributes (VSA's) to include in a RADIUS access accept message to obtain MSAP parameters in the RADIUS authentication phase.

| Attribute name | Type | Purpose and format |
|---|---|---|
| Alc-MSAP-Serv-Id [26-6527-31] | Integer | Service ID of the service context in which the M-SAP will be created. |
| Alc-MSAP-Policy [26-6527-32] | String | Name of the policy that defines the M-SAP parameters. |
| Alc-MSAP-Interface [26-6527-33] | String | Name of the group-interface context in which the MSAP will be created. |

MSAP parameters that are not obtained from a local user database lookup, and that are not returned from RADIUS can be specified in the default-msap section of the capture-sap context (last resort):

```
>config>service>vpls>sap# msap-defaults ?
  - msap-defaults

 [no] group-interface - Configure the group interface
 [no] policy          - Configure the MSAP policy
 [no] service         - Configure the service
```

While M-SAPs are supported in both routed-co and VPLS TPSDA models, the triggering SAP can be configured only in VPLS.

The managed SAP configuration can be persistent. The template MSAP policy is stored with the subscriber host which in turn can be made persistent.

If RADIUS does not provide all the information required to install the host (lacking an IP address), the MSAP is created with a short timer while waiting for the host acquire the necessary information and install the host. Default SAP policies are used unless the profiles are known.

In most cases, M-SAPs are allowed to have multiple subscribers to share an MSAP. In architectures that provide service access using a shared SAP, multiple subscribers can share the SAP. These environments require few SAPs and therefore are not supported. Multiple leases for the same subscriber are allowed. Only a single M-SAP policy is allowed. If an M-SAP was defined by a host and a new host installation is attempting to change the policy, the installation fails and an event is raised.

All trigger types can be combined on a SAP supporting DHCP and PPPoE hosts.

The authentication policy is defined in the M-SAP policy. Based on the configuration, the system will re-authenticate. If not used, or for PPPoE, the MSAPs remains active if the session renews.

When PPPoE is used with M-SAPs, the authentication-policy cannot use the username for the M-SAP creation.

The authentication policy used in the capture SAP is the same as the policy used for the managed SAP. In a Layer 3 scenario, the authentication policy is defined in the group-interface context. The managed SAP will not be created if the group-interface name returned from RADIUS points to a different authentication policy other than the policy defined by the capture SAP.

## ESM Identification Process

### SAP-ID ESM Identifier

Providers migrating from Basic Subscriber Management (BSM) can assign a subscriber to a SAP. The SAP ID ESM identifier makes the transition easier by allowing the operator to continue using the *sap-id* as a subscriber-ID.

An ESM SAP ID provides the system the ability to:

- Provide access to the SAP ID string in the Python script.
- Allow the automatic assignment of the SAP-ID to a static subscriber or subscriber host.

### DSLAM-ID

A DSLAM ID provides the system the ability to define a DSLAM-ID string provided through the Python script, RADIUS, or local user database. If the DSLAM-ID was provided, but the subscriber host is instantiated on a regular MDA (a non-HSMDA), the DSLAM-ID will be ignored.

The HSMDA and the ability to aggregate subscribers into DSLAMs for the purpose of QoS, can use the SAP ID to identify subscribers and associated DSLAMs.

## Default-Subscriber

This feature provides a default subscriber definition under the SAP. If the object was configured the operator may use ESM without enabling a processing script or a RADIUS authentication policy. In the event both have been disabled any host that was installed for the SAP will be installed with the configured default subscriber ID. If a RADIUS policy was used or if a script was enabled but a subscriber ID was not returned the default subscriber ID will be used.

## Multicast Management

The multicast-management CLI node contains the bandwidth-policy and multicast-info-policy definitions. The bandwidth-policy is used to manage the ingress multicast paths into the switch fabric. The multicast-info-policy is used to define how each multicast channel is handled by the system. The policy may be used by the ingress multicast bandwidth manager, the ECMP path manager and the egress multicast CAC manager.

## Subscriber Mirroring

This section describes mirroring based on a subscriber match. Enhanced subscriber management provides the mechanism to associate subscriber hosts with queuing and filtering resources in a shared SAP environment. Mirroring used in subscriber aggregation networks for lawful intercept and debugging is required. With this feature, the mirroring capability allows the match criteria to include a subscriber-id.

Subscriber mirroring provides the ability to create a mirror source with subscriber information as match criteria. Specific subscriber packets can be mirrored mirror when using ESM with a shared SAP without prior knowledge of their IP or MAC addresses and without concern that they may change. The subscriber mirroring decision is more specific than a SAP. If a SAP (or port) is placed in a mirror and a subscriber host of which a mirror was configured is mirrored on that SAP packets matching the subscriber host will be mirrored to the subscriber mirror destination.

The mirroring configuration can be limited to specific forwarding classes used by the subscriber. When a forwarding class (FC) map is placed on the mirror only packets that match the specified FCs are mirrored. A subscriber can be referenced in maximum 2 different mirror-destinations: 1 for ingress and 1 for egress.

# Volume and Time Based Accounting

Time and volume-based accounting includes the following components:

- Metering function performing stateful monitoring of the service delivery to the subscriber.
- Communication with an external management system that gets and updates credit per subscriber, notifications of credit exhaustion, etc.
- Action on credit exhaustion takes pre-defined action when the credit has been exhausted

## Metering

Metering represents the core of time and volume-based accounting. Service usage is typically measured by performing an accounting of the traffic passing through corresponding subscriber-host queues (volume usage) or by keeping lease-state while the given subscriber-host is connected to the network (time usage).

- Statefullness — The accounting information is compared with pre-defined credit expressed in terms of time or volume to monitor service usage.
- Sensitivity — Defining so called activity-threshold allows distinction between subscriber-host being connected and subscriber-host effectively using the service. This is particularly of interest in cases of time based charging.
- Aggregated usage per-category per-subscriber-host — Accounting information can be reported on per-queue per-sla instance of the given subscriber. In many situations, a certain level of aggregation (such as a per-subscriber or HSI ingress and egress traffic) is required to perform meaningful mechanism for pre-paid services.

## Categories Map and Categories

This feature introduces a new object category-map which defines individual aggregates (such as data in and out, video and data, etc.) and their mapping to individual forwarding queues.

The following output depicts a category-map configured in the subscriber management context.

```
*A:ALA-48>config>subscr-mgmt# info
----------------------------------------------
...
        category-map "triple-play" create
            category "data" create
                queue 1 ingress-egress
            exit
            category "video" create
                queue 2 egress-only
            exit
            category "voice" create
                queue 3 ingress-egress
            exit
        exit
        category-map "aggr-subscriber-service" create
            category "data-services" create
                queue 1 ingress-egress
                queue 3 egress-only
            exit
        exit
...
----------------------------------------------
*A:ALA-48>config>subscr-mgmt#
```

Based on a category-map the system gathers usage information (volume/time) on a per-sla-instance-per-category basis. In order to do so, statistics of all queues forming the category of the given sla-instance are aggregated.

- Single subscriber host (routed CPE) — Single SLA instance.
- Multiple subscriber hosts on the same SAP (bridged CPE) — Single SLA instance. Note that several hosts use the same credit and the renewal of one will cause renewal for all.
- Multiple subscriber hosts on different SAP (bridged CPE) — SLA instance per host.

The per-category usage gathered as described above is compared with per-subscriber-host-per-category credit and when credit is exhausted several actions can be taken.

There are several category-maps pre-configured on the system. The category-map applicable to a given subscriber-host will be derived at the host creation from the RADIUS VSA in an authentication-response, Python script, or static configuration in the local-user-database. All subscriber-hosts belonging to the same subscriber and created on the same SAP (hence, sharing the same sla-instance) must use the same category-map. In case of conflict, (an existing subscriber host has a different category-map than the one derived for the new host) the category-map of the last host will be applied to a given sla-instance. As a consequence, all previous information related to the status of the credit will be lost.

There can be multiple queues aggregated into one category. There can be up to three categories in a category map.

## Quota Consumption

There are two types of quota (credit), volume and time. In volume usage monitoring, the system accumulates byte counters per category-sla-instance and compares it with the assigned quota. Once the credit is exhausted (or threshold for renewal is met) the system attempts to renew it with corresponding management system.

In time-based credit, the distinction between active-usage and active-connection is made by defining an activity-threshold, where an object defines an average data rate under which the subscriber-host is considered silent.

As long as the effective rate of the application usage does not exceed the rate defined by the activity-threshold, the given subscriber host will be considered silent and its corresponding credit will not be used. As long as the application usage exceeds the rate, the application-credit will be consumed (in terms of time).

## RADIUS VSA Credit-Control-Quota

The quota in the RADIUS VSA Credit-Control-Quota uses this fixed format:

Alc-Credit-Control-Quota = "<volume quota>|<time quota>|<category name>"

- Where Volume: in bytes (B), kilobytes (K or KB), megabytes (M or MB), gigabytes (G or GB)
- Where Time: in seconds (s), in minutes (m), in hours (h), in days (d) or a combination (5m30s) but there is a restriction; a lower unity may never exceed the higher unity (5m60s is not allowed)

For example, Alc-Credit-Control-Quota = "1G|1h30m|cat1"

Volume quota, as well as time quota, needs to be specified.

- The minimum volume quota is 100 megabytes.
- The minimum time quota is 15 minutes.

# Credit Negotiation Mechanisms

The per-subscriber per-category credit can be obtained by several ways:

- RADIUS during authentication process.
- Static configuration - configured in the **config>subscr-mgmt>category-map>category** context.

Credit can be expressed by either

- Volume
- Time

The renewal of the credit using RADIUS authentication is triggered by credit exhaustion or (if configured) by depletion of the credit to exhausted-credit-threshold level. If this occurs, the system will send a RADIUS authentication message indicating the corresponding category and usage. The following are several possibilities for the RADIUS server response (as shown in figure below):

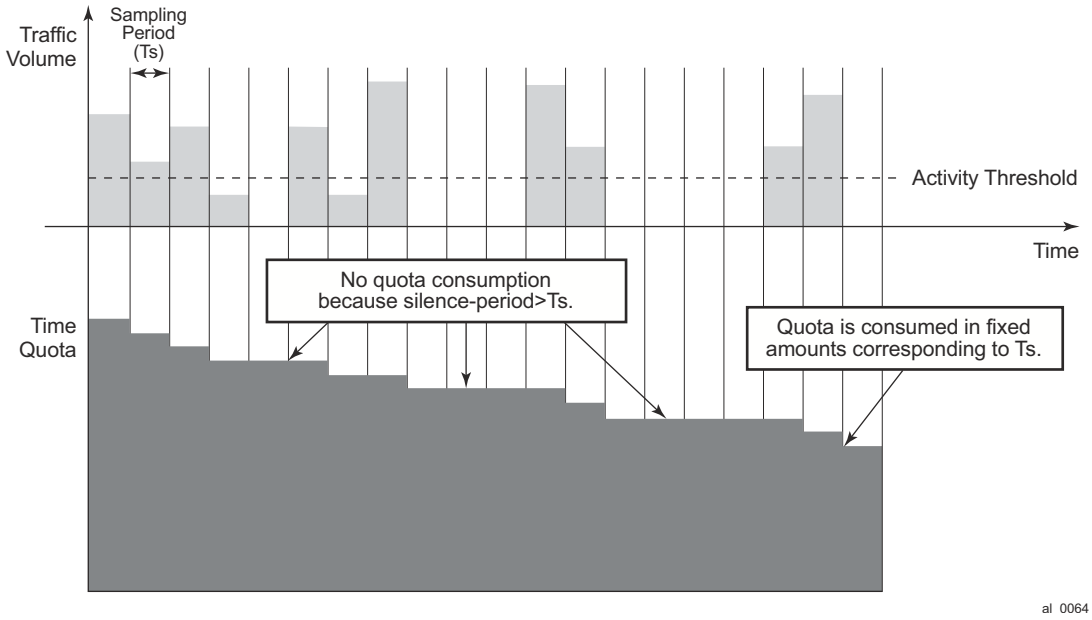



**Figure 129: Threshold Configured/Not Configured**

1. No authentication response — The system will install out-of-credit action after the original credit has been used.
2. Authentication response with reject — The corresponding host is removed after the original credit has been used.

3. Authentication response with accept and no credit VSA included — The system will install out-of-credit action.
4. Authentication response with accept and credit VSA included — The out-of-credit will be installed.
   Note that the new credit is always reduced by the amount of credit consumed in time between renewal has been initiated and authentication-respond has been received. In case of a negative result (the newly receive credit is smaller than the amount consumed in the mean-time) the test cr
5. is installed.

In order to identify that the given RADIUS-auth request is related to credit renewal rather than to plain authentication, the node will include empty credit VSAs, depending on categories which has been exhausted. The RADIUS server can identify which category has requested credit renewal.

## Action on Credit Exhaustion

System supports configurable actions once the credit for given subscriber is exhausted:

- Sends an SNMP trap and continue (the credit-usage counter is reset).
- Disconnect.
- Changes to a pre-defined service level (such as adjusting the queue rate).
- Blocks the category.

## Action on Error-Conditions

During credit negotiation, the number of errors can occur which can lead to a given subscriber-host category with no new credit renewed. This is different from credit exhaustion where a separate configurable action will be taken. The following occurs:

- Sends an SNMP trap and continues.
- Sends a trap and blocks the category.

# Applicability of Volume and Time Based Accounting

Volume and time based accounting is applicable to the ESM mode of operation only. Using credit control concept is not mutually exclusive with other accounting methods. In many network implementations the more traditional accounting methods such as XML file or RADIUS accounting will be still used in a combination with the credit concept but with larger intervals. This is helpful when providing overviews of the average usage and service utilization.

# Subscriber Host Idle Timeout

An idle timeout is the maximum time that a subscriber session can be idle before the session is terminated or a connectivity check is started. Idle timeout applies to PPPoE, PPPoEoA, PPPoA and IPoE hosts.

The time/volume based accounting model is used to configure an idle timeout:

- Create a category-map (Categories Map and Categories on page 1111)
  - → Define a category with queues and/or policers to be monitored for activity (packets being forwarded).
  - → An activity threshold (in kbps) must be configured for idle timeout to take effect. The activity threshold suppresses background traffic (for example control flows) from activity monitoring.

Example:

```
config>subscr-mgmt
        category-map "idle-timeout" create
            activity-threshold 25
            category "cat-1" create
                queue 1 ingress-egress
            exit
        exit
```

- In the sla-profile, associate the category-map and optionally define
  - → An idle-timeout (60..15552000 seconds). The default is infinite (no idle-timeout).

  The idle-timeout can also be specified from RADIUS in an access-accept or CoA message with the [28] Idle-Timeout attribute. A RADIUS specified idle-timeout overrides the CLI-configured value. The values outside the limits are accepted but rounded to these boundaries.

| Attribute ID | Attribute name | Type | Limits | Purpose and Format |
|---|---|---|---|---|
| 28 | Idle-Timeout | integer | [60..15552000] seconds | 0 = infinite (no idle-timeout)<br>[60..15552000] in seconds<br>For example:<br>Idle-Timeout = 3600 |

  - → An idle-action:
    - – **shcv-check** — Perform a subscriber host connectivity check (IPoE hosts only). Host connectivity verification should be enabled on the corresponding group-interface for the **idle-action shcv-check** to take effect:

      **configure service ies|vprn** *service-id* **subscriber-interface** *ip-int-name* **group-interface** *ip-int-name* **host-connectivity-verify**

If the shcv check is successful, the subscriber host is not disconnected and the idle-timeout timer is reset to zero. If the shcv check fails, the subscriber host is disconnected (same as terminate).

For PPP hosts, the **idle-action shcv-check** is ignored and has the same effect as "idle-action terminate"

- Terminate (default): disconnect the subscriber hosts
  - IPoE:
    - Delete the subscriber host
    - Send a DHCP release message to the DHCP server
    - Send an Accounting Stop message to the RADIUS accounting server
  - PPP:
    - Delete the subscriber host
    - Send a terminate request message to the CPE
    - Send an Accounting Stop message to the RADIUS accounting server

Example

```
config>subscr-mgmt
        sla-profile "sla-profile-1" create
            category-map "idle-timeout"
                category "cat-1" create
                    idle-timeout 3600
                    idle-timeout-action terminate
                exit
            exit
        exit
```

At host instantiation, a timer is initialized to the idle-timeout value (one timer per sla-profile instance). Each queue or policer in the category is monitored for activity over a fixed polling interval:

- During the polling interval:
  - → if the forwarding rate falls below the configured activity threshold then the timer is deducted by the polling interval (time elapsed).
  - → If the forwarding rate is above the configured activity threshold then the timer is initialized to the idle-timeout value.

When the timer becomes zero, the idle-timeout-action is performed for all hosts associated with the SLA-profile-instance (all hosts from a subscriber on a single sap and that share the same sla-profile).

# Web Authentication Protocol (WPP)

The Web Authentication Protocol (WPP) is a protocol running between BNG and Web portal server. WPP is used for web portal authentication of WLAN users (DHCP Host). It can function like a web portal that can trigger BNG to do RADIUS authentication for WLAN users, or send user disconnection notification to BNG.

The Figure 130 illustrates high level of call flow of WPP authentication.

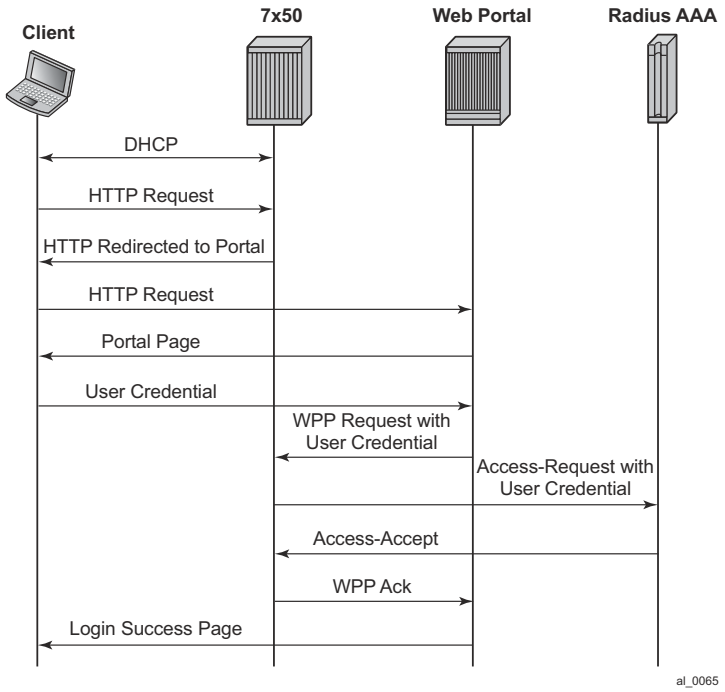



**Figure 130: WPP Authentication**

The following describes WPP authentication call flow:

1. When WLAN user start DHCP exchange with 7750, 7750 will create a DHCP host from following configuration:
   - → Sub-id is the default sub-id configured in the **sap>sub-sla-mgmt** context.
   - → sla-profile/sub-profile/aa-profile will take the configuration from CLI command **grp-if>wpp>initial-sla-profile/initial-sub-profile/initial-app-profile**.
   - → IP address from local or external DHCP server will be assigned to the host.
2. When user sends HTTP request to visit a web site by browser, 7750 redirect the HTTP request to the web portal.
3. Portal Server sends authentication page to WLAN user.

4. WLAN user enters username and password in the authentication page and submit to the Portal Server.
5. Portal server sends WPP request to router together with the user credentials.
6. 7750 sends a access-request to RADIUS server with user credentials.
7. RADIUS returns access-accept if authentication succeeds.
8. 7750 returns WPP ack to portal server.
9. If it was access-accept then 7750 could optionally override following host properties:
   → Sub-id: sub-id from RADIUS. If there is NO sub-id from RADIUS, the host will keep using current sub-id.
   → Sla-profile/sub-profile/aa-profile: system will use RADIUS server returned values; if RADIUS server did not return these then system will try to use LUDB (in local DHCP server) return values if they are available. If not, the system will try to use default values configured under SAP.

## WPP Configurations

A minimal WPP configurations must include the following:

- WPP portal server: specifies the name and ip address of WPP portal server.
- Enable WPP under group-interface:
  → WPP portal server that system should listen to.
  → **authentication-policy** on **group-interface** that specifies address of RADIUS server.
  → **def-sub-id** under sap>sub-sla-mgmt that is used for DHCP host before user is authenticated by portal server.
  → **initial-sla-profile** and **initial-sub-profile** that are used for the DHCP host before user is authenticated by portal server.
  **Note: initial-sla-profile** should include a ingress filter that has **http-redirection** entry.

Following is an example configuration:

```
#--------------------------------------------------
echo "Web Portal Protocol Configuration"
#--------------------------------------------------
        wpp
            portals
                portal "portal-1" address 9.9.9.9 create
                    no shutdown
                exit
            exit
            no shutdown
        exit
config>service>vprn# info
----------------------------------------------------
```

```
            subscriber-interface "sub-if" create
                address 192.168.10.1/24
                group-interface "grp-if" create
                    dhcp
                        server 1.1.1.1
                        gi-address 192.168.10.1
                        no shutdown
                    exit
                    authentication-policy "radius-auth"
                    sap 1/1/9 create
                        sub-sla-mgmt
                            def-sub-id "WLAN-User-Unauth"
                            no shutdown
                        exit
                    wpp
                        initial-sla-profile "webportal"
                        initial-sub-profile "webportal"
                        portal router "Base" name "portal-1"
                        no shutdown
                    exit
                exit
            exit
…
-----------------------------------------------
```

# One-time HTTP Redirection Overview

With this feature enabled, after an ESM host is created, only the FIRST HTTP request from the host will be redirected to a configured URL with specified parameters. Subsequent HTTP request will go through without being redirected.

This feature could be used by service providers to push a web-page to broadband users for purpose of advertisement, announcements, and such.

A **one-time-http-redirection** filter could be configured in **sla-profile**, this filter will be replaced by ingress filter in **sla-profile** after 1st HTTP request is redirected. There is also a RADIUS VSA ( ALC-Onetime-Http-Redirection-Filter-Id) that could be included in access-accept or CoA request to override CLI configuration. The format of ALC-Onetime-Http-Redirection-Filter-Id is **Ingr-v4:***filter-id*; for example, **Ingr-v4:1000**. If the the filter-id is 0, then system will replace the current **one-time-http-redirection** filter with ingress filter.

**Note:** In case of CoA, if the host's **one-time-http-filter** has already been replaced then system will just ignore the ALC-Onetime-Http-Redirection-Filter-Id.

If a 7750 SR receives filter insertion via CoA or access-accept when **one-time-http-redirection** filter is still active then the received filter entries will only be applied to the ingress filter. And after 1st http redirection, the update ingress filter will replace the one-time-http-redirection filter.

This feature only supports IPv4 filter.

# ESM over MPLS Pseudowires

This feature allows IPoE and PPPoE (terminated or L2TP tunneled) subscriber sessions to be backhauled through an Ethernet aggregation network using MPLS pseudowires terminating directly on the BNG. 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 BNG. Multiple subscriber sessions from a given access-port on the Access-PE can be backhauled over a single P2P MPLS pseudowire towards the BNG. This capability allows the network to scale and does not require a MPLS pseudowire per subscriber between Access-PE and the BNG. The access-port on the Access-PE can be dot1q, q-in-q or NULL encapsulated. The BNG terminates the MPLS pseudowire, decapsulates the received frames, and provides ESM 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. A PW-port can be configured with capture SAP. Both static and managed SAPs are supported. The underlying Ethernet port is required to be in hybrid mode. The feature set is supported for IOM3-XP and HSMDAv2. This feature is supported on the 7750 SR and 7450 ESS in mixed mode.

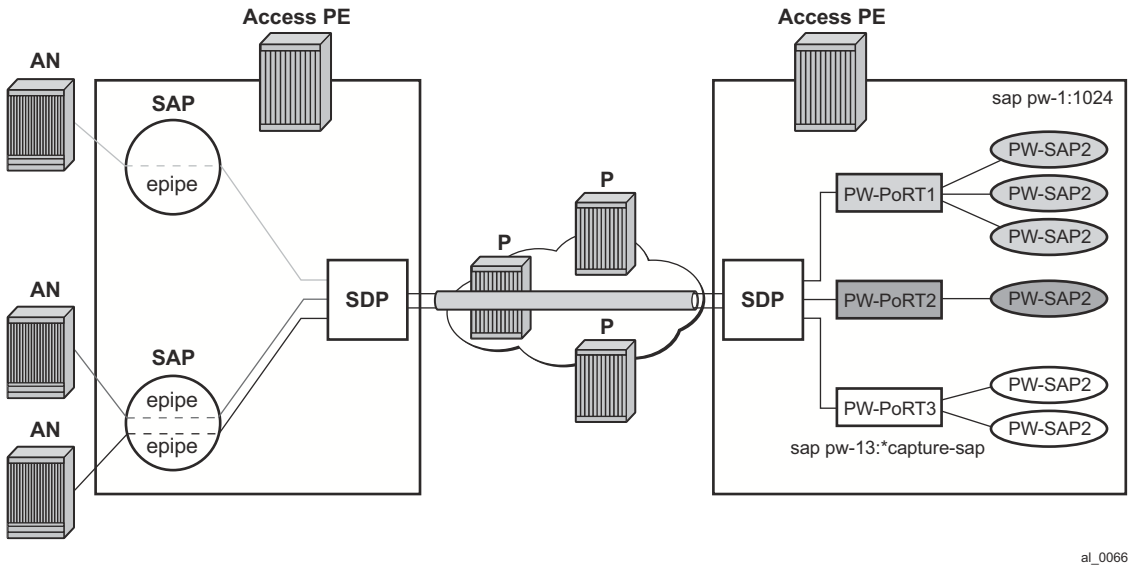



**Figure 131: ESM over MPLS Pseudowire Example**

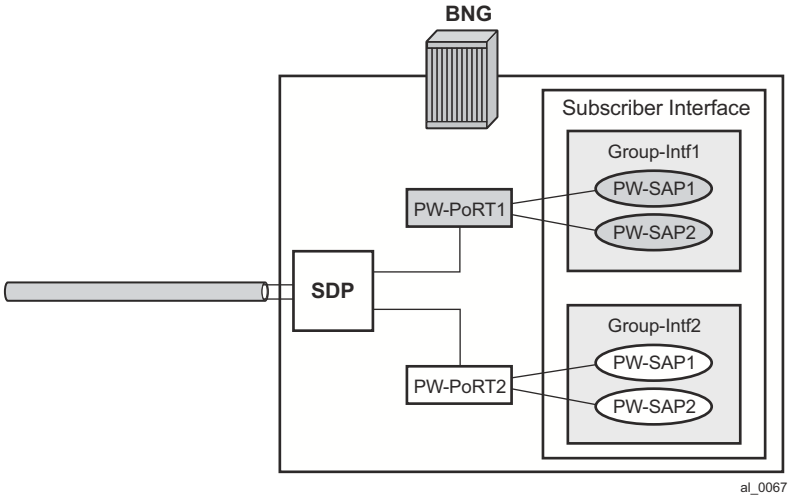



**Figure 132: Group Interface Example**

## Encapsulation

The subscriber frame encapsulated within the pseudowire is shown in Figure 133. Optional control word is not supported. The SDP could be LDP, RSVP or LDP over RSVP. Hash labels are not supported. SDP is bound to a port or a LAG. In case the SDP is re-routed, the corresponding PW-ports are operationally brought down. The PW-ports are associated with the SDP by configuration.

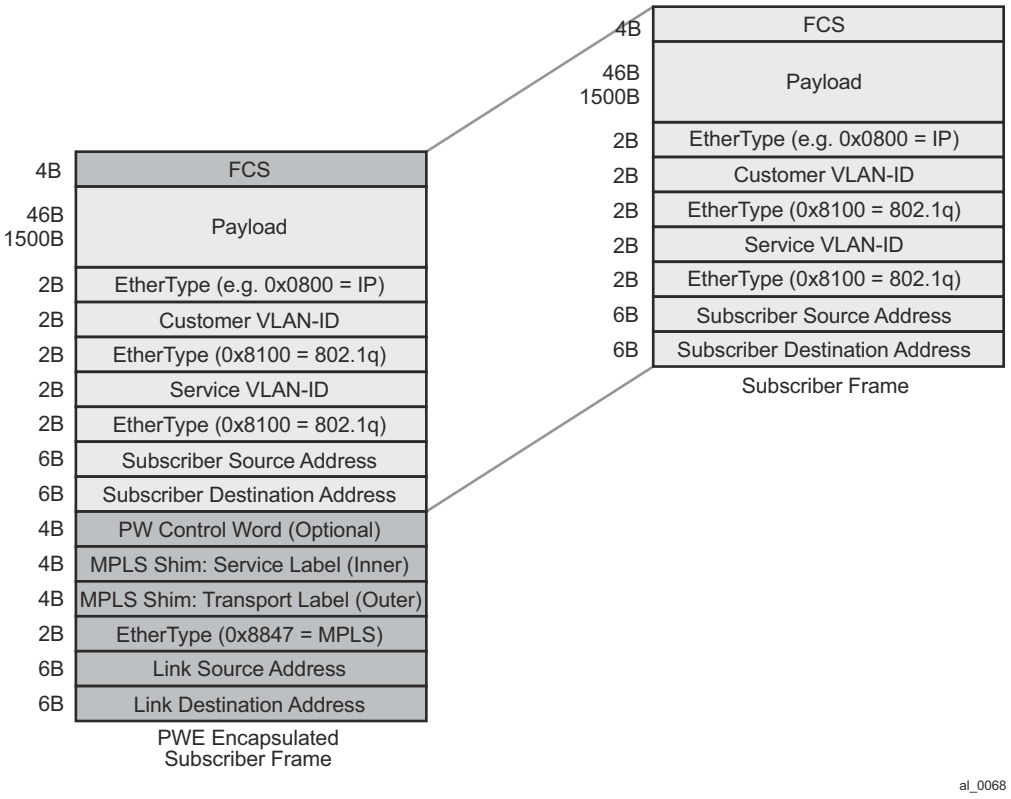



**Figure 133: Subscriber Frame with PWE Encapsulation**

## ESM Configuration with PW-Ports and PW-SAPs

BNG requires configuration for PW-ports. The VC-label for configured PW-port is dynamically signaled using T-LDP with the far-end A-PE. The configuration for the PW-port includes the port-id (unique identifier within the chassis), vc-id (virtual circuit identifier, which is signaled to the peer), and the vc-type (Ether or VLAN, which is signaled to the peer). The vc-id and vc-type must match with the configuration of the PW on the far-end. The vc-id MUST be unique across PW-ports. The encapsulation type (dot1Q or q-in-q) on the PW-port is configurable. The default value

for vc-type is Ether, and the default encapsulation type is dot1Q. With vc-type vlan, the vc-vlan-tag can be configured. vc-type vlan forwarding mode can only be set if encapsulation type is dot1Q. On the BNG, the vc-vlan-tag is only relevant for transport, and not service delineation and ESM. On the BNG, with vc-type vlan configured on the PW-port, the configured vc-type-vlan tag is inserted when forwarding traffic into the PW (i.e. in downstream direction), and top dot1Q tag is stripped when forwarding traffic out of the PW (i.e. in upstream direction). On the BNG, with vc-type ether configured on the PW-port, the received tags (max two, including any provider tag inserted by the far-end) are preserved and passed for PW-SAP lookup or creation. In the downstream direction the PW-SAP tags are inserted and passed back to the far-end.

The following output displays an ESM configuration with PW-ports and PW-SAPs.

```
config>service#
  customer 1 create
     description "Default Customer"
  exit
  sdp 1 mpls create
       description "Default sdp description"
       far-end 10.20.1.2
       ldp
       keep-alive
          shutdown
       exit
       binding
          port 1/1/3
          pw-port 11 vc-id 11 create
             vc-type vlan        #### default encaps-type dot1Q
             no shutdown
          exit
          pw-port 44 vc-id 2 create #### default vc-type Ether, encaps-type dot1Q
             no shutdown
          exit
        exit
        no shutdown
    exit
    vpls 1 customer 1 vpn 1 create
       sap pw-11:* capture-sap create
           trigger-packet arp dhcp dhcp6 pppoe #
           msap-defaults
               group-interface "grpif-pw-11"
               policy "msap-policy1"
               service 3
           exit
           authentication-policy "base_authpolicy"
        exit
        no shutdown
     exit

     ies 3 customer 1 vpn 3 create
         description "Default ies description for sevice id 3"
         subscriber-interface "subif" create
              address 11.11.1.1/16
              address 44.44.1.1/16
              group-interface "grpif-pw-11" create
                 arp-populate
                 dhcp
                    server 10.20.1.2
```

```
            gi-address 11.11.1.1
            no shutdown
         exit
         authentication-policy “base_authpolicy”
         sap pw-11:11 create
             sub-sla-mgmt
                  def-sub-profile “sub_prof_1”
                  def-sla-profile “sla_prof_1”
                  no shutdown
              exit
          exit
      exit
group-interface "grpif-pw-44" create
     arp-populate
     dhcp
        server 10.20.1.2
        gi-address 11.11.1.1
        no shutdown
     exit
     sap pw-44:44 create
          sub-sla-mgmt
              def-sub-profile “sub_prof_1”
              def-sla-profile “sla_prof_1”
              no shutdown
      exit
   exit
   no shutdown
exit
```

## QoS Support

QoS is supported for ESM over PW-SAPs as with ESM over regular SAPs, and includes currently supported models.

- FC to queue mapping
- H-QOS
  - → Per-subscriber HQOS (service scheduler child to port-scheduler parent).
  - → PW-SAP queues attached to H-QOS scheduler by "port-parent" statement.
  - → Scheduler attached to port scheduler by "port-parent" statement.
- Direct service queue to port-scheduler.
  - → Aggregate-rate-limit.

---

### Bandwidth Control at PW-Port Level via Vport

Bandwidth control per PW-port (per AN or per AN/ per service), via Vport.

- The vport can be created on the binding port.
- The vport can be associated with the PW-port either via static assignment or dynamic selection via inter-dest-id (returned from RADIUS or DHCP for a host).
- Aggregate-rate-limit can be configured to shape the egress traffic across all hosts associated with the vport via inter-dest-sting match or via static association of underlying PW-port with the vport.

The following output displays a dynamic Vport selection based on an inter-dest-id configuration.

```
config>
   Port 1/1/1
      ethernet
         mode hybrid
         encap-type dot1Q
         mtu 1540
         access
            egress
            vport "v1" create
               agg-rate-limit 1000
               host-match dest "dslam-1"    #### hosts will be associated with
            exit                            #### vport based on inter-dest-id
          exit
      exit
   exit

config>service>sdp>binding
   pw-port 11 vc-id 11 create
       egress
          shaping int-dest-id "dslam-1"   #### dynamic vport selection based on
                                          #### int-dest-id.
```

The following output displays a static assignment of PW-port to Vport configuration.

```
config>
   Port 1/1/2
      ethernet
         mode hybrid
         encap-type dot1Q
         access
            egress
            vport "v2" create
               agg-rate-limit 1000
            exit
          exit
       exit
   exit

config>service>sdp>binding
    pw-port 20 vc-id 20 create
       egress
          shaping vport "v2"    #### static assignment of pw-port to vport.
       exit
    exit
```

## Last Mile Shaping

With normal Ethernet aggregation in the next-mile, when last-mile shaping is on, fixed encapsulation-offset is calculate based on the last-mile encapsulation type and the next-mile encapsulation (26 Bytes with q-in-q). This offset is applied to the frame, and the ATM overhead is then dynamically calculated on the adjusted size. The resulting dynamically calculated overhead in the data-path is then applied to the queue-rates and the subscriber aggregate-rate.

With this feature of backhauling subscriber sessions using MPLS PW in the aggregation network, the encapsulation is shown in Fig 3. The last mile does not see any MPLS PW overhead. The next-mile includes overhead due to the PW encapsulation shown in Figure 133. Therefore, when last mile shaping is enabled, the fixed encapsulation-offset is calculated based on the difference between last-mile encapsulation type and next-mile encapsulation, The next-mile encapsulation takes into account the additional PW overhead, which includes:

```
14B Ethernet header + [4B] (optional network interface Q-tag) + MPLS
Labels (variable)
```

In the data-path the actual PW encapsulation overhead, taking into account the MPLS labels which could be variable (with FRR or PHP) is tracked, and is applied to the computed "encapsulation offset". This adjusted "encapsulation offset" is applied to the frame. The ATM overhead is then dynamically calculated on the adjusted size, and applied for last mile shaping (to queue-rates and subscriber-aggregate-rate). Note that there is no change from ESM over normal SAPs, in how last-mile shaping is triggered or how the last mile encapsulation type is determined (via configuration in egress context of subscriber profile or dynamically learned from Access-Loop-Encapsulation sub-TLV in vendor specific PPPoE tags).

# BNG Redundancy with ESM over Pseudowire

This feature provides support for stateful BNG redundancy (when the far-end aggregation PE (A-PE) is dual-homed to two BNGs terminating subscriber sessions over MPLS pseudowires (pws) that are initiated from the A-PE and provides ESM). Subscriber state between BNGs is synced using MCS.

## EPIPE Based Aggregation Service

For an EPIPE based aggregation service, the redundancy is based on active/standby PWs from A-PE to dual BNGs. A-PE signals active/standby pseudowire status to peer BNGs. An SRRP instance per PW-Port (group-interface) is required on the BNG with messaging SAP on each PW-Port. BNG terminating active PW assumes the mastership for the SRRP instance on the corresponding PW-Port. SRRP state is tied to the state of the messaging SAP. The messaging SAP goes down when the underlying PW-Port goes down, based on PW status bit signaled by the A-PE.

In this model, there is no SRRP message exchange between the two BNGs, as there is no L2 path between the BNGs. The purpose of SRRP is to get SRRP-aware routing for subscriber routes and managed routes, and/or to be able to use the redundant (shunt) interface. Downstream traffic for a subscriber that ingresses the backup BNG can only be shunted to the active BNG, if the corresponding subscriber-interface on the backup BNG is operationally UP. This can be achieved by creating a second empty group-interface (without SAPs) on the same subscriber-interface with the knob 'oper-up-while-empty' configured. Multiple PWs with endpoint configuration is not supported on the BNG.

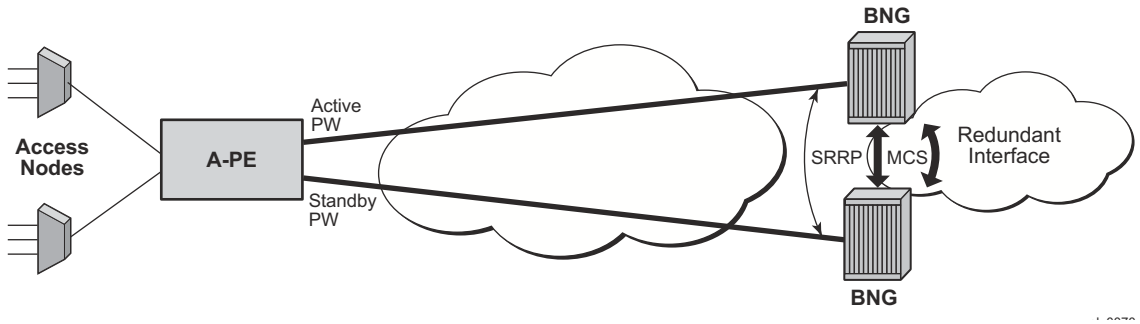



**Figure 134: BNG Redundancy Based on Active/Standby PW Signaling**

## Sample Configuration on Master BNG

```
config>
    pw-port 2 create
    exit

config>redundancy#
     multi-chassis
         peer 10.20.1.3 create
             source-address 10.20.1.2
                sync
                    srrp
                    sub-mgmt ipoe pppoe
                    port pw-2 sync-tag "tag2" create
                    exit
                    no shutdown
                exit
                no shutdown
            exit
        exit
     exit

config>service>ies#
       redundant-interface "redundant-interface" create
          address 10.10.30.2/24 remote-ip 10.10.30.3
          spoke-sdp 23:1000 create
             no shutdown
          exit
       exit
config>service#
    sdp 1 mpls create
       far-end 10.20.1.2
       ldp
       keep-alive
          shutdown
       exit
       binding
          port 1/1/1
          pw-port 2 vc-id 2 create
             vc-type vlan        #### default encaps-type dot1Q
             no shutdown
          exit
       exit
       no shutdown
    exit
config>service#
       subscriber-interface "subif" create
           address 11.11.1.2/16 gw-ip-address 11.11.1.1 populate-host-routes
                group-interface "grpif" create
                    authentication-policy "base_authpolicy"
                    redundant-interface "redundant-interface"
                    sap pw-2:1000 create
                        description "sap-grp-3"
                    exit
                    srrp 1 create
                        message-path pw-2:1000
```

```
                    no shutdown
                exit
                arp-host
                    host-limit 8000
                    min-auth-interval 1
                    no shutdown
                exit
            exit
        exit
    exit
```

---

## Sample Configuration on Slave BNG

```
config>
    pw-port 2 create
    exit
config>redundancy#
     multi-chassis
         peer 10.20.1.2 create
             source-address 10.20.1.3
                sync
                    srrp
                    sub-mgmt ipoe pppoe
                    port pw-2 sync-tag "tag2" create
                    exit
                exit
                no shutdown
            exit
        exit
     exit
config>service>ies#
       redundant-interface "redundant-interface" create
          address 10.10.30.3/24 remote-ip 10.10.30.2
          spoke-sdp 32:1000 create
             no shutdown
          exit
       exit
config>service#
    sdp 1 mpls create
       far-end 10.20.1.2
       ldp
       keep-alive
          shutdown
       exit
       binding
          port 1/1/1
          pw-port 2 vc-id 2 create
             vc-type vlan        #### default encaps-type dot1Q
             no shutdown
          exit
       exit
       no shutdown
    exit
config>service#
       subscriber-interface "subif" create
           address 11.11.1.3/16 gw-ip-address 11.11.1.1 populate-host-routes
```

```
            group-interface "grpif" create
                authentication-policy "base_authpolicy"
                redundant-interface "redundant-interface"
                sap pw-2:1000 create
                    description "sap-grp-3"
                exit
                srrp 1 create
                    keep-alive-interval 1
                    message-path pw-2:1000
                    no shutdown
                exit
                arp-host
                    host-limit 8000
                    min-auth-interval 1
                    no shutdown
                exit
            exit
            group-interface "dummy" create
                oper-up-while-empty
            exit
        exit
    exit
```

## Sample Configuration on A-PE

```
config>service>epipe#
     description "Default epipe description for service id 103"
     service-mtu 1492
     service-name "XYZ Epipe 103"
     endpoint "x" create
         standby-signaling-master
     exit
     sap 1/1/3 create
          description "Default sap description for service id 103"
     exit
     spoke-sdp 1:2 endpoint "x" create
          description "Description for Sdp Bind 1 for Svc ID 103"
          precedence primary
          no shutdown
     exit
     spoke-sdp 2:2 endpoint "x" create
          description "Description for Sdp Bind 2 for Svc ID 103"
          no shutdown
     exit
     no shutdown
```

## VPLS Based Aggregation Service

With VPLS based aggregation service from A-PE, normal SRRP message exchange can take place between the primary and backup BNGs. Master-ship decision and switch-over is based on SRRP. SRRP instance is configured per group-interface corresponding to PW-Port. Fate-sharing groups (FSG) can be configured for a set of SRRP instances (for example, SRRP instances corresponding to PW-Ports sharing the same subnet). Standard **oper-group** *grp-id* would need to be configured with messaging SAPs for all PW-Ports that are in the same FSG, and **monitor-oper-group** *grp-id* would need to be configured under each SRRP instance in same FSG. Existing SRRP support defined in Triple-play services guide for ESM over regular group-interfaces and subscriber SAPs is applicable identically to ESM over PW-Ports and PW-SAPs.

**Note:** With PW over ESM, redundancy in the aggregation network based on MC-LAG between A-PE and dual BNGs is not supported.

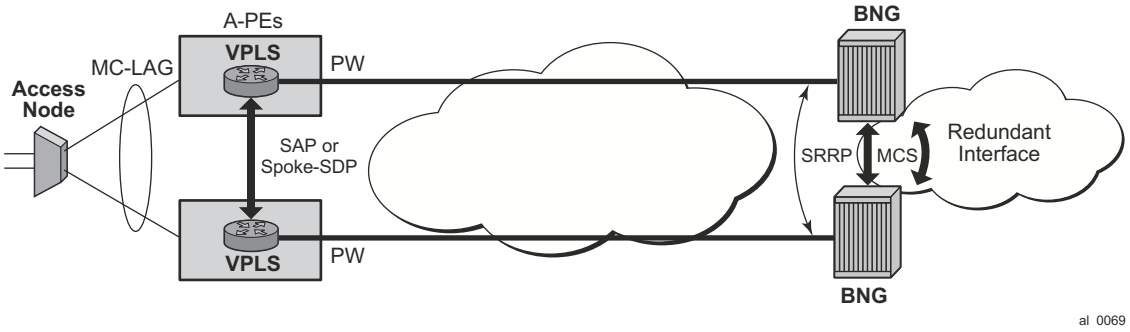



**Figure 135: BNG Redundancy with VPLS Based Aggregation Service**

## Sample BNG Redundancy (SRRP/MCS) Configuration with VPLS Service on A-PE

```
config>
    pw-port 1 create
    exit

config>redundancy#
     multi-chassis
         peer 10.20.1.2 create
             source-address 10.20.1.3
                sync
                    srrp
                    sub-mgmt ipoe pppoe
                    port pw-1 sync-tag "tag1" create
                    exit
                exit
                no shutdown
            exit
        exit
     exit

 config>service>ies
       redundant-interface "red-1-1" create
           address 1.1.1.2/24 remote-ip 1.1.1.1
               spoke-sdp 1:1 create
                  no shutdown
               exit
           exit

       subscriber-interface "sub-1-1" create
            address 20.1.2.2/16 gw-ip-address 20.1.255.254 track-srrp 1
            address 20.2.2.2/16 gw-ip-address 20.2.255.254 track-srrp 2
            dhcp
                gi-address 20.1.2.2
            exit
            group-interface "grp-1-1-1" create
                 srrp-enabled-routing
                 arp-populate
                 dhcp
                     server 10.20.1.2
                     trusted
                     lease-populate 32767
                     client-applications dhcp ppp
                     gi-address 20.1.2.2
                     no shutdown
                 exit
                    authentication-policy "iesAuthPol"
                    redundant-interface "red-1-1"

                   sap pw-1:1.1 create
                       sub-sla-mgmt
                           def-sub-profile “sub_prof_1”
                           def-sla-profile “sla_prof_1”
                           no shutdown
                   exit
                   sap pw-1:4000.1 create
                        oper-group "1"
```

```
                exit

                srrp 1 create
                     gw-mac 00:00:5e:00:01:01
                     keep-alive-interval 50
                     message-path pw-1:4000.1
                     monitor-oper-group "1" priority-step 10
                     no shutdown
                 exit
              exit
```

## A-PE configuration with VPLS Aggregation Service (A-PE1)

```
config>service
     customer 1 create
         description "Default customer"
     exit
     sdp 1000 mpls create
         far-end 10.20.1.2
         lsp "lsp_1"
         path-mtu 1600
         keep-alive
         no shutdown
      exit
      sdp 1002 mpls create
         far-end 10.20.1.3
         lsp "lsp_3"
         path-mtu 1600
         keep-alive
         no shutdown
      exit
      vpls 1 customer 1 create
          service-mtu 1600
          stp
          sap 1/1/2 create   // to Access-Node
          exit
        sap 1/1/3 create; //to A-PE2
        exit
          spoke-sdp 1000:1 create  // to BNG1
              no shutdown
          exit
          no shutdown

        exit
    exit
```

## A-PE Configuration with VPLS Aggregation Service (A-PE2)

```
config>service
     customer 1 create
         description "Default customer"
     exit
     sdp 1002 mpls create
         far-end 10.20.1.3
         lsp "lsp_2"
```

```
        path-mtu 1600
        keep-alive
        no shutdown
    exit

    vpls 1 customer 1 create
        service-mtu 1600
        stp
        sap 1/1/2 create   // to Access-Node
        exit
      sap 1/1/3 create; //to A-PE1
      exit
        spoke-sdp 1002:1 create  // to BNG2
            no shutdown
        exit
        no shutdown
      exit
exit
```

## Show Commands Related to Active/Standby Pseudowire on Dual BNGs

The following example shows SRRP status, subscriber host, and routing information on master BNG:

```
A:Dut-B>config>redundancy# show srrp 1

===============================================================================
SRRP Instance 1
===============================================================================
Description         : (Not Specified)
Admin State         : Up                Oper State       : master
Preempt             : yes               One GARP per SAP : no
Monitor Oper Group  : None
System IP           : 10.20.1.2
Service ID          : VPRN 3
Group If            : grpif             MAC Address      : 1c:85:ff:00:00:00
Grp If Description  : N/A
Grp If Admin State  : Up                Grp If Oper State: Up
Subscriber If       : subif
Sub If Admin State  : Up                Sub If Oper State: Up
Address             : 11.11.1.2/16      Gateway IP       : 11.11.1.1
Redundant If        : redundant-interfa*
Red If Admin State  : Up                Red If Oper State: Up
Address             : 10.10.30.2/24
Red Spoke-sdp       : 23:1000
Msg Path SAP        : pw-2:1000
Admin Gateway MAC   :                   Oper Gateway MAC : 00:00:5e:00:01:01
Config Priority     : 100               In-use Priority  : 100
Master Priority     : 100
Keep-alive Interval : 1 deci-seconds    Master Since     : 05/29/2012 07:22:26
Fib Population Mode : all
VRRP Policy 1       : None              VRRP Policy 2    : None
===============================================================================
* indicates that the corresponding row element may have been truncated.

A:Dut-B>config>redundancy# show service id 3 arp-host

===============================================================================
ARP host table, service 3
===============================================================================
IP Address      Mac Address       Sap Id              Remaining           MC
                                                      Time                Stdby
-------------------------------------------------------------------------------
11.11.1.11      00:80:00:00:00:01 [pw-2:11]           03h35m47s
11.11.1.12      00:80:00:00:00:02 [pw-2:12]           03h35m47s
-------------------------------------------------------------------------------
Number of ARP hosts : 2
===============================================================================

A:Dut-B>config>redundancy# show router 3 route-table 11.11.1.11

===============================================================================
Route Table (Service: 3)
===============================================================================
Dest Prefix[Flags]                          Type    Proto    Age         Pref
      Next Hop[Interface Name]                                 Metric
-------------------------------------------------------------------------------
```

```
11.11.1.11/32                                    Remote  Sub Mgmt 00h24m26s   0
       [grpif]                                                             0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: L = LFA nexthop available    B = BGP backup route available
       n = Number of times nexthop is repeated
===============================================================================
A:Dut-B>config>service>vprn#
```

The following shows SRRP status, subscriber host, and routing info in slave BNG:

```
A:Dut-C>config>redundancy# show srrp 1

===============================================================================
SRRP Instance 1
===============================================================================
Description         : (Not Specified)
Admin State         : Up                Oper State       : initialize
Preempt             : yes               One GARP per SAP : no
Monitor Oper Group  : None
System IP           : 10.20.1.3
Service ID          : VPRN 3
Group If            : grpif             MAC Address      : 1c:87:ff:00:00:00
Grp If Description  : N/A
Grp If Admin State  : Up                Grp If Oper State: Down
Subscriber If       : subif
Sub If Admin State  : Up                Sub If Oper State: Up
Address             : 11.11.1.3/16      Gateway IP       : 11.11.1.1
Redundant If        : redundant-interfa*
Red If Admin State  : Up                Red If Oper State: Up
Address             : 10.10.30.3/24
Red Spoke-sdp       : 32:1000
Msg Path SAP        : pw-2:1000
Admin Gateway MAC   :                   Oper Gateway MAC : 00:00:5e:00:01:01
Config Priority     : 1                 In-use Priority  : 1
Master Priority     : 1
Keep-alive Interval : 1 deci-seconds    Master Since     : 05/29/2012 07:22:26
Master Down Interval: 0.000 sec (Expires in 0.000 sec)
Fib Population Mode : all
VRRP Policy 1       : None              VRRP Policy 2    : None
===============================================================================
* indicates that the corresponding row element may have been truncated.


A:Dut-C>config>redundancy# show service id 3 arp-host

===============================================================================
ARP host table, service 3
===============================================================================
IP Address      Mac Address       Sap Id              Remaining           MC
                                                      Time                Stdby
-------------------------------------------------------------------------------
11.11.1.11      00:80:00:00:00:01 [pw-2:11]           03h38m01s           Yes
11.11.1.12      00:80:00:00:00:02 [pw-2:12]           03h38m02s           Yes
-------------------------------------------------------------------------------
Number of ARP hosts : 2
===============================================================================

A:Dut-C>config>redundancy# show router 3 route-table 11.11.1.11
```

```
===============================================================================
Route Table (Service: 3)
===============================================================================
Dest Prefix[Flags]                              Type    Proto    Age         Pref
      Next Hop[Interface Name]                                     Metric
-------------------------------------------------------------------------------
11.11.1.11/32                                   Remote  Sub Mgmt 00h22m03s   0
       [redundant-interface]                                          0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: L = LFA nexthop available    B = BGP backup route available
       n = Number of times nexthop is repeated
===============================================================================
```

# 3GPP-based Diameter Credit Control Application (DCCA) – Online charging

On-line charging applications allow to control subscriber access to services based on a pre-paid credit. The volume and time accounting in 7750 SR supports online charging using the Diameter Credit-Control Application (DCCA). The 7750 SR supports Session Charging with Unit Reservation (SCUR) allowing the 7750SR to reserve volume and time quota for rating-groups. Furthermore, the 7750 SR supports centralized unit determination and centralized rating: it requests quota and reports usage against the quota provided by the Online Charging Server (OCS). Credit control is always on a per rating group basis. A rating group maps to a category inside a category-map of the 7750SR volume and time based accounting function.

The following are the basic configuration steps:

1. Configure a diameter policy

   In the diameter-base CLI context, configure one or multiple Diameter peers.

```
config>subscr-mgmt
        diameter-policy "diameter-1" create
            diameter-base
                origin-host "bng.alcatel-lucent.com"
                origin-realm "alcatel-lucent.com"
                source-address 10.0.0.1
                connection-timer 5
                peer "peer-1" create
                    address 10.1.0.1
                    destination-host "server.alcatel-lucent.com"
                    destination-realm "alcatel-lucent.com"
                    no shutdown
                exit
            exit
        exit
```

   Optionally, configure additional parameters for the Diameter Credit Control Application in the dcca CLI context.

2. Create a category-map in which you define:
   → the credit type (time or volume)
   → a category defining the queues to monitor for quota consumption and the rating-group this category maps to in DCCA.

```
config>subscr-mgmt
        category-map "cat-map-1" create
            credit-type time
            category "cat-1" create
                rating-group 1
                queue 1 ingress-egress
                exhausted-credit-service-level
                    pir 64
                exit
            exit
        exit
```

3. Create a credit control policy

   Define the credit control servers to use by specifying the diameter policy. Optionally specify the default-category-map and an out-of-credit-action.

```
        credit-control-policy "cc-policy-1" create
            credit-control-server diameter "diameter-1"
            default-category-map "cat-map-1"
            out-of-credit-action change-service-level
        exit
```

4. Configure the diameter credit-control-policy to use for a subscriber host in the corresponding sla-profile.

```
        sla-profile "sla_prof_1" create
            credit-control-policy "cc-policy-1"
        exit
```

The following are examples of Diameter on-line charging flows:

Scenario 1 — Depicts a redirect use-case:

When the quota is depleted, the subscriber is redirected to a web portal. When the credit is refilled, the OCS server will notify the BNG and provide new quota. Note that 7750SR will install the configured out-of-credit-action when receiving a Final Unit Indication with action different from Terminate.

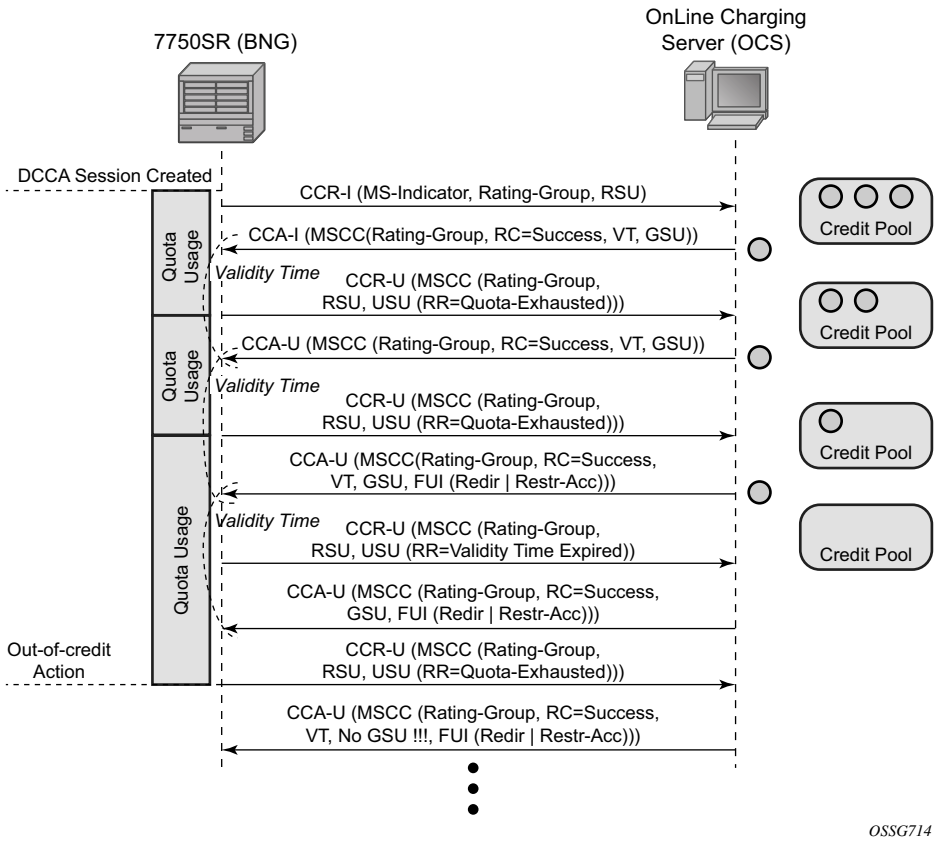



**Figure 136: On-Line Charging Scenario 1 - Redirect (1/2)**

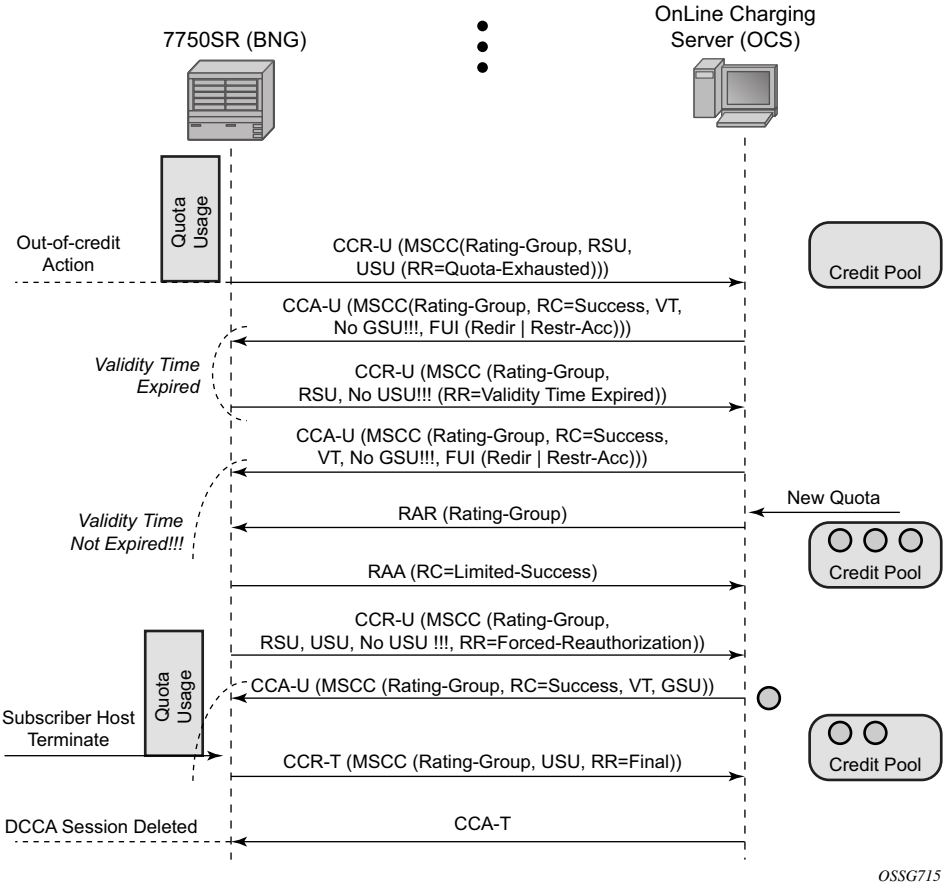



**Figure 137: On-Line Charging Scenario 1 - Redirect (2/2)**

Scenario 2 — Depicts a terminate use case:

When the quota is depleted after reception of a Final Unit Indication with action set to Terminate, the subscriber host is disconnected. The configured out-of-credit-action is ignored in this case.

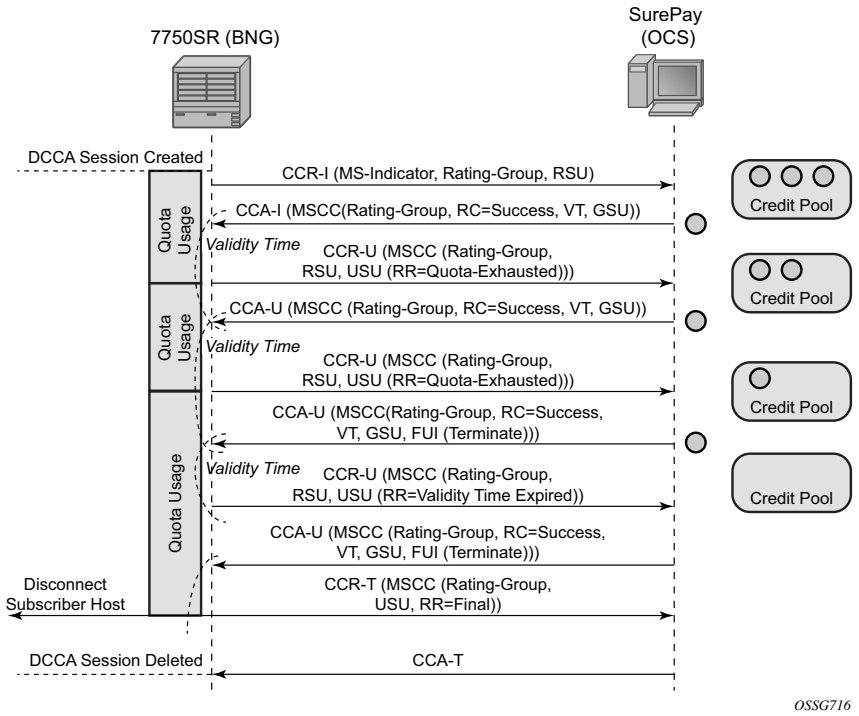



**Figure 138: On-Line Charging Scenario 2 – Terminate**

Abbreviations used in the previous drawings:

| | |
|---|---|
| CCR | Credit Control Request (-Initial, -Update, -Terminate) |
| CCA | Credit Control Answer (-Initial, -Update, -Terminate) |
| RAR | Re-Authentication Request |
| RAA | Re-Authentication Answer |
| MSCC | Multiple Services Credit Control |
| GSU | Granted Service Unit |
| RSU | Requested Service Unit |
| USU | Used Service Unit |
| RC | Result Code |
| RR | Reporting Reason |
| VT | Validity Time |

# On-Demand Subnet Allocation (ODSA)

## DHCP pool *subnet-binding-key*

To share a DHCP pool among BNG, the DHCP server will first require a **gi-address**. This feature is used in conjunction with **use-gi-address** from the **local-dhcp-server** and the scope should be *pool* to allow allocation of all subnets within the pool. The DHCP discovery must contain any of the three vendor specific options in Option 82: service ID, service-plus-system ID, or custom string. The intent is to bind a subnet from the shared pool to one of the three required parameters from the vendor specific option (VSO).

When starting to use the shared pool for the first time, a DHCP request should arrives with one of the three required VSOs. The DHCP server looks for a free subnet to bind to the VSO, and then an address is offered to the subscriber. Only subscribers utilizing the same DHCP VSO are allowed to request addresses from the same registered subnet. All new DHCP discovery VSOs are matched against VSOs bound to existing subnets. For the case where a DHCP discovery VSO matches a VSO bound to a subnet, an address from the subnet is offered to subscriber until exhaustion. Once exhausted, the DHCP looks for a free available subnet (if any) for binding and the same process continues until the new subnet is exhausted. Subscribers with non-matching DHCP VSOs are prohibited to request address from the bounded subnet. For the case where a DHCP discovery VSO fails to match any of the current binding, a new unbound subnet is searched for binding (if any).

Each subscriber interface sharing a DHCP pool, must utilize a unique **gi-address**. This is to ensure that the DHCP offer can correctly route back to the subscriber interface as shown in Figure 139. The number of subscriber interfaces sharing the DHCP pool, must have at least the equivalent number of subnets in the DHCP pool. The number of subnets available for sharing should always be equal or greater then the number of subscriber interfaces.

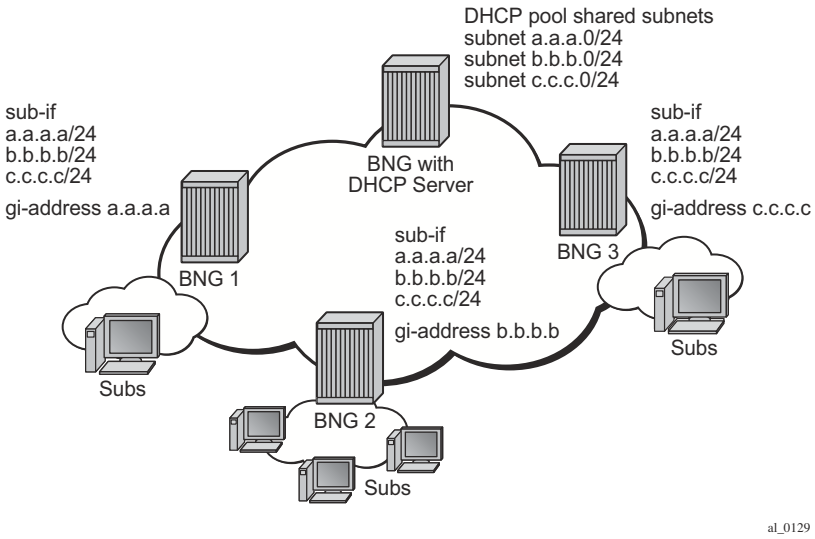



**Figure 139: Subscriber Interfaces Sharing a DHCP Pool**

# ODSA Subnet Advertisement and Routing

To avoid advertising the same IP subnet from multiple subscriber interfaces. A combination of router aggregation and route policies are used to ensure BNGs do not advertisement overlaps subnets. If BNG1, BNG2, and BNG N are sharing a DHCP pool with subnet A,B,C,…,Z. All BNGs must first provision the same subnets on their subscriber interfaces. These subscriber interfaces must not be included in any IGP interfaces, because that will lead to multiple BNG advertising the same subnets which is undesirable. The key is to advertise a subnet only if a subscriber has successfully allocated an address from that particular subnet. Aggregate route can be used to accomplish this. First, the aggregation route will specify a smaller (more specific) subnet of the subscriber interface. For example if Subnet A is 1.1.1.1/24 then the aggregation route would be split into the next smallest subnet 1.1.1.0/25 and 1.1.1.128/25 respectively. A Route policy is then used to advertise the two /25 into IGP only if these aggregate routes become active. A subscriber that has successfully obtained an IP address from the shared pool will activate the aggregate route in the RIB. Once activated in the RIB, the route policy exports these routes by means of IGP or BGP. This in conjunction with ODSA, allows a subnet to be shared between BNGs and provides proper routing for all subscriber traffic.

# ODSA with SRRP

For PPPoE SRRP setups, it is mandatory to use “string” as the subnet-binding key. For IPoE SRRP setups, “string” is preferred but not a mandatory requirement. The benefits of using “string” as a binding key are discussed later in the failover and recovery sections.

For an ODSA PPPoE SRRP setup, there are two mandatory requirements. They are:

- Configure a DHCP vendor specific string under the group interface PPPoE option.
- ODSA DHCP pools must use “string” as the subnet-binding key.

Each group interface can have its own customized string. This will result in each group-interface requesting for its own subnet. Another possibility is to share subnets among all group interfaces under a single subscriber interface. This is accomplished by using the same string on all group-interface under the single subscriber interface. A pair of SRRP group interface between two BNGs should use the same string.

For an IPoE setup, both SRRP and non SRRP, it is possible to insert a custom string via the DHCP relay (configured under the dhcp-relay option). Remember to also use “string” as the subnet-binding.

## ODSA SRRP Failover DHCP Behavior

ODSA allows the option of binding a subnet to any of 3 following keys: system-id, system-id + svc-id, or string. For SRRP setups, while PPPoE only allow the use of “string”, IPoE allow the use of any of the three keys for subnet binding. Notice that first two keys both use the element “system-id” for binding. During a SRRP failover, the slave system-id takes over and the new system-id will in turn bind to new subnets. Existing subnet bounded by the old master system-id are non-accessible even if there are still free addresses available. Subscribers created before the failover still requires DHCP renews, requests, and releases of their addresses. When the DHCP server receives these, renews, requests, and releases, the gi-address and system-id might no longer match. The DHCP server will still answer these DHCP messages as long as the IP and MAC address matches the one registered on the server. This allows the subscriber to experience a seamless connection during a SRRP failover.

The other option is subnet-binding with “string” which is mandatory for ODSA PPPoE SRRP setup and optional for ODSA IPoE SRRP setup. Unlikely the system-id, the “string” is customizable and can match between two SRRP group interfaces across two different BNGs. During a failover, the subnet-binding key will remain the same. New subscribers can reuse existing subnet already bound by the group interface. No free addresses are wasted. The key requirement is that each pair of SRRP group interface must have a custom unique string.

## ODSA SRRP Recovery DHCP Behavior

When SRRP is repaired, one of the nodes will become the slave. All DHCP relays will come from the SRRP master. For example, if node 1.1.1.3 becomes the slave, then all DHCP relay messages will have the gi-address of 1.1.1.2. Old subscribers that utilize 1.1.1.3 to retrieve DHCP address still require DHCP requests, renews, and releases from the DHCP pool. In the case of IPoE, DHCP server will allow old subscribers with matching IP and MAC to perform DHCP requests, renews, and releases even though the system-id does not match the previous one. This is to ensure subscribers will have uninterrupted services during a recovery.

To recover the addresses/subnets from the slave node, DHCP drain can be used to ensure that IP addresses are released back to the pool when the leases expire. Otherwise, it is best to wait for the DHCP leases to expire for subscribers to ensure that services are uninterrupted. Once expired, subscribers will route through the proper Master SRRP using the correct system-id to retrieve DHCP addresses.

In the case where the subnet-binding key is “string”, the SRRP recovery is more seamless. Although utilizing a different gi-address, DHCP relays will utilize the same “string” for all DHCP transaction. The same subnets will continue to supply, renew, and release DHCP addresses. There is no need to manually drain unused subnets because the same subnets will be used.

# Logical Link Identifier (LLID)

This feature enables service providers to track subscribers on the basis of a virtual-port known as logical line ID (LLID). The LLID (an alphanumeric string) is a logical identification of a subscriber line. Mapping of physical line of a subscriber to LLID is performed via pre-authentication with a separate AAA server than the AAA server used for authenticating the subscriber session during normal access authentication.

LLID serves the purpose of abstracting the physical line of the user from the ISP. If the user moves to a new physical line, the RADIUS server database maintaining the physical line of the subscriber to LLID is updated. Because a subscriber's LLID remains same regardless of subscriber's physical location, using LLID gives service provider a stable and secure identifier for tracking subscriber.

The local user database assigned to the PPPoE node under the group interface can have both a pre-authentication policy and an authentication policy. The purpose of the pre-authentication policy is to retrieve the LLID from the AAA server. The pre-authentication will only extract the calling-station-id attribute (0x31) which is used as the LLID, anything else returned during pre-authentication are simply ignored. If the pre-authentication is missing the LLID, the session will move on to the authentication policy. In the authentication policy that follows, it is possible to use the LLID as the calling-station ID.

It is possible to convey LLID from the LAC to the LNS. The LLID is retrieved through PPPoE pre-authentication where the returned radius attribute calling station ID is used as the LLID. This LLID is selectable attribute in L2TP as a calling-number (AVP 22) to be passed from LAC to LNS. At the LNS, the subscriber calling station number is retrieved from AVP 22 and can be included as an attribute during authentication.

# Open Authentication Model for DHCP and PPPoE Hosts

## Terminology

LUDB – Local User Database configured within 7x50

- IP Address Assignment via DHCP Relay — IP address assignment request (DHCP or IPCP) from the host is relayed to an internal or external DHCP server. Gi-address must be present in this relayed request while the pool name is optional. The internal 7x50 DHCP server may select the IP address from is local pool based on the gi-address or based on the pool-name present in the request. The IP address selection method is configuration dependent. Third party DHCP servers may consider additional fields in IP address selection process (mac address, circuit-id, etc).
- IP Address Assignment via DHCP Proxy — A preconfigured IP address in LUDB or RADIUS server is handed out to the host via a 7x50 DHCP proxy function. This proxy function responds natively using DHCP protocol to the IPoE host. Although PPPoE hosts are not utilizing DHCP protocol, the DHCP proxy functionality within 7x50 is still needed for successful IP address delegation to PPPoE hosts.

## LUDB and RADIUS Access Models

During the subscriber-host instantiation phase in 7x50, various parameters for the hosts are gathered from a single or multiple sources. These parameters represent the level of service within 7x50 to which the host is entitled. Some of the parameters are mandatory for subscriber instantiation while others are optional. The following lists the parameter sources in the order of priority:

- LUDB
- RADIUS
- DHCP Server => DHCP server directly queries LUDB
- DHCP option processed on DHCP ACK that is indicated in subscriber identification policy.
- Extraction from the DHCP Ack via Python (IPv4 only)
- Defaults that are statically configured on the 7x50 node (SAP, msap-policy, capture-sap, and subscriber-identification-policy).

In most cases, the host IP address assignment process is controlled by the parameters returned via LUDB or RADIUS. As such, the IP address delegation is integral part of the host instantiation process and will consequently be described in the following sections.

## No Authentication

IPoE and PPPoE v4/v6 hosts on static SAPs can be instantiated without the need to access LUDB or RADIUS server. In this case, the default subscriber host parameters (sla-profile, sub-profile, subscriber-id) must be provisioned statically under the SAP. The IP address assignment is provided by internal or external DHCP server. The IP address selection on 7x50 based DHCP server is based on the gi-address while third party DHCP servers may provide additional means to select the IP address (*mac-address*, *circuit-id*, etc.).

A DHCP pool name cannot be provided by 7x50 DHCP relay agent, since the LUDB and/or RADIUS are not utilized.

This model does not support IP address delegation via DHCP Proxy function since there is no LUDB or RADIUS server available that can supply pre-configured IP address.

Host instantiation without LUDB or RADIUS access on dynamic VLANs (capture SAP and consequently mSAP) is not supported.

---

## LUDB Only Access

Subscriber-host authentication, identification and IP address assignment can be performed via LUDB without the need to access the RADIUS server.

The LUDB is normally configured under the group-interface>ppp/dhcp hierarchy and can provide subscriber-identification parameters as well as IP addressing parameters:

Pool names for DHCP relay function (IPv4, IPv6 IA-NA, IPv6 IA-PD)

Fixed IP addresses – IPv4, IPv6 IA-NA, IPv6 IA-PD and IPv6 SLAAC prefix.

In case of capture SAP, the LUDB name configured under the capture SAP must much the LUDB name under the group-interface>dhcp/ppp hierarchy. If the LUDB names do not mach, the subscriber-host instantiation will fail.

---

## LUDB Access via DHCPv4 Server

In case that the IPv4 addressing assignment is facilitated by the DHCPv4 relay and an internal DHCPv4 server, the DHCPv4 server itself can query the LUDB for IPv4 address information. LUDB can provide a v4 pool name and IPv4 DHCP options to the DHCPv4 server or it can instruct it to use the gi-address as the IPv4 address selection mechanism.

ESM strings can also be provided via LUDB queried by the DHCPv4 server.

If LUDB access via DHCPv4 server is provided in addition to other authentication means (another LUDB under the group-interface, or RADIUS server), the ESM strings from the LUDB under the grp-interface or from the RADIUS server will have priority over the ESM strings configured under the LUDB accessed by the DHCPv4 server. On the other hand, the IPv4 addressing information will have the highest priority from the LUDB accessed directly by the DHCPv4 server.

Accessing LUDB directly via DHCPv4 server should be used in rare and exceptional cases.

LUDB access under the group-interface, possibly complemented by the RADIUS server will provide necessary means for subscriber-host instantiation in majority of use cases.

## RADIUS Only Access

Similar to LUDB-only access, RADIUS server can provide all the necessary information for subscriber-host instantiation, including the IP addressing parameters (pool names or IP addresses/prefixes). Authentication-policy which defines the RADIUS access must be applied to the group-interface.

In case of capture SAP, the authentication policy must be applied under the capture SAP. This authentication policy name must match the authentication policy name that is configured under the group-interface. Otherwise, the host instantiation will fail.

## Consecutive Access to LUDB and RADIUS

LUDB and RADIUS access can be combined during subscriber-host instantiation phase.

Configuration wise, LUDB must be referenced under the **group-interface>dhcp/ppp/pppoe** hierarchy (and possibly under the capture SAP), while the authentication-policy is specified within the LUDB. In this fashion, LUDB access is followed by RADIUS access. The subscriber-host parameters retrieved from both sources are combined with LUDB parameters being prioritized over RADIUS parameters in case that both sources return the same parameters.

In case that LUDB and authentication policy are configured simultaneously under the group-interface (and possibly under the capture SAP), the RADIUS authentication policy will be evaluated and LUDB will be ignored.

## RADIUS Fallback

In case that RADIUS server is not accessible (non-responsive), the host instantiation phase can be:

Terminated in the case the there is no fallback action within authentication policy specified.

Continued within LUDB if the fallback action within the authentication-policy references LUDB.

Continued without any response from RADIUS. Subscriber-host will be instantiated if defaults parameters are statically configured or the instantiation will fail in case that the defaults are not available.

The fallback action takes effect once the preconfigured RADIUS timeout period expires.

RADIUS fallback is currently not supported for DHCPv6 hosts.

## Subscriber Services

Subscriber services enable an operational model to activate and deactivate subscriber functions from RADIUS through an Access-Accept or CoA message. Using the flexible RADIUS Python script interface, the operator defines the subscriber service functionality by populating a data structure using a parameter list received in a RADIUS Vendor Specific Attribute (VSA). The format and content of the parameter list VSA is defined by the operator. Each subscriber service instance can have a dedicated RADIUS accounting session; an accounting start/stop is sent when the subscriber service is activated/deactivated. Optionally, interim updates are sent with an interim update interval that can be specified per subscriber service instance. Accounting interim update and stop messages contain the subscriber service related statistics (time or volume-and-time).

Subscriber services can be activated on a dual-stack PPPoE session or a single stack IPv4 host. Subscriber service functionality is supported for subscriber QoS overrides: changing queues or policer parameters like rate or burst sizes and adapting root arbiter or subscriber aggregate rates. For example, an operator defines a service to boost the downstream rate using the parameters ("rate-limit":downstream-rate-in-mbps). A subscriber service is activated through a subscriber service activate VSA with value “rate-limit:20” that is received for a PPPoE session. This triggers the operator-defined RADIUS Python script to populate the subscriber-service data-structure variable that changes the subscriber aggregate downstream rate to 20 Mbps. Optionally an accounting start is sent for the subscriber service. Later, when a subscriber service deactivate VSA with the same value “rate-limit:20” is received for the same PPPoE session, the original subscriber aggregate downstream rate is restored, and optionally an accounting stop sent.

# Flexible Subscriber-Interface Addressing (Unnumbered Subscriber-Interfaces)

## Terminology

Subscriber host — A 7x50 representation of an external host requesting a service. Each such host is fully instantiated within the 7x50 for the purpose of providing traffic control and billing services (for example, QoS, filtering, antispoofing, accounitng). The external hosts may represent variety of devices such as regular PCs, STBs, residential gateways, CPEs, VoIP devices. In most cases, the external host will run a DHCPv4/v6 or PPPoEv4/v6 client. DHCP and PPPoE initiation messages from such clients will trigger host instantiation within 7x50. For this the subscriber host term can be interchangeablye used with a term DHCP client or PPPoE client.

## Flexible Subscriber-Interface Addressing for IPOE/PPPOE v4/v6 Subscribers

In certain wholesale/retail environments, the wholesale provider that own the 7x50-BNG does not know the IP addresses that the retailers will assign to their clients in advance. For this reason, wholesaler's 7x50-BNG must accept any IP address from retailers and consequently pass it to the client during subscriber-host initiation phase.

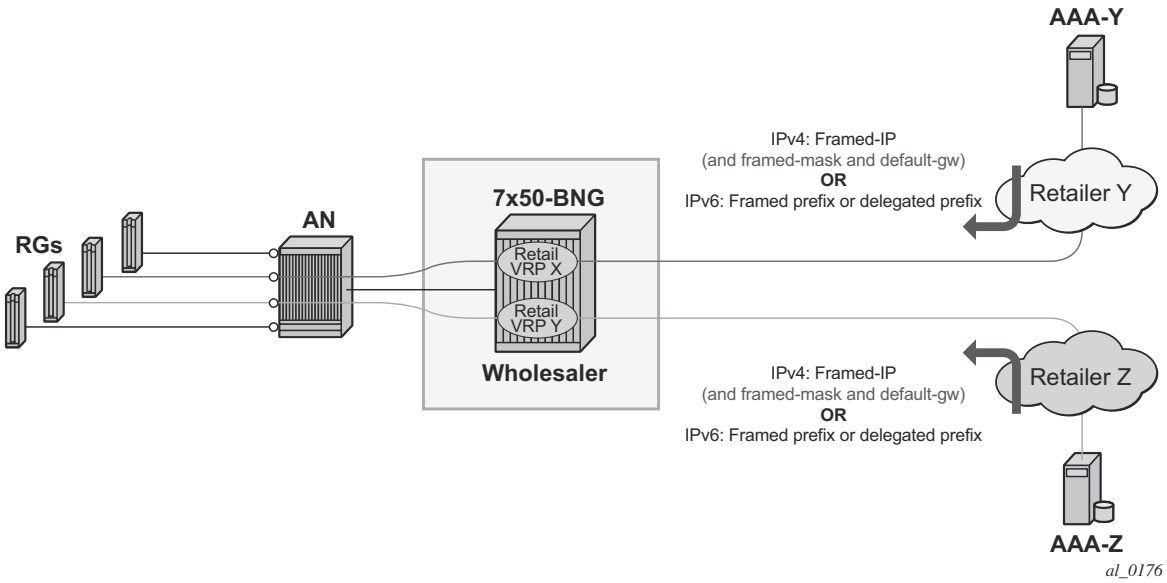



**Figure 140: Use Case for Flexible IP Addressing Model**

Flexible addressing of the subscriber-interface assumes two deployment scenarios:

1. Subscriber-interface is unnumbered — For example, there is no explicit assigned IP address. Instead the subscriber-interface borrows the IP address from an existing interface that is operationally UP and is located in the same routing instance (router | vprn)[2].

   In this case any IP address can be assigned to the subscriber host under the unnumbered subscriber-interface. The subscriber IPv4 address will be installed in the FIB as /32 route while IPv6 address will be installed as an entry of the length anywhere between 64 and 128 bits.

2. Subscriber-interface is numbered — The IP address/prefix is explicitly configured and solely owned by the subscriber-interface.

   In this case, all subscriber IP addresses/prefixes that fall under the subnet/prefix dictated by the configured subscriber-interface IP address/prefix will be directly aggregated under the subscriber-interface subnet. As such they will occupy a single entry in the FIB. The rest of the subscriber hosts with IP addresses/prefixes that fall outside of the configured range will be installed in the FIB as individual entries (/32 for IPv4 and an entry of the length anywhere between 64 and 128 bits for IPv6 hosts).

## Default Gateway in IPv4 Flexible Addressing

In scenarios where subscriber host IPv4 address lies within the configured subscriber-interface subnet, the default-gw IPv4 address for the host will be one of the subscriber-interface IPv4 addresses. In this case, the service provider is aware of the IPv4 addressing scheme in the 7x50-BNG and as such it will supply the DHCP client with the appropriate default-gw IPv4 address via LUDB, RADIUS or DHCP Server (in that order of priority).

In scenarios where the retail service provider wants to maintain independence from the IPv4 addressing scheme deployed in the 7x50-BNG (that is controlled by wholesaler), the retailer can always supply its own IPv4 address, the subnet mask and the default-gw IPv4 address. But if the default-gw IPv4 address and/or subnet mask is not supplied by the retailer, then they will be auto-generated by 7x50-BNG. Once the default-gw IPv4 address is auto-generated, it will be sent to the requesting DHCP client via DHCP offer in option 3 (RFC 2132, Router Option, section 3.5). There is no additional configuration needed for this action. 7x50-BNG will automatically detect whether the default-gw IPv4 address is supplied via LUDB, RADIUS or DHCP server and it will act correspondingly.

The default-gw IPv4 address is auto-generated based on the assigned IPv4 address/mask by setting the last bit of the assigned host IPv4 address to binary 01 or binary 10. For example if the subscriber host's assigned IPv4 address is 10.10.10.10 255.255.255.0, then the default-gw IPv4

2. Note that an interface must have an IP address assigned in order to be operationally UP. Therefore, an unnumbered subscriber-interface must reference another existing interface that is operationally UP in the same routing instance. The subscriber-interface will borrow the IP address from the referenced interface.

address is set to 10.10.10.1. If the assigned IPv4 address is 10.10.10.1 255.255.255.0 , then the auto-generated default gateway IPv4 is set to 10.10.10.2.

The default gateway IPv4 address will always have to be within the subscriber's subnet. If it is not, the behavior might be inconsistent. For example:

1. RADIUS (or DHCP) returns IP@, mask and def-gw:
   - → IP 10.10.10.1
   - → Subnet mask 255.255.255.0
   - → Def-gw 10.10.0.254

The subscriber will be successfully instantiated in 7x50-BNG but the client may not ARP for a default-gw outside of its configured subnet. Whether the client will or will not ARP for a default-gw outside of its configured subnet will depend on the implementation in the RG and CPE.

2. RADIUS returns IP@ and subnet mask.
   - → In this case the auto-generated default-gw IPv4 address will always be within subscriber's subnet.

Flexible IPv4 addressing with auto-generated default-gw is supported only in Routed Central Office (RCO) model with routed residential gateways (RGs) or CPEs. In RCO model with bridged residential gateways or CPEs, the default-gw IPv4 addresses and the assigned IPv4 addresses may overlap. Once the IPv4 address of the default-gw is auto-generated, it is possible that the second host behind the bridged residential gateway or CPE is assigned the same IPv4 address as the IPv4 address of the default gateway of the first host. Such hosts would not be able to communicate with outside world.

For example:

RADIUS or DHCP server assigns IPv4 address and subnet mask to the first host in a bridged environment:

IP1: 10.10.10.1

Auto-generated default-gw IPv4 address: 10.10.10.2

Since the RADIUS and DHCP Server are not aware of the auto-generated default-gw, they may assign the following IPv4 address to the second host that comes on-line:

IP 2: 10.10.10.2 (same IPv4 address as the default-gw IPv4 address of the first host)

Auto-generated default-gw IPv4 address: 10.10.10.1

Now the first host will forward all traffic outside of the configured subnet to the second hosts which will discard this traffic, effectively rendering this operation model non-deployable. And vice versa.

## IPv4 Subnet Sharing

Subnet sharing between the hosts in flexible IPv4 addressing model is supported. In other words, in flexible IPv4 addressing model the operator can assign all IPv4 addresses (minus one – the default-we IPv4 address) from a given subnet. In this fashion, all subscribers (routed RGs or CPEs) within a single subnet can share the same default gateway.

For example if the operator owns the IPv4 subnet 10.10.10.0/24, then one IPv4 address can be set aside for the default-gw (for example 10.10.10.254) and the remaining addresses can be assigned to the subscriber (routed RGs or CPEs). An example would be:

RG1: IP=10.10.10.1/24 def-gw 10.10.10.254

RG2: IP=10.10.10.2/24 def-gw 10.10.10.254

RG3: IP=10.10.10.3/24 def-gw 10.10.10.254

:

RG100: IP=10.10.10.100/24 def-gw 10.10.10.254

The subnet sharing is also supported in conjunction with auto-generated default-gw IPv4 address. The implication of this is that the IPv4 address of the default-gw can collide with the same IPv4 address already assigned to an existing subscriber. This is not an issue for routed RGs or CPEs since 7x50-BNG will always answers ARPs for the IPv4 address of the default-gw with its own (7x50) MAC address. However, local-proxy ARP functionality in 7x50-BNG MUST be enabled to support this.

This behavior can be further clarified with the following example.

Let’s assume that we have scenario with two routed RGs:

RG-1, IP=10.10.10.0/24, default-gw IP=10.10.10.1

RG-2, IP=10.10.10.1/24, default-gw IP=10.10.10.0

Once RG-1 ARPs for its default gateway of 10.10.10.1, 7x50-BNG will reply with its own MAC address.

Now that host RG-1 has resolved ARP for it default-gw (mac address pointing to 7x50), it can send traffic to the outside world via 7x50-BNG. When such traffic arrives to 7x50, the destination IPv4 address of the received packet will determine the forwarding decision within 7x50. If the destination IPv4 address matches the IPv4 address of any subscriber (RG) instantiated within 7x50, the traffic will be forwarded to the that RG. This also includes the case where the destination IPv4 address is the default-gw IPv4 address (10.10.10.1), which represents just another RG within 7x50. The traffic will be consequently passed from RG-1 via 7x50 to RG-2.

# IPv4 Subnet Mask Auto-Generation

The subnet mask corresponding to the IPv4 address assigned to the subscriber is auto-generated in case that the IPv4 addressing authority (LUDB, RADIUS or DHCP Server) does not supply it. The subnet mask is derived from the IPv4 address of the subscriber and possibly the default-gw IPv4 address and it is the smallest subnet that contains both, the IPv4 address of the subscriber and the default-gw.

For example if the RADIUS received IPv4 address is 10.10.10.138 and the received default –gw IPv4 address is 10.10.10.170 , then the subnet mask will be auto-generated and set to 255.255.255.192 ( /26).

138 = 10001010

170 = 10101010

192 = 11000000

In case that neither the subnet mask nor the default-gw are returned, then both would be auto-generated:

1. Subnet mask would be set to /31
2. Default-gateway which must belong to the subscriber’s subnet would be set to 10.10.10.139.

In cases where the host IPv4 address and the default-gw are directly supplied by the addressing authority but the subnet mask is missing, the subnet mask auto-generation may cause the host part of the default-gw IPv4 address to become a broadcast IPv4 address. If this is an issue, then it can be avoided by directly providing the subnet mask via the addressing authority.

## Local-proxy-arp and arp-populate

Local-proxy-arp and arp-populate are two commands that are relevant only to IPoEv4 hosts.

Local-proxy-arp command ensures that 7x50 answers ARP Requests with its own MAC address for any ‘active’ IPv4 address under the subnet on which the ARP request arrived. The ‘active’ IPv4 address is considered the one that is assigned to an already instantiated hosts or the default-gw (even auto-generated).

In absence of local-proxy-arp command, the only ARP Request that 7x50 will answer is the one for the statically configured IPv4 addresses of the subscriber-interface. In flexible IPv4 addressing, the IPv4 address of the default-gw does not necessarily match any of the configured subscriber-interface IPv4 addresses. The ARP Request for such default-gw IPv4 address would go unanswered. Consequently, the subscriber hosts would not be able to communicate with outside world. Therefore, the flexible IPv4 addressing requires that the local-proxy-arp command is configured.

Arp-populate command disables dynamic learning of ARP entries (IPv4<->MAC mapping) on an interface based on the ARP protocol. In this case, the ARP table is populated based on the DHCPv4 lease state table which contains IPv4<->MAC mappings obtained via DHCP processing during the host instantiation phase. Arp-populate functionality is highly desirable in case of flexible IPv4 addressing.

When arp-populate command is disabled the ARP entries are dynamically learned based on the ARP protocol. This, in conjunction with flexible IPv4 addressing may cause certain issues. Consider the following example:

- The subscriber-host is instantiated in 7x50
- The subscriber-interface is unnumbered
- The ARP table does not contain an ARP entry for the subscriber-host

In this case, downstream traffic towards the subscriber host will trigger 7x50 to send ARP Request for the subscriber host IPv4 address. 7x50 needs to know the MAC address of the subscriber-host in order to forward traffic. Since the subscriber-interface is unnumbered , the source IPv4 address of the ARP request is unknown and consequently the ARP Request will not be sent. As a result, downstream traffic will be dropped.

Note however that the above example is an unlikely scenario. If the subscriber host sends the ARP Request for the default-gw first, 7x50 would create an entry in the ARP table for it and the issue would be resolved. This is the most likely outcome since the subscriber host will always try to initiate communication with the outside and therefor ARP for the IPv4 address of the default-gw (which is 7x50).

## Gi-address Configuration Consideration

With flexible IPv4 address assignment, the gi-address can be configured as any IPv4 address that is already assigned to an interface (loopback interface, regular interface attached to physical port or subscriber interface) within the same routing instance (VRF or GRT).

## PPPoE Considerations

PPPoE subscriber hosts do not have the concept of default-gw. Consequently the default-gw auto-generation concept does not apply to PPPoE hosts.

## IPoEv6 Considerations

The default-gw for IPoEv6 hosts is link-local IPv6 address. Since this address is always present, there is no need for auto-generation during the subscriber instantiation time.

SLAAC hosts are installed as /64 entries, the length of the installed DHCP-PD prefix is dictated by the prefix-length and the DHCP-NA hosts are installed as /128 entries.

## General Configuration Guidelines for Flexible IP Address Assignment

Flexible IP addressing for IPoE/PPPoE v4 and v6 hosts is by default disabled. In other words, the subscriber hosts will be instantiated in 7x50-BNG with ability to forward traffic only if their assigned IP addresses belong to one of the configured subnets/prefixes that are associated with the subscriber-interfaces. IPv4 and IPv6 cases will be examined separately:

**IPv4:**

By default, IPoE and PPPo subscriber host creation will fail in the following two cases:

1. The subscriber-interface does not have an IPv4 address configured, and therefore it will be operationally down. This configuration is also known as unnumbered subscriber-interface.
2. The subscriber-interface does have an IPv4 address configured but the IPv4 address assigned to the subscriber host itself is outside of the subscriber-interface configured subnet(s). In such case the host will be instantiated but the forwarding will be disabled.

Subscriber host instantiation and forwarding can be explicitly enabled for both cases above with flexible IP addressing functionality.

For case 1, this can be achieved by borrowing an IP address for the subscriber-interface from any interface that is operationally up within the given routing context. This functionality can be enabled with the following command:

```
configure service ies <id>
configure service vprn <id>
subscriber-interface <intf-name>
     unnumbered <ip-addr | interface-name >
```

To enable forwarding for the subscribers whose IP address falls outside of the configured subnet under the subscriber-interface (case 2), the following command must be entered:

```
configure service ies <id>
configure service vprn <id>
subscriber-interface <intf-name>
     allow-unmatching-subnets
```

The above commands (**unnumbered** and **allow-unmatching-subnets**) are mutually exclusive. In addition, the unnumbered command can be configured only if the subscriber-interface does not have an IP address already configured. Otherwise the execution of this command will fail.

In both of these cases the host will be installed in the routing table as /32.

**IPv6:**

For IPv6 there is a single command that will enable flexible IP addressing for both cases:

1. IPv6 prefixes are not configured under the **subscriber-interface>ipv6** node
2. IPv6 prefixes are configured but the actual address or prefix assigned to the subscriber (via DHCP, LUDB or RADIUS) is outside any prefix that is configured under the **subscriber-interface>ipv6** hierarchy.

This single command is:

```
configure service ies <id>
configure service vprn <id>
         subscriber-interface <name>
              ipv6
                   allow-unmatching-prefixes
```

To summarize, the following scenarios are possible:

- PPPoEv4
  - → An IPv4 address under the subscriber-interface is configured
    - By default hosts outside of the sub-intf subnet are instantiated but they are in a non-forwarding-state. Traffic is dropped.
    - **allow-unmatching-subnets** is configured. This command is allowed only if subscriber-interface has also configured its own IPv4 address(es). In this case the IP address for IPCP negotiation is one of the sub-intf addresses. Hosts outside of the sub-intf subnets are instantiated and forwarded.

- The **unnumbered** <*ip-address* | *intf*> command is not allowed in this scenario.

→ An IPv4 address under the subscriber-interface is not configured

- By default, the subscriber-interface is operationally down. Subscribers cannot be instantiated.
- The **allow-unmatching-subnets** command has no effect since subscriber-interface does not have an IPv4 address configured and is therefore operationally down. No subscribers can be instantiated.
- The **unnumbered** <*ip-address* | *intf*> command is the only viable option in this case. The subscriber-interface borrows an IPv4 address from another interface that is operationally UP and consequently this allows subscribers to be instantiated. This command is mutually exclusive with **allow-unmatching-subnets**. In addition, this command can only be configured if the subscriber-interface itself does not have explicitly configured an IPv4 address.

- IPoEv4

  Similar to the PPPoE case above.

- IPoEv6 and PPPoEv6 — the **allow-unmatching-prefixes** command is independent of any IPv4 command related to flexible IP address assignment (**unnumbered** or **allow-unmatching-subnets**). This command can always be enabled, regardless of the v6 prefixes configured under the **subscriber-interface>ipv6** hierarchy. Any subscriber, regardless of the subscriber-interface prefix configuration will be instantiated and forwarded.

---

## Caveats

- Auto-generation of the default-gw IPv4 address is supported only in RCO model with routed RGs/CPEs. Bridged RGs/CPEs are not supported.
- Dual homing redundancy is not supported for unnumbered subscriber-interfaces or allow-unmatching-subnets/prefixes.
- A configured IPv4 address cannot be removed from the subscriber-interface when DHCPv4 hosts under the corresponding subnet are instantiated in the system.
- An IPv4 address cannot be configured under the subscriber-interface while (unnumbered) DHCPv4 hosts under that subnet are already instantiated.
- Executing the **no allow-unmatching-subnets** command is only allowed when there are no unnumbered DHCPv4 hosts instantiated under the subscriber-interface.

# uRPF for Subscriber Management

uRPF is supported for IPv4 and IPv6 dual-stack subscribers with framed routes.

For IPv4, uRPF is supported on group interfaces using anti-spoofing filters. A group interface configured for NATed subscribers will be configured with MAC/IP/PPPoE Session-ID anti-spoofing filters.

IPv6 subscribers, which are non-NAT, are always treated as being on a local subnet. For such subscribers, a 7x50 BNG will install a FIB entry for local routes that match either the wan-host prefix, or the delegated prefix, or both. In strict mode for IPv6 ESM, the uRPF check will check not just that the route matching the SA (which should be a local route, i.e. a subnet) would route the packet back out of the interface it came in on, but in addition that we would route the packet out to the same SAP it was received on.

SROS supports the ability to configure a NH-MAC anti-spoof type for non-NATed subscribers. When configured, the datapath performs ingress anti-spoofing based on source MAC address and egress anti-spoof (also referred to as egress subscriber-host look-up) based on the nh-ip address.

The NH-MAC anti-spoof type is configured under the following context:

**config>service>vprn>if>sap**

**config>service>ies>sub-if>grp-if>sap**

**config>service>vprn>sub-if>grp-if>sap**

**config>subscr-mgmt>msap-policy**

A uRPF check is also performed that prefixes delegated to a subscriber on that MAC address exist in the FIB.