7450 Ethernet Service Switch
7750 Service Router
Virtualized Service Router

Releases up to 24.10.R2

Multiservice ISA and ESA Advanced Configuration Guide for Classic CLI

3HE 20798 AAAB TQZZA
Edition: 01
March 2025

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

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

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

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

No part of this document may be modified.

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

Table of contents

List of tables

List of figures

Preface

About This Guide

Each Advanced Configuration Guide is organized alphabetically and provides feature and configuration explanations, CLI descriptions, and overall solutions. The Advanced Configuration Guide chapters are written for and based on several Releases, up to 24.10.R2. The Applicability section in each chapter specifies on which release the configuration is based.

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

Audience

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

Application Assurance — Application Identification and User-Defined Applications

This chapter describes Application Assurance (AA) Application Identification and User-Defined Applications configurations.

Topics in this chapter include:

Applicability

The information and configuration in this chapter were initially based on SR OS Release 11.0.R3, but updates were made based on SR OS Release 19.10.R2.

There are no specific prerequisites for this example.

Overview

This chapter is intended for Application Assurance (AA) network architects and engineers. It provides best practice information to customize the AA policy and classify any type traffic to meet the service provider reporting, charging or control requirements.

In addition to the signatures built and supported by Nokia, service providers can create their own application signatures based on various criteria. This customization capability can be used to classify traffic hosted on the provider network (web portal, streaming service) or hosted on the Internet and not yet covered by the default AA signature set.

Basics and terminology

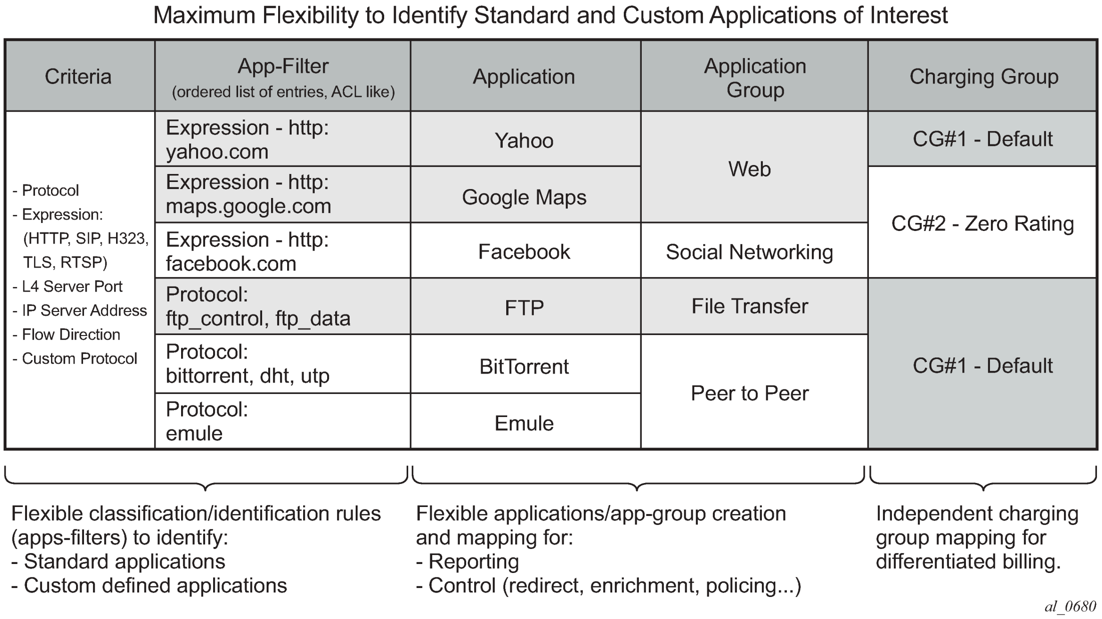
The following main components are used for AA classification:

- **Application Filters** — App-filters are used to define applications based on Layer 3 to Layer 7 criteria. They provide a mapping between one or more protocol signatures or customized traffic patterns into an application of interest.
- **Application** — Such as BitTorrent®, Netflix®. Traffic is classified into applications using app-filters.
- **Application Group** — Such as peer-to-peer, multimedia streaming. For the purpose of reporting and control, applications of similar type/function can be grouped together in Application Groups (App-Group).

- **Charging Group** — Such as zero rating, default. For the purpose of charging or control, applications and app-group can be grouped together in charging groups.

The following table is a high-level example to illustrate how app-filters are used to defined applications and show their logical grouping into app-group and charging group.

Figure 1: App-Filters/Applications/AppGroup



Maximum Flexibility to Identify Standard and Custom Applications of Interest

Criteria	App-Filter (ordered list of entries, ACL like)	Application	Application Group	Charging Group
- Protocol - Expression: (HTTP, SIP, H323, TLS, RTSP) - L4 Server Port - IP Server Address - Flow Direction - Custom Protocol	Expression - http: yahoo.com	Yahoo	Web	CG#1 - Default
	Expression - http: maps.google.com	Google Maps		CG#2 - Zero Rating
	Expression - http: facebook.com	Facebook	Social Networking	
	Protocol: ftp_control, ftp_data	FTP	File Transfer	CG#1 - Default
	Protocol: bittorrent, dht, utp	BitTorrent	Peer to Peer	
	Protocol: emule	Emule		

Flexible classification/identification rules (apps-filters) to identify:
- Standard applications
- Custom defined applications

Flexible applications/app-group creation and mapping for:
- Reporting
- Control (redirect, enrichment, policing...)

Independent charging group mapping for differentiated billing.

al_0680

- BitTorrent® and Emule® applications are defined using their protocol signature and grouped in the P2P app-group.
- FTP application is defined using both ftp_data and ftp_control protocol signatures, the app is mapped in the file transfer app-group.
- Google Maps® and Yahoo® web sites are defined using http expression and grouped together in the Web app-group.

Configuration

Classification criteria (App-filter)

The operator can take full advantage of the flexible AA policy configuration to classify traffic from any application of interest using various criteria ranging from Layer 3 to Layer 7 expressions.

Expression match criteria allows to further refine traffic classification by identifying traffic from HTTP, HTTPS (SSL/TLS), SIP, H323, RTSP, Citrix protocol signatures.

The different app-filter match criteria are listed below:

- L7 expression
  - HTTP: host, URI, user agent, referer

  - SSL/TLS: certificate org name, common name, SNI
  - H323: product ID
  - SIP: URI, user agent, media type
  - RTSP: host, URI, user agent
  - Citrix: application published name
  - RTMP: page-host, page-uri, swf-host, swf-uri
- IP protocol number
- IP server address
- TCP/UDP server port
- Custom protocol
- Protocol signature

The following operators are supported to define expression-based app-filters:

^ :	Expression starts with
$:	Expression ends with
*:	Wildcard - anything before or after
\I:	Forces case sensitivity
\d:	Any single decimal digit [0-9]
\.:	Any single character
*:	Asterisk character

Examples of expression match combinations:

^abcd*:	match ‘abcd’ at beginning, can end with anything
abcd:	match ‘abcd’ anywhere
*abcd$:	match ‘abcd’ at the end
^abcd$:	exact expression match ‘abcd’
^ab*cd$:	string starts with 'ab', ends with 'cd' (anything else in between)
^ab\dcd$:	string starts with 'ab', followed by a decimal digit, ends with 'cd'

Note:
It is possible to combine different criteria or expressions within the same filter in which case an implicit AND operation between the criteria within the same filter is done by the system.

Application definition example

The following example provides a basic configuration example with the application FTP made of two protocol signatures ftp_control and ftp_data; the application is mapped into the application group file transfer:

Create the application group.

```
configure application-assurance group 1:1 policy
    app-group "File Transfer"
    exit
```

Create the application.

```
configure application-assurance group 1:1 policy
    application "FTP"
        app-group "File Transfer"
    exit
```

Create the app-filters.

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            protocol eq "ftp_data"
            application "FTP"
            no shutdown
        exit
        entry <1..65535> create
            protocol eq "ftp_control"
            application "FTP"
            no shutdown
    exit
```

Note:
Once the application is created, the operator is expected to configure the collection of statistics at the subscriber level for this new application (usually only for business VPNs).

User-defined applications

General recommendations

In order to classify traffic properly, it is recommended to follow the guidelines and best practices defined in this section before creating a new application:

- Analyze the application traffic
  - Identify what type traffic is used (Wireshark®).
  - Use the application the same way the end user would use it, the same application can create various flows.
- Configure the appropriate App-Filters
  - Following the analysis of the application done above, create the application.

- Follow the App-filter best practices chapter to understand in which range to add the filters.
- More than one App-filter can be required to identify a single application.

AppDB/default AA policy

The default AA policy called AppDB (Application Database) is provided by Nokia and should be used on most deployments. Contact your regional support organization for more details on how to obtain it.

This configuration includes applications and application-groups most providers can use by default and is designed to allow the addition of any custom entries required by service providers to identify additional services/applications.

The following customization options exist:

- Generating a configuration for a particular region (such as APAC)
- Generating a new configuration (or updating a configuration) containing specific applications

Before adding new entries to the template and customizing the configuration, it is recommended to follow the next guidelines on app-filters and ranges. These guidelines are key to allow an easy upgrade path from the policy configuration provided by Nokia.

App-filters

App-filters are an ordered list of entries. It is important to keep the order of this list consistent with the classification objective.

For instance, a common configuration mistake is to configure a filter rule for the HTTP protocol signature before HTTP expression filters. If that were the case, then app-filters using HTTP expressions would not be used as the system would find an acceptable match with the protocol signature before walking the list of expressions configured. This mistake is described in the following example:

```
entry 100 create
  description "Default HTTP Protocol"
  protocol eq "http"
application "HTTP"
  no shutdown
exit
entry 110 create
  description "Google"
  expression 1 http-host eq "*.google.com$"
  application "Google"
  no shutdown
exit
```

This is an incorrect App-filter order. App-filter entry #100 will always match before the http expression entry #110.

Note:
It is not necessary to specify a protocol when defining an expression filter, the protocol is implicit based on the type of expression match criteria used (for instance, HTTP, SIP, H323).

App-filter ranges

The App-filter list is an ordered list, it is key to configure each app-filter in the right order and in the proper range.

The operator can customize the policy and create applications and app-filters by using the following ranges shown in Table 1: Customer reserved App-filter ranges (other ranges are used by the default policy):

Table 1: Customer reserved App-filter ranges

Range Name	Description	Start	End
Extended top range	Top range, matches before any other filters	1	1499
High priority	Top range for high priority matches	2000	4999
Expression range A	HTTP Host, Host+URI; optionally with IP/Port match	19000	22999
Expression range B	Other Expression Match; optionally with IP/ Port match	33000	34999
Extended protocols	Protocol-signature + Port|IP|Dir. match	40000	58999
Custom protocols	Custom protocol signature match	61000	61999
Trusted/validate ports	1st packet validate, 1st packet trusted match	62000	63999

Ordering basics:

- Layer 7 expression-based filters are located before their parent protocol signature (for example, expression matches on http are located before the HTTP protocol app-filter; the same applies to TLS, SIP, H323, RTSP, Citrix).
- HTTP host and URI are located before the HTTP referer for accounting accuracy (for example, YouTube® from within Facebook® is classified as YouTube®)
- App-filters combining protocol signatures with Layer 4 port, IP protocol, IP address, or flow direction are always located before the protocol signature only filter range.

HTTP

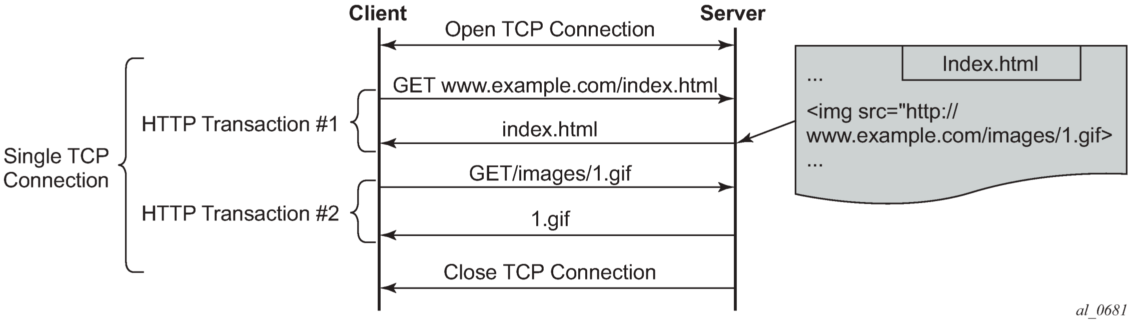
Protocol

HTTP is a client/server protocol using TCP/IP at the transport layer to deliver resources such as HTML files, images, videos and more.

HTTP 1.1 enables HTTP clients to use a persistent connection to a server allowing them to reuse the same TCP session for multiple HTTP transactions. Text, images, video, scripts and other objects can be downloaded individually in different transactions through the same TCP session.

Figure 2: HTTP persistent connection describes a typical persistent HTTP connection between a web client and a server with multiple HTTP transactions within the same TCP session:

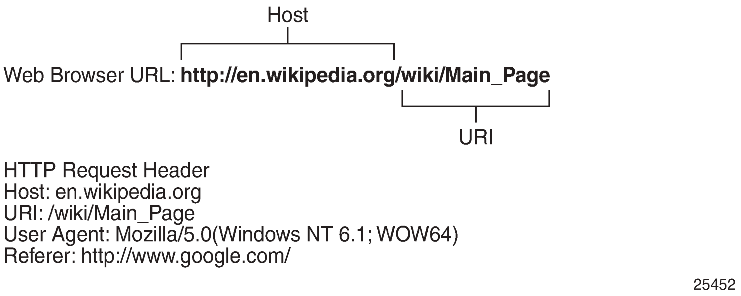
Figure 2: HTTP persistent connection



User-defined expression-based HTTP applications will use the first HTTP transaction to classify the flow (optionally this behavior can be modified).

HTTP request

The following example shows the content of a typical HTTP request to wikipedia.org which includes the following header fields: HTTP Host, HTTP URI, HTTP User Agent and HTTP referer fields:



Web Browser URL: **http://en.wikipedia.org/wiki/Main_Page**

HTTP Request Header
Host: en.wikipedia.org
URI: /wiki/Main_Page
User Agent: Mozilla/5.0(Windows NT 6.1; WOW64)
Referer: http://www.google.com/

- HTTP Host — Represents the domain name (does not include "http://").
- HTTP URI — The URL trailer after the host domain name (begins with slash "/").
- HTTP Referer — The address of the previous web page from which a link to the currently requested page was followed (in this example, the referer is www.google.com which means the user clicked on a link from a Google search pointing to wikipedia.org).
- HTTP User Agent — This identifies the web browser or application making the HTTP request.

Configuration examples

HTTP host (Wikipedia)

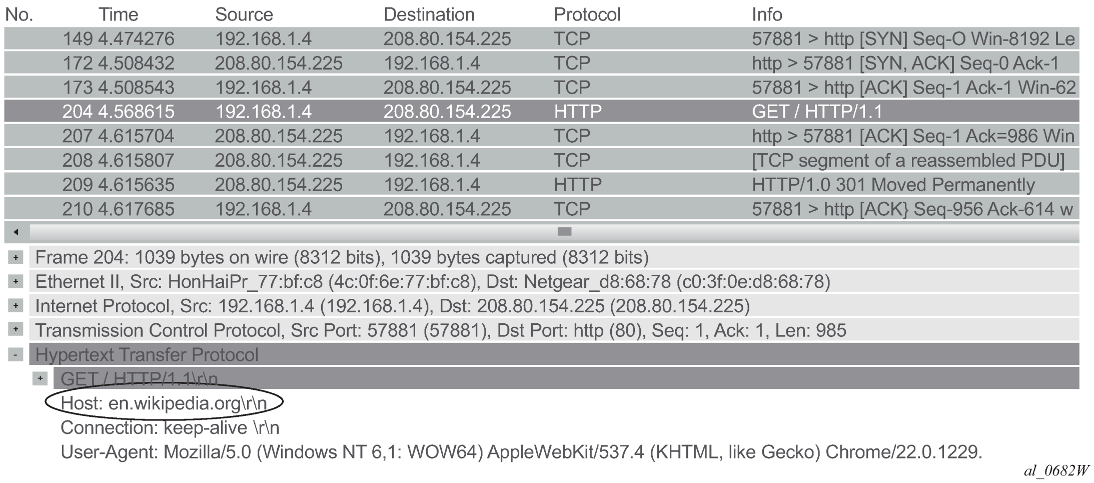
Classifying HTTP traffic from this web site can be done using a single expression tail anchored on the HTTP host:

```
configure application-assurance group 1:1 policy app-filter
    entry <1..65535> create
```

```
        description "Wikipedia Web Access" expression 1 http-host eq "*.wikipedia.org$"
        application "Wikipedia"
        no shutdown
    exit
```

This can be confirmed using Wireshark®.

Figure 3: Wireshark® www.wikipedia.org



```
No.      Time        Source            Destination       Protocol    Info
     149 4.474276    192.168.1.4       208.80.154.225    TCP         57881 > http [SYN] Seq-O Win-8192 Le
     172 4.508432    208.80.154.225    192.168.1.4       TCP         http > 57881 [SYN, ACK] Seq-0 Ack-1
     173 4.508543    192.168.1.4       208.80.154.225    TCP         57881 > http [ACK] Seq-1 Ack-1 Win-62
     204 4.568615    192.168.1.4       208.80.154.225    HTTP        GET / HTTP/1.1
     207 4.615704    208.80.154.225    192.168.1.4       TCP         http > 57881 [ACK] Seq-1 Ack=986 Win
     208 4.615807    208.80.154.225    192.168.1.4       TCP         [TCP segment of a reassembled PDU]
     209 4.615635    208.80.154.225    192.168.1.4       HTTP        HTTP/1.0 301 Moved Permanently
     210 4.617685    192.168.1.4       208.80.154.225    TCP         57881 > http [ACK} Seq-956 Ack-614 w

+ Frame 204: 1039 bytes on wire (8312 bits), 1039 bytes captured (8312 bits)
+ Ethernet II, Src: HonHaiPr_77:bf:c8 (4c:0f:6e:77:bf:c8), Dst: Netgear_d8:68:78 (c0:3f:0e:d8:68:78)
+ Internet Protocol, Src: 192.168.1.4 (192.168.1.4), Dst: 208.80.154.225 (208.80.154.225)
+ Transmission Control Protocol, Src Port: 57881 (57881), Dst Port: http (80), Seq: 1, Ack: 1, Len: 985
- Hypertext Transfer Protocol
    + GET / HTTP/1.1\r\n
      Host: en.wikipedia.org\r\n
      Connection: keep-alive \r\n
      User-Agent: Mozilla/5.0 (Windows NT 6,1: WOW64) AppleWebKit/537.4 (KHTML, like Gecko) Chrome/22.0.1229.
```

Classification per URI within the same host

Operators may need to apply different charging rules to different content located on the same HTTP domain (different URI, same HOST).

Table 2: Classification rules for the ISP ON-NET content services displays an example of classification rules for the ISP ON-NET content services:

Table 2: Classification rules for the ISP ON-NET content services

URL	Charging rule	AA application
www.ispdomain.com/video	Rule #1 – 0 Rating	ISP-Portal-Video
www.ispdomain.com/images	Rule #2 – Charge X	ISP-Portal-Images
www.ispdomain.com/*	Rule #3 – Charge Y	ISP-Portal-Default

HTTP 1.1 can reuse the same TCP connection for many transactions to the same server. Classifying each HTTP transaction to www.ispdomain.com independently requires a specific AA configuration.

SR OS allows to selectively enable “http-match-all-requests” in app-filters to improve the system performance and limit the HTTP analysis per domain.

The following configuration example allows traffic classification of different URIs of the same domain (www.ispdomain.com) independently, therefore allowing differentiated charging and control:

- http-match-all-req is enabled on all host+uri app-filters to www.ispdomain.com
- default app-filter required to match any traffic to www.ispdomain.com

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Zero rated content"
            expression 1 http-host eq "^www.ispdomain.com$"
            expression 2 http-uri eq "^/video*"
            http-match-all-req
            application "ISP Portal Video"
            no shutdown
        exit
        entry <1..65535> create
            description "Image charging"
            expression 1 http-host eq "^www.ispdomain.com$"
            expression 2 http-uri eq "^/images*"
            http-match-all-req
            application "ISP Portal Images"
            no shutdown
        exit
        entry <1..65535> create
            description "Default charging"
            expression 1 http-host eq "^www.ispdomain.com$"
            http-match-all-req
            application "ISP Portal Default"
            no shutdown
        exit
```

SSL/TLS (HTTPs)

Protocol

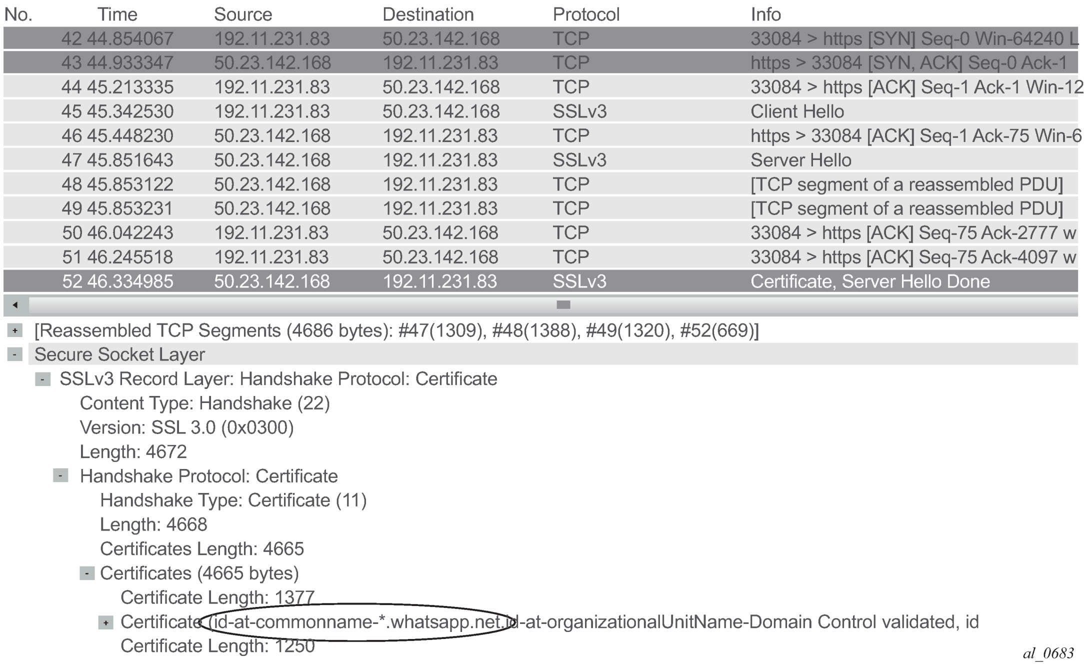
HTTPS uses SSL/TLS to encrypt traffic between the client and the server. Because this communication is encrypted, it is not possible to identify the HTTP Host or URI. However, AA can still identify the service requested by the subscriber by looking at the TLS certificate information or Server Name Indication exchanged in the clear before the TLS session is established.

Note:
SSL/TLS expression-based app-filters are not limited to HTTPS. HTTPS is not a protocol in itself, but it is HTTP traffic, tunneled encrypted into SSL/TLS on port 443.

SSL/TLS certificates

The following snapshot (Figure 4: Wireshark® HTTPS www.whatsapp.com) from Wireshark shows the SSL/TLS certificate exchanged using the mobile application **whatsapp**®.

Figure 4: Wireshark® HTTPS www.whatsapp.com



```
No.       Time          Source            Destination       Protocol    Info
      42 44.854067     192.11.231.83     50.23.142.168     TCP         33084 > https [SYN] Seq-0 Win-64240 L
      43 44.933347     50.23.142.168     192.11.231.83     TCP         https > 33084 [SYN, ACK] Seq-0 Ack-1
      44 45.213335     192.11.231.83     50.23.142.168     TCP         33084 > https [ACK] Seq-1 Ack-1 Win-12
      45 45.342530     192.11.231.83     50.23.142.168     SSLv3       Client Hello
      46 45.448230     50.23.142.168     192.11.231.83     TCP         https > 33084 [ACK] Seq-1 Ack-75 Win-6
      47 45.851643     50.23.142.168     192.11.231.83     SSLv3       Server Hello
      48 45.853122     50.23.142.168     192.11.231.83     TCP         [TCP segment of a reassembled PDU]
      49 45.853231     50.23.142.168     192.11.231.83     TCP         [TCP segment of a reassembled PDU]
      50 46.042243     192.11.231.83     50.23.142.168     TCP         33084 > https [ACK] Seq-75 Ack-2777 w
      51 46.245518     192.11.231.83     50.23.142.168     TCP         33084 > https [ACK] Seq-75 Ack-4097 w
      52 46.334985     50.23.142.168     192.11.231.83     SSLv3       Certificate, Server Hello Done

+ [Reassembled TCP Segments (4686 bytes): #47(1309), #48(1388), #49(1320), #52(669)]
- Secure Socket Layer
    - SSLv3 Record Layer: Handshake Protocol: Certificate
          Content Type: Handshake (22)
          Version: SSL 3.0 (0x0300)
          Length: 4672
        - Handshake Protocol: Certificate
              Handshake Type: Certificate (11)
              Length: 4668
              Certificates Length: 4665
            - Certificates (4665 bytes)
                  Certificate Length: 1377
                + Certificate (id-at-commonname-*.whatsapp.net,id-at-organizationalUnitName-Domain Control validated, id
                  Certificate Length: 1250
```

The certificate information can be found in the Server Hello message sent by the server, capturing SSL/ TLS (HTTPS) traffic from this application can be done using a single app-filter entry tail anchored on the TLS Common Name Certificate:

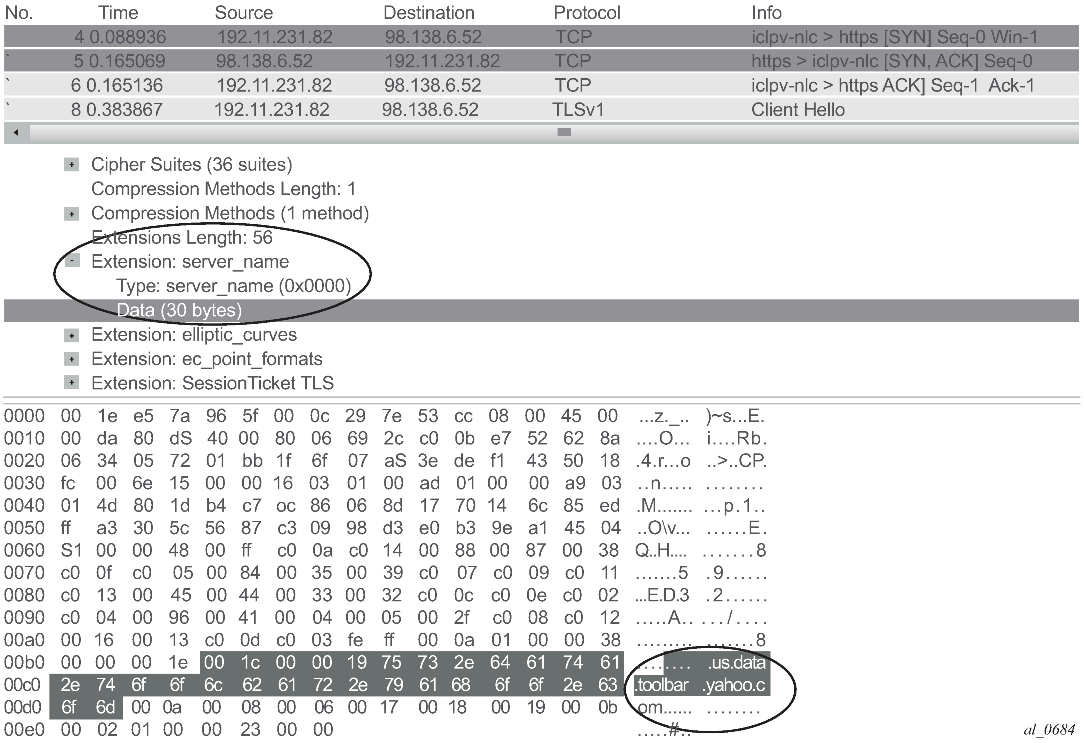
```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Whats App tls and image/voice/video traffic"
            expression 1 tls-cert-subj-common-name eq
                         "*.whatsapp.net$"
            application "Whats App"
            no shutdown
          exit
```

Server name indication

SSL/TLS traffic can optionally be identified using the Server Name Indication (SNI) which is an extension to the TLS protocol.

The SNI is found in the TLS Client Hello, the http-host expression in the app-filter is reused to classify this traffic:

Figure 5: HTTPS SNI



```
No.        Time         Source           Destination      Protocol         Info
         4 0.088936     192.11.231.82    98.138.6.52      TCP              iclpv-nlc > https [SYN] Seq-0 Win-1
`        5 0.165069     98.138.6.52      192.11.231.82    TCP              https > iclpv-nlc [SYN, ACK] Seq-0
`        6 0.165136     192.11.231.82    98.138.6.52      TCP              iclpv-nlc > https ACK] Seq-1  Ack-1
`        8 0.383867     192.11.231.82    98.138.6.52      TLSv1            Client Hello
```

```
    + Cipher Suites (36 suites)
      Compression Methods Length: 1
    + Compression Methods (1 method)
      Extensions Length: 56
    - Extension: server_name
          Type: server_name (0x0000)
          Data (30 bytes)
    + Extension: elliptic_curves
    + Extension: ec_point_formats
    + Extension: SessionTicket TLS
```

```
0000  00 1e e5 7a 96 5f 00 0c 29 7e 53 cc 08 00 45 00   ...z._..  )~s...E.
0010  00 da 80 dS 40 00 80 06 69 2c c0 0b e7 52 62 8a   ....O...  i....Rb.
0020  06 34 05 72 01 bb 1f 6f 07 aS 3e de f1 43 50 18   .4.r...o  ..>..CP.
0030  fc 00 6e 15 00 00 16 03 01 00 ad 01 00 00 a9 03   ..n.....  ........
0040  01 4d 80 1d b4 c7 oc 86 06 8d 17 70 14 6c 85 ed   .M......  ...p.1..
0050  ff a3 30 5c 56 87 c3 09 98 d3 e0 b3 9e a1 45 04   ..O\v...  ......E.
0060  S1 00 00 48 00 ff c0 0a c0 14 00 88 00 87 00 38   Q..H....  .......8
0070  c0 0f c0 05 00 84 00 35 00 39 c0 07 c0 09 c0 11   .......5  .9......
0080  c0 13 00 45 00 44 00 33 00 32 c0 0c c0 0e c0 02   ...E.D.3  .2......
0090  c0 04 00 96 00 41 00 04 00 05 00 2f c0 08 c0 12   .....A..  .../....
00a0  00 16 00 13 c0 0d c0 03 fe ff 00 0a 01 00 00 38   ........  .......8
00b0  00 00 00 1e 00 1c 00 00 19 75 73 2e 64 61 74 61   ........  .us.data
00c0  2e 74 6f 6f 6c 62 61 72 2e 79 61 68 6f 6f 2e 63   .toolbar  .yahoo.c
00d0  6f 6d 00 0a 00 08 00 06 00 17 00 18 00 19 00 0b   om......  ........
00e0  00 02 01 00 00 23 00 00                           .....#..
```

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Yahoo HTTP or TLS SNI"
            expression 1 http-host eq "*.yahoo.com$"
            application "Yahoo"
            no shutdown
          exit
```

SIP

Protocol

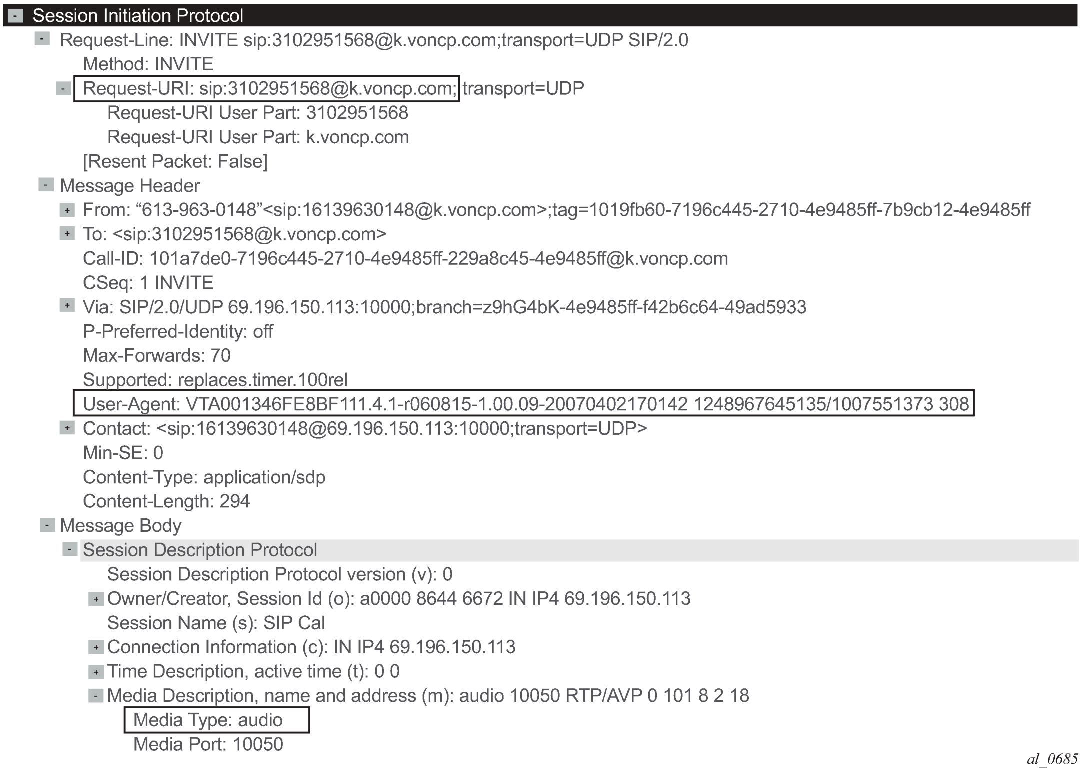
SIP is a signaling protocol used for controlling multimedia communication sessions such as voice and video over RTP. AA automatically monitors SIP control flows and associates RTP/RTCP media flows accordingly in the sip_rtp protocol signature.

The operator can use a SIP expression match criteria in app-filter to further refine traffic classification and identify any additional application on top of the default AA policy. This can be particularly useful in business VPNs to identify voice and telepresence applications.

AA supports SIP expression match criteria on SIP URI, SIP user agent and SIP media type. The following snapshot from Wireshark® shows a SIP control exchange using the voice-video application Vonage®

followed by the RTP media audio flow; the expression fields that can be matched using AA app-filters are highlighted:

Figure 6: SIP Wireshark® capture



```
- Session Initiation Protocol
  - Request-Line: INVITE sip:3102951568@k.voncp.com;transport=UDP SIP/2.0
      Method: INVITE
    - Request-URI: sip:3102951568@k.voncp.com;transport=UDP
        Request-URI User Part: 3102951568
        Request-URI User Part: k.voncp.com
      [Resent Packet: False]
  - Message Header
    + From: “613-963-0148”<sip:16139630148@k.voncp.com>;tag=1019fb60-7196c445-2710-4e9485ff-7b9cb12-4e9485ff
    + To: <sip:3102951568@k.voncp.com>
      Call-ID: 101a7de0-7196c445-2710-4e9485ff-229a8c45-4e9485ff@k.voncp.com
      CSeq: 1 INVITE
    + Via: SIP/2.0/UDP 69.196.150.113:10000;branch=z9hG4bK-4e9485ff-f42b6c64-49ad5933
      P-Preferred-Identity: off
      Max-Forwards: 70
      Supported: replaces.timer.100rel
      User-Agent: VTA001346FE8BF111.4.1-r060815-1.00.09-20070402170142 1248967645135/1007551373 308
    + Contact: <sip:16139630148@69.196.150.113:10000;transport=UDP>
      Min-SE: 0
      Content-Type: application/sdp
      Content-Length: 294
  - Message Body
    - Session Description Protocol
        Session Description Protocol version (v): 0
      + Owner/Creator, Session Id (o): a0000 8644 6672 IN IP4 69.196.150.113
        Session Name (s): SIP Cal
      + Connection Information (c): IN IP4 69.196.150.113
      + Time Description, active time (t): 0 0
      - Media Description, name and address (m): audio 10050 RTP/AVP 0 101 8 2 18
          Media Type: audio
          Media Port: 10050
```

Configuration example

The following configuration example provides the configuration to classify Vonage® SIP/RTP desktop traffic using SIP URI expression:

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Vonage"
            expression 1 sip-uri eq "*voncp.com*"
            application "Vonage"
            no shutdown
        exit
```

H323

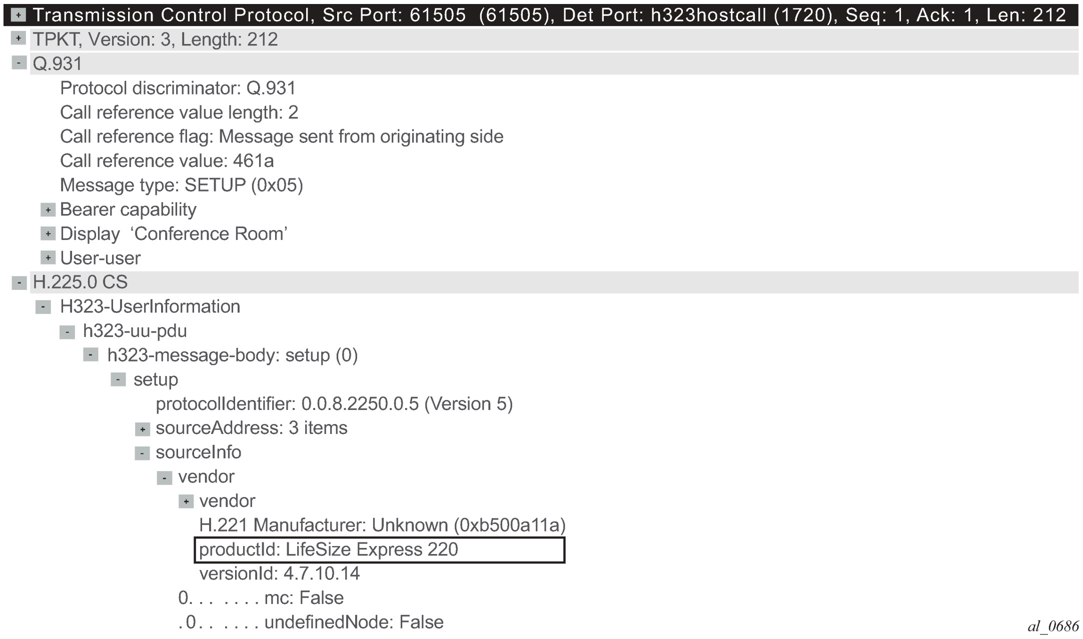
Protocol

Similar to SIP, H323 is a signaling protocol used for controlling multimedia communication sessions such as voice and video over RTP. AA automatically monitors H323 control flows and associates the RTP media flow accordingly in the h323_rtp protocol signature.

The operator can use an H323 expression match criteria app-filter to further refine traffic classification and identify any additional application on top of the default AA policy. This can be particularly useful in business VPNs to identify voice and telepresence applications.

AA supports H323 expression match criteria on the H323 product ID. The following snapshot from Wireshark shows an H323 control exchange using the Telepresence application LifeSize® followed by the RTP media audio flow; the expression field that can be matched using AA app-filters is highlighted:

Figure 7: H323 Wireshark® capture



```
+ Transmission Control Protocol, Src Port: 61505  (61505), Det Port: h323hostcall (1720), Seq: 1, Ack: 1, Len: 212
+ TPKT, Version: 3, Length: 212
- Q.931
      Protocol discriminator: Q.931
      Call reference value length: 2
      Call reference flag: Message sent from originating side
      Call reference value: 461a
      Message type: SETUP (0x05)
    + Bearer capability
    + Display  'Conference Room'
    + User-user
- H.225.0 CS
    - H323-UserInformation
        - h323-uu-pdu
            - h323-message-body: setup (0)
                - setup
                      protocolIdentifier: 0.0.8.2250.0.5 (Version 5)
                    + sourceAddress: 3 items
                    - sourceInfo
                        - vendor
                            + vendor
                              H.221 Manufacturer: Unknown (0xb500a11a)
                              productId: LifeSize Express 220
                              versionId: 4.7.10.14
                          0. . .  . . . . mc: False
                          . 0 . .  . . . . undefinedNode: False
```

Configuration example

The following configuration example provides the configuration to classify LifeSize® H323/RTP traffic using the H323 product ID expression:

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "LifeSize H323 traffic"
            expression 1 h323-product-id eq "^LifeSize*"
            application "LifeSize"
            no shutdown
        exit
```

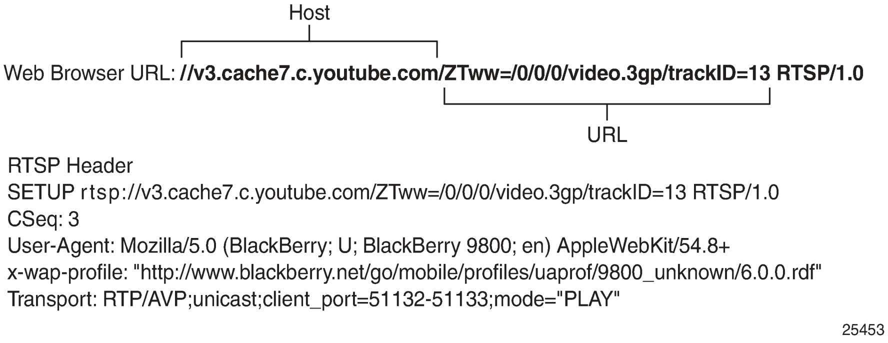
RTSP

Protocol

RTSP is a signaling protocol used for controlling media streaming content such as audio and video over RTP/RDT. AA automatically monitors the RTSP control flows and associates its RTP/RDT media flow with the rtp_rtsp protocol signature.

The operator can use an RTSP expression match criteria app-filter to further refine traffic classification and identify any additional application on top of the default AA policy. This can be particularly useful to identify specific streaming applications.

AA supports RTSP expression match criteria on the RTSP Host, URI, User Agent. The following snapshot from Wireshark® shows an RTSP setup request to YouTube® followed by the RTP media audio flow; the expression fields that can be matched in RTSP SETUP request using AA app- filters are highlighted:


Host
Web Browser URL: //v3.cache7.c.youtube.com/ZTww=/0/0/0/video.3gp/trackID=13 RTSP/1.0
URL
25453

```
RTSP Header
SETUP rtsp://v3.cache7.c.youtube.com/ZTww=/0/0/0/video.3gp/trackID=13 RTSP/1.0
CSeq: 3
User-Agent: Mozilla/5.0 (BlackBerry; U; BlackBerry 9800; en) AppleWebKit/54.8+
x-wap-profile: "http://www.blackberry.net/go/mobile/profiles/uaprof/9800_unknown/6.0.0.rdf"
Transport: RTP/AVP;unicast;client_port=51132-51133;mode="PLAY"
```

Configuration example

The following configuration example provides the configuration to classify YouTube® RTSP/RTP traffic using RTSP Host expression:

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "YouTube RTSP/RTP Video"
            expression 1 rtsp-host eq "*.youtube.com$"
            application "YouTube"
            no shutdown
        exit
```

Citrix

Protocol

Independent Computing Architecture (ICA) is a Citrix Systems® protocol used in Citrix’s WinFrame, Citrix XenApp (formerly called MetaFrame/Presentation Server), and Citrix XenDesktop products.

Citrix makes it possible to run applications remotely on large servers, thus making better use of server resources while at the same time allowing people using other platforms to use the applications, for example, run Microsoft® Word on a UNIX workstation.

Citrix_ica protocol signature will detect any remote application using Citrix (the protocol needs to be unencrypted and configured to non-seamless). The Citrix ICA session is started from a client and can be anything from Remote Desktop, SAP to Microsoft® Word.

The Citrix expression match app-filter is used to classify traffic based on the Citrix-published application. This published application is configured on the server and in the preceding example, it can be for instance RDP, SAP, Word, XLS or Microsoft® Word depending how the server is configured.

Configuration example

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Citrix SAP Application"
            expression 1 citrix-app eq "SAP"
            application "Citrix SAP"
            no shutdown
        exit
```

IP address and TCP/UDP port

Traffic from specific servers can be classified using IPv4/v6 server-address app-filter rules. It is used usually to identify traffic from an internal (on-net) server as opposed to an Internet (off-net) server.

The server-address app-filter automatically detects the client from the server by identifying which side opens the connection. It implicitly classifies traffic based on the server IP address or port number. For example, if A initiates a TCP connection to B, then flows from A to B and from B to A can be classified with a match on server-address B. Similarly, a flow initiated from B to A can be classified using a match on server-address A.

Server address

The following configuration example uses a server-address app-filter to classify traffic from server 10.1.1.1 in the application called Application-1:

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Server #1 10.0.0.1"
            server-address eq 10.0.0.1/32
            application "Application-1"
            no shutdown
        exit
```

Server address and server port

The following configuration example uses server-address and server-port app-filters to classify traffic from server 10.0.0.2 on port 1234 in the application called Application-2. It is particularly useful when the same server is used to provide different services that need to be classified separately:

```
configure application-assurance group 1:1 policy
        entry <1..65535> create
            description "Server #2 10.0.0.2 port 1234 Only"
            server-address eq 10.0.0.2/32
            server-port eq 1234
            application "Application-2"
            no shutdown
        exit
```

Server port and protocol signature

It is possible to combine a protocol signature with a port number in the same app-filter, this is typically done in business VPNs for specific internal applications not detected using existing AA protocol signatures.

The following configuration example classifies a business VPN application running on TCP port 4000 and not detected by any other signatures. It combines the protocol signature unknown_tcp with the desired port number. This allows keeping the classification untouched for the rest of the protocols/applications and is the recommended approach:

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "Business VPN Application X Port 4000"
            server-port eq 4000
            protocol eq unknown_tcp
            application "Busines VPN Application X"
            no shutdown
        exit
```

Note:
It is important to follow the app-filter range recommendations for a proper classification of traffic using IP address or port number.

Flow setup direction

Traffic can be classified based on flow-setup-direction app-filter. The flow setup direction can be either subscriber-to-network or network-to-subscriber.

Network side and subscriber side is AA terminology related to where AA is enabled:

- In broadband and mobile networks, AA is enabled per subscriber. This means the subscriber side represents the ESM/mobile/transit subscriber while the network side represents Internet or other subscribers.
- In business VPNs, AA is enabled on a VPN SAP/spoke SDP and the subscriber side represents the local VPN site (SAP/spoke/transit).

The following example shows the configuration to classify http traffic hosted by AA subscribers (for example, broadband subscribers running a web server):

```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "HTTP Server on the subscriber side"
            flow-setup-direction network-to-subscriber
            protocol eq http
            application "HTTP Server"
            no shutdown
        exit
```

IP protocol

Traffic can be classified using an IP protocol number for non TCP/UDP traffic.

The following example provides the configuration to classify ICMP IPv4/v6 traffic:

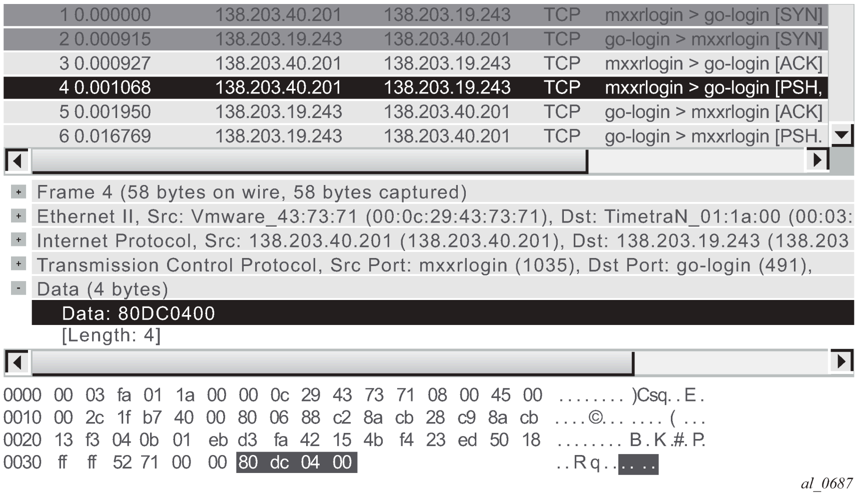
```
configure application-assurance group 1:1 policy
    app-filter
        entry <1..65535> create
            description "ICMP v4"
            protocol eq "non_tcp_udp"
            ip-protocol-num eq icmp
            application "ICMP"
            no shutdown
        exit
        entry <1..65535> create
            description " ICMP v6"
            protocol eq "non_tcp_udp"
            ip-protocol-num eq ipv6-icmp
            application "ICMP"
            no shutdown
        exit
```

Custom protocol

Custom protocols can be used to classify TCP/UDP applications using hexadecimal string matching (up to 16 hex octets) at a configurable payload offset in the data payload. The expression string length and offset must not exceed 128 bytes.

To illustrate this feature the Solaris® application GoGlobal is used. It provides remote access to a server (similar to VNC®). The following snapshot (Figure 6: SIP Wireshark® capture) from Wireshark® shows a TCP SYN/ACK session establishment followed by the first data exchange:

Figure 8: Wireshark® GoGlobal



```
1 0.000000   138.203.40.201   138.203.19.243   TCP   mxxrlogin > go-login [SYN]
2 0.000915   138.203.19.243   138.203.40.201   TCP   go-login > mxxrlogin [SYN]
3 0.000927   138.203.40.201   138.203.19.243   TCP   mxxrlogin > go-login [ACK]
4 0.001068   138.203.40.201   138.203.19.243   TCP   mxxrlogin > go-login [PSH,
5 0.001950   138.203.19.243   138.203.40.201   TCP   go-login > mxxrlogin [ACK]
6 0.016769   138.203.19.243   138.203.40.201   TCP   go-login > mxxrlogin [PSH.
```

```
+ Frame 4 (58 bytes on wire, 58 bytes captured)
+ Ethernet II, Src: Vmware_43:73:71 (00:0c:29:43:73:71), Dst: TimetraN_01:1a:00 (00:03:
+ Internet Protocol, Src: 138.203.40.201 (138.203.40.201), Dst: 138.203.19.243 (138.203
+ Transmission Control Protocol, Src Port: mxxrlogin (1035), Dst Port: go-login (491),
- Data (4 bytes)
    Data: 80DC0400
    [Length: 4]
```

```
0000  00 03 fa 01 1a 00 00 0c 29 43 73 71 08 00 45 00   ........)Csq..E.
0010  00 2c 1f b7 40 00 80 06 88 c2 8a cb 28 c9 8a cb   ....©.......(...
0020  13 f3 04 0b 01 eb d3 fa 42 15 4b f4 23 ed 50 18   ........B.K.#.P.
0030  ff ff 52 71 00 00 80 dc 04 00                     ..Rq......
```

al_0687

Wireshark® shows that each TCP session payload starts with 80DC0400 (no offset) after the three-way TCP handshake. As a result, the configuration required to classify this traffic is as follows:

```
configure application-assurance group 1:1 policy
    custom-protocol 1 ip-protocol-num tcp create
        description "goglobal tcp"
        expression 1 eq "\x80\xdc\x04\x00" offset 0 direction client-to-server
        no shutdown
    exit
    app-filter
        entry <1..65535> create
            description "GoGlobal "
            protocol eq "custom_01"
            application "GoGlobal"
            no shutdown
        exit
```

Typical configuration mistakes

An operator creating new user-defined applications can make a few typical mistakes which are listed below:

- App-filters in shutdown state — The default app-filter state is shutdown. A **no shutdown** command must be executed in order for it to be enabled.
- App-filters with no match criteria — This is a more troublesome mistake as it will catch all the traffic entering the filter in a particular application.

Troubleshooting application identification

Show commands

Router/partition statistics

Partition level statistics are not updated in real time. Instead, statistics for a particular flow are updated either at flow closure or every five minutes. The five-minute sliding window interval is a common interval for all flows in an ISA MDA. Different ISA MDAs will have a different five-minute windows as this interval is set at the MDA boot time.

The following command can be used to view the statistics for all applications configured in the ISA Group 1, Partition 1:

```
show application-assurance group 1:1 application count
```

Alternatively, it is possible to sort the display by octets, packets, flows:

```
show application-assurance group 1:1 application count top [octets | packets | flows] [max-
count <max-count>]
```

The operator can also identify which app-filters are being hit by the AA policy per partition (this command is not available per subscriber), it is particularly useful to identify which filters are used and optionally prune unnecessary app-filters from user-defined applications:

```
show application-assurance group 1:1 policy app-filter
```

Note:
The app-filter policy is usually relatively large, in which case additional 7750 SR CLI functionality can be used to filter out the output and only show the relevant information.

The following example was created for the application FTP:

```
A:PE# show application-assurance group 1:1 policy app-filter | match "application \"FTP\""
                                 pre-lines 3 post-lines 2
                    exit
                    entry 44300 create (2 flows, 1205 B)
                        protocol eq "ftp_control"
                        application "FTP"
                        no shutdown
                    exit
                    entry 44301 create (2 flows, 1401 B)
                        protocol eq "ftp_data"
                        application "FTP"
                        no shutdown
                    exit
```

Because partition level statistics are not updated in real time, it is recommended for troubleshooting purposes to use subscriber statistics or sub-study statistics.

Subscriber statistics

Subscriber-level statistics can be updated in real time. AA is usually configured by the operator to collect subscriber-level statistics for all application groups in residential and Wifi, while business VPNs typically collect Application group and all applications for each site with AA enabled.

The following commands can be used to view per subscriber statistics for all app-groups or applications configured in ISA group 1, partition 1 for the ESM subscriber "Bob" or business VPN SAP 1/1/1:10:

```
show application-assurance group 1:1 aa-sub esm "bob" app-group count
show application-assurance group 1:1 aa-sub sap 1/1/1:10 application count
```

In case only app-group statistics are collected per subscriber, the aa-sub-study feature can be used to collect per application-level statistics for selected subscribers, as follows:

```
A:PE# configure application-assurance group 1:1 statistics aa-sub-study application
A:PE>config>app-assure>group>statistics>aa-sub-study# aa-sub esm "bob"
```

Once done, the system will show all application level statistics for this subscriber:

```
show application-assurance group 1:1 aa-sub-study esm "bob" application count
```

Similar to partition-level statistics, aa-sub and aa-sub-study statistics can be sorted by octets, packets, flows:

```
show application-assurance group 1:1 aa-sub-study esm "bob" application count top [octets |
 packets | flows] [max-count <max-count>]
```

Note:
When the number of flows per ISA card reaches a threshold then per subscriber statistics are not available in real time anymore and only the snapshot command can be used to display the statistics recorded in the previous five-minute interval window:

show application-assurance group 1:1 aa-sub-study esm "bob" snapshot application count

AppFilterMiss

The default policy configuration provides a failsafe application at the very end of the app-filter list to classify any remaining traffic in the AppFilterMiss application. There should never be any traffic in this application. This failsafe filter is used as a debug to make sure that there are no major issues in the configuration.

Note:
Traffic can typically be classified as AppFilterMiss when not all protocol signatures are mapped to a particular application. This could happen when upgrading to a new ISA software and enabling new protocol signature detection while not ensuring first that the correct application was provisioned. See the Release Note upgrade section for more details on AA signature upgrade.

Tools

Flow-record-search

Traditional show commands may not provide enough information when troubleshooting flow identification and the operator can use the ISA flow-record-search tool to dump the ISA flow table for more information. This feature comes with a large number of filtering options documented in the user guide.

Each flow gives visibility into: Flow ID, Sub-Type, Sub-Name, Initiator, Direction, Source IP, Dest. IP, IP Protocol, Source Port, Dest. Port, FC, DSCP, Classified, Protocol, Application, App- Group, Charging Group, Packets tx, Bytes Tx, Packets-discarded, Bytes-discarded etc.

See below for the most commonly used commands.

The following command shows all the flows in an ISA card per ISA group:partition (can be a very long output, up to 3M entries):

```
application-assurance group 1:1 flow-record-search isa 1/2
```

The following command shows all the flows per AA subscriber in a given group:partition:

```
tools dump application-assurance group 1:1 flow-record-search aa-sub esm "bob"
```

The following command shows all the active flows per AA subscriber in a given group:partition:

```
tools dump application-assurance group 1:1 flow-record-search aa-sub esm "bob" flow- status
 active
```

The flow-record-search command is also available with additional details by adding search-type detail at the end of the command line. Note that due to the length of the output it is recommended to paste the CLI output content in a notepad file.

HTTP host recorder

AA cflowd allows operators to export the HTTP domain extracted from HTTP flows to the NSP cflowd collector. This allows the operator to understand which HTTP hosts are visible in the network.

However, in case a cflowd collector is not deployed, AA provides the HTTP host recorder tool command to record HTTP hosts seen by AA. See the *Multiservice ISA and ESA Guide* for more details.

```
A:PE# show debug
debug
    application-assurance
        group 1:1
            http-host-recorder
                filter
                    default-filter-action record
                    record http-host-app-filter-candidates
                exit
                rate 100
                no shutdown
            exit
        exit
    exit
```

```
exit

A:PE# tools dump application-assurance group 1:1 http-host-recorder top bytes
```

Port recorder

This function is particularly useful in business VPN (it can also be used in residential networks). The port-recorder AA tool function is similar to the http-recorder. It allows the operator to record which ports are used on selected applications.

It is most commonly used with the applications Unidentified TCP and Unidentified UDP but it can be configured to record any other applications:

```
A:PE# show debug
debug
    application-assurance
        group 1:1
            port-recorder
                application "Unidentified TCP"
                application "Unidentified UDP"
                rate 100
                shutdown
            exit
        exit
    exit
exit

A:PE# tools dump application-assurance group 1:1 port-recorder top bytes
```

Conclusion

This example, which is intended for Application Assurance network architects and engineers, provides the information required to modify an existing AA policy following AA best practices and guidelines, and provides the necessary troubleshooting information to better understand application classification using Application Assurance.

Application Assurance — App-Profile, ASO and Control Policies

This chapter provides information about Application Assurance (AA) app-profile, Application Service Options (ASOs) and control policy configurations.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This information and configuration in this chapter are based on SR OS Release 12.0.R4.

It is recommended to use the AppDB prior to configuring traffic control policies. The AppDB is a default configuration file to define all of the applications of interest, including all of the relevant application-groups, applications and app-filters to classify traffic, and can be obtained through Nokia's support organization.

Overview

In addition to providing valuable traffic analysis and statistics information using the 7750 Service Router (SR) or 7450 Ethernet Service Switch (ESS) and Application Assurance (AA), one of the key objectives of the AA solution is to provide the tools to manage subscriber traffic at the application level. Examples of traffic management actions include:

- Throttling low priority bandwidth hungry applications during peak hours.
- Prioritizing and remarking selected applications.
- Implementing a walled-garden environment providing open access to selected free web services only, redirecting all other requests from unregistered subscribers to a registration portal with payment services.
- Enrich HTTP Header with subscriber identification parameters to offer subscribers transparent access to premium content.
- In browser notification which triggers the display of administrative, informational or promotional messages in selected browser-sessions.
- Stateful session filtering with Application Level Gateway (ALG) support to protect subscribers against unsolicited flows.
- Parental control services interworking with an external Internet Content Adaptation Protocol (ICAP) server for rating the requested web sites.

Application traffic control policies can be applied as global policies for all subscribers, or they can be activated for individual subscribers or groups of subscribers.

This example describes the basics of activating Application Assurance on a given subscriber through the use of App-Profile and demonstrates the use of static or dynamic traffic control policies using Application Service Options (ASOs) and Application QoS Policies (AQP). It also provides detailed information for configuring Bandwidth, Flow-Count and Flow-Rate Policing including Time of Day (ToD) policing. Other policy control actions can be found in the Advanced Configuration Guide or in the 7450 ESS, 7750 SR, and VSR Multiservice ISA and ESA Guide.

Configuration

Activation of AA services

App-profile

Application profiles (app-profile) enable application assurance services for a given Enhanced Subscriber Management (ESM), Distributed Subscriber Management (DSM), or transit subscriber, or for a SAP or spoke SDP which are commonly referred to as AA subscribers (**aa-sub**). Each app-profile is unique in the system and defines the services that the AA subscriber will receive.

Assigning an app-profile to an ESM subscriber affects every host of that subscriber. Similarly, applying an app-profile to a SAP/spoke SDP will affect all traffic within that SAP/spoke SDP.

App-profiles are defined at the AA group partition level (in case of a partitioned ISA-AA group) as in the following configuration example:

```
A:BNG# configure
    application-assurance group 1:1 policy
        app-profile "1-1/15M" create
            description "App-Profile Description"
            divert
            characteristic "Parental Control" value "enabled"
            capacity-cost 15
        exit
```

The app-profile parameters are:

- **divert** — Diverts all traffic from and to this subscriber to an ISA-AA. Configuring **no divert** effectively disables all AA services for subscribers using this app-profile.

  Default value: **no divert**.
- **characteristic [**<*characteristic-name*> **value** <*value-name*>**]** — one or more optional ASO service characteristics can be used to apply an AA control policy to the subscriber.
- **capacity-cost** <*cost*> — An application profile capacity cost is used to load balance AA subscribers across multiple ISA-AA cards. A common practice is to define a cost proportional to the expected peak BW for the subscribers using this profile (in Kbps or Mbps). The capacity cost is out of the scope of this example. The range is 1 to 65535, default 1.

This app-profile example uses the following naming convention:

<*group-id*>-<*partition-id*>/<*BW*>M where

- *<group-id>* — The ISA-AA group ID on which this profile is created.
- *<partition-id>* — The AA partition ID on which this profile is created.
- *<BW>* — Defines the maximum bandwidth used by the subscriber, which is used for aa-subscriber cost load balancing and subscriber rate limiting. The *M* stands for Mbps.

In general the operator can choose to use either ASO characteristics override or multiple app-profiles to apply different AA QoS policies to ESM Subscribers or Business VPN sites. For flexibility and scale, it is recommended to use ASO overrides whenever possible. This is described in more detail in the following sections.

Note:
Prior to using special characters in a policy object name the operator should verify the list of special characters supported by the 5620 SAM; for instance the 5620 SAM does not support the use of “:” in the app-profile name therefore it should be avoided.

Residential and Wi-Fi services

The app-profile can be assigned or modified for ESM, DSM or Transit IP subscribers either at subscriber creation time or while the subscriber is in service:

- Subscriber creation — An app-profile can be assigned at subscriber creation time through RADIUS, DHCP Option 82, Local User Database, static configuration or through a default app-profile.
- In service app-profile modification — An app-profile can be dynamically modified in service through a RADIUS Change of Authorization (CoA). From software Release 12.0.R1 an app-profile can also be dynamically modified in service through Gx.

In case no app-profile is returned at subscriber creation by RADIUS, LUDB or DHCP, or when no static configuration is present, the system can apply a default app-profile if configured within the subscriber group-interface (or MSAP policy) sub-sla-mgmt:

```
sub-sla-mgmt
 def-app-profile "1-1/15M"
exit
```

Business VPN and other service interfaces

App-profiles are statically assigned to a given SAP, spoke SDP or transit prefix VPN site via the 5620 SAM or CLI.

The following configuration shows how to enable application assurance on a SAP or spoke SDP in a business VPRN service:

```
A:PE>config>service# vprn 100 customer 1 create
 description "L3 Service Customer 1"
 interface "to-site1" create
    address 192.168.1.1/24
    sap 1/1/10:11 create
        app-profile "1-1/15M"
    exit
interface "to-site2" create
    address 192.168.2.1/24
    spoke-sdp 12:100 create
        app-profile "1-1/15M"
    exit
no shutdown
```

Defining application service options

ASOs for trafficcontrol - introduction

To determine which application control policies need to be applied to a AA-subscriber, an app-profile with a number of service characteristics (ASOs) is associated with each subscriber. These service characteristics are then used as match criteria in AQP policy rules to determine which rules to apply.

Therefore ASOs are service characteristics assigned to a subscriber and are used to identify the traffic control policy rule (AQP) applicable to a subscriber or a group of subscribers.

Most policy rules will be applicable to multiple subscriber profiles; nevertheless it is possible that a specific subscriber requires a dedicated policy.

ASO characteristics and values

For each service option that can be used by one or more subscribers, an ASO characteristic should be defined with a number of values that represent all available choices for that service characteristic. The names and values of the ASO characteristics are configurable string values; best practice is to use strings that provide a meaningful description of the service characteristic they represent.

Each ASO characteristic requires a default value and each app-profile inherits the default value of all the ASO characteristics created in a given partition unless a characteristic is referenced directly in the app-profile or overwritten as described below.

ASOs are defined at the AA group partition level (in case of a partitioned ISA-AA group). In the configuration example below two different ASO characteristics are defined: “Parental Control” and “P2P-Sub-DL”:

```
BNG>config>app-assure# group 1:1 policy
app-service-options
   characteristic "Parental Control" create
     value "disabled"
     value "enabled"
     default-value "disabled"
   exit
   characteristic "P2P-Sub-DL" create
     value "500k"
     value "1M"
     value "unlimited"
     default-value "unlimited"
   exit
```

The ASO values and default value of a characteristic can be displayed using a show command:

```
A:BNG# show application-assurance group 1:1 policy app-service-option "P2P-Sub-DL"
===============================================================================
Application-Assurance Application Service Options
===============================================================================
Characteristic "P2P-Sub-DL"
Value                            Default
```

```
-------------------------------------------------------------------------------
1M                              No
500k                            No
unlimited                       Yes
===============================================================================
```

When configuring service characteristics for optional service options, it is recommended to configure a default value which will not trigger any AQP policy action (the default value does not match any AQP match criteria) such that the behavior of existing subscribers and app-profiles will not change until the operator specifically configures or signals a non-default characteristic value for the subscriber or the app-profile. In the preceding example, "Parental Control disabled" and "P2P-Sub-DL unlimited" would have no corresponding AQP by design; therefore if these particular service options were applied to a subscriber they would not match a QoS policy entry.

How to specify service options for AA subscribers

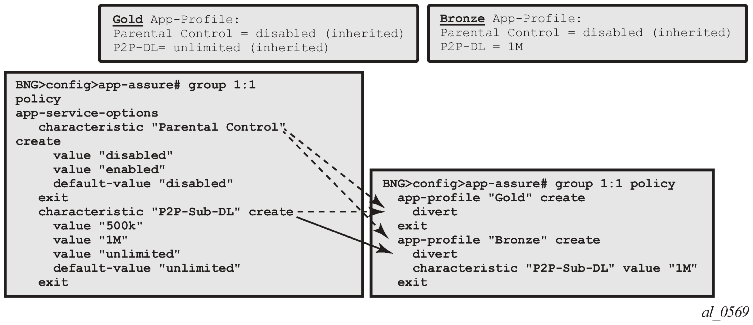
ASO assignment in app-profile

ASOs can be statically assigned in the app-profile; this type of ASO characteristic assignment is typically reserved to the default service options enabled on a large number of subscribers.

Figure 9: Service tier example using ASO, app-profile and AQP shows an example of AA service definition (ASO and app-profile) for a Gold and Bronze service tier definition with the following characteristics:

- Two app-profiles *Gold* and *Bronze*
- *Gold* app-profile — No specific policy actions or ASO characteristics are configured statically in the app-profile.
- *Bronze* app-profile — A specific ASO characteristic and value is assigned to the profile to limit peer to peer download traffic to 1Mbps (this example does not show the app-qos-policy nor policer configuration, this will be described later).

Figure 9: Service tier example using ASO, app-profile and AQP



```
Gold App-Profile:
Parental Control = disabled (inherited)
P2P-DL= unlimited (inherited)
```

```
Bronze App-Profile:
Parental Control = disabled (inherited)
P2P-DL = 1M
```

```
BNG>config>app-assure# group 1:1
policy
app-service-options
   characteristic "Parental Control"
create
     value "disabled"
     value "enabled"
     default-value "disabled"
   exit
   characteristic "P2P-Sub-DL" create
     value "500k"
     value "1M"
     value "unlimited"
     default-value "unlimited"
   exit
```

```
BNG>config>app-assure# group 1:1 policy
  app-profile "Gold" create
    divert
  exit
  app-profile "Bronze" create
    divert
    characteristic "P2P-Sub-DL" value "1M"
  exit
```

Each app-profile inherits the default values of all the ASO characteristics defined in a AA group-partition; in the preceding example, this is the reason why the app-profile Gold inherits "Parental control disabled" and "P2P-Sub-DL unlimited". The app-profile Bronze inherits "Parental control disabled" while "P2P-Sub-DL 1M" is assigned to this profile statically.

The operator can identify per app-profile which characteristics values are inherited from their default value and which are statically assigned using the following show command:

```
*A:BNG# show application-assurance group 1:1 policy app-profile "Gold"
    app-profile "Gold" create
        divert
        characteristic "P2P-Sub-DL" inherits default-value "unlimited"
        characteristic "Parental Control" inherits default-value "disabled"
    exit

A:BNG# show application-assurance group 1:1 policy app-profile "Bronze"
    app-profile "Bronze" create
        divert
        characteristic "P2P-Sub-DL" value "1M"
        characteristic "Parental Control" inherits default-value "disabled"
    exit
```

Note:
Using ASO overrides, described later, it is possible to implement the same choice of AA service options using a single app-profile.

ASO overrides per subscriber via RADIUS or Gx

Prior to SR OS 12.0.R1 the operator can assign (and modify: CoA) the app-profile per ESM or Transit-IP subscribers using the "Alc-App-Prof-Str" [26-6527-45] RADIUS attribute.

SR OS 12.0.R1 added support for ASO characteristic overrides for ESM and Transit-IP subscribers via RADIUS using the attribute "Alc-AA-App-Service-Options" [26-6527-193].This attribute can be returned during the subscriber creation process or while the subscriber is in service through RADIUS CoA. Refer to 7450 ESS, 7750 SR, and VSR RADIUS Attributes Reference Guide for more details related to the use of the AA RADIUS attributes.

An example of a RADIUS CoA message returned to the system to modify both the app-profile and one ASO characteristic is as follows:

```
    NAS-Port-Id = "1/1/5:4088"
    Framed-IP-Address = 192.168.211.30
    Alc-App-Prof-Str = "1-1/15M"
    Alc-AA-App-Service-Options = "P2P-Sub-DL=1M"
```

The ASO characteristics and values assigned to a given subscriber (statically via app-profile or overridden) can be displayed using the following show command:

```

A:BNG# show application-assurance group 1:1 aa-sub esm "sub1" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub1 (esm)
ISA assigned            : 1/2
App-Profile             : 1-1/15M
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : 2014/08/07 12:07:47
-------------------------------------------------------------------------------
Traffic                        Octets              Packets                Flows
```

```
--------------------------------------------------------------------------------
...
...
--------------------------------------------------------------------------------
Application Service Options (ASO)
--------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
--------------------------------------------------------------------------------
P2P-Sub-DL                       1M                               dyn-override
Parental Control                 disabled                         default
================================================================================
```

In the preceding show command output, the **derived from** field describes how the characteristics and values are assigned to the subscriber:

- app-profile — The characteristic’s value statically configured in the app-profile.
- dyn-override — The characteristic’s value received from RADIUS or Gx.
- default — The characteristic’s default value inherited (not statically configured in the app-profile nor dynamically modified).

SR OS 12.0.R1 also introduced support for signaling the app-profile or ASO characteristics override via Gx, see Application Assurance — App-Profile, ASO and Control Policies for more details.

ASO overrides for business VPN and other services

ASO characteristic override values can be statically assigned to business VPN SAP, spoke SDP and transit prefix subscribers.

The operator can provision the AA policy override parameters, multiple characteristics overrides per AA-sub can be defined per override policy, as in the following configuration example:

```
A:BNG>config>app-assure# group 1:1 policy-override
    policy aa-sub sap 1/1/5:210 create
        characteristic "P2P-Sub-DL" value "1M"
        characteristic "Parental Control" value "enabled"
    exit
```

Application control policies

App-QoS policy (AQP)

App-profile / ASO / AQP workflow summary

App-profiles enable application assurance services for a given AA-subscriber. Each app-profile is unique in the system and defines the service that the AA subscriber will receive.

To determine which control policies need to be applied to an AA-subscriber, a number of service characteristics (ASO) are associated with each AA-subscriber.

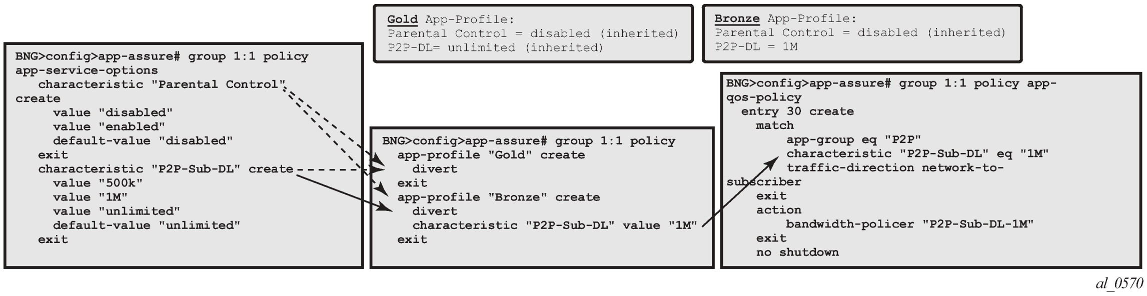
As described earlier, these service characteristics can either be configured directly within the app-profile or assigned using overrides and they are then used as match criteria in AQP policy rules to determine which application policy rules to apply.

The app-qos-policy (AQP) is an ordered list of entries defining policy actions for flows diverted to Application Assurance. Each AQP entry is composed of match criteria and action(s).

Flows are evaluated against all entries of the AA QoS policy defined in the AA group partition that the subscriber app-profile belongs to (in case of a partitioned AA group).

Figure 10: App-Profile, ASO, AQP workflow summary provides a configuration example summary with app-profile, ASO, AQP and policers:

Figure 10: App-Profile, ASO, AQP workflow summary



```
BNG>config>app-assure# group 1:1 policy
app-service-options
   characteristic "Parental Control"
create
     value "disabled"
     value "enabled"
     default-value "disabled"
   exit
   characteristic "P2P-Sub-DL" create
     value "500k"
     value "1M"
     value "unlimited"
     default-value "unlimited"
   exit
```

```
Gold App-Profile:
Parental Control = disabled (inherited)
P2P-DL= unlimited (inherited)
```

```
BNG>config>app-assure# group 1:1 policy
  app-profile "Gold" create
    divert
  exit
  app-profile "Bronze" create
    divert
    characteristic "P2P-Sub-DL" value "1M"
  exit
```

```
Bronze App-Profile:
Parental Control = disabled (inherited)
P2P-DL = 1M
```

```
BNG>config>app-assure# group 1:1 policy app-
qos-policy
  entry 30 create
    match
        app-group eq "P2P"
        characteristic "P2P-Sub-DL" eq "1M"
        traffic-direction network-to-
subscriber
    exit
    action
        bandwidth-policer "P2P-Sub-DL-1M"
    exit
    no shutdown
```

Match and action criteria

AQP match criteria

Multiple match criteria can be specified per AQP entry in which case the action will only apply to flows that match all criteria. The most common match criteria are: characteristic, application, app-group and charging-group.

The following AA match criteria can be used in an AQP:

- **app-group {eq | neq}** <*app-group name*>
- **application {eq | neq}** <*app name*>
- **charging-group {eq | neq}** <*charging-group-name*>
- **traffic-direction {subscriber-to-network | network-to-subscriber | both}**
- **characteristic** <*characteristic-name*> **eq** <*value-name*>: up to 4 characteristics and values per AQP
- **ip-protocol-num {eq | neq}** <*protocol-id*>
- **src-ip {eq | neq}** <*ip-address*> or **ip-prefix-list** <*ip-prefix-list-name*>
- **dst-ip {eq | neq}** <*ip-address*> or **ip-prefix-list** <*ip-prefix-list-name*>
- **src-port {eq | neq}** <*port-num*> or **range** <*start-port-num*> <*end-port-num*>
- **dst-port {eq | neq}** <*port-num*> or **range** <*start-port-num*> <*end-port-num*>
- **dscp {eq | neq}** <*dscp-name*>
- **aa-sub** <*aa-sub-name*>

AQP actions

The following AA traffic control policies can be specified in an AQP:

- **drop**
- **bandwidth-policer** <*policer-name*>
- **flow-count-limit** <*policer-name*>
- **flow-rate-limit** <*policer-name*>
- **remark dscp in-profile** <*dscp-name*> **out-profile** <*dscp-name*>
- **remark fc** <*fc-name*>
- **remark priority** <*priority-level*>
- **http-error-redirect** <*redirect-name*>
- **http-redirect** <*redirect-name*> **flow-type** <*flow-type*> — Redirect traffic to a landing page
- **mirror-source [all-inclusive]** <*mirror-service-id*>
- **session-filter** <*session-filter-name*> — Session filter firewall
- **url-filter** <*url-filter-name*>: category-based URL filtering using ICAP
- **http-notification** <*http-notification-name*>
- Additional drop actions:
  - **error-drop**: configure a drop action for packets cut-through due to IP packet errors (bad IP checksums, tcp/udp port 0, etc.)
  - **overload-drop**: configure a drop action for packets cut-through due to overload
  - **fragment-drop**: configure a drop action for IP fragmented packets

Default versus application-specific AQP policies

Application QoS policy

It usually requires the examination of a few packets to identify the protocol/application of a flow. When AQP entries are defined to match on IP header criteria (IP address, IP prefix list, TCP/UDP port number, IP protocol, DSCP) or application criteria (application, app-group or charging group), the AQP action will only be applied to matching application flows after a flow has been classified as a given application.

Default QoS policy

If the AQP entry does not include match criteria against application (application, app-group and charging-group) or IP header information (IP address, IP prefix list, TCP/UDP port number, IP protocol, DSCP) then the AQP policy will be applied to all matching flows starting with the first packet of a flow before protocol and application identification is complete. Such AQPs are called default subscriber policies.

For an AQP to be qualified as a default subscriber policy, the match criteria must be limited to any combination of ASO characteristic values, traffic direction and optional AA subscriber name.

AQP match and actions for the default QoS policy and application QoS policy are summarized in Table 3: Default QoS policy, application QoS policy table :

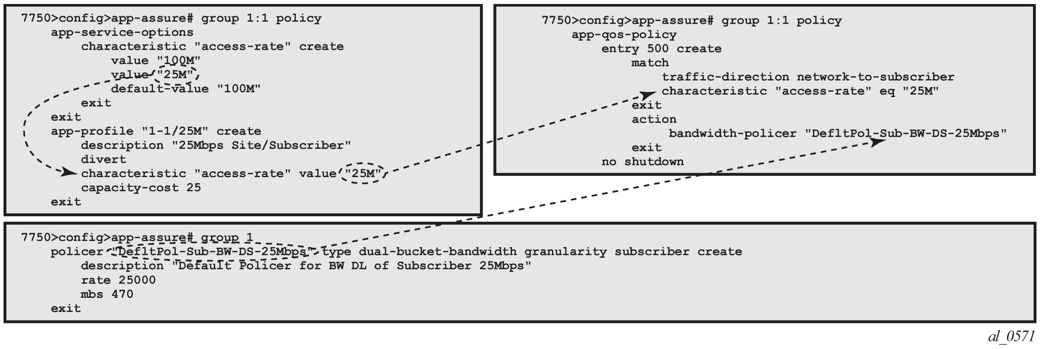
Table 3: Default QoS policy, application QoS policy table

Policy	AQP match	AQP action
Default QoS	ASO characteristic/values traffic direction aa-sub	Remark FC, DSCP, priority Bandwidth, flow-count, flow-rate policing Session-filter Url-filter Mirror error-drop, overload-drop, fragment-drop Drop
Application QoS	ASO characteristic/values traffic direction aa-sub application app-group charging-group IP address, IP prefix list TCP/UDP port number DSCP IP protocol number	Remark FC, DSCP, priority Bandwidth, flow-count, flow-rate policing HTTP notification HTTP redirect HTTP enrichment mirror Drop

To ensure fair access to the ISA-AA bandwidth and flow resources, it is recommended to configure default AQP policy entries limiting bandwidth and flow resources per AA sub.

Figure 11: Default downstream bandwidth policing shows a default subscriber policy limiting the downstream bandwidth (network-to-subscriber direction) to 25Mbps per subscriber:

Figure 11: Default downstream bandwidth policing



```
7750>config>app-assure# group 1:1 policy
    app-service-options
        characteristic "access-rate" create
            value "100M"
            value "25M"
            default-value "100M"
        exit
    exit
    app-profile "1-1/25M" create
        description "25Mbps Site/Subscriber"
        divert
        characteristic "access-rate" value "25M"
        capacity-cost 25
    exit
```

```
7750>config>app-assure# group 1:1 policy
    app-qos-policy
        entry 500 create
            match
                traffic-direction network-to-subscriber
                characteristic "access-rate" eq "25M"
            exit
            action
                bandwidth-policer "DefltPol-Sub-BW-DS-25Mbps"
            exit
        no shutdown
```

```
7750>config>app-assure# group 1
    policer "DefltPol-Sub-BW-DS-25Mbps" type dual-bucket-bandwidth granularity subscriber create
        description "Default Policer for BW DL of Subscriber 25Mbps"
        rate 25000
        mbs 470
    exit
```

Implicit default subscriber policy

Session-filter, url-filter, overload-drop, fragment-drop and error-drop can only be used as part of a default subscriber policy; therefore these actions are not compatible with application or IP header match criteria within the same AQP.

AQP entries evaluation

Multiple AQP match entries per flow

A single flow can match multiple AQP entries, in which case multiple actions can be selected based on the AQP entry’s order (the lowest number entry has the highest priority); the drop action takes precedence over any other AQP entry. The maximum numbers of actions that can be applied on a single flow are:

- 1 drop action
- Any combination of (applied only if no drop action is selected)
  - Up to 1 mirror action
  - Up to 1 FC, 1 priority and 1 DSCP remark action
  - Up to 4 BW policers (1 single rate AA-sub, 1 dual rate AA-sub, 2 single rate system level)
  - Up to 12 flow policers (3 subscriber flow-count, 3 subscriber flow-rate, 3 system flow-count, 3 system flow-rate)
  - Up to 1 HTTP redirect
  - Up to 1 HTTP error redirect
  - Up to 1 HTTP enrichment
  - Up to 1 URL-filter
  - Up to 1 HTTP-notification
  - Up to 1 session-filter firewall
- 1 error drop
- 1 overload drop
- 1 fragment drop

An AQP entry match that would cause the preceding limits to be exceeded is ignored (no actions from that rule are selected) and the conflict counter for this AQP is incremented.

The operator can display hits and potential conflicts per AQP entry using the following show command:

```
A:BNG# show application-assurance group 1:1 policy app-qos-policy
================================================================================
Application QOS Policy Table
================================================================================
Entry         Admin State                      Flow Hits          Flow Conflicts
--------------------------------------------------------------------------------
30            in-service                               0                       0
--------------------------------------------------------------------------------
No. of AQP entries: 1
================================================================================
```

AQP evaluation

Flows are evaluated against all entries of the AA QoS policy at different steps during the lifetime of the flow:

- **Flow creation** — The default subscriber policy AQP entries for matching flows are applied starting with the first packet of a flow so before application identification completes.
- **Application identification completion**— The application QoS policies are applied once flow identification has been completed.

  Note:
  The default QoS policy entries are applied to the subscriber's flows for packets received before and after application identification is completed.

- **Policy change** — When a configuration change is applied to the AA policy by executing the commit command on the AA group:partition policy, all diverted flows for subscribers using this policy partition will be evaluated again against all AQP entries. This re-evaluation happens as a paced background task; hence AQP control changes may not be applied immediately to all existing flows.

Policing

Policers

AA policer templates are configured as part of the AA group configuration by specifying the policer name, type and granularity. Policers are unidirectional by definition so that separate policers must be defined per flow direction if the traffic needs to be policed in both directions (a separate AQP for each flow direction is therefore required as well).

The operator can configure the following types of policers:

- Bandwidth policers
  - Single bucket system level
  - Single bucket flow level
  - Single bucket AA subscriber level
  - Dual bucket AA subscriber level
- Flow count policer: system or AA subscriber level
- Flow setup-rate policer: system or AA subscriber level

Subscriber level policers are instantiated per AA sub, meaning:

- The system automatically uses a dedicated policer for every single subscriber, even when multiple subscribers match the same AQP entry.
- The same policer can be referenced in different AQP entries; in this case all subscribers' flows matching any of these AQP entries are policed by the same subscriber policer. Example: if the same subscriber level policer '1Mbps' is referenced in AQP entry 100 matching application BitTorrent and in AQP entry 110 matching application EDonkey, then the sum of both the BitTorrent and EDonkey traffic cannot exceed 1Mbps.

System level policers on the other hand are shared by all AA subscribers matching a given AQP entry. These policers are typically used in residential and Wi-Fi service deployments to limit the total bandwidth for an application or application group, for all subscribers or for a group of subscribers on the system or partition. An example would be a system level 500Mbps policer to limit the aggregated downstream bandwidth of peer-to-peer applications for all subscribers with a "Bronze" app-profile to 500Mbps.

Note:
In case multiple ISA-AA cards are used per system, the overall maximum throughput using a system level policer is equal to the policer rate limit times the number of ISA cards in the system.

Bandwidth policing

Single bucket subscriber/system bandwidth policer

Single bucket policers police the matching traffic against a configured peak information rate (PIR). Traffic above the PIR can be marked as out of profile or dropped.

The configuration template for a single rate bandwidth policer is as follows:

```
BNG>config>app-assure# group 1
    policer <policer-name> type single-bucket-bandwidth
                                    granularity {subscriber|system} create
        description <string>
        rate <pir-rate-in-Kbps>
        mbs <max-burst-size-in-Kbytes>
        adaptation-rule pir {max|min|closest}
        tod-override <tod-override-id>
        action permit-deny|priority-mark
```

where:

- **action** — defines the action that must be taken by the policer for non-conforming traffic.
- **permit-deny** — non-conforming packets will be dropped.
- **priority-mark** — non-conforming traffic will be marked as out of profile (increasing the chances that non-conforming packets will be discarded in case of congestion on the egress queues).
- **rate** — peak information rate in Kbps.
- **mbs** — maximum burst size in Kbytes.
- **adaptation-rule pir <max|min|closest>** — The policers work at discrete operational rates supported by the hardware. The adaptation rule specifies how the actual operational policer rate (supported by the hardware) must be selected as compared to the configured PIR. During operation, both the operational and configured rate can be displayed using the operational **show application-assurance group** <*n*> **policer** <*policer-name*> **detail** command.
- **tod-override** — defines a time of day override policy applicable to a policer, this is described in more detail at the end of the policing section.

The following shows a single bucket subscriber level policer configuration example:

```
BNG>config>app-assure# group 1
    policer "P2P-Sub-DL-1M" type single-bucket-bandwidth granularity subscriber create
        rate 1000
        mbs 19
    exit
```

The following shows a single bucket system level policer configuration example:

```
BNG>config>app-assure# group 1
    policer "P2P-Sys-DL-100M" type single-bucket-bandwidth granularity system create
        rate 100000
        mbs 1875
    exit
```

Dual bucket subscriber bandwidth policer

Dual-bucket policers police the matching traffic against a configured peak information rate (PIR) and committed information rate (CIR). Traffic below CIR is marked in profile, traffic between CIR and PIR is marked as out of profile, and traffic above the PIR is dropped.

Dual-bucket policers can only be used as subscriber policers; system policers cannot be defined as dual-bucket policers.

The configuration is similar to the single-bucket policer, but adds the configuration of a CIR and a committed burst size (CBS), and the action cannot be configured:

```
BNG>config>app-assure# group 1
    policer <policer-name> type dual-bucket-bandwidth
                                    granularity {subscriber|system} create
        description <string>
        rate <pir-rate-in-Kbps> cir <cir-rate-in-Kbps>
        mbs <max-burst-size-in-Kbytes>
        cbs <committed-burst-size-in-Kbytes>
        adaptation-rule pir {max|min|closest} cir {max|min|closest}
```

The following shows a dual-bucket subscriber level policer configuration example:

```
BNG>config>app-assure# group 1
    policer "P2P-Sub-DL-2M-DB" type dual-bucket-bandwidth granularity subscriber create
        rate 2000 cir 1000
        cbs 19
        mbs 38
    exit
```

MBS/CBS calculation for bandwidth policers

The default MBS/CBS value of a bandwidth policer is set to 0, but the operator must modify this value to allow proper interworking with TCP-based applications. The network operator must carefully consider the values to be used in production networks based on the applications in the network and several other factors like traffic patterns,traffic volume, bursts, and so on. Nokia recommends to configure a minimum MBS for a bandwidth policer, as follows:

MBS = 2 * MTU or 0.00025 * peak rate (whichever is larger)

The formula to calculate the MBS or CBS buffer size, as documented in RFC 6349 is:

Buffer (B) = RTT (s) * Rate (bps) / 8

For Internet applications it is recommended to use a common Round Trip Time (RTT) of 150 msec.

An example using a single bucket subscriber level policer rate of 1000 Kbps:

MBS (B) = 1,000,000 / 8 * 0.150 = 18750 Bytes or 190 KB

Note: These policer values may need to be further adjustment depending on the application.

Flow rate limit policer

The default MBS/CBS value of a flow rate limit policer is set to 0, but Nokia recommends to configure a non-zero MBS value for a flow rate limit policer.

Flow rate limit policers police the maximum number of new flows that are accepted per second for matching traffic. The configuration is similar to the single-bucket bandwidth policer, with the rate and MBS now expressed in flows/sec and flows, respectively.

```
BNG>config>app-assure# group 1
    policer <policer-name> type flow-rate-limit granularity {subscriber|system} create
        description <string>
        rate <flow-rate-in-flows/sec>
        mbs <max-burst-size-in-flows>
        adaptation-rule pir  {max|min|closest}
        action permit-deny|priority-mark
```

This type of policer is primarily used for the default subscriber AQP policy in order to limit the maximum number of flow/seconds allocated per AA subscriber.

Note: In case the policer is used as part of the default AA subscriber policy then the **priority-mark** action has the effect to cut-through non conformant traffic in the ISA instead of drop using **permit-deny**.

Flow count limit policer

Flow count limit policers police the maximum number of concurrent flows for matching traffic:

```
BNG>config>app-assure# group 1
    policer <policer-name> type flow-count-limit granularity {subscriber|system} create
        description <string>
        action permit-deny|priority-mark
        flow-count <max-number-of-flows>
```

This type of policer is primarily used for the default subscriber AQP policy in order to limit the maximum number of concurrent flows allocated per AA subscriber.

Note:
The **priority-mark** has the effect to cut-through non conformant traffic in the ISA instead of drop using **permit-deny**.

Time of day policing

For time-of-day (ToD) policer override, up to 8 override rates with time of day specifications can be defined per policer, this time of day override using the system local time.

ToD overrides are supported for all policer types described in the previous section (bandwidth, flow count, and flow rate) and can be configured using either daily or weekly patterns.

The configuration of ToD override on daily or weekly basis is shown in the following template:

```
BNG>config>app-assure# group 1
    policer "P2P-Sub-DL-1M-TOD" type single-bucket-bandwidth
                                           granularity subscriber create
        action permit-deny
        rate 1000
        mbs 19
        adaptation-rule pir closest
        tod-override <override-id>
            description <string>
            time-range daily start <start-time> end <end-time>
                                        [on <day> [<day>...(upto 7 max)]]
            time-range weekly start <day,start-time> end <day,end-time>
            rate 2000
            mbs 38
```

where:

- **tod-override** *<override-id>* — up to 8 override IDs (with value 1-255) can be configured per policer.
- **time-range** — can be configured to be triggered.
  - On a daily basis at the indicated start/end-time on the specified days.
  - On a weekly basis at the indicated start day and time and end day and time.
  - Times can be indicated as <hh>:<mm> with a 15-minute granularity for the minutes (mm = 0 **|** 15 **|** 30 **|** 45).

A configuration example for a single bucket system level bandwidth policer with the following ToD-override patterns follows:

- Default rate limit: 300Mbps
- Rate limit override to 100Mbps between 5PM and 10PM
- Rate limit override to 200Mbps between 10PM and 12PM

```
BNG>config>app-assure# group 1
     policer "P2P-Sys-DL-300M-TOD" type single-bucket-bandwidth
                                       granularity system create
        description "Peer to Peer Policer System level Policer"
        rate 300000
        mbs 5625
        tod-override 1 create
            description "Override busy hour #1"
            time-range daily start 17:00 end 22:00
            rate 100000
            mbs 1875
            no shutdown
        exit
        tod-override 2 create
            description "Override busy hour #1"
            time-range daily start 22:00 end 24:00
            rate 200000
            mbs 3750
            no shutdown
        exit
```

The operator can display which policing rate is applied at any moment in time together with all configured override rates using the following command:

```
show application-assurance group <n> policer <policer-name> detail
```

Design and configuration examples

Default AA QoS policy

To ensure fair access for all subscribers to the ISA-AA resources, and avoid that a disproportionate amount of ISA-AA resources are used by one or more subscribers which are misbehaving or receiving large traffic bursts from the Internet, it is recommended to configure the following three types of subscriber-level default AA QoS policies:

- **A default bandwidth policer** to limit the downstream bandwidth per subscriber (upstream bandwidth is already limited by ESM/SAP access ingress IOM QoS).
- **A default flow count policer** to limit the maximum number of active flows per traffic direction per subscriber. The operator can choose to drop or cut-through non conforming traffic.
- A **default flow rate policer** to limit the maximum flow setup rate per traffic direction per subscriber. The operator can choose to drop or cut-through non conforming traffic.

The minimum set of app-profiles used in a network is typically determined by the different access bandwidth rates; services characteristics are then used for each profile to apply a default QoS policy to limit bandwidth and flow resources accordingly.

In theory, it is possible to configure a set of default policers for every individual access bandwidth rate that is offered to a subscriber. This would however result in a large number of policers and corresponding ASO values plus app-profiles that need to be configured. Therefore, a best practice guideline is to define a small number of bandwidth ranges (not more than five to ten) that cover the full offered access bandwidth spectrum, and define for each bandwidth range a default bandwidth policer plus flow policers with appropriate limits.

As an example, assuming a residential deployment with 2 bandwidth ranges of up to 25Mbps and 100Mbps, the following configuration provides:

- Complete ASO and app-profile configuration.
- Default QoS policy for subscribers in the 25Mbps range including bandwidth.
- Flow count and flow rate policers are configured by default as permit-deny. Non-conforming traffic is dropped which is common for residential deployments; alternatively the operator can decide to configure these policers as priority-mark to cut-through traffic in the ISA-AA.

In this example the resources are limited per subscriber based on their access rate maximum speed from which flow count and flow rate are derived.

App-profile and ASO

The following configuration provides the app-profile and ASO characteristics used for the default subscriber AQP policy for the 25Mbps and 100Mbps access bandwidth range:

```
BNG>config>app-assure# group 1:1 policy
    app-service-options
        characteristic "access-rate" create
            value "100M"
            value "25M"
            default-value "100M"
        exit
    exit
    app-profile "1-1/25M" create
        description "25Mbps Site/Subscriber"
        divert
        characteristic "access-rate" value "25M"
        capacity-cost 25
    exit
    app-profile "1-1/100M" create
        description "100Mbps Site/Subscriber"
        divert
        characteristic "access-rate" value "100M"
        capacity-cost 100
    exit
```

Default bandwidth policing – 25Mbps AA-sub

```
BNG>config>app-assure# group 1
    policer "DefltPol-Sub-BW-DS-25Mbps" type dual-bucket-bandwidth
                                                     granularity subscriber create
        description "Deflt downstream BW policer for 25Mbps Subs"
        rate 25000
    mbs
```

The following AQP entry will act as a default AQP policy because it does not include application or IP header match criteria:

```
BNG>config>app-assure# group 1:1 policy
    app-qos-policy
        entry 500 create
            description "Deflt downstream BW policer for 25Mbps Subs"
            match
                traffic-direction network-to-subscriber
                characteristic "access-rate" eq "25M"
            exit
            action
                 bandwidth-policer "DefltPol-Sub-BW-DS-25Mbps"
            exit
            no shutdown
        exit
```

Note:
A similar configuration can be implemented for the 100Mbps access rate service option.

Default flow-count-limit policing – 25Mbps AA-sub

```
BNG>config>app-assure# group 1
    policer "DefltPol-Sub-FlowCount-US-25Mbps" type flow-count-limit
                                                 granularity subscriber create
        description "Deflt policer to limit active upstream flows for 25Mbps Subs”
        flow-count 10000
        action permit-deny
    exit
    policer "DefltPol-Sub-FlowCount-DS-25Mbps" type flow-count-limit
                                                 granularity subscriber create
        description "Deflt policer to limit active downstream flows for 25Mbps Subs”
        flow-count 10000
        action permit-deny
    exit
```

The following AQP entry acts as a default AQP policy because it does not include application or IP header match criteria:

```
BNG>config>app-assure# group 1:1 policy app-qos-policy
    entry 510 create
        description " Deflt policer to limit active upstream flows for 25Mbps Subs"
        match
            traffic-direction subscriber-to-network
            characteristic "access-rate" eq "25M"
        exit
        action
            flow-count-limit "DefltPol-Sub-FlowCount-US-25Mbps"
        exit
        no shutdown
    exit
    entry 515 create
        description " Deflt policer to limit active downstream flows for 25Mbps Subs"
        match
            traffic-direction network-to-subscriber
            characteristic "access-rate" eq "25M"
        exit
        action
            flow-count-limit "DefltPol-Sub-FlowCount-DS-25Mbps"
        exit
        no shutdown
    exit
```

Note:
A similar configuration can be implemented for the 100Mbps access rate service option.

Default flow-rate-limit policing – 25Mbps AA-sub

```
BNG>config>app-assure# group 1
    policer "DefltPol-Sub-FlowRate-US-25Mbps" type flow-rate-limit
                                                 granularity subscriber create
        description "Deflt policer to limit upstream flow setup rate for 25Mbps Subs”
        rate 200
        action permit-deny
    exit
    policer "DefltPol-Sub-FlowRate-DS-25Mbps" type flow-rate-limit
                                                granularity subscriber create
        description "Deflt policer to limit downstr flow setup rate for 25Mbps Subs"
        rate 200
        action permit-deny
    exit
```

The following AQP entry acts as a default AQP policy because it does not include application or IP header match criteria:

```
BNG>config>app-assure# group 1:1 policy app-qos-policy
    entry 520 create
        description "Deflt policer to limit upstream flow setup rate for 25Mbps Subs"
        match
            traffic-direction subscriber-to-network
            characteristic "access-rate" eq "25M"
        exit
        action
            flow-rate-limit "DefltPol-Sub-FlowRate-US-25Mbps"
        exit
        no shutdown
    exit
    entry 525 create
        description "Deflt policer to limit downstr flow setup rate for 25Mbps Subs"
        match
            traffic-direction network-to-subscriber
            characteristic "access-rate" eq "25M"
        exit
        action
            flow-rate-limit "DefltPol-Sub-FlowRate-DS-25Mbps"
        exit
        no shutdown
    exit
```

Note:
A similar configuration can be implemented for the 100Mbps access rate service option.

Application BW policing (per subscriber)

The following configuration example provides a per AA subscriber peer-to-peer rate limit of 1Mbps. It does not include the app-profile configuration because the ASO characteristic and values can be either statically configured within the app-profile or dynamically signaled through RADIUS or Gx using ASO overrides.

AA subscribers with service characteristic "P2P-Sub-DL" value of "1M" will have a bandwidth policer of 1Mbps applied to peer to peer traffic in the network to subscriber direction:

```
BNG>config>app-assure# group 1
    policer "P2P-Sub-DL-1M" type single-bucket-bandwidth granularity subscriber create
        description "Per-subscr BW policer to limit P2P downstream traffic to 1Mbps"
        rate 1000
        mbs 19
        action permit-deny
    exit

BNG>config>app-assure# group 1:1 policy
    app-service-options
        characteristic "P2P-Sub-DL" create
            value "10M"
            value "1M"
            value "unlimited"
            default-value "unlimited"
        exit

BNG>config>app-assure# group 1:1 policy app-qos-policy
    entry 30 create
        description "Per-subscr BW policer to limit P2P downstream traffic to 1Mbps"
        match
            app-group eq "Peer to Peer"
            traffic-direction network-to-subscriber
            characteristic "P2P-Sub-DL" eq "1M"
        exit
        action
            bandwidth-policer "P2P-Sub-DL-1M"
        exit
        no shutdown
    exit
```

Conclusion

This example provides detailed information to properly configure and use app-profiles, ASOs, and AQPs to successfully configure application policy control rules using Application Assurance.

Application Assurance — Asymmetry Removal

This chapter describes Application Assurance asymmetry removal configurations.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This chapter was originally written for and configured on SR OS Release 11.0.R1. The CLI in the current edition corresponds to SR OS Release 14.0.R4.

The prerequisites for this chapter are a base understanding of AA configuration and operation for single homed deployments. This chapter applies to dual-homed SAPs and spoke SDPs configurations, in a business or residential AA context. AARP is not used for ESM AA subscribers.

Overview

This chapter is intended for Application Assurance (AA) network architects and engineers. It provides best practices recommendations to configure AA asymmetry removal.

Asymmetry means that the two directions of a traffic flow (to-sub and from-sub) take different paths through the network. Asymmetry removal is a means of eliminating traffic asymmetry between a set of dual-homed SAP or spoke SDP endpoints. This can be across endpoints within a single node or across a pair of inter-chassis link connected routers, which is the topology explained in this chapter. Asymmetry removal ensures all packets of a dual-homed AA subscriber are diverted to an AA ISA in order to achieve accurate per subscriber traffic identification and policy enforcement.

Traffic asymmetry is created when there are dual-homed links for a service, and the links are simultaneously carrying traffic. Asymmetry removal for transit subscribers must be implemented in the first routed hop on the network side of the subscriber management point, so there will be a deterministic and fixed SAP/spoke SDP representing the downstream subscriber management node. This ensures there are no more than two paths that the flows can take, both covered by the asymmetry removal solution.

Configuration

Application Assurance Redundancy Protocol (AARP) provides the data plane connectivity for dynamically keeping a dual-homed AA subscriber’s traffic on the same ISA-AA for AA processing. An AARP instance is configured between the dual-homed routers to establish connectivity with the same AARP instance number on each node.

When asymmetry exists between dual-chassis redundant systems, Ipipe spoke SDPs are used to interconnect these services between peer nodes over an Inter-Chassis Link (ICL). The following sections explain the configuration and operation of the services for use with the Application Assurance Redundancy Protocol.

AARP service configuration

The following services must be configured to establish communications between the AARP instances in each of the paired nodes.

- Network topology is a VPRN (or IES) service configured in each node, with a dual-homed SAP from each node to a downstream access element.
- Assumes starting point with AA ISAs installed with identical AA policy and divert enabled in each node.
- Also, the system needs basic routing and LDP configuration for the SDP and the spoke SDPs to be established.

Figure 12: AA asymmetry removal topology

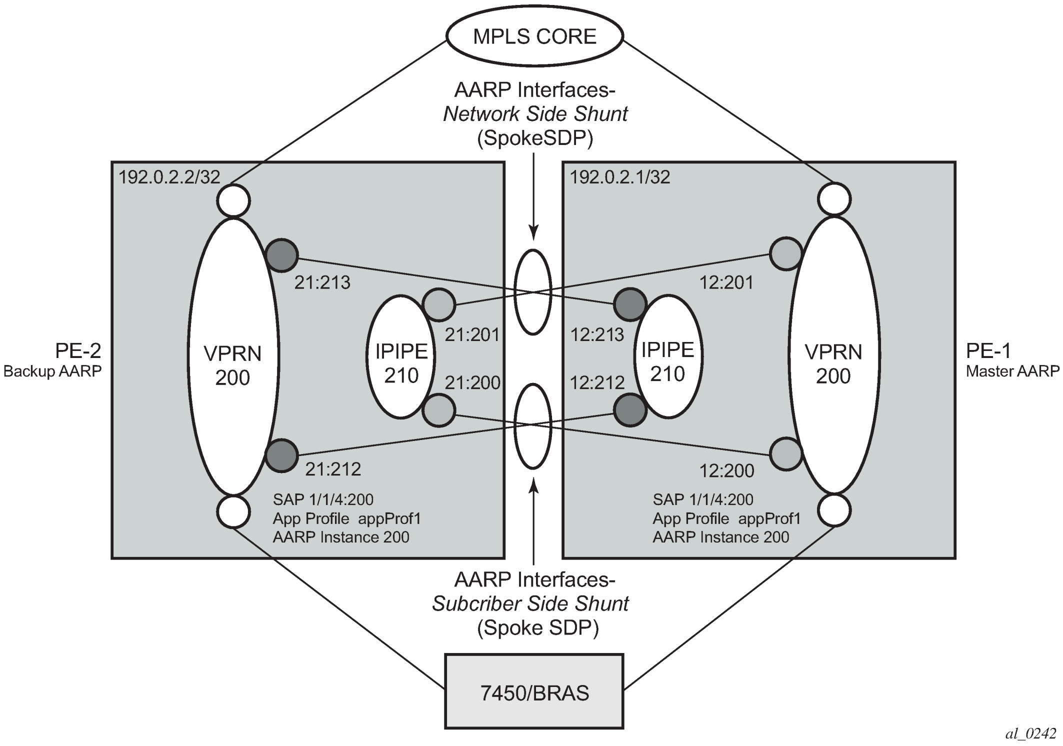


Table 4: AA asymmetry removal topology

On PE-2	On PE-1
system ip: 192.0.2.2	system ip: 192.0.2.1
dual-homed service: 200	dual-homed service: 200

On PE-2	On PE-1
dual-homed sap: 1/1/4:200	dual-homed sap: 1/1/4:200
app-profile diverting: yes	app-profile diverting: yes

Configuration commands for AARP

To enable AARP, AARP instances and AARP interfaces on both nodes must be configured. AARP operation has the following dependencies between the nodes:

- Shunt links configured and operationally up, both subscriber side shunt and network side shunt.
- Peer communications established between nodes, AARP instance operational status will be up when peers are communicating.
- Dual-homed sap/spoke SDP configured with a unique AARP instance (matched by dual-homed interface).
- App-profile configured against sap/spoke SDP with divert enabled (making the sub an aa-sub). The app-profile is the trigger to divert the traffic in the node with the active AARP instance to one of the ISAs in that node, per normal AA divert behavior.

Begin with PE-2:

```
configure
    application-assurance
        aarp 200 create
            description "aarp protecting a dual-homed sap"
            priority 100
            peer 192.0.2.1
            no shutdown
        exit
    exit
exit
```

Ipipe shunt configuration

```
configure
    service
        sdp 21 mpls create
            far-end 192.0.2.1
            ldp
            keep-alive
                shutdown
            exit
            no shutdown
        exit
        ipipe 210 customer 1 vc-switching create
            service-mtu 1556
            spoke-sdp 21:200 create
                aarp 200 type subscriber-side-shunt
                no shutdown
            exit
            spoke-sdp 21:201 create
                aarp 200 type network-side-shunt
                no shutdown
            exit
            no shutdown
        exit
    exit
exit
```

Dual-homed and interface shunt configuration

```
configure
    service
        vprn 200 customer 1 create
            description "VPRN 200 Dual Homed Routed Service"
            aarp-interface "subside_1" create
                spoke-sdp 21:212 create
                    aarp 200 type subscriber-side-shunt
                    no shutdown
                exit
            exit
            aarp-interface "netside_1" create
                spoke-sdp 21:213 create
                    aarp 200 type network-side-shunt
                    no shutdown
                exit
            exit
            interface "int-BRAS-1" create
                sap 1/1/4:200 create
                    aarp 200 type dual-homed
                    app-profile "app-prof-1"
                exit
            exit
            no shutdown
        exit
    exit
exit
```

Then similarly configure the associated AARP configuration on PE-1:

```
configure
    application-assurance
        aarp 200 create
            description "aarp protecting a dual-homed sap"
            priority 200
            peer 192.0.2.2
            no shutdown
        exit
    exit
exit
```

Ipipe shunt configuration

```
configure
    service
        sdp 12 mpls create
            far-end 192.0.2.2
            ldp
            keep-alive
                shutdown
            exit
            no shutdown
        exit
        ipipe 210 customer 1 vc-switching create
            service-mtu 1556
            spoke-sdp 12:212 create
                aarp 200 type subscriber-side-shunt
                no shutdown
            exit
            spoke-sdp 12:213 create
                aarp 200 type network-side-shunt
                no shutdown
            exit
            no shutdown
        exit
    exit
exit
```

Dual-homed and interface shunt configuration

```
configure
    service
        vprn 200 customer 1 create
            aarp-interface "subside_1" create
                spoke-sdp 12:200 create
                    aarp 200 type subscriber-side-shunt
                    no shutdown
                exit
            exit
            aarp-interface "netside_1" create
                spoke-sdp 12:201 create
                    aarp 200 type network-side-shunt
                    no shutdown
                exit
            exit
            interface "int-BRAS-1" create
                sap 1/1/4:200 create
                    description "AA enabled SAP"
                    aarp 200 type dual-homed
                    app-profile "app-prof-1"
                exit
            exit
            no shutdown
        exit
    exit
exit
```

Show commands for AARP

Verify correct configuration on each node. The following output displays the example configuration for PE-1.

Starting with the AARP instance in each node, verify that the AARP instance operational state is up (if everything is properly configured as intended):

```
*A:PE-1# show application-assurance aarp 200
===============================================================================
AARP Instance 200
===============================================================================
Description    : aarp protecting a dual-homed sap
Admin State    : Up                     Oper State     : Up

Local IP       : 192.0.2.1              Peer IP        : 192.0.2.2
Local State    : master                 Peer State     : backup
Local Priority : 200                    Peer Priority  : 100

Local Flags    : none
```

```
Peer Flags     : none
Peer End-Point : none

Master Selection Mode       : minimizeSwitchovers

-------------------------------------------------------------------------------
Service References
-------------------------------------------------------------------------------
Service             Reference           Reference Type
-------------------------------------------------------------------------------
VPRN 200            1/1/4:200           Dual-Homed
Ipipe 210           12:212              Subscriber-Side Pipe Shunt
Ipipe 210           12:213              Network-Side Pipe Shunt
VPRN 200            12:200              Subscriber-Side AARP-Interface Shunt
VPRN 200            12:201              Network-Side AARP-Interface Shunt
-------------------------------------------------------------------------------
No. of service references: 5
-------------------------------------------------------------------------------
===============================================================================
*A:PE-1#
```

Verifying that the AARP instance is up is an indication that the dual-node communications for AARP is working (instance, shunts, etc.). In addition, in the preceding output, verify on both PE nodes that the intended SAPs are dual-homed for that instance.

Now a detailed review of the configured AARP shunt infrastructure services can be shown to make sure they are all properly configured with intended AARP parameters (such as AARP ID and Type on the network and subscriber side shunts) as displayed in the following output:

```
*A:PE-1# show service id 210 all

===============================================================================
Service Detailed Information
===============================================================================
Service Id        : 210                 Vpn Id            : 0
Service Type      : Ipipe
Name              : (Not Specified)
Description       : (Not Specified)
Customer Id       : 1                   Creation Origin   : manual
Last Status Change: 10/03/2016 11:45:51
Last Mgmt Change  : 10/03/2016 11:45:51
Admin State       : Up                  Oper State        : Up
MTU               : 1556
Vc Switching      : True
SAP Count         : 0                   SDP Bind Count    : 2
CE IPv4 Discovery : n/a                 Keep address      : No
CE IPv6 Discovery : n/a                 Stack Cap Sig     : n/a

Eth Legacy Fault Notification
-------------------------------------------------------------------------------
Recovery Timer    : 10.0 secs           Admin State       : outOfService

-------------------------------------------------------------------------------
ETH-CFM service specifics
-------------------------------------------------------------------------------
Tunnel Faults     : ignore


-------------------------------------------------------------------------------
Service Destination Points(SDPs)
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
 Sdp Id 12:212  -(192.0.2.2)
```

```
-------------------------------------------------------------------------------
Description     : (Not Specified)
SDP Id             : 12:212                   Type              : Spoke
Spoke Descr     : (Not Specified)
Split Horiz Grp    : (Not Specified)
VC Type            : Ipipe                    VC Tag            : 0
Admin Path MTU     : 0                        Oper Path MTU     : 1556
Delivery           : MPLS
Far End            : 192.0.2.2
Tunnel Far End     : 192.0.2.2                LSP Types         : LDP
Hash Label         : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label    : Disabled
Entropy Label      : Disabled

Admin State        : Up                       Oper State        : Up
MinReqd SdpOperMTU : 1556
Acct. Pol          : None                     Collect Stats     : Disabled
Ingress Label      : 262141                   Egress Label      : 262139

---snip---

Application Profile: None
Transit Policy     : None
AARP Id            : 200
AARP Type          : subscriber-side-shunt

---snip---

-------------------------------------------------------------------------------
IPIPE Service Destination Point specifics
-------------------------------------------------------------------------------
Configured CE IPv4 Addr: n/a                  Peer CE IPv4 Addr : 0.0.0.0

-------------------------------------------------------------------------------
 Sdp Id 12:213  -(192.0.2.2)
-------------------------------------------------------------------------------
Description     : (Not Specified)
SDP Id             : 12:213                   Type              : Spoke
Spoke Descr     : (Not Specified)
Split Horiz Grp    : (Not Specified)
VC Type            : Ipipe                    VC Tag            : 0
Admin Path MTU     : 0                        Oper Path MTU     : 1556
Delivery           : MPLS
Far End            : 192.0.2.2
Tunnel Far End     : 192.0.2.2                LSP Types         : LDP
Hash Label         : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label    : Disabled
Entropy Label      : Disabled

Admin State        : Up                       Oper State        : Up
MinReqd SdpOperMTU : 1556
Acct. Pol          : None                     Collect Stats     : Disabled
Ingress Label      : 262140                   Egress Label      : 262138

---snip---

Application Profile: None
Transit Policy     : None
AARP Id            : 200
AARP Type          : network-side-shunt

---snip---

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

Next, the configuration of the VPRN service of the dual-homed SAP can be reviewed to ensure it reflects the attached endpoints for the shunt Ipipe spoke SDPs:

```
*A:PE-1# show service id 200 all
===============================================================================
Service Detailed Information
===============================================================================
Service Id        : 200                 Vpn Id            : 0
Service Type      : VPRN
Name              : (Not Specified)
Description       : (Not Specified)
Customer Id       : 1                   Creation Origin   : manual
Last Status Change: 10/03/2016 11:45:51
Last Mgmt Change  : 10/03/2016 11:45:51
Admin State       : Up                  Oper State        : Up

Route Dist.       : 64496:200           VPRN Type         : regular
Oper Route Dist   : 64496:200
Oper RD Type      : configured
AS Number         : None                Router Id         : 192.0.2.1
ECMP              : Enabled             ECMP Max Routes   : 1
Max IPv4 Routes   : No Limit
Auto Bind Tunnel
Resolution        : disabled
Max IPv6 Routes   : No Limit
Ignore NH Metric  : Disabled
Hash Label        : Disabled
Entropy Label     : Disabled
Vrf Target        : target:64496:200
Vrf Import        : None
Vrf Export        : None
MVPN Vrf Target   : None
MVPN Vrf Import   : None
MVPN Vrf Export   : None
Car. Sup C-VPN    : Disabled
Label mode        : vrf
BGP VPN Backup    : Disabled
BGP Export Inactv : Disabled

SAP Count         : 1                   SDP Bind Count    : 2
VSD Domain        : <none>

---snip---

-------------------------------------------------------------------------------
Service Destination Points(SDPs)
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
 Sdp Id 12:200  -(192.0.2.2)
-------------------------------------------------------------------------------
Description     : (Not Specified)
SDP Id             : 12:200                   Type              : Spoke
Spoke Descr     : (Not Specified)
VC Type            : n/a                      VC Tag            : n/a
Admin Path MTU     : 0                        Oper Path MTU     : 1556
Delivery           : MPLS
Far End            : 192.0.2.2
Tunnel Far End     : 192.0.2.2                LSP Types         : LDP
Hash Label         : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label    : Disabled
Entropy Label      : Disabled

Admin State        : Up                       Oper State        : Up
```

```
---snip---

Application Profile: None
Transit Policy     : None
AARP Id            : 200
AARP Type          : subscriber-side-shunt

---snip---

-------------------------------------------------------------------------------
IPIPE Service Destination Point specifics
-------------------------------------------------------------------------------
Configured CE IPv4 Addr: n/a                  Peer CE IPv4 Addr : 0.0.0.0

-------------------------------------------------------------------------------
 Sdp Id 12:201  -(192.0.2.2)
-------------------------------------------------------------------------------
Description     : (Not Specified)
SDP Id             : 12:201                   Type              : Spoke
Spoke Descr     : (Not Specified)
VC Type            : n/a                      VC Tag            : n/a
Admin Path MTU     : 0                        Oper Path MTU     : 1556
Delivery           : MPLS
Far End            : 192.0.2.2
Tunnel Far End     : 192.0.2.2                LSP Types         : LDP
Hash Label         : Disabled                 Hash Lbl Sig Cap  : Disabled
Oper Hash Label    : Disabled
Entropy Label      : Disabled

Admin State        : Up                       Oper State        : Up

---snip---

Application Profile: None
Transit Policy     : None
AARP Id            : 200
AARP Type          : network-side-shunt

---snip---
```

Continuing deeper into the same VPRN service show output, or using the following show command, it can be verified that the dual-homed SAP itself is properly configured and associated with that service and AARP instance:

```
*A:PE-1# show service id 200 sap 1/1/4:200 detail

===============================================================================
Service Access Points(SAP)
===============================================================================
Service Id         : 200
SAP                : 1/1/4:200                Encap             : q-tag
Description        : AA enabled SAP
Admin State        : Up                       Oper State        : Up
Flags              : None
Multi Svc Site     : None
Last Status Change : 10/03/2016 11:45:51
Last Mgmt Change   : 10/03/2016 11:45:51
Sub Type           : regular
Dot1Q Ethertype    : 0x8100                   QinQ Ethertype    : 0x8100
Split Horizon Group: (Not Specified)

Admin MTU          : 1518                     Oper MTU          : 1518
```

```
Ingr IP Fltr-Id    : n/a                      Egr IP Fltr-Id    : n/a
Ingr Mac Fltr-Id   : n/a                      Egr Mac Fltr-Id   : n/a
Ingr IPv6 Fltr-Id  : n/a                      Egr IPv6 Fltr-Id  : n/a
qinq-pbit-marking  : both
                                              Egr Agg Rate Limit: max
Q Frame-Based Acct : Disabled                 Limit Unused BW   : Disabled

Acct. Pol          : None                     Collect Stats     : Disabled

Anti Spoofing      : None                     Dynamic Hosts     : Enabled
Avl Static Hosts   : 0                        Tot Static Hosts  : 0
Calling-Station-Id : n/a

Application Profile: app-prof-1
Transit Policy     : None
AARP Id            : 200
AARP Type          : dual-homed

Oper Group         : (none)                   Monitor Oper Grp  : (none)
Host Lockout Plcy  : n/a
Lag Link Map Prof  : (none)
Bandwidth          : Not-Applicable

---snip---

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

Network to subscriber traffic flow

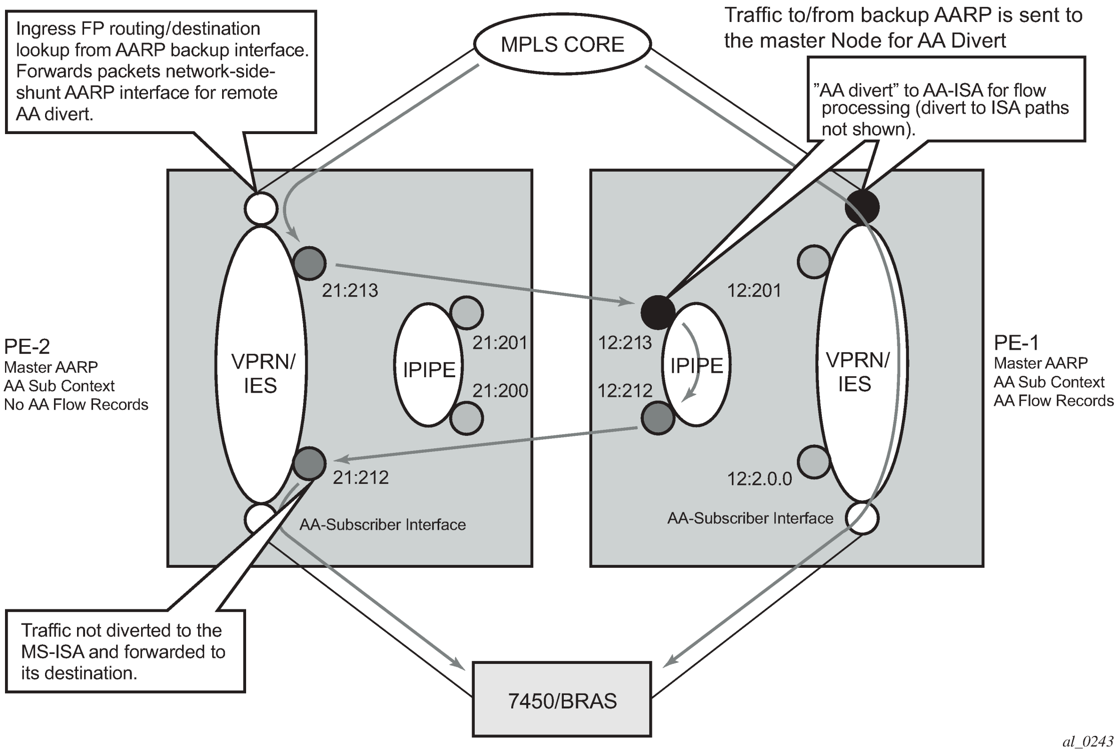
When the AARP is operationally up, AARP tracks which ISA is the master ISA for each dual-homed AARP instance and uses the inter-chassis services (spoke SDP AARP shunts) to move all traffic for each instance traffic to the node with the Master ISA.

Looking at traffic in the network to subscriber direction (Figure 13: Network to subscriber traffic flow):

- Traffic arriving on PE-1 is diverted to the local master ISA, processed, then proceeds to the egress SAP.
- Traffic arriving on PE-2 with the backup AARP interface is sent to the master node for AA processing. The ingress FP forwards packets to network-side-shunt AARP interface for remote AA divert.
- Arriving on PE-1, the packets on the AARP Ipipe are diverted to the master ISA where the packets are processed as if this traffic was traveling in the to-sub direction towards the dual-homed endpoint on PE-1, then returned to PE-2.
- Entering PE-2, the traffic from the subscriber side shunt interface is not diverted to ISAs in that node and egresses on the AARP instance SAP.

With this behavior, traffic always returns to the original ingress node before egressing toward the subscriber (network path for the flows are not modified).

Figure 13: Network to subscriber traffic flow

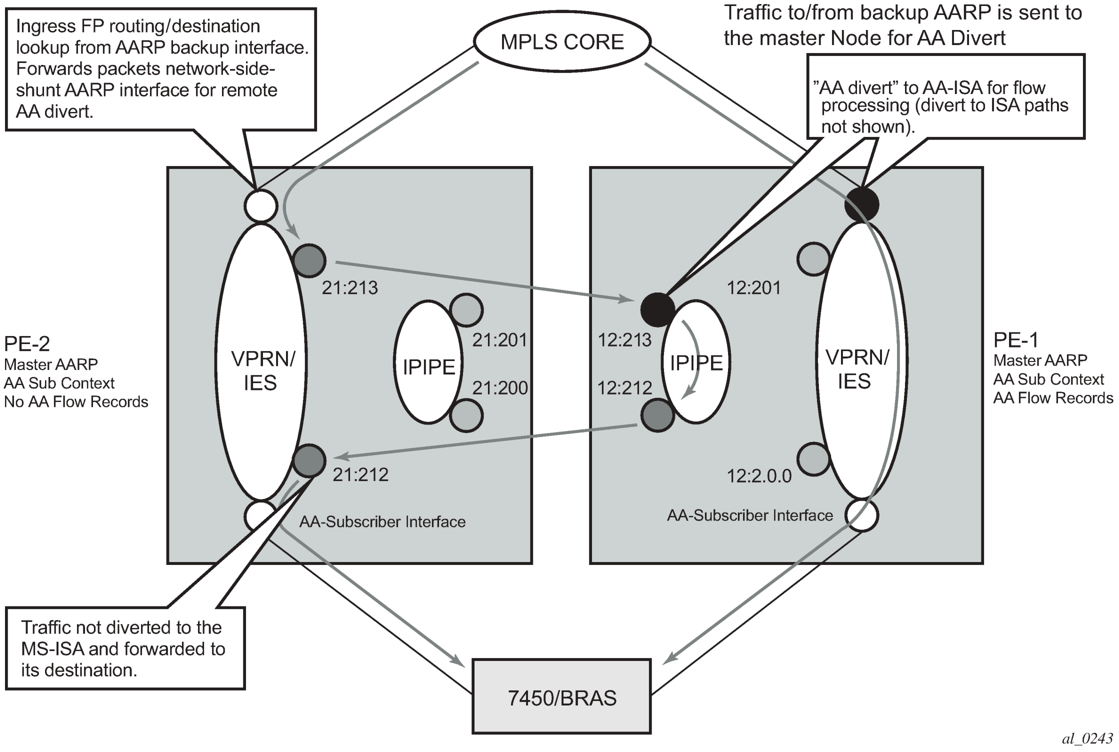

Ingress FP routing/destination lookup from AARP backup interface. Forwards packets network-side-shunt AARP interface for remote AA divert.
MPLS CORE
Traffic to/from backup AARP is sent to the master Node for AA Divert
"AA divert" to AA-ISA for flow processing (divert to ISA paths not shown).
21:213
12:201
PE-2
Master AARP
AA Sub Context
No AA Flow Records
VPRN/
IES
IPIPE
21:201
21:200
12:213
12:212
IPIPE
VPRN/
IES
PE-1
Master AARP
AA Sub Context
AA Flow Records
21:212
12:2.0.0
AA-Subscriber Interface
AA-Subscriber Interface
Traffic not diverted to the MS-ISA and forwarded to its destination.
7450/BRAS
al_0243

Subscriber to network traffic flow

Looking at traffic in the subscriber to network direction (Figure 14: Subscriber to network traffic flow):

- Traffic arriving on PE-1 is diverted to the local master ISA, processed, then proceeds to the egress SAP.
- Traffic arriving on PE-2 with the backup AARP ISA is sent to the master node for AA processing (not diverted to an ISA in PE-2). The ingress FP forwards packets to subscriber-side-shunt AARP interface for remote AA divert.
- Arriving on PE-1, the packets on the AARP Ipipe are diverted to the master ISA where the packets are processed as if the traffic was flowing in the from-sub direction on the dual-homed endpoint, then returned to PE-2 over the Ipipe's AARP subscriber-side-shunt.
- Entering PE-2, the traffic from the network side shunt interface is forwarded by the IES/VPRN service to its destination.

Figure 14: Subscriber to network traffic flow


Ingress FP routing/destination lookup from AARP backup interface. Forwards packets network-side-shunt AARP interface for remote AA divert.
MPLS CORE
Traffic to/from backup AARP is sent to the master Node for AA Divert
"AA divert" to AA-ISA for flow processing (divert to ISA paths not shown).
21:213
12:201
PE-2
Master AARP
AA Sub Context
No AA Flow Records
VPRN/
IES
IPIPE
21:201
21:200
12:213
12:212
IPIPE
VPRN/
IES
PE-1
Master AARP
AA Sub Context
AA Flow Records
21:212
12:2.0.0
AA-Subscriber Interface
AA-Subscriber Interface
Traffic not diverted to the MS-ISA and forwarded to its destination.
7450/BRAS
al_0243

Typical configuration mistakes

Operators configuring AARP can make some typical mistakes listed below that will keep the AARP instance in operational state down:

- The spoke SDP AARP shunt instances' IDs must be aligned with the respective spoke SDP on the peer node: if not, it will result in a flag indicating **shunt mismatch** in the show output.
- Ipipe service MTU alignment — The Ipipe service MTU values must be the same in both nodes, otherwise it will result in the services be in operational status UP, but the AARP instance will remain down.

Conclusion

This chapter is intended for Application Assurance (AA) network architects and engineers to provide the information required to understand and configure dual-node asymmetry removal following the intended service configuration as used by the AARP implementation.

Application Assurance — Best Practices for ISA and Host IOM Overload Protection

This chapter provides information about Application Assurance best practices for ISA and host IOM overload protection.

Topics in this chapter include:

Applicability

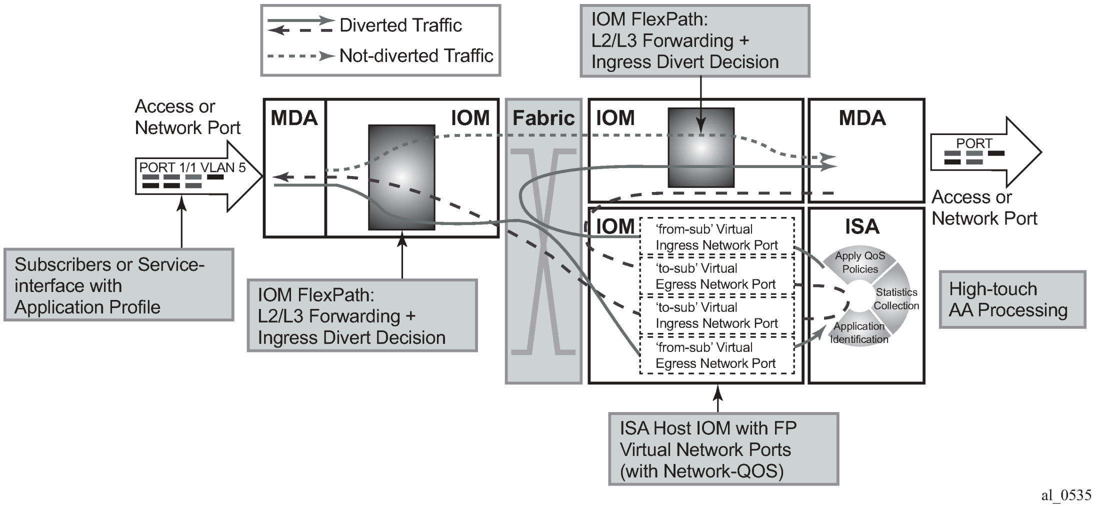
The information and configuration in this chapter is based on SR OS Release 12.0.R4.

Overview

The multiservice integrated services adapter (MS-ISA) is a processing resource module installed on an ISA host IOM. This example describes the best practices for configuration and monitoring of the system to ensure proper engineering of the system resources involved in AA ISA capacity planning.

As shown in Figure 15: System packet datapath to AA ISA, traffic is diverted to an AA ISA by provisioning of an application profile (app-profile) for a subscriber or SAP service context. SR OS then automatically handles traffic diversion for both directions of traffic for that AA subscriber context, through one of the AA ISAs in the AA group where that app-profile is defined.

Figure 15: System packet datapath to AA ISA



The following elements in the SR OS node must be properly engineered for any given AA deployment. Each element is described in this section:

1. ISA capacity cost and load balancing across ISAs.
2. ISA host IOM network egress QoS. Host IOM egress network ports weighted-average shared buffer pool thresholds (within the egress QoS configuration for each group) are used for overload cut-through processing.
3. ISA resources and statistics collection.
   - Flows
   - Traffic volume (bandwidth)
   - Subscribers
   - Flow setup rate
   - ISA overload cut-through
   - ISA default subscriber policies

ISA capacity planning approach

This example illustrates an approach to the configuration of the AA system to address these considerations:

- IOM/ISA-AA network egress QoS configuration should be designed to treat the ISA as a network port with normal network port maximum delay (by MBS).
- Within the ISA, fair access to the ISA-AA bandwidth and flow resources must be ensured: it is recommended that default application QoS policy (AQP) policy entries be configured limiting bandwidth and flow resources per AA subscriber.
- Thresholds for SNMP alerts that indicate a high load on ISA processing should be configured: capacity cost, flow, bandwidth.

- Capacity tracking in live deployments should be performed for parameters that can affect overload: flow setup rate, bandwidth, and subscriber-count per ISA.
- Use of other scale related consumable AA resources against system maximum limits. This includes parameters such as statistics records, transit-ip table entries, and transit-prefix TCAM entries, which should be planned and periodically tracked. These limits will not affect overload of the ISA but may affect intended service operation.
- For recommendations of the specific parameters to watch in a given deployment as well as the values of the system limits for a given release, contact your regional support organization.

AA overload and resource monitoring

Overload is a condition where the total packet processing requirements for traffic arriving on a given ISA exceeds the available resources, resulting in the host IOM egress buffers reaching a configured "overload" threshold. Above this threshold, the ISA can be configured to forward excess traffic (called overload cut-through). If cut-through is not enabled and the overload condition continues, the egress queue MBS threshold will eventually be reached, after which packets will be discarded. Even if overload cut-through is enabled, any egress traffic that exceeds the maximum bus capacity of the ISA queue discard threshold will be discarded.

ISA capacity overload events are supported within the system resource monitoring and logging capabilities if the traffic and resource load crosses any of the following high and low load thresholds on a per-ISA basis. Exceeding one of these thresholds does not in itself indicate an overload state.

1. Host IOM egress network ports weighted-average shared buffer pool thresholds (within the egress QoS configuration for each AA group) are used for triggering and removing overload cut-through processing. Care should be taken in the configuration of these buffers, as the IOM flexpath has significant buffer capacity that can result in latency larger than the network SLA acceptable guidelines. A properly engineered configuration will have large enough buffering to not trigger ISA overload unnecessarily (due to normal bursts with a reasonable traffic load) but will not incur excessive latency prior to triggering the overload state.
2. ISA capacity cost: if the capacity cost of all subscribers on the ISA exceeds a threshold, an event is raised but the overload condition is not set (unless other resources are exhausted). ISA overload or traffic cut-through does not occur simply if capacity cost is exceeded. It is used to capacity plan an intended load for the ISA, proportional to resource use per subscriber, in order to generate events prior to overload to allow appropriate action to mitigate the resource consumption (such as provisioning more ISAs).
3. Flow table consumption (number of allocated flow resources in use): the flow table high-watermark threshold warnings are for proactive notification of a high load. The ISA will cut-through new flows when the "flow resources in use" is at the maximum flow limit. Reaching the flow limit does not generate backpressure to the IOM, nor is the ISA considered in an overload state. Flow usage thresholds are different from bit-rate/packet-rate/flow-setup-rate thresholds in that when the flow table high-watermark threshold is exceeded, the ISA will no longer be operating as application-aware for the flows with no context. The default subscriber policy is applied to traffic that required a flow record but was unable to allocate one, which is a similar behavior to overload cut-through.

   The following terms are used to describe flow resources:

   - Maximum flows: the maximum AA flow table size for a given release.
   - Flows: on the show screens, the "flows" field is an indication of the number of unique 5-tuple entries in the flow table. This includes active and inactive flows; inactive will age out of the table after a period of inactivity that is dependent on the protocol used.

  - Active flows: the number of flows with traffic in the current reporting interval.
  - Flow resources in use: the number of allocated flows in the flow table. This number is greater than the number of active flows, reflecting inactive flows and flows pre-allocated for some dynamic protocols (control + data channels) and for some UDP traffic.

4. Traffic volume: traffic rate in bytes/sec and packets/sec is the dominant cause of ISA overload in most network scenarios, when the ISA is presented with more traffic than it can process. This results in the ISA internal ingress buffers reaching a threshold that causes backpressure to the IOM egress queues (toward the ISA), allowing the ISA to process the packets it already has. This internal backpressure mechanism is normal behavior, allowing burst tolerance at the IOM-to-ISA interface; thus backpressure is not in itself an indication of overload. Overload occurs when the bursts or the load of traffic is sustained long enough to reach the ISA host IOM network port egress weighted-average shared buffer threshold. The actual amount of traffic that can be passed through an ISA is dependent on the application traffic mix, flow density, and AA policy configurations and will vary by network type and by region. The bit-rate and packet-rate watermarks can be used to provide event notification when the traffic rates exceed planning expectations.
5. Flow setup rate: this is generally proportional to total traffic volume, and as such can be a factor in ISA overload. The flow setup rate is the rate at which new flows are presented to the ISA, each resulting in additional tasks that are specific to flow state creation; thus the ISA has a sensitivity to flow setup rates as fewer cycles are available for other datapath tasks when the flow setup rate is high. In residential networks, flow setup rates of 3 k to 5 k flows/sec per Gbps of traffic are common. The flow setup watermarks can be used to provide event notification when the rate exceeds planning expectations.

ISA Overload Models

For an ISA overload strategy, there are two design options for configuring the overload behavior of the system:

- Host IOM egress discards: in this model, the philosophy is to treat AA packet processing resources in the same way as a network interface (of somewhat variable capacity depending on the traffic characteristics). When too much traffic is presented to the ISA, it backpressures the host IOM egress, which will buffer packets. If the egress buffer thresholds are exceeded, the ISA will discard according to the egress QoS slope policy. This is configured by not enabling **isa-overload-cut-through** and use of appropriate egress QoS policies. Firewall or session filter deployments may use this model.
- Overload cut-through: the ISA group can be enabled to cut-through some traffic if an overload event occurs, triggered when the IOM network port weighted-average queues depth exceeds the weighted-average shared high-watermark threshold. In this ISA state, some packets are cut-through from application analysis but retain subscriber context with the default subscriber policy applied. This mode of deployment is intended for situations where it is preferable to forward packets even if not identified by AA than to drop/discard the packet. For example, if AA is providing value-added services (VAS) such as In-Browser Notification (IBN), analytics, or traffic rate limiting, this would usually be the preferred model as the underlying service should be preserved even if capacity to provide the VAS is not available.

Note that even with overload cut-through enabled, there is a hardware-based maximum ISA throughput of approximately 11 Gbps for MS-ISA and 40 Gbps for MS-ISA2. If this is exceeded on a sustained basis, IOM egress discards may still occur.

Understanding Packet and Protocol Cut-Through

Traffic can be cut-through the ISA-AA card on a packet-by-packet basis, in which case packets do not go through AA identification and subscriber application policy. The conditions that trigger cut-through include:

- Overload (IOM egress network port weighted-average shared buffer threshold): excess traffic bypasses all AA processing except for the default subscriber policy
- Non-conformant IP packet: traffic bypasses all AA processing except IP protocol checks and the default subscriber policy. Optionally, these packets can be discarded in AA.
- Flow table full: for new 5-tuples sent to the ISA, if the flow table is full, the packets are cut-through the ISA and only the default subscriber policy is applied.

Note:
The default subscriber policy is a set of AQP rules that apply AQP match criteria limited to Application Service Options (ASO), aa-sub, and traffic-direction starting with the first packet of a flow, with no match conditions based on AA identification (application, app-group, charging-group, IP header). Packets will be either denied_by_default_policy or cut_through_by_default_policy, depending on the policer action configuration in the AQP rules.

For cut-through traffic, no flow records exist but it is counted under per-subscriber protocol statistics as one of the following counters, depending on the case:

- cut_through — Statistics for any packet that could not map to a flow, but that has a valid subscriber ID. This can be an error packet, fragmented out-of-order, no flow resource, invalid TCP flags, etc. This is the most important count for indicating overload cut-through, as it counts all traffic in overload cut-through mode (when the weighted-average threshold has been crossed).
- denied_by_default_policy — Packets that are dropped due to a default policy with a flow-based policer (flow rate or flow count) with action discard.
- cut_through_by_default_policy — Packets that failed to pass flow-based policers with an action of priority-mark.

An example of overload cut-through statistics in the CLI is shown below:

```
A:BNG# show application-assurance group 1 protocol count
===============================================================================
Application-Assurance Protocol Statistics
===============================================================================
Protocol                         Disc          Octets       Packets      Flows
-------------------------------------------------------------------------------
advanced_direct_connect            0%               0             0          0
aim                                0%               0             0          0
amazon_video                       0%               0             0          0
ares                               0%               0             0          0
bbm                                0%               0             0          0
betamax_voip                       0%               0             0          0
bgp                                0%               0             0          0
bittorrent                         0%       678428534       5322929    1036129
cccam                              0%               0             0          0
citrix_ica                         0%               0             0          0
citrix_ima                         0%               0             0          0
cnnlive                            0%               0             0          0
cups                               0%               0             0          0
cut_through                        0%      5299435739      10603771          0
cut_through_by_default_policy      0%               0             0          0
cvs                                0%               0             0          0
```

```
daap                               0%              0             0          0
dcerpc                             0%              0             0          0
denied_by_default_policy           0%              0             0          0
```

Configuration

This example illustrates a typical configuration of an SR OS node for AA for each of the configuration topics.

AA traffic load test environment

Application assurance identifies every byte and every packet of hundreds of real-world applications using per-flow stateful analysis techniques. It is a challenge to find test equipment that can accurately emulate full scale (10 Gbps to 40 Gbps) with traffic mixes and flow behaviors representing hundreds of thousands of end users with application clients across a range of devices. Some specialized stateful test equipment can emulate large traffic rates, but even the best will have equipment-specific patterns and behaviors not representative of live traffic. Therefore, the best scenario to engineer the AA overload configuration is by iteration in live deployments: setting an initial target and modifying the configuration based on ISA performance under load.

For a lab test of ISA throughput and loading, Nokia uses stateful test equipment which supports emulation of various service provider traffic mix profiles suitable for generating overload conditions; however, it is outside the scope of this document to configure AA throughput tests.

The operator should be aware that use of unrealistic, non-stateful traffic generators can result in a high level of unknown traffic, with the ISA performance impacted by continually trying to identify large numbers of packets of no real application type. This, combined with cut-through for invalid IP packets, can result in ISA overload and traffic cut-through (due to overload or invalid IP packets) at traffic levels not representative of actual ISA performance on real traffic.

ISA capacity cost and load balancing across ISAs

These AA group-level commands define the load balancing parameters within an ISA group.

```
*A:BNG# configure isa
    application-assurance-group 1 aa-sub-scale residential create
        no description
        no fail-to-open
        isa-capacity-cost-high-threshold 304000
        isa-capacity-cost-low-threshold 272000
        partitions
        divert-fc be
        no shutdown
    exit
```

The following should be noted related to this configuration:

- Up to 7 primary and 1 backup ISAs are allowed. If the AA services are considered “value added” and not part of a paid service, backups are usually not used because the “fail to fabric” capability keeps the underlying service running.

- The default behavior in case of ISA failure is "no fail-to-open", which means "fail-to-wire"; if an ISA fails, traffic is forwarded as if **no divert** was configured
- Threshold for sending capacity-cost SNMP traps: the unit used for capacity cost is a variable defined in the network design; in this example, it is expressed in Mbps of the subscriber total BW UP+DOWN with a high watermark set to 7600 Mbps × 40 = 304000 (where 40 is an oversubscription ratio). The low watermark is equal to 6800 Mbps × 40 = 272000.
- Partitions should always be enabled to configure additional policies in the future (for example, wifi/business)
- **divert-fc** configuration applies to the AA group: in this example, FC BE Internet is the only diverted FC; this is typical for AA residential and WLan-GW deployments. For VPN services, typically all datapath FCs are diverted to AA.

ISA-AA host IOM - network egress shared memory and QoS

The amount of shared memory allocated per port, along with the network port egress QoS policy, determine the maximum delay for traffic diverted to Application Assurance.

This maximum network port delay is typically determined by the operator and must be used to define the proper QoS configuration to apply to the ISA-AA ports; this QoS configuration may be the same (typically) as what is applied to regular network ports on the SR OS node.

On the line cards there is shared network egress memory per ISA-AA port, with the ISA-AA is represented by two network ports on the host IOM:

- "from-sub": for traffic sent from the subscriber to the network
- "to-sub": for traffic sent from the network to the subscriber

```
configure isa application-assurance-group 1
            qos
                egress
                    from-subscriber
                        pool
                            slope-policy "default"
                            resv-cbs default
                        exit
                        queue-policy "network-facing-egress"
                        port-scheduler-policy "network-facing"
                    exit
                    to-subscriber
                        pool
                            slope-policy "default"
                            resv-cbs default
                        exit
                        queue-policy "network-facing-egress"
                        port-scheduler-policy "network-facing"
                    exit
            exit
            no shutdown
```

The amount of shared memory reserved for each egress network port is determined by the speed of the port (10 Gbps for MS-ISA and 40 Gbps for MS-ISA2) and the **egr-percentage-of-rate** ratio configuration.

MS-ISA uses by default 1000% and 500% of the rate respectively for to-sub and from-sub ports, while MS-ISA2 uses by default 100% for both to-sub and from-sub ports.

It is typically recommended that these values be adjusted when MS-ISA and a high-speed Ethernet MDA are mixed on the same IOM, because in this context the amount of shared memory allocated to the Ethernet MDA should be increased by reducing the MS-ISA network ports memory allocation ratio. If two MS-ISAs are installed on the same IOM, the system will by default allocate 50% of the network egress shared memory to each ISA. In addition, an operator may adjust these values in case the actual network-to-subscriber versus subscriber-to-network ratio is significantly different in the production network, in order to achieve the expected maximum tolerated network delay.

The operator can modify the **egr-percentage-of-rate** per port using the following command:

```
A:BNG# configure port 1/2/fm-sub
A:BNG>config>port# info detail
----------------------------------------------
        modify-buffer-allocation-rate
            egr-percentage-of-rate 500
        exit
----------------------------------------------
A:BNG# configure port 1/2/to-sub
A:BNG>config>port# info detail
----------------------------------------------
        modify-buffer-allocation-rate
            egr-percentage-of-rate 1000
        exit
```

Network egress scheduling/queuing priority is for all ISAs within a group defined at the AA ISA group level

An example below with ISA-AA and 2 x 10G Eth MDA:

```
7750# configure port <slot>/<isa-aa-mda>/fm-sub
    modify-buffer-allocation-rate
        egr-percentage-of-rate 65

7750# configure port <slot>/<isa-aa-mda>/to-sub
    modify-buffer-allocation-rate
        egr-percentage-of-rate 130
```

In this example, the configuration defines:

- from-sub — Approximately 190 msec worth of buffer at 2500 Mbps.
- to-sub— Approximately 190 msec worth of buffer at 5000 Mbps.
- The buffer can be further refined from the network QoS policy.

For MS-ISA2, each MS-ISM flexpath will default the buffer allocation rate to 100%, which is a suitable value assuming that both modules in a slot are MS-ISA2 (which is the MS-ISM configuration), or that the I/O module has a similar traffic rate as the MS-ISA2 (which is also the case in the 10x10GE and 1x100GE versions of the MS-ISA2 line cards).

Configuring ISA resources and stats collection

The following are the key consumable resources in an AA ISA:

- Flows
- Bandwidth
- Subscribers
- Flow setup rate

The AA group should be configured with watermark thresholds where each ISA will generate SNMP events when resources reach this level.

- Per-ISA-card resource usage watermarks trigger SNMP traps to the management system (5620 SAM)
- The values defined below can be refined based on the network characteristics in term of flows and bandwidth per ISA after the initial deployment

```
7750# configure application-assurance
----------------------------------------------
        flow-table-low-wmark 90
        flow-table-high-wmark 95
        flow-setup-high-wmark 66500
        flow-setup-low-wmark 63000
        bit-rate-high-wmark 7600
        bit-rate-low-wmark 6800
```

In this example, the usage SNMP watermarks are configured for:

- Flow table: 95%/90% (maximum 4M flows on MS-ISA)
- Flow setup rate: configured to 95%/90% (of maximum 70k fps on MS-ISA)
- Bit rate/total diverted throughput

The **show app-assure group status detail** command is used to display basic ISA health status:

- # aa-sub, active aa-sub, bitrate, flows in use, flow setup rate
- statistics for all ISAs combined or per ISA

```
A:BNG# show application-assurance group 1 status detail
  ===============================================================================
  Application-Assurance Status
  ===============================================================================
  Last time change affecting status : 05/30/2014 17:18:34
  Number of Active ISAs             : 4
  Flows                             : 214007945881
  Flow Resources In Use             : 2955164
  AA Subs Created                   : 70567
  AA Subs Deleted                   : 10544
  AA Subs Modified                  : 0
  Seen IP Requests Sent             : 0
  Seen IP Requests Dropped          : 0
  -------------------------------------------------------------------------------
                                      Current    Average    Peak
  -------------------------------------------------------------------------------
  Active Flows                      : 2911508    2769454    4582522
  Flow Setup Rate (per second)      : 33923      29400      67865
  Traffic Rate (Mbps)               : 7620       7238       22628
  Packet Rate (per second)          : 1254138    1182571    3044376
  AA-Subs Downloaded                : 69887      66129      70567
  Active Subs                       : 23131      19737      38114
  -------------------------------------------------------------------------------
                                      Packets               Octets
  -------------------------------------------------------------------------------
  Diverted traffic                  : 7437950197613         5530634242355947
  Diverted discards                 : 0                     0
      Congestion                    : 0                     0
      Errors                        : 0                     N/A
  Entered ISA-AAs                   : 7437950180191         5530634229794634
  Buffered in ISA-AAs               : 22                    29849
```

```
Discarded in ISA-AAs              : 97790                 47801217
    Policy                        : 0                     0
    Congestion                    : 0                     0
    Errors                        : 97790                 47801217
Modified in ISA-AAs
    Packet size increased         : 0                     0
    Packet size decreased         : 0                     0
Errors (policy bypass)            : 28283549              21160338635
Exited ISA-AAs                    : 7437950082379         5530634181963568
Returned discards                 : 0                     0
    Congestion                    : 0                     0
    Errors                        : 0                     N/A
Returned traffic                  : 7437950054070         5530634162337570
===============================================================================
```

This can also be run on a per-ISA basis:

```
show application-assurance group 1 status isa <slot/port> detail
```

Note that for MS-ISA2, there is a maximum AA packet rate of 7 M pps; under most known traffic mix scenarios, the ISA should be safely below this packet rate when at maximum bandwidth throughput. However, it is worth periodically checking this value, because if the maximum packet rate is exceeded, and overload cut-through will result. (For MS-ISA, the maximum packet rate supported is high enough to not be feasible with realistic application-based traffic mixes).

The ISA aa-performance record should always be enabled in a network for capacity planning purposes in order to properly plan when to add new ISA cards if required and to monitor the network health:

```
*A:BNG>config>isa# info
----------------------------------------------
        application-assurance-group 1 aa-sub-scale residential create
            no description
            primary <slot/port>
            backup <slot/port>
            no fail-to-open
            isa-capacity-cost-high-threshold 304000
            isa-capacity-cost-low-threshold 272000
            partitions
            statistics
                performance
                    accounting-policy 7
                    collect-stats
                exit
            exit
            divert-fc be
            no shutdown
        exit
```

The commands highlighted in bold above will export information on the total traffic load and resource utilization of the ISA card:

- Flows — active flows, setup rates, resource allocation
- Traffic rates — bandwidth, packets
- Subscribers — active, configured, statistics resource allocation in use

The AA statistics collection configuration refers to accounting policies that are also defined in the SR OS node:

```
*A:BNG>config# log
    file-id 7
        description "ISA Performance Stats"
        location cf2:
        rollover 15 retention 12
    exit
    accounting-policy 7
        description "ISA Performance Stats“
        collection-interval 15
        record aa-performance
        to file 7
        no shutdown
    exit
```

From the AA performance record the following fields in Table 5: Tracking ISA load in the reporting interval can be used as to tracking ISA load in the reporting interval (typically a 15 to 60 minute period):

Table 5: Tracking ISA load in the reporting interval

Record name	Type	Description	Load planning use
dco	cumulative	octets discarded due to congestion in MDA	Should be 0; ISA internal congestion
dcp	cumulative	packets discarded due to congestion in MDA	Should be 0; ISA internal congestion
dpo	cumulative	octets discarded due to policy in MDA	Not related to load planning
dpp	cumulative	packets discarded due to policy in MDA	Not related to load planning
pbo	cumulative	octets policy bypass	Not used. Traffic was for an invalid subscriber and the group was "no fail-to-open"
pbp	cumulative	packets policy bypass	Not used. Traffic was for an invalid subscriber and the group was "no fail-to-open"
nfl	cumulative	number of flows	informative
caf	intervalized	current active flows	informative
aaf	intervalized	average active flows	informative
paf	intervalized	peak active flows	Check vs max
cfr	intervalized	current flow setup rate	informative
afr	intervalized	average flow setup rate	Check meets expected norms; increasing over time increases load

Record name	Type	Description	Load planning use
pfr	intervalized	peak flow setup rate	informative
ctr	intervalized	current traffic rate	informative
atr	intervalized	average traffic rate	Check meets expected norms; increasing over time increases load
ptr	intervalized	peak traffic rate	Check vs max
cpr	intervalized	current packet rate	informative
apr	intervalized	average packet rate	informative
ppr	intervalized	peak packet rate	informative
cds	intervalized	current diverted subscribers	informative
ads	intervalized	average diverted subscribers	informative
pds	intervalized	peak diverted subscribers	Check vs max and expected norms; increasing over time increases load
rfi	intervalized	flows in use	Check vs max and expected norms; increasing over time increases load
rcc	cumulative	ISA capacity cost	Check meets expected norms; increasing over time increases load

The intended deployment model is for this statistic record to be collected by a centralized service-aware management system along with all other AA records and to be stored in a reporting and analysis management database for subsequent analytics purposes, such as trending charts or setting thresholds of key values. It is recommended that a CRON script be used to export the AA performance record to a storage server for post processing if no reporting and analysis management tool is deployed:

- If no reporting and analysis management tool is deployed in the network, it is possible to automatically collect the XML accounting files and provide high-level reporting through an XML-to-CSV conversion.
- The simplest approach is to configure a CRON script on the SR OS node to automatically retrieve the CF accounting file (alternatively, any other scripting mechanism with an interval smaller than the retention period can be used)
- It is recommended that the rollover interval of the file-id policy be modified to 6H or above in order to collect fewer files while keeping the same collection interval.

```
*A:BNG# file type cf2:/script
file copy cf2:/act/*.gz ftp://login:password@IP-ADDRESS/acct/router1/

*A:BNG>config>cron# info
----------------------------------------------
        script "test-ftp-act"
            location "cf2:/script"
            no shutdown
        exit
        action "cron1"
```

```
            results "ftp://login:password@IP-ADDRESS/results/router1-result.log"
            script "test-ftp-act"
            no shutdown
        exit
        schedule "schedule1"
            interval 36000
            action "cron1"
            no shutdown
        exit
```

The schedule : interval 36000 is in seconds (10 hours).

With this XML to CSV export mechanism, a spreadsheet can be used by the network engineer to periodically track the ISA resource utilization.

ISA overload cut-through

The system can be configured to react to overload based on the weighted-average (WA) queue depth of the shared network port buffer pool from-sub and to-sub. Overload cut-through is typically recommended for use of AA for value-added services where, in the event of overload, the preference is for the ISA to continue to pass packets without AA processing. However, firewall use cases will prefer to drop excess traffic in the event of overload, in which case overload cut-through may not be desired.

In addition to triggering an alarm, further packets sent to the ISA after the WA high-watermark threshold is reached are cut-through immediately by the ISA card without application identification or subscriber policy processing, if the **isa-overload-cut-through** command is enabled.

The WA queue depth is typically configured based on the maximum tolerated delay for the service diverted and the amount of shared buffer space allocated from the IOM.

AA deployment recommended settings:

- high watermark — 33% of the maximum MBS for all diverted network queues
- low watermark — 5% of the maximum MBS for all diverted network queues

The recommended high and low watermarks assume that the sum of the network port egress queues MBS size is 100% of the shared buffer. If this network queue maximum size is further reduced in the network QoS policy, the watermark values must be adapted proportionally; for example, if the total MBS size cannot exceed 50% of the shared buffer, then the watermark values would be divided by 2: the high watermark = 33% / 2 = 16%, the low watermark = 5% / 2 = 2%. Adjusting the MBS and the **wa-shared-high-wmark** and **wa-shared-low-wmark** values proportionately ensures that the MBS point (after which discards occur) is above the WA shared high-watermark threshold; otherwise, the ISA will not ever overload if MBS discards are occurring first.

```
A:BNG# configure isa application-assurance-group 1
            isa-overload-cut-through
            qos
                egress
                    from-subscriber
                        wa-shared-high-wmark 16
                        wa-shared-low-wmark 2
                    exit
                    to-subscriber
                        wa-shared-high-wmark 16
                        wa-shared-low-wmark 2
                    exit
                exit
            exit
```

The **show isa group** commands can be used to verify that overload cut-though is enabled.

```
*A:BNG>show isa application-assurance-group 1
=============================================================
ISA Application-assurance-groups
=============================================================
ISA-AA Group Index          : 1
Description                 : (Not Specified)
Subscriber Scale            : residential
WLAN GW Group Index         : N/A
Primary ISA-AA              : 1/2 up/active
Backup ISA-AA               : 2/1 down
Last Active change          : 07/02/2014 12:17:45
Admin State                 : Up
Oper State                  : Up
Diverted FCs                : be
Fail to mode                : fail-to-wire
Partitions                  : enabled
QoS
  Egress from subscriber
    Pool                    : default
      Reserved Cbs          : default
      Slope Policy          : default
    Queue Policy            : default
    Scheduler Policy        :
  Egress to subscriber
    Pool                    : default
      Reserved Cbs          : default
      Slope Policy          : default
    Queue Policy            : default
    Scheduler Policy        :
Capacity Cost
    High Threshold          : 4294967295
    Low Threshold           : 0
Overload Cut Through        : enabled
Transit Prefix
    Max IPv4 entries        : 0
    Max IPv6 entries        : 0
    Max IPv6 remote entries : 0
HTTP Enrichment
    Max Packet Size         : 1500 octets
========================================================================
```

To monitor the load status of an ISA, enter the following CLI command.

```
*A:BNG>show application-assurance group 1 status isa 5/1 cpu
==========================================
Application-Assurance ISA CPU Utilization
(Test time 993791 uSec)
==========================================
Management CPU Usage
------------------------------------------
Name               CPU Time    CPU Usage
                  (uSec)
------------------------------------------
System               14277        1.43%
Management           61101        6.15%
Statistics           69850        7.02%
Idle                848563       85.39%
==========================================
Datapath CPU Usage
------------------------------------------
Name              CPU Time     CPU Usage
```

```
                  (uSec)
------------------------------------------
System              14277         1.43%
Packet Processing   61101         6.15%
Application ID      69850         7.02%
Idle               848563        85.39%
```

Additionally, the system log files can be used to examine the AA overload history to determine when the overload state was entered and exited. It can be helpful to send AA events to a separate log using the following configuration:

```
    log
        filter 45
            default-action drop
            entry 10
                action forward
                match
                    application eq "application_assurance"
                exit
            exit
        exit
        log-id 45
            description "application-assurance log"
            filter 45
            from main
            to memory 500
        exit
```

The log files can then be examined to see if overload has occurred, and how frequently. If overload occurs with any regularity, it is a situation that should be addressed. Below is an example of a log file showing AA overload:

```
A:BNG# show log log-id 45
===============================================================================
Event Log 45
===============================================================================
Description : application-assurance log
warning: 13 events dropped from log
Memory Log contents  [size=500   next event=16  (not wrapped)]

15 2014/08/14 17:00:32.66 EST WARNING: APPLICATION_ASSURANCE #4433 Base
"ISA AA Group 1 MDA 5/1 exiting overload cut through processing."

14 2014/08/14 17:00:32.55 EST WARNING: APPLICATION_ASSURANCE #4431 Base
"ISA-AA group 1 MDA 5/1 wa-shared buffer use is less than or equal to 1% in the to-subscriber
 direction or corresponding tmnxBsxIsaAaGrpToSbWaSBufOvld notification has been disabled."

13 2014/08/14 17:00:32.06 EST WARNING: APPLICATION_ASSURANCE #4432 Base
"ISA AA Group 1 MDA 5/1 entering overload cut through processing."

12 2014/08/14 17:00:32.05 EST WARNING: APPLICATION_ASSURANCE #4430 Base
"ISA-AA group 1 MDA 5/1 wa-shared buffer use is greater than or equal to 35% in the to-
subscriber direction."
```

The primary indicator to look at in CLI statistics for ISA load indication is datapath CPU usage. Regardless of the configuration and traffic profiles in use, datapath CPU usage gives a consistent indication of whether the ISA is under heavy load (the cause of overload is the inability of the ISA to perform more tasks). The idle datapath time is not proportionate to bandwidth throughput, but if idle datapath CPU usage is under 5%, this indicates an approaching maximum processing load.

At an average datapath use of 95-100% (less than 5% idle) the ISA is creating latency and backpressuring the host IOM egress. It is the best way to know how close to overload the ISA has been. Attempting to examine data throughput statistics such as bit rate, flow setup rate and packet rate to predict overload is not recommended, as these are quite variable under normal circumstances and are not directly correlated to overload. Once in overload, the data statistics (volume, setup rate, and so on) are useful for determining what threshold traps to put in place for the future, but the needed thresholds will always be specific to the live deployment traffic mix and policy configuration.

Below is an example of the status for an ISA that is fully loaded but not yet in overload:

```
*A:BNG>show application-assurance group 1 status isa 5/1 cpu
==========================================
Application-Assurance ISA CPU Utilization
==========================================

----------------------------------------------
Management CPU Usage (Test time 999636 uSec)
----------------------------------------------
Name                  CPU Time     CPU Usage
                      (uSec)
----------------------------------------------
System                    1540         0.15%
Management                  14        ~0.00%
Statistics              643955        64.42%
ICAP Client                603         0.06%
Idle                    353524        35.37%
----------------------------------------------

----------------------------------------------
Datapath CPU Usage   (Test time 999735 uSec)
----------------------------------------------
Name                  CPU Time     CPU Usage
                      (uSec)
----------------------------------------------
System                  188374        18.84%
Packet Processing       534203        53.43%
Application ID          277158        27.72%
Idle                         0         0.00%
----------------------------------------------
```

In this example, 0% idle datapath CPU means the ISA is fully used. When the Datapath CPU Usage Idle average is in the 5-10% range consistently, the ISA should be considered "full"; to add new subscribers, more ISAs are required.

If the excessive traffic condition persists, backpressure from the ISA to the IOM will buffer packets in the egress buffers, and when the egress MBS is exceeded, the ISA host IOM will indicate diverted discards due to congestion if cut-through is not enabled:

```
*A:BNG>show application-assurance group 1 status detail
===============================================================================
Application-Assurance Status
===============================================================================
Last time change affecting status : 08/12/2014 13:16:15
Number of Active ISAs             : 1
Flows                             : 235754165
Flow Resources In Use             : 12000000
AA Subs Created                   : 14224
AA Subs Deleted                   : 0
AA Subs Modified                  : 1
Seen IP Requests Sent             : 0
Seen IP Requests Dropped          : 0
```

```
-------------------------------------------------------------------------------
                                    Current    Average    Peak
-------------------------------------------------------------------------------
Active Flows                      : 8452434    3786948    10632607
Flow Setup Rate (per second)      : 246578     65104      298677
Traffic Rate (Mbps)               : 33702      13229      35813
Packet Rate (per second)          : 6847697    2466118    6945936
AA-Subs Downloaded                : 14224      13710      14224
Active Subs                       : 14224      9934       14224
-------------------------------------------------------------------------------
                                    Packets               Octets
-------------------------------------------------------------------------------
Diverted traffic                  : 8924242848            5983284952320
Diverted discards                 : 752486                729147667
    Congestion                    : 752486                729147667
    Errors                        : 0                     N/A
Entered ISA-AAs                   : 8923417360            5982508976617
Buffered in ISA-AAs               : 57                    19277
Discarded in ISA-AAs              : 0                     0
    Policy                        : 0                     0
    Congestion                    : 0                     0
    Errors                        : 0                     0
Modified in ISA-AAs
    Packet size increased         : 0                     0
    Packet size decreased         : 0                     0
Errors (policy bypass)            : 0                     0
Exited ISA-AAs                    : 8923417303            5982508957340
Returned discards                 : 0                     0
    Congestion                    : 0                     0
    Errors                        : 0                     N/A
Returned traffic                  : 8923285123            5982432640249
===============================================================================
```

ISA default subscriber policy

Default subscriber policy — AQP with match criteria not using App-ID or 5-tuple. Match **only** includes traffic direction and/or ASO characteristic and/or subscriber name.

It is recommended that each ISA be configured with some default subscriber policies that get applied to all subscribers at all times, independent of application flow ID, and even when an ISA is in overload cut-through. These policies protect the ISA resources and provide fairness of resource allocation between subscribers by limiting the ISA resources that can be consumed by a single subscriber. A starting point for the recommended policers is (in all cases, network-specific tuning is recommended):

- Per-subscriber flow rate policer: value more than expected maximum peak per-subscriber rate for active subscribers. The policer protects one subscriber from attacking the network with an excessive flow rate and affecting ISA flow rate resources used by other customers. A typical rate for residential networks could be 100 fps per subscriber.
- Per-subscriber flow count policer: value more than expected maximum per-subscriber flow count for active subscribers. The policer protects one subscriber from consuming excessive flow counts and affecting ISA flow resources used by other customers.
- Downstream bandwidth per subscriber: to a value more than the maximum rate supported by the service, or to less than the maximum per-subscriber capability of the ISA, whichever is lower. For fixed networks, several default policer rates are recommended using a per-sub ASO value for low, medium and large rate ranges set at a rate related to the subscriber access speed. For example, for an FTTH service the per-sub policers could be set at 3 value ranges: below 25Mbps, with another at 100Mbps sub policer for services between 25Mbps and 100Mbps, and another sub-policer for rates between

100Mbps and 1Gbps. The settings for a mobile 3G network rate may be 1Mbps and in an LTE network the rate may be vhcc10Mbs.

For a CLI example of a default subscriber policy, see Application Assurance — App-Profile, ASO and Control Policies.

Conclusion

Any deployment of Application Assurance should include careful capacity planning of the ISA resources, with an appropriate ISA overload strategy, whether for overload cut-through to keep excess traffic flowing, or with a discard policy engineered in the host IOM egress QoS policies.

ISA resource use should be monitored via appropriately configured resource thresholds, events, log files, XML records and show screens to ensure that sufficient ISA resources are available as required.

Application Assurance — DNS IP Cache

This chapter describes the Application Assurance - DNS IP Cache.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 15.0 R5.

Overview

This chapter is intended for network architects and engineers working with Application Assurance (AA). It provides an example to configure the DNS IP Cache, which can be used by operators to ensure that traffic will be correctly classified to prevent charging fraud.

This chapter assumes that the reader has a basic understanding of AA policy (application-filters, applications, application-qos-policies). Additionally, the reader is expected to know the basics of HTTP and DNS.

Nokia recommends using the AA policy configuration templates provided by the Nokia Application Database (AppDB). The AppDB contains an optimized configuration to classify all supported applications and includes all relevant application groups, applications, and app-filters. It can be obtained through Technical Support.

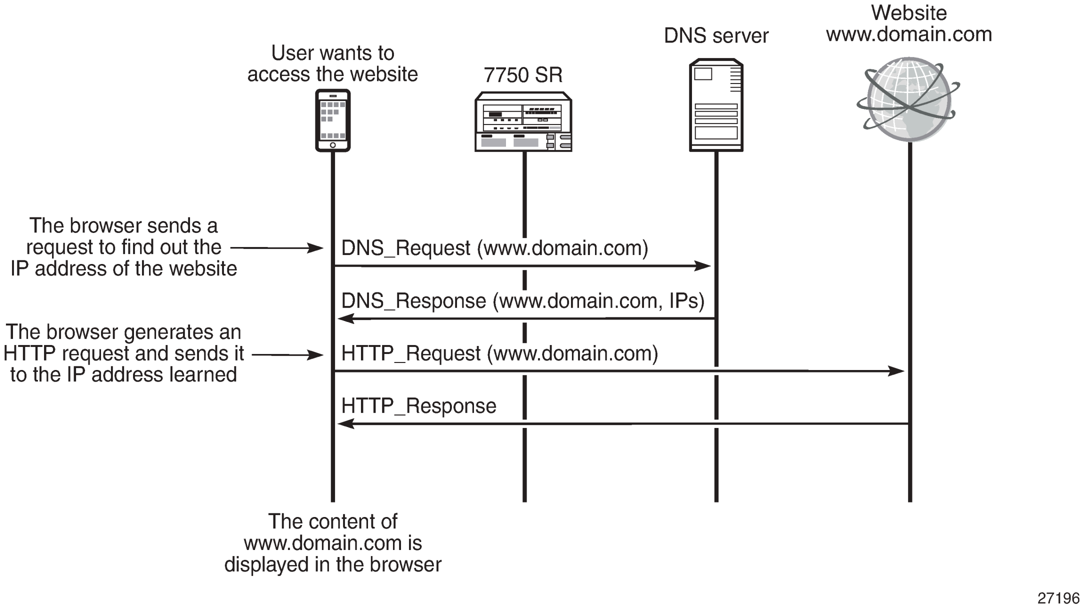
AA uses various methods to classify traffic. Criteria include L3/4 information (IP addresses and ports), or L7 information (HTTP hostnames or other HTTP headers). Classification of Internet services using L3/4 information alone is not very common because IP addresses may change (service providers may add or remove IP addresses), or IP addresses may be region-specific. A more reliable method is to use L7 information, usually the HTTP hostname.

To access a website, after receiving the URL, a browser will perform two basic tasks:

1. Generate a DNS request to resolve the hostname to an IP address
2. Send an HTTP request to the web server, using the IP address from the preceding step, and display the web server response to the user browser

Figure 16: Basic Message Flow When Accessing Website shows the basic message flow.

Figure 16: Basic Message Flow When Accessing Website


Website
www.domain.com
DNS server
User wants to access the website
7750 SR
The browser sends a request to find out the IP address of the website
DNS_Request (www.domain.com)
DNS_Response (www.domain.com, IPs)
The browser generates an HTTP request and sends it to the IP address learned
HTTP_Request (www.domain.com)
HTTP_Response
The content of www.domain.com is displayed in the browser
27196

HTTP is not a strict protocol. Several web server implementations have not implemented the standard completely and may allow requests that appear invalid. The most common case is an HTTP request in which the destination IP address of the packet does not correspond to the hostname.

For example, consider that:

- The IP address of www.domain1.com is a.b.c.d (IP1).
- The IP address of www.domain2.com is e.f.g.h (IP2).

If the web server running at IP1 receives an HTTP request containing the hostname www.domain2.com , it will reply and provide the content of www.domain1.com . This may happen because several web server implementations (for example, HTTP1.0) ignore the hostname requested.

This may lead to misclassification and incorrect handling of the flow. More specifically, some users may take advantage of this, as follows:

- Assume that www.domain2.com is a free-of-charge (zero rated) site and www.domain1.com is a chargeable site. Any classification and charging function (including AA) that uses only the hostname will classify the traffic as traffic toward www.domain2.com .
- As a result, the user will receive the content of www.domain1.com for free.

There are Internet applications that use this technique to bypass the operator charging policies. For malicious purposes, some applications manipulate (forge) the HTTP headers. Subscribers appear to be browsing a zero-rated site, while traffic is directed to a chargeable site. The applications tunnel traffic to a proxy server. The HTTP host header is a zero-rated site and the host that the subscriber wants to visit is inserted in another header (usually the x-online-host header).

To protect against such host header attacks, AA has implemented the DNS IP Cache feature.

AA will populate a cache with the DNS responses generated when the user requests to access a website, by inspecting DNS traffic. During traffic classification, apart from checking the hostname included in the

HTTP request, AA will ensure that the destination IP of the packet matches an entry contained in the cache table.

For example, consider that the cache currently contains the following:

Domain	IP address
www.domain1.com	172.20.1.135 172.20.1.136

The operator can configure an app-filter with match criteria, as follows:

- Hostname equal to www.domain1.com
- Server address contained in the cache

The system will then verify that the packet contains the hostname configured and the packet destination is an IP address contained in the cache. Only if both criteria are true, the app-filter will match.

Regardless of how access to www.domain1.com is charged, the operator will be certain that this traffic will be correctly classified and charged and no other traffic will be charged at this rate (which may be zero-rated).

For additional protection, the operator can also configure a list of trusted DNS servers. The cache will be populated only with the responses from those DNS servers. This ensures that the cache will not be corrupted with responses from malicious DNS servers.

Operators are advised to use the DNS IP Cache feature for any web traffic that the operator charges at a lower rate or for web traffic that is free. This includes operator websites or partner sites.

Some users will attempt to exploit the operator configuration, trying to obtain content at reduced rates or free of charge. Enabling the feature in the configuration will validate the hostname included in the HTTP request, preventing charging fraud.

Configuration

The example configuration comprises a 7750 SR configured with ISA-AA. The configuration requires a web client and Internet access through SR. The same general DNS IP cache configuration also applies to AA deployed in VSR, MG, and CMG systems.

To pass traffic between two endpoints and verify the configuration, an Epipe service can be created.

The following example assumes:

- The user has the IP address 192.168.2.11.
- The zero-rated website is "www.nokia.com".
- The IP address of the trusted DNS server is 192.0.2.1.

The following sections show how to configure a DNS IP cache to contain the IP addresses of the website www.nokia.com . Only the IP addresses provided by the trusted DNS server will be added in the cache. Finally, an app-filter will be configured whose match criteria will be both the domain and the IP addresses contained in the cache.

To verify the configuration, the operator needs to generate DNS traffic and HTTP traffic to www.nokia.com (as shown in Figure 16: Basic Message Flow When Accessing Website).

Creating a DNS IP Cache

Create a **dns-ip-cache** entry and populate it with DNS responses for "*.nokia.com$", only from the DNS server 192.0.2.1:

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
dns-ip-cache "nokia-cache" create
    description "dns cache for nokia"
    dns-match
        domain "domain1" expression "*.nokia.com$"
        server-address 192.0.2.1
    exit
no shutdown
----------------------------------------------
```

Create an app-filter

Nokia recommends using the AA AppDB for a complete library of application and app-filter definitions. The AppDB contains verified configurations for all common applications. Operators are advised to consult the AppDB, then edit one of the predefined app-filters for the application to be verified.

It is common for operators to define app-filters for domains that are not included in the AppDB, such as on-net websites:

```
app-filter
        entry 100 create
            expression 1 http-host eq "*.sponsor1-operator.com$"
            server-address eq dns-ip-cache "dns-ip-cache1"
            application "Sponsor Content #1"
            no shutdown
```

Create the Application for the Nokia Website

The configuration is as follows:

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
policy
    begin
    application "Nokia NET" create
        description "Nokia website"
        app-group "Web"
    exit
    commit
exit
----------------------------------------------
```

Create an App-filter for Nokia Traffic, Which Uses the DNS IP Cache

In the example, it is assumed that Nokia content is fetched from the host www.nokia.com .

```
*A:Dut-C>config>app-assure>group# info
---------------------------------------------
policy
   begin
    app-filter
        entry 61000 create
            description "Nokia Web Access"
            expression 1 http-host eq "*.nokia.com$"
            server-address eq dns-ip-cache "nokia-cache"
            application "Nokia NET"
            no shutdown
        exit
    exit
    commit
exit
----------------------------------------------
```

The preceding app-filter will match traffic if the hostname matches *.nokia.com$ and if the destination IP address of the packet matches an entry in the cache.

Create application-qos-policy entries

Create a default subscriber policy to inspect DNS responses. This application QoS policy (AQP) is used to populate the cache. No match criteria are needed. AA will only analyze DNS traffic to populate the cache.

```
*A:Dut-C>config>app-assure>group# info
---------------------------------------------
policy
    begin
    app-qos-policy
        entry 50 create
            action
                dns-ip-cache "nokia-cache"
             exit
             no shutdown
        exit
    exit
    commit
exit
----------------------------------------------
```

Operational Considerations

1. When configuring the DNS IP cache (chapter 3.1), configuration of the following parameters was omitted:
    - size (10)
    - high-wmark (90)
    - low-wmark (90)

The default values—shown in parentheses—were used.

- **size** refers to the maximum number of entries in the cache (it can be between 10 and 32000). The cache size is set to values above 1000 in live deployments, depending on the operator use cases. If the cache becomes full, new IP addresses will be ignored.
- **high-wmark** defines the high watermark value for the DNS IP cache (percentage). When the number of IP addresses stored in the cache exceeds the threshold defined, the system will generate a trap.
- **low-wmark** defines the low watermark value for the DNS IP cache (percentage). Assuming the high watermark value was reached and a trap was generated, the system will clear the trap if the number of IP addresses stored in the cache drops below the low watermark value.

2. Apart from the number of IP addresses that can be stored in the cache (10 in our example), there are also a maximum number of domains that can be stored. These maximum values are indicated in the **show** command outputs described later.
3. The AQP configured to populate the cache uses a default subscriber policy. A default subscriber policy does not contain any match criteria against application (application, app-group, charging-group) or L3/4 information (IP address or prefix, TCP/UDP port, IP protocol, DSCP). The AQP will be applied to all matching flows starting with the first packet of a flow. Match criteria can only be by application service options, traffic direction, and optional AA subscriber name.
4. The app-filter was created by consulting app-filters defined in the AppDB. The entry ranges have been defined in the AppDB. The ranges defined ensure that after Nokia upgrades the AppDB, the custom protocols (defined by the operator) will not be affected. Therefore, the operator should use an entry in the correct range, so that the configuration is not removed after a policy upgrade done with the policy sync tool. User-defined application groups must prefix their description with the text "Custom", thus facilitating future policy upgrades.
5. If the cache becomes full, new DNS responses will still increase the "Cache full count" and will still get a hit on "DNS Total responses", "Domain name matched", and "Domain & server matched", assuming that they match (see the next section for the **show** command output). An IP address will expire and be removed from the cache if it is not included in any DNS response for a specified period of time.

Troubleshooting and Debugging

Using the **tools** and **show** commands, the operator can verify whether DNS responses are present in the cache and that the configuration is valid.

Show Commands

The following show commands can be used to evaluate the configuration created and verify whether traffic has matched.

The following command shows the DNS IP cache:

```
*A:Dut-C# show application-assurance group 1 dns-ip-cache "nokia-cache"

================================================================================
Application Assurance Group 1 dns-ip-cache "nokia-cache"
================================================================================
Admin Status                : Up
AQP Ref                     : Yes
```

```
Domain expressions          : 1 (out of 32)
Server addresses            : 1 (out of 64)
High watermark              : 90%
Low watermark               : 80%
Cache size                  : 10

-------------------------------------------------------------------------------
ISA                        Usage         (%)        Alarm             Hit Count
                                                    State
-------------------------------------------------------------------------------
1/2                            0        0.00        clear                     8
3/1                            0        0.00        clear                     8
===============================================================================
```

The preceding output provides the following information:

- The administrative status of the DNS IP cache "nokia-cache" is "Up". An administrative status of "Down" indicates that the DNS IP cache is not in use.
- The DNS IP cache "nokia-cache" is referenced by an AQP.
- The number of configured domain expressions (in our example, "*.nokia.com$") equals 1 and the maximum value (shown in brackets) is 32.
- The number of server addresses is 1 and the maximum value (shown in brackets) is 64.
- The number of times a cache lookup was successful (hit count) is 8. A hit count of 0 indicates a possible configuration error.

The following command displays the content of each ISA card:

```
*A:Dut-C# show application-assurance group 1 dns-ip-cache "nokia-cache" isa 1/2

===============================================================================
Application Assurance Group 1 dns-ip-cache "nokia-cache" ISA 1/2
===============================================================================
Admin Status                : Up
AQP Ref                     : Yes
Domain expressions          : 1 (out of 32)
Server addresses            : 1 (out of 64)
High watermark              : 90%
Low watermark               : 80%
Cache size                  : 10

-------------------------------------------------------------------------------
ISA 1/2 DNS Stats
-------------------------------------------------------------------------------
DNS
    Total responses         : 5
    Domain name matched     : 5
    Domain & server matched : 5
Cache
    Total entries added     : 2
    Total entries removed   : 1
Usage                       : 1 (10.00%) threshold alarm clear
    Full count              : 0
    Hit count               : 8
    Miss count              : 0
===============================================================================
```

The preceding output provides additional information about the:

- total number of DNS responses that were analyzed (Total responses)

- number of times the domain name defined in the DNS match criteria matched a DNS response (Domain name matched)
- number of times both the domain name and server address defined in the DNS match criteria matched a DNS response (Domain & server matched)
- total number of entries added in the cache (2 in the preceding example)
- total number of entries removed from the cache (1 in the preceding example). An entry will be removed from the cache if an internal timer expires. This ensures that the cache does not maintain very old IP addresses, which may no longer be valid. If the IP address is re-learned, the timer will be reset.
- successful/unsuccessful IP address lookups in the cache (Hit/Miss count)

If one or more of the preceding values indicates an error, the configuration should be checked. The domain/server configured may not match the hostname requested (for example, the operator configured the domain "*.nokia.com" while the user requested the hostname "*.nokia.co.uk").

The following command shows the created app-filter:

```
*A:Dut-C# show application-assurance group 1 policy app-filter 61000
                app-filter
                    entry 61000 create (0 flows, 0 B)
                        description "Nokia Web Access "
                        expression 1 http-host eq "*.nokia.com$"
                        server-address eq dns-ip-cache "nokia-cache"
                        application "Nokia NET"
                        no shutdown
                    exit
                exit
```

The following command shows the AQP used to populate the cache and verifies that traffic is matched.

```
*A:Dut-C# show application-assurance group 1 policy app-qos-policy 50

===============================================================================
Application QOS Policy Entry 50 (Default Subscriber Policy)
===============================================================================
Description : (Not Specified)
Admin State : in-service
Hits        : 2 flows
Conflicts   : 0 flows

Match :

Action :
    DNS IP Cache                 : nokia-cache
===============================================================================
```

The Admin State "in-service" indicates that the AQP is in use. If no traffic has matched and the number of hits remains zero, the configuration should be checked.

Tools Commands

The following **tools** command dumps the initial status of the DNS IP cache (assuming that the server with IP address 192.168.1.100 can be accessed via FTP):

```
*A:Dut-C# tools dump application-assurance group 1 dns-ip-cache "nokia-cache"
url ftp://username:password@192.168.1.100/tmp/mylog.log
```

This command dumps the contents of the cache in a file. Adding a URL is optional; however, for cache tables that contain a large number of entries, it is a better option. The command will generate a text file that contains the following:

```
=====================================================
Application-Assurance dns-ip-cache "nokia-cache"
Current Time:          "11/08/2017 12:45:32" (UTC)
  group:               1
  isa:                 3/2
  admin state:         no shutdown
  max-entries:         10
=====================================================
ip-address  creationTime(UTC)  lastUpdated(sec)  numDNSResponses
lastMatchTime(UTC)  numTimesMatched

Total entries in-use:   0

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

DNS traffic is generated to populate the cache and HTTP messages are sent to verify that the filter matched correctly. Afterward, the following command is launched to dump the cache.

```
*A:Dut-C# tools dump application-assurance group 1 dns-ip-cache "nokia-cache"
url ftp://username:password@192.168.1.100/tmp/mylog.log

=====================================================
Application-Assurance dns-ip-cache "nokia-cache"
Current Time:          "11/13/2017 09:02:25" (UTC)
  group:               1
  isa:                 3/1
  admin state:         no shutdown
  max-entries:         10
=====================================================
ip-address  creationTime(UTC) lastUpdated(sec)  numDNSResponses  lastMatchTime(UTC)
numTimesMatched
8.1.17.21  "11/13/2017 08:31:35"  136             4        "11/13/2017 09:00:09"
    7

Total entries in-use:   1
=====================================================
```

The preceding log shows that the cache entry matched four times (numDNSResponses). Traffic was sent to this address seven times (numTimesMatched).

Conclusion

This chapter describes how an operator can ensure that traffic will match the correct filter and ensure that fraudulent traffic will not be misclassified. By using the DNS IP Cache feature, operators can ensure that traffic will be correctly charged.

Application Assurance — GTP Roaming Firewall

This chapter describes Application Assurance GTP roaming firewall.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

The configuration and information in this chapter are based on SR OS Release 20.2.R2.

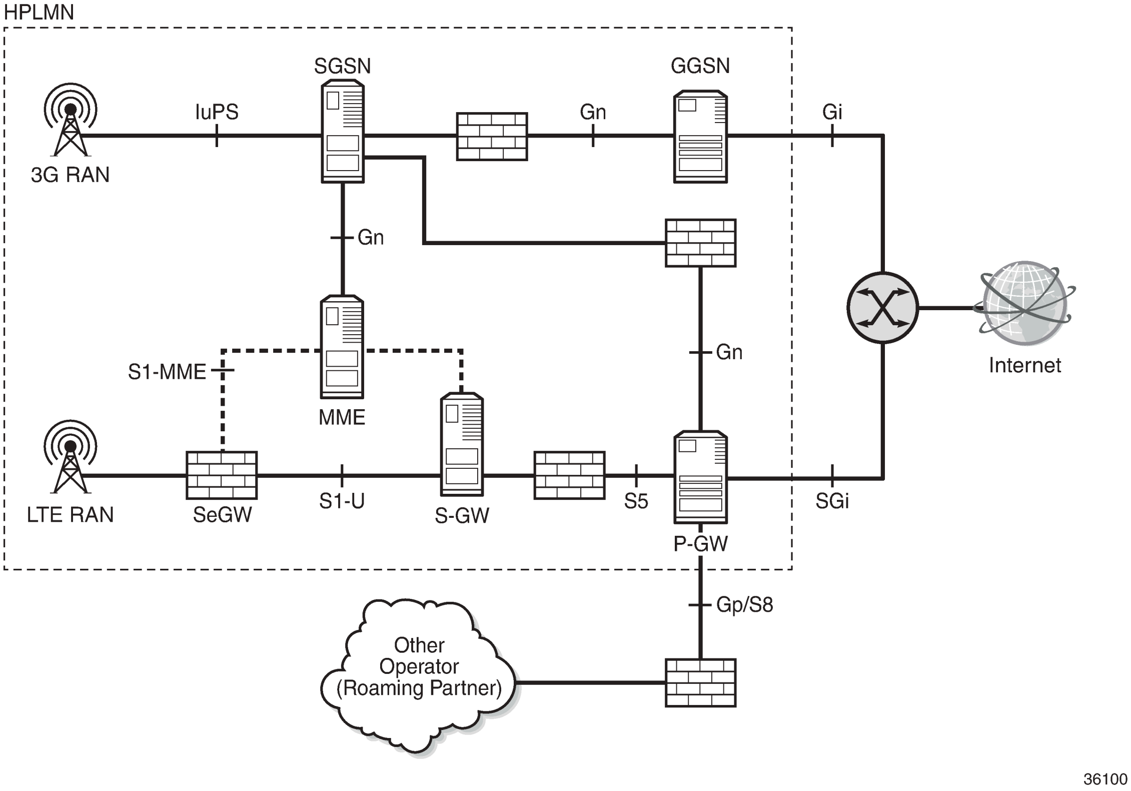
Overview

Wireless network operators rely on GPRS Tunneling Protocol (GTP) for the delivery of mobile data services across the access network. However, GTP is not designed to be secure, exposing the mobile access network to attacks from both its own subscribers and its partner networks.

The Application Assurance (AA) SR OS 20.2.R2 firewall feature extends AA-Integrated Service Adapter (AA-ISA) application-level analysis to provide an in-line stateful service that integrates into a 7750 Service Router. The feature provides protection for mobile operator infrastructure against attacks from compromised mobile gateways: Serving Gateways (SGWs) or Packet data network Gateways (PGWs).

AA stateful packet filtering, combined with AA L7 classification and control, provides operators with advanced, next-generation firewall (FW) functionality. This AA stateful FW feature runs on AA-ISA and, using stateful inspection, not only inspects packets at Layers 3 to 7, but also monitors and keeps track of the connection state. Figure 17: AA GTP roaming FW deployment shows an example AA GTP roaming FW deployment.

Figure 17: AA GTP roaming FW deployment



S8/Gp AA FW deployment

AA FW is deployed as a GTP FW on S8/Gp (or S5/Gn) interfaces, either as part of a *7750 SR* router in the form of an AA-ISA hardware module or as a separate Virtual SR (VSR) appliance. AA FW provides operators with network security, such as:

- GTP protocol validation, which checks for anomaly attacks that involve malformed, corrupt, or spoofed traffic:
  - header length checks
  - Information Element (IE) length validation
  - invalid reserved field validation
  - reserved IE validation
  - missing mandatory IE validation
  - sequence number validation
  - Tunnel Identification (TEID) validation - blocks GTP tunnel creations that have not been signaled correctly
- PGW and SGW redirection protection
- GTP-in-GTP check

- Handover control to prevent session hijacking
- Source address (User Equipment (UE)) anti-spoofing protection
- Protection against unauthorized Public Land Mobile Network (PLMN) and/or Access Point Name (APN) access:
  - filter message-based APN, International Mobile Subscriber Identity (IMSI) prefix
- Protection against unsupported GTP message types:
  - filter messages, based on message types and/or message length
- Protection against flooding attack:
  - GTP traffic bandwidth policing, which limits the GTP bandwidth from a roaming partner SGW/PGW
  - GTP tunnel limiting, which limits the number of concurrent GTP tunnels and/or the setup rate of these tunnels from a roaming partner SGW/PGW
- Protection against IP fragmentation-based attacks:
  - drop rules for IP fragmentation of GTP messages

AA FW supports both GTPv1 and GTPv2. It is typically deployed as an L3/VPRN service. SAPs/spokes are diverted to AA for a GTP FW. L2/VPLS connectivity is supported by AA. AA transit subscribers (identified by SGW IPs) are auto-created under the parent-diverted SAPs/spokes.

UE IP address anti-spoofing

Source address spoofing is initiated by a malicious UE that hijacks (spoofs) an IP address of another UE and invokes a download from a malicious server on the Internet. After the download begins, the malicious UE exits the session. The UE under attack (receiving the download traffic) gets charged for traffic it did not solicit.

AA FW associates the GTP-c messages of the UE IP address IE with the GTP-u packets to ensure that the packets carried in the upstream have the correct source IP address (inner IP within the GTP-u tunnel). Because the UE address is negotiated within the PDP context creation handshake, any packets originating from the UE that contain a different source address are detected by AA FW and dropped.

To enable UE IP address anti-spoofing protection, the operator needs to enable "validate-source-ip-addr", as follows:

```
*A:Dut-C>config>app-assure>group>
+---gtp-filter <gtp-filter-name> [create]
|      +---gtp-tunnel-database
|          +---validate-source-ip-addr
```

GTP TEID validation

Compromised mobile gateways (GSNs) can send storms of GTP traffic with invalid GTP TEIDs to cause a denial of service (DoS) attack. By inspecting GTP-c messages, AA FW supports stateful correlation of upstream and downstream GTP flows (DstIP + TEID) of the same PDN session. AA drops packets with TEIDs that have not been negotiated correctly.

The operator can enable AA to drop GTP traffic with an invalid TEID using:

```
*A:Dut-C>config>app-assure>group>
```

```
+---gtp-filter <gtp-filter-name> [create]
|       +---gtp-tunnel-database
|           +---validate-gtp-tunnels
```

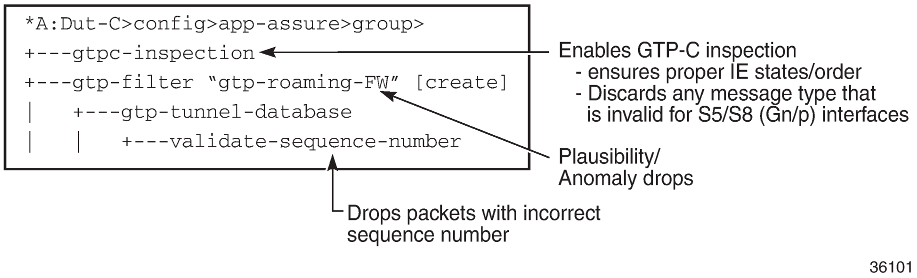
GTP anomaly prevention/sequence number checks

Protocol anomaly attacks involve malformed or corrupt packets that typically fall outside of the protocol specifications. Packets are denied by AA FW if they fail the sanity check. The following are some examples of GTP sanity checks:

- invalid GTP header length
- invalid IE length
- invalid reserved fields
- invalid sequence number
- missing mandatory IEs

Also, AA FW performs sequence number validation whereby it ensures no out-of-sequence GTP packets. By default, sequence number validation is disabled. To enable it, the following CLI command can be used:

Figure 18: CLI command



```
*A:Dut-C>config>app-assure>group>
+---gtpc-inspection
+---gtp-filter "gtp-roaming-FW" [create]
|   +---gtp-tunnel-database
|   |   +---validate-sequence-number
```

Enables GTP-C inspection
- ensures proper IE states/order
- Discards any message type that is invalid for S5/S8 (Gn/p) interfaces
Plausibility/ Anomaly drops
Drops packets with incorrect sequence number
36101

GTP packets with wrong sequence numbers are dropped when validate-sequence-number is enabled.

GTP message-type filtering

AA FW performs GTP message validation, in which packets with invalid message types (that are not applicable to the roaming interfaces) are denied by the AA GTP-c inspection command:

Table 6: Denied GTP message types for roaming interface

	GTP-u port	GTP-c port
Denied GTPv1 message types	None	GTPU_PDU GTPV1_END_MARKER GTPV1_MSG_ERR_IND GTPV1-ALL-MBMS message-types GTPV1-ALL-Location management message-types

	GTP-u port	GTP-c port
Denied GTPv2 message types	N/A	GTP_PKT_ERROR_INDICATION GTP_PKT_DNLK_DATA_FAIL_INDICATION GTP_PKT_STOP_PAGING_INDICATION GTP_PKT_CRE_INDR_TNL_REQ GTP_PKT_CRE_INDR_TNL_RSP GTP_PKT_DEL_INDR_TNL_REQ GTP_PKT_DEL_INDR_TNL_RSP GTP_PKT_RELEASE_BEARERS_REQ GTP_PKT_RELEASE_BEARERS_RSP GTP_PKT_DNLK_DATA GTP_PKT_DNLK_DATA_ACK GTP_PKT_MOD_ACCESS_BEARERS_REQ GTP_PKT_MOD_ACCESS_BEARERS_RSP

Also, AA FW allows the operator to further restrict allowed message types (shown in the following table) by configuring GTP message type filter entries to deny (or allow) the following message types:

Table 7: Allowed GTP message types (Cat-1)

	GTP-u port	GTP-c port
Allowed GTPv1 message types	GTPV1_MSG_ECHO_REQ GTPV1_MSG_ECHO_RESP GTPV1_SUPP_EXT_HDR_NOTIF GTPV1_MSG_ERR_IND GTPV1_END_MARKER GTPU_PDU	GTPV1_MSG_ECHO_REQ GTPV1_MSG_ECHO_RESP GTPV1_SUPP_EXT_HDR_NOTIF GTPV1_MSG_VER_NOT_SUPP_IND GTPV1_MSG_PDP_CREATE_REQ GTPV1_MSG_PDP_CREATE_RESP GTPV1_MSG_PDP_UPD_REQ GTPV1_MSG_PDP_UPD_RESP GTPV1_MSG_PDP_DEL_REQ GTPV1_MSG_PDP_DEL_RESP GTPV1_MSG_NET_INIT_REQ GTPV1_MSG_NET_INIT_RESP GTPV1_MSG_MSINFO_REQ GTPV1_MSG_MSINFO_RESP
Allowed GTPv2	N/A	GTP_PKT_ECHO_REQ GTP_PKT_ECHO_RSP

	GTP-u port	GTP-c port
message types		GTP_PKT_VERSION_NOT_SUPPORTED GTP_PKT_CRE_SES_REQ GTP_PKT_CRE_SES_RSP GTP_PKT_MOD_BEARER_REQ GTP_PKT_MOD_BEARER_RSP GTP_PKT_DEL_SES_REQ GTP_PKT_DEL_SES_RSP GTP_PKT_CHG_NOT_REQ GTP_PKT_CHG_NOT_RSP GTP_PKT_MOD_BEARER_CMD GTP_PKT_MOD_BEARER_FAIL_INDICATION GTP_PKT_DEL_BEARER_CMD GTP_PKT_DEL_BEARER_FAIL_INDICATION GTP_PKT_BEARER_RESOURCE_CMD GTP_PKT_BEARER_RESOURCE_FAIL_INDICATION GTP_PKT_SUSPEND_NOTIFICATION GTP_PKT_SUSPEND_ACK GTP_PKT_RESUME_NOTIFICATION GTP_PKT_RESUME_ACK GTP_PKT_CRE_BEARER_REQ GTP_PKT_CRE_BEARER_RSP GTP_PKT_UPD_BEARER_REQ GTP_PKT_UPD_BEARER_RSP GTP_PKT_DEL_BEARER_REQ GTP_PKT_DEL_BEARER_RSP GTP_PKT_TRACE_SESSION_ACTIVATION GTP_PKT_TRACE_SESSION_DEACTIVATION GTP_PKT_UPDATE_PDN_CONNECTION_SET_REQ GTP_PKT_UPDATE_PDN_CONNECTION_SET_RSP GTP_PKT_DELETE_PDN_CONNECTION_SET_REQ GTP_PKT_DELETE_PDN_CONNECTION_SET_RSP

By default, the GTP message filter allows all GTP messages.

To configure GTPv2 message filtering, the following command is used:

```
*A:Dut-C>config>app-assure>group>
+---gtp-filter <gtp-filter-name> [create]
```

```
    +---gtpc-inspection
|   +---message-type-v2
|   |   |
|   |   +---default-action {permit|deny}
|   |   |
|   |   +---entry <entry-id> value <gtpv2-message-value> action {permit|deny}
|   |   |   no entry <entry-id>
```

To configure GTPv1 message filtering, the following command is used:

```
*A:Dut-C>config>app-assure>group>
|   +---message-type
|   |   |
|   |   +---default-action {permit|deny}
|   |   |
|   |   +---entry <entry-id:1..255> value <gtpv1-message-value> action {permit|deny}
|   |   |   no entry <entry-id>
```

Note:
If the operator configures a message type that is invalid for the roaming interface to be denied, it will be dropped and counted under that filter entry (and not tagged as dropped due to "wrong-interface" in the event log). However, configuring the message-type filter to "permit" a message type that is invalid for the roaming interface will not take effect, because the packet with the specified message type will be dropped by the GTP-c protocol inspection process.

Unauthorized APN attack – APN filtering

APN filtering checks GTP-c messages to determine if a roaming subscriber is allowed to access a specified external network (/aka APN). The "create-session-request" and "create pdp request" GTP message types contain an APN IE in the header of a GTP packet. An APN IE consists of an external network ID (for example, nokia.com) and, optionally, a unique ID that identifies the operator PLMN.

APN filtering prevents malicious UEs from initiating a "create PDP/session request" flood attack toward the PGW/GGSN for invalid or disallowed APNs. The operator can configure an AA GTP filter to perform APN filtering to restrict roaming subscribers access to specific external networks.

An APN filter, an IMSI prefix, and an SGSN address pool can be used together to filter GTP packets, as follows:

```
*A:Dut-C>config>app-assure>group>
+---gtp-filter <gtp-filter-name> [create]
|   +---imsi-apn-filter //NEW and all its children attributes
|   |   +---default-action {permit|deny} //default permit
|   |   +---entry <entry-id: 1031..2030> create
|   |   |       + apn <string 0|1..32 characters>
|   |   |       |---no apn
|   |   |       + src-gsn ip-prefix-list <ip-prefix-list>
|   |   |       + src-gsn <ip address prefix>
|   |   |       |---no src-gsn
|   |   |       + action {permit|deny} //default permit
|   |   +---no entry <entry-id>
```

By default, AA FW permits all APNs.

Unauthorized PLMN access – IMSI prefix filtering

The PLMN of a subscriber home network is identified by combining the Mobile Country Code (MCC) and Mobile Network Code (MNC). MCC-MNC is also known as the International Mobile Station Identity (IMSI) prefix. The IMSI prefix acts as a PLMN identifier.

GTP IMSI prefix filters can be configured to deny GTP incoming traffic from invalid roaming partners. Conversely, GTP IMSI prefix filters can allow only incoming traffic from those network operators that have signed roaming agreements. Any GTP packets with IMSI prefixes not matching the configured prefixes are dropped.

As shown in the Unauthorized APN attack – APN filtering section, IMSI filter entry can also be optionally combined with an SGSN/SGW IP address (or IP address prefix list) to further restrict allowed IMSI prefix traffic to specific SGSN/SGW nodes.

Unauthorized network access

An attacker, using an unauthorized GSN, can cause a DoS attack using spoofed PDP context delete messages (DoS attack) or spoofed update PDP context requests to hijack existing sessions. Such attacks can also spoof a create PDP context request to gain unlawful Internet access. Session hijacking can come from either the SGW/SGSN or the PGW/GGSN. An unauthorized GSN can hijack GTP tunnels or cause a DoS attack by intercepting another GSN and redirecting traffic to it.

Operators can use AA FW to configure pools of trusted GSN IP addresses (using AA IP-prefix-list) to stop spoofed requests from untrusted GSNs. AA IP prefix lists can be configured to model GSN groups, as follows:

```
*A:Dut-C>config>app-assure>group#
   ip-prefix-list ip-prefix-list-name [create]
            prefix ip-prefix/ip-prefix-length [name prefix-name]
```

These lists are then referenced in session filters, such that only sessions that match the lists can be "permitted", as follows:

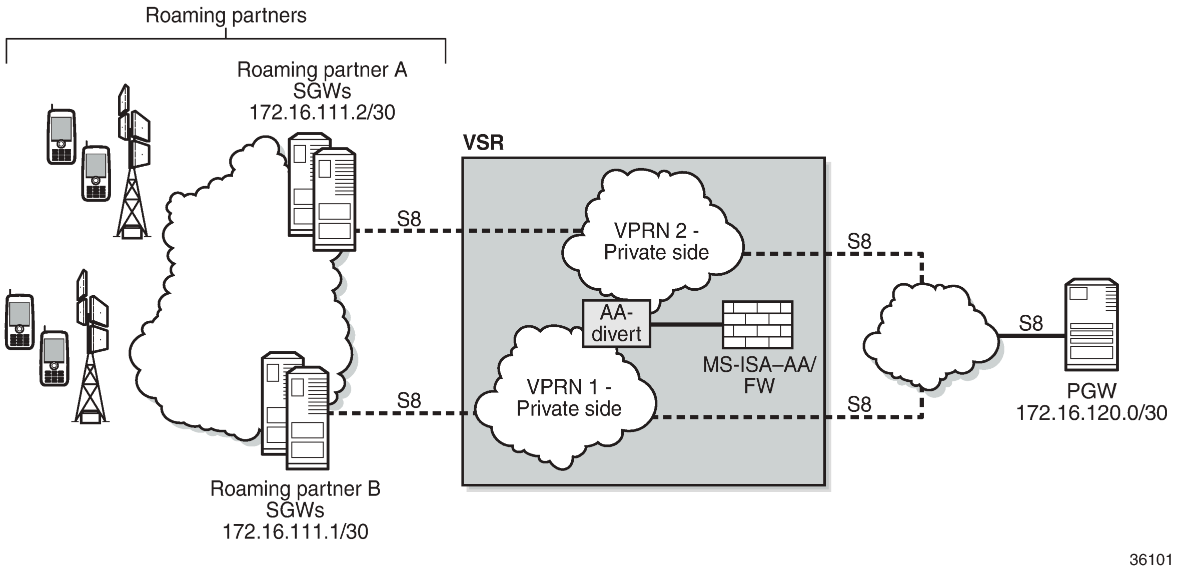
```
*A:Dut-C>config>app-assure>group# session-filter
      default-action  deny
      entry           # Configure an entry in the session filter
          match
              src-ip  # Configure IPs that correspond to authorized SGW/SSGN
          action
              permit
```

Configuration

AA GTP filtering functionality is enhanced in SR OS Release 20.2.R2 to include support for the AA FW feature related to the GTP roaming interface. The GTP filters are optional Application QoS Policy actions (AQP actions). AQPs have partition-level scope, which allows different FW policies to be implemented by using AA partition concepts within the same AA-ISA.

The configuration topology in Figure 19: Configuration topology shows how the VSR equipped with AA FW functionality provides protection for the S8 interfaces.

Figure 19: Configuration topology



Pre-setup requirements

Configuration of a VSR router is required if a 7750 SR is not already used in the access network on the S8 interfaces. If a 7750 SR is already deployed, AA-ISA must be configured.

See the Application Assurance — Stateful Firewall chapter for basic knowledge about AA-FW functionality.

Platform-dependent configuration

VSR

For GTP FW deployment in VSR, the following configuration supports load balancing of traffic across multiple CPU cores:

```
*A:7750-1>config> config>isa>aa-grp# vm-traffic-distribution-by-teid
```

7750 hardware

GTP FW deployment in 7750 hardware is only supported by ISA2:

```
*A:7750-1>config> config>isa>aa-grp# minimum-isa-generation 2
```

Allocation of memory for stateful GTP processing

To support stateful GTP processing (for example, TEID, sequence number, and UE IP validation FW operations), the operator must configure the system to allocate sufficient memory resources, as follows:

```
*A:7750-1>config>isa>aa-grp# shared-resources
                             gtp-tunnel-database 100
```

Configuration to divert SAPs/VPRN traffic into AA-ISA

In this configuration example, one VPRN is used per wireless roaming partner network. In the example, two roaming partner networks are used for illustration. In real networks, this number is much bigger.

However, before configuring SAPs for diversion, the operator can optionally define some Application Service Option (ASO) characteristics to provide different FW policies for different roaming partners, as follows:

```
*A:7750-1>config>app-assure>group 1:1 policy
                begin
                app-service-options
                    characteristic "FW-Protection" persist-id 1 create
                        value "OFF" persist-id 1
                        value "ON" persist-id 2
                        default-value "OFF"
                    exit
                    characteristic "strict-FW-Protection" persist-id 2 create
                        value "OFF" persist-id 1
                        value "ON" persist-id 2
                        default-value "OFF"
                    exit
                exit
                commit
```

For more information about ASO configuration, see the Application Assurance — App-Profile, ASO and Control Policies chapter.

After configuring any ASO characteristics, define an application profile and transit IP policy; for example:

```
*A:7750-1>config>app-assure>group$ info
----------------------------------------------
            policy
                begin
                app-profile "default" create
                    description "App profile that applies to the whole SAP"
                    divert
                    characteristic "FW-Protection" value "ON"
                exit
                app-profile "strict-FW" create
                    description "App profile that applies strict FW rules to the SAP"
                    divert
                    characteristic "FW-Protection" value "ON"
                    characteristic "strict-FW-Protection" value "ON"
                exit
                commit
            exit
            transit-ip-policy 1 create
                def-app-profile "strict-FW"
                detect-seen-ip
                transit-auto-create
                    no shutdown
                exit
            exit
```

Traffic of these two VPRNs needs to be diverted into AA-ISA to provide firewall protection. Apply the following policies to the SAPs of the two VPRNs:

```
*A:7750-1>config#
    service
        customer 1 name "1" create
            description "Default customer"
        exit
        customer 2 name "2" create
        exit
        vprn 100 name "100" customer 2 create
            interface "to-NetA" create
            exit
        exit
        vprn 200 name "200" customer 1 create
            interface "to-site1" create
            exit
        exit
        vprn 100 name "100" customer 2 create
            description "L3 Service roaming partner 2"
            route-distinguisher 100:2
            interface "to-NetA" create
                address 192.168.1.1/24
                sap 1/2/3 create
                    app-profile "default"
                exit
            exit
        exit
        vprn 200 name "200" customer 1 create
            description "L3 Service roaming partner 1"
            route-distinguisher 200:1
            interface "to-site1" create
                address 192.168.2.1/24
                static-arp 1.1.1.2 00:ff:02:00:00:01
                sap 1/2/1 create
                    transit-policy ip 1
                    app-profile "strict-FW"
                exit
            exit
        exit
    exit
```

This configuration achieves the following.

1. Roaming traffic is diverted to AA-ISA for FW protection.
2. Customer 1 traffic will have a "strict" FW rule attribute, while customer 2 traffic will be subject to basic FW rules.
3. Within AA-ISA, the customer 1 diverted SAP is treated as a parent SAP.
    - Instead of treating the whole SAP as a single subscriber, subscribers are auto-created within this SAP, based on the IP address of the SGWs/SSGNs.

Note:
If the operator does not require per SGW/SSGN control (such as limiting the total bandwidth of a SGW to prevent DoS attack), the "transit IP policy" from the SAP configuration can be removed. This will cause AA to treat the whole SAP as a single subscriber, as in the case of the customer 2 SAP.

Configuration FW events log

To configure a log that captures events related to various AA firewall actions:

```
*A:7750-1# configure application-assurance group 1:1
*A:7750-1>config>app-assure>group# event-log "FW_events_log" create
*A:7750-1>config>app-assure>group>evt-log$ buffer-type circular
*A:7750-1>config>app-assure>group>evt-log$ max-entries 100000
*A:7750-1>config>app-assure>group>evt-log$ no shutdown
*A:7750-1>config>app-assure>group>evt-log$ exit
*A:7750-1>config>app-assure>group# info
----------------------------------------------
                ---snip---
            event-log "FW_events_log" create
                buffer-type circular
                max-entries 100000
                no shutdown
            exit
```

Note:
Alternatively, due to the limited size of the log and the large amount of traffic AA can handle, it is recommended that the operator use the syslog mechanism instead of local logging, as follows:

`config>app-assure>group>event-log>syslog`

The event log can be referenced in various FW actions that are configured later in this chapter.

```
*A:7750-1# tools dump application-assurance group 1:1 event-log "FW_event_log"
                                                                           isa 1/1
=====================================================
Application-Assurance event-log "FW_events_log"
Current Time:          "05/19/2020 19:03:58" (UTC)
  group[:partition]:   1:1
  isa:                 1/1
  admin state:         no shutdown
  buffer-type:         circular
  max-entries:         100000
=====================================================
Event-source
Action        SubType     SubName                            Direction Src-ip
Dst-ip                           Ip-protocol Src-port Dst-port Timestamp
"gtp filter gtp-filter-partner1 reason: filtered-gtp-message-type, teid: 0x0001d100,
 MT: 36, version: 2"                                                                deny
   transit     "1_10.10.68.1/32"                  from-sub  10.10.68.1
10.10.68.3                       udp         2123     2123     "05/19/2020 18:54:50"
"gtp filter gtp-filter-partner1 reason: filtered-gtp-message-type, teid: 0x00019100,
 MT: 37, version: 2"                                                                deny
   transit     "1_10.10.68.1/32"                  to-sub    10.10.68.3
10.10.68.1                       udp         2123     2123     "05/19/2020 18:54:55"

Total Records:   2
=====================================================
```

The following command clears all the entries within the specified log:

```
*A:7750-1# clear application-assurance group 1:1 event-log "FW_events_log"
```

Configuration to limit total traffic from SGWs

Nokia recommends that a total limit be placed on how much bandwidth and how many flows an SGW/SGSN can generate toward the network. The exact limit values are a function of the number of end devices that are served by the roaming partner SGW/SGSN and capacity limits of the HPLMN PGW/GGSN, plus some additional margin.

In the following example, it is assumed that traffic from each roaming SGW will not exceed 1200 concurrent flows/second (serving about 200 roaming UEs) and 50 Mb/s. These need to be replaced in actual deployments with appropriate values that reflect the specific network deployment.

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
           policer "limit_roamingSGW_Flows" type flow-count-limit
                                            granularity subscriber create
                flow-count 1200
                gtp-traffic
            exit
            policer "limit_roamingSGW_bw" type single-bucket-bandwidth
                                          granularity subscriber create
                rate 50000
                mbs 500
            exit
----------------------------------------------
```

Apply the configured policers as actions from within the default subs-policy AQP entry:

```
*A:7750-1>config>app-assure>group>policy>aqp#
                    entry 500 create
                        description "limit per SGW flow and b/w- partner 1"
                        match
                            traffic-direction subscriber-to-network
                            characteristic "strict-FW-Protection" eq "ON"
                        exit
                        action
                            bandwidth-policer "limit_roamingSGW_bw"
                            flow-count-limit "limit_roamingSGW_Flows"
                        exit
                        no shutdown
                    exit
```

For GTP traffic flow count policing, it is important that "aqp-initial-lockup" is enabled:

```
*A:7750-1# configure application-assurance group 1:1 aqp-initial-lookup
```

Note:
All the preceding actions apply to the traffic direction "subscriber-to-network". These actions do not apply to traffic in the other direction (downlink), because the purpose of the AA FW is to protect the network resources from upstream traffic from compromised roaming partner SGWs.

Note:
No policers are placed for the traffic of customer 2, because its profile does not have "strict policing" enabled. A policer can be configured to limit the total bandwidth and flows from all SGWs served to the customer 1 SAP, as follows:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            policer "limit_roamingSGWs_total_Flows" type flow-count-limit
                                                    granularity subscriber create
                flow-count 12000
                gtp-traffic
            exit
            policer "limit_roamingSGWs_total_bw" type single-bucket-bandwidth
                                                 granularity subscriber create
                rate 500000
                mbs 5000
            exit
----------------------------------------------
```

Apply the configured policers as actions from within the default subs-policy AQP entry, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp#
                    entry 501 create
                        no shutdown
                        description "limit total SGW flow and b/w- partner 2 "
                        match
                            traffic-direction subscriber-to-network
                            characteristic "strict-FW-Protection" eq "OFF"
                        exit
                        action
                            bandwidth-policer "limit_roamingSGWs_total_bw"
                            flow-count-limit "limit_roamingSGWs_total_Flows"
                        exit
                    exit
```

GTP filtering – disallow traffic from unauthorized SGWs

To use GTP filtering to disallow traffic from unauthorized SGWs, perform the following steps:

1. Create AA IP lists
2. Use AA IP lists in session filters and AQPs
3. Reference session filters within AQPs

1. Create AA IP lists, by creating an AA IP prefix list that contains SGW IP addresses or range of addresses for each customer, as follows:

   Roaming partner 1

   ```
   *A:7750-1>config>app-assure# group 1:1
              ip-prefix-list "Roaming1_ALL_SGWs" create
                   description "SGWs subnet-partner 1"
                   prefix 172.16.100.0/24
               exit
               ip-prefix-list "Roaming2_ALL_SGWs" create
                   description "SGWs subnet for roaming partner2"
                   prefix 172.16.110.100/30
               exit
   ```

Roaming partner 2

```
*A:7750-1>config>app-assure>group#
            ip-prefix-list "Roaming2_ALL_SGWs" create
                description "SGWs subnet for roaming partner2"
                prefix 172.16.110.100/30
            exit
```

2. The AA IP prefix lists can be referenced and used in AA FW rules using session filters and AQPs, as follows:

```
*A:7750-1>config>app-assure>group#
           session-filter "restricted_access_partner1" create
                description "SGWs_allowed_partner1"
                default-action deny
                entry 10 create
                    description "allow GTP-u from authorized subnets"
                    match
                        ip-protocol-num udp
                        src-ip ip-prefix-list "Roaming1_ALL_SGWs"
                        dst-port eq 2152
                    exit
                    action permit
                exit
                entry 11 create
                    description "allow GTP-c from authorized subnets"
                    match
                        ip-protocol-num udp
                        src-ip ip-prefix-list "Roaming1_ALL_SGWs"
                        dst-port eq 2123
                    exit
                    action permit
                exit
                entry 20 create
                    description "allow DNS"
                    match
                        ip-protocol-num *
                        src-ip ip-prefix-list "Roaming1_ALL_SGWs"
                        dst-port eq 53
                    exit
                    action permit
                exit
            exit
            session-filter "restricted_access_partner2" create
                description "SGWs_allowed_partner2"
                default-action deny event-log "FW_events_log"
                entry 10 create
                    description "allow GTP-u from authorized subnets"
                    match
                        ip-protocol-num udp
                        src-ip ip-prefix-list "Roaming2_ALL_SGWs"
                        dst-port eq 2152
                    exit
                    action permit
                exit
                entry 11 create
                    description "allow GTP-c from authorized subnets"
                    match
                        ip-protocol-num udp
                        src-ip ip-prefix-list "Roaming2_ALL_SGWs"
                        dst-port eq 2123
                    exit
                    action permit
                exit
                entry 20 create
                    description "allow DNS"
                    match
                        ip-protocol-num *
                        src-ip ip-prefix-list "Roaming2_ALL_SGWs"
                        dst-port eq 53
                    exit
                    action permit
                exit
            exit
```

Note:
Optionally, you can combine the session filter entries for the two roaming partners into a single session filter (for scale reasons). AA supports a total of 300 session filters. If there are less than 300 roaming partners, you can use a session filter per partner for customization purposes (related to, for example, IP subnets). If the number of partners is greater than the maximum number of session filters, you need to aggregate entries into a fewer number of session filters. Be aware of overlapping IP addresses from different roaming partner networks/VPRNs.

3. The configured session filters need to be referenced within AQPs, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp#
                    entry 510 create
                        description "apply FW rules for roaming partner 1"
                        match
                            traffic-direction subscriber-to-network
                            characteristic "strict-FW-Protection" eq "ON"
                        exit
                        action
                            session-filter "restricted_access_partner1"
                        exit
                        no shutdown
                    exit
                    entry 511 create
                        description "apply FW rules for roaming partner 2"
                        match
                            traffic-direction subscriber-to-network
                            characteristic "strict-FW-Protection" eq "OFF"
                        exit
                        action
                            session-filter "restricted_access_partner2"
                        exit
                        no shutdown
                    exit
```

Restrict downstream traffic (optional)

Operators can optionally restrict downstream traffic to specific destinations and protocols, as follows:

```
*A:7750-1>config>app-assure>group#
           session-filter "restricted_downstream_traffic_1" create
                description "allow only traffic to only signed up partners"
                default-action deny
                entry 10 create
                    description "allow GTP-u from authorized subnets"
                    match
```

```
            ip-protocol-num udp
            dst-ip ip-prefix-list "Roaming1_ALL_SGWs"
            dst-port eq 2152
          exit
          action permit
        exit
        entry 11 create
          description "allow GTP-c to authorized subnets"
          match
            ip-protocol-num udp
            dst-ip ip-prefix-list "Roaming1_ALL_SGWs"
            dst-port eq 2123
          exit
          action permit
        exit
        entry 20 create
          description "allow DNS"
          match
            ip-protocol-num *
            dst-ip ip-prefix-list "Roaming1_ALL_SGWs"
          exit
          action permit
        exit
      exit
      session-filter "restricted_downstream_partner2" create
        description "SGWs_allowed_partner2"
        default-action deny event-log "FW_events_log"
        entry 10 create
          description "allow GTP-u to authorized subnets"
          match
            ip-protocol-num udp
            dst-ip ip-prefix-list "Roaming2_ALL_SGWs"
            dst-port eq 2152
          exit
          action permit
        exit
        entry 11 create
          description "allow GTP-c to authorized subnets"
          match
            ip-protocol-num udp
            dst_ip ip-prefix-list "Roaming2_ALL_SGWs"
            dst-port eq 2123
          exit
          action permit
        exit
      exit
```

Note:
The preceding configuration provides the most flexibility and allows IP addresses to overlap between different partner networks. However, it comes at the cost of creating separate session filters for each partner. If IP addresses do not overlap, a single session filter is sufficient.

The session filters need to be referenced from AQP, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp#
          entry 514 create
            description "apply FW rules for roaming partner 1"
            match
              traffic-direction network-to-subscriber
              characteristic "strict-FW-Protection" eq "ON"
            exit
            action
```

```
                            session-filter "restricted_downstream_partner1"
                        exit
                        no shutdown
                    exit
                    entry 513 create
                        description "apply FW rules for roaming partner 2"
                        match
                            traffic-direction network-to-subscriber
                            characteristic "strict-FW-Protection" eq "OFF"
                        exit
                        action
                            session-filter "restricted_downstream_partner2"
                        exit
                        no shutdown
                    exit
```

Configuration to protect against malformed packets

It is always recommended in FW deployments that overload-drop, error-drop, and fragment-drop are enabled within the default sub-policy, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp#
                    entry 50 create
                        description "drop error and fragmented packets"
                        action
                            overload-drop event-log "FW_events_log"
                            error-drop event-log "FW_events_log"
                            fragment-drop all event-log "FW_events_log"
                        exit
                        no shutdown
                    exit
```

Note:

- The overload-drop action ensures that AA-ISA, if it gets overloaded, drops the excess traffic instead of cutting it through without applying FW rules.
- The error-drop action ensures that AA-ISA drops malformed IP packets.
- The fragment-drop all action allows the operator to drop all fragmented traffic, drop out-of-order fragments only, or allow fragments through. Because many network DoS attacks use IP fragmentation to initiate attacks, allowing fragments through is not recommended for firewall deployments. As a minimum, if fragmentation is used, the operator is recommended to configure AA to drop out-of-order fragmented packets.
- The preceding actions are applied to all traffic. Therefore, there are no AQP match conditions configured.

Plausibility of GTP messages and GTP message validation

To protect the network from malformed GTP packets and associated attacks as described in the overview section, a GTP filter needs to be created and referenced from an AQP entry.

1. Configure the GTP filter object to:
    a. Enable GTP-c inspection so that the FW:

   - **i.** Ensures the correct IE states and order
   - **ii.** Discards any GTP packet that contains an invalid message type for S5/S8 (Gn/Gp) interface

- **b.** Enable sequence number checking for GTP-c traffic (for partner 1 traffic)
- **c.** Enable the GTP filter to check and drop errored GTP packets (anomalies)
- **d.** Enable GTP message length checking (to minimize exposure to code injection attacks). The maximum is set here (for example, 1250 bytes). The value is operator dependent, and should be replaced with the figure used by the operator.
- **e.** Drop GTP-in-GTP encapsulated packets

```
*A:7750-1>config>app-assure>group>gtp#
                event-log "FW_events_log"
                gtpc-inspection
                gtp-filter "gtp-filter-partner1" create
                    description "gtp-filter for partner 1"
                    max-payload-length 1250
                    event-log "FW_events_log" action deny
                    gtp-in-gtp deny
                    gtp-tunnel-database
                        validate-sequence-number
                    exit
                exit
                gtp-filter "gtp-filter-partner2" create
                    description "gtp-filter for partner 2"
                    max-payload-length 1250
                    event-log "FW_events_log" action deny
                    gtp-in-gtp deny
                no shutdown
```

2. The configured GTP filters need to be referenced from AQP entries, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp#
                   entry 512 create
                        description "apply SGW GTP filter rules"
                        match
                            characteristic "strict-FW-Protection" eq "ON"
                        exit
                        action
                            gtp-filter "gtp-filter-partner1"
                        exit
                        no shutdown
                    exit
                    entry 513 create
                        description "apply SGW GTP filter rules"
                        match
                            characteristic "strict-FW-Protection" eq "OFF"
                        exit
                        action
                            gtp-filter "gtp-filter-partner2"
                        exit
                        no shutdown
                    exit
```

Filtering of GTP message types

In this configuration, traffic from partner 2 is considered "safe/trusted". Therefore, unlike traffic from partner 1, no additional GTP message-type filtering is applied to it, beyond the GTP Cat-1 (see Table 7: Allowed GTP message types (Cat-1)) message filtering applied as a result of enabling GTP-c inspection.

For roaming partner 1 traffic, the GTP filter is configured to block some Cat-1 optional message types (GTPv1 and GTPv2):

- GTPv2: Trance session activation/deactivation (this is optional for S8)
- GTPv1: Allows only the message types used by GTP-u and blocks GTPv1 message types used by GTP-c

Note:
By configuring GTP-c inspection, only Cat-1 message types (see Table 7: Allowed GTP message types (Cat-1)) are allowed and all others are denied. Therefore, there is little to no need for additional GTP message filtering configuration.

The GTP filter is configured as follows:

```
*A:7750-1>config>app-assure>group>gtp#
                event-log "FW_events_log" action deny
                gtpc-inspection
                gtp-filter "gtp-filter-partner1" create
                    description "gtp-filter for partner 1"
                    max-payload-length 1250
                    event-log "FW_events_log" action deny
                    message-type
                        default-action deny
                        entry 1 value "echo-request" action permit
                        entry 2 value "echo-response" action permit
                        entry 3 value "error-indication" action permit
                        entry 4 value "supported-extension-headers-notification"
                                                                    action permit
                        entry 5 value "end-marker" action permit
                        entry 6 value "g-pdu" action permit
                    exit
                    message-type-gtpv2
                        default-action permit
                        entry 524 value "trace-session-activation" action deny
                        entry 525 value "trace-session-deactivation" action deny
                    exit
                    gtp-in-gtp deny
                    imsi-apn-filter
                        default-action deny
                        entry 1031 create
                            apn ANY_APN
                            mcc-mnc-prefix 161379
                            action permit
                        exit
                    exit
                    gtp-tunnel-database
                        validate-sequence-number
                    exit
                exit
```

TEID validation

For roaming partner 1, to protect the network resources from spoofed TEIDs, the FW is recommended to verify that the TEIDs used in the GTP-u traffic are valid (that is, correctly negotiated via GTP-c), as follows:

```
*A:7750-1>config>app-assure>group>gtp#
                gtp-filter "gtp-filter-partner1"
                    gtp-tunnel-database
                        validate-gtp-tunnels
                    exit
                exit
```

Since roaming partner 2 network is trusted, no TEID validation is needed.

UE IP address anti-spoofing

It is a good practice to protect the network against UEs spoofing a different IP address, as follows:

```
*A:7750-1>config>app-assure>group>gtp#
                gtp-filter "gtp-filter-partner1"
                    gtp-tunnel-database
                        validate-source-ip-addr
                     exit
                exit
```

This example applied to partner 1 traffic. Validation of source IP requires the use of a GTP tunnel database.

APN and IMSI filtering

In this example, only Home-Routed (HR) traffic from partner 1 is allowed, regardless of the APN. The rest is denied. This is achieved by configuring an IMSI prefix (="1613797") that corresponds to the Home network.

For roaming partner 2, MVNO traffic is allowed as well as HR traffic. This MVNO traffic (specific IMSI prefix = "1613400" in this example) is only allowed to attach to the mvnoguest.com APN, as follows:

```
*A:7750-1>config>app-assure>group>gtp#
                gtp-filter "gtp-filter-partner1"
                    imsi-apn-filter
                        default-action deny
                        entry 1031 create
                            apn ANY_APN
                            mcc-mnc-prefix 161379
                            action permit
                        exit
                    exit
                exit
                gtp-filter "gtp-filter-partner2"
                    imsi-apn-filter
                        default-action deny
                        entry 1040 create
                            apn mvnnoguest.com$
                            mcc-mnc-prefix 161340
                            action permit
                        exit
                        entry 1041 create
                            apn ANY_APN
                            mcc-mnc-prefix 161379
                            action permit
                        exit
                    exit
                 exit
```

Limiting concurrent session creations

To further lower the risk of DoS attacks using massive amounts of session/PDN create messages, it is recommended that the operator configure the maximum concurrent number of endpoints (TIEDs) that an SGW can create.

In this example, the limit that is configured in the GTP filter corresponds to the maximum concurrent TEIDs that can be created by any SGW IP address, as follows:

```
*A:7750-1>config>app-assure>group>gtp#
                gtp-filter "gtp-filter-partner1"
                    gtp-tunnel-database
                        default-tunnel-endpoint-limit 400
                    exit
```

Configuring FW statistics

To gain visibility into the traffic passing through the FW and the FW actions taken, it is highly recommended to enable "deny-admit" statistics, as follows:

```
A:7750-1>config>#
    log
        file-id 5
            location cf3:
        exit
        accounting-policy 5
            description "LogFileforAAFirewallAccounting"
            record aa-admit-deny
            collection-interval 10
            to file 5
            no shutdown
        exit
    exit
```

```
A:7750-1>config>app-assure>group#
            statistics
                aa-admit-deny
                    accounting-policy 5
                    collect-stats
                    gtp-filter-stats
                    session-filter-stats
                    policer-stats-resources
                    policer-stats
                exit
                protocol
                    shutdown
                exit
```

Configuring threshold crossing alerts

As well as admit-deny statistics, the operator can optionally enable the FW to generate Threshold Crossing Alerts (TCAs) against the collected statistics.

Note:
NSP also supports TCAs. The operator has a choice to enable TCAs on both the FW and/or NSP. The advantage of TCAs generated directly from the FW is that they tend to be more real-time relative to NSP TCAs. However, NSP supports a larger TCA scale than the FW.

The operator needs to set the low- and high-water marks according to the conditions of their networks. The following values are for illustration purposes only.

```
A:7750-1> config>app-assure>group>stats#
            threshold-crossing-alert
                gtp-sanity-drop direction from-sub create
                        high-wmark 100 low-wmark 60
                exit
                threshold-crossing-alert
                    gtp-sanity-drop direction from-sub create
                        high-wmark 100 low-wmark 60
                    exit
                    gtp-filter "gtp-filter-partner1"
                        validate-gtp-tunnels direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        validate-sequence-number direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        validate-src-ip-addr direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        missing-mandatory-ie direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        tunnel-resource-limit direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        tunnel-endpoint-limit direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        message-type
                            default-action direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                            header-sanity direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                        exit
                        message-type-gtpv2
                            entry 524 direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                            entry 525 direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                        exit
                        imsi-apn-filter
                            default-action direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                        exit
                    exit
                    gtp-filter "gtp-filter-partner2"
                        missing-mandatory-ie direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        tunnel-resource-limit direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        tunnel-endpoint-limit direction from-sub create
                            high-wmark 100 low-wmark 60
                        exit
                        imsi-apn-filter
                            default-action direction from-sub create
                                high-wmark 100 low-wmark 60
                            exit
                        exit
                    exit
                exit
```

Configuring GTP and GTP-c applications

By configuring AA app-filters to define GTP-u and GTP-c applications, the operator can gain further visibility into the volume of traffic of these applications, as follows:

```
A:7750-1>config> config>app-assure>group>
            policy
                begin
                application "GTP_c" create
                exit
                application "GTP_other" create
                exit
                application "GTP_u" create
                exit
                app-filter
                    entry 40000 create
                        protocol eq "gtp"
                        server-port eq 2152
                        application "GTP_u"
                        no shutdown
                    exit
                    entry 40010 create
                        protocol eq "gtp"
                        server-port eq 2123
                        application "GTP_c"
                        no shutdown
                    exit
                    entry 40020 create
                        protocol eq "gtp"
                        application "GTP_other"
                        no shutdown
                    exit
                exit
                commit
```

The export and display of statistics related to these applications use standard AA per application per partition or per application per subscriber XML records and "show" routines.

Relevant debug routines

CLI show routines

```
*A:7750-1>config>app-assure>group>gtp>gtp-fltr# show application-assurance
group 1:1 session-filter "restricted_access_partner1"
===============================================================================
AA Session Filter Instance "restricted_access_partner1"
===============================================================================
Description    : SGWs_allowed_partner1
Default Action : deny
    Event Log  : (Not Specified)
AQP Entries    :
          510
-------------------------------------------------------------------------------
Filter Match Criteria
-------------------------------------------------------------------------------
Entry          : 10
Description    : allow GTP-u from authorized subnets
IP Protocol    : udp
Source IP List : Roaming1_ALL_SGWs
Dest Port      : eq 2152
Action         : permit
    Event Log  : (Not Specified)
Hits           : 1 flows
-------------------------------------------------------------------------------
Entry          : 11
Description    : allow GTP-c from authorized subnets
IP Protocol    : udp
Source IP List : Roaming1_ALL_SGWs
Dest Port      : eq 2123
Action         : permit
    Event Log  : (Not Specified)
Hits           : 2 flows
-------------------------------------------------------------------------------
Entry          : 20
Description    : allow DNS
IP Protocol    : *
Source IP List : Roaming1_ALL_SGWs
Dest Port      : eq 53
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
No. of entries   : 3
===============================================================================
```

```
*A:7750-1>show application-assurance group 1:1 gtp gtp-filter "gtp-filter-partner1"
===============================================================================
Application Assurance Group 1:1 GTP Filter "gtp-filter-partner1"
===============================================================================
Description                 : gtp-filter for partner 1
Maximum payload length      : 1250
Event log                   : FW_events_log
Event log action            : deny
```

```
Default action              : deny
Default GTPv2 action        : permit
Default IMSI-APN action     : deny
GTP in GTP action           : deny
Validate GTP tunnels        : enabled
Validate sequence number    : enabled
Validate source IP address  : enabled
GTP tunnel endpoint limit   : 400
Configured messages         : 6
Configured GTPv2 messages   : 2
Configured IMSI-APN filters : 1
Packets arrived             : 18
Packets denied
  Payload length            : 0
  Message type              : 0
  GTPv2 message type        : 2
  IMSI-APN filter           : 0
  Mandatory header          : 0
  Extension header          : 0
  Information element       : 0
  Invalid TEID              : 0
  Invalid sequence number   : 0
  Invalid source IP address : 0
  Missing mandatory IE      : 0
  GTP in GTP                : 0
  No tunnel resource        : 0
  Tunnel endpoint limit     : 0
Packets permitted           : 16
===============================================================================
```

```
*A:7750-1>show application-assurance group 1:1 gtp
===============================================================================
Application Assurance Group 1:1 GTP
===============================================================================
Admin status      : Up
Event log         : FW_events_log
Event log action  : deny
Mode              : filtering
GTP-C inspection  : Enabled

-------------------------------------------------------------------------------
GTP Statistics                                sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Incoming packets                                       9                      9
Packets denied
  UDP packet length                                    0                      0
  GTP message length                                   0                      0
  GTP version                                          0                      0
-------------------------------------------------------------------------------
Packets permitted                                      9                      9
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
GTP Policing Statistics                       sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Packets arrived                                        9                      9
Packets denied
  gtp-traffic flow-count policer                       0                      0
  Other                                                0                      0
-------------------------------------------------------------------------------
Packets permitted                                      9                      9
-------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------
GTP Filter Statistics                         sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Packets arrived                                        9                      9
Packets denied                                         1                      1
Packets permitted
  gtp-filter                                           8                      8
  no gtp-filter                                        0                      0
-------------------------------------------------------------------------------
Total GTP packets permitted                            8                      8
===============================================================================
```

```
*A:7750-1>show application-assurance group 1 aa-sub-list
==============================================================================
Application-Assurance Subscriber List for Group 1
==============================================================================
type      aa-sub                    ISA       App-Profile              divert
                                    assigned
------------------------------------------------------------------------------
group 1:1
------------------------------------------------------------------------------
sap       1/2/1                     1/1       strict-FW                Yes
sap       1/2/3                     1/1       default                  Yes
transit   1_10.10.68.1/32           1/1       strict-FW                Yes
------------------------------------------------------------------------------
Number of aa-subs found in group 1:1               : 3
Total number of aa-subs found                      : 3
==============================================================================
```

CLI tools dump routines

```
*A:7750-1>tools dump application-assurance group 1:1 flow-record-search isa 1/1

=====================================================
Application-Assurance flow record search
Search Start Time:     "05/19/2020 19:04:37" (UTC)
 Search Criteria:
  group[:partition]:   1:1
  isa:                 1/1
  protocol name:       none specified
  application name:    none specified
  app-group name:      none specified
  flow-status:         none specified
  start-flowId:        none specified
  classified:          none specified
  server-ip:           none specified
  server-port:         none specified
  client-ip:           none specified
  bytes-tx:            none specified
  flow-duration:       none specified
  max-count:           none specified
  flow-modified:       none specified
  search-type:         default
=====================================================
FlowId     Init  Src-ip           Dst-ip           Ip-prot     Src-prt Dst-prt
 Protocol               Application
   Pkts-tx    Bytes-tx             Pkts-disc  Bytes-disc
      Time-ofp(UTC)         Time-olp(UTC)
2          no    10.10.68.3       10.10.68.1       udp         2123    2123
 "gtp"                  "GTP_c"
```

```
    1         46                  0          0
      "05/19/2020 18:54:50" "05/19/2020 18:54:50"
3          yes   10.10.68.1      10.10.68.3        udp         2123    2123
 "gtp"                  "GTP_c"
    2         359                0          0
      "05/19/2020 18:54:50" "05/19/2020 18:54:50"
4          yes   10.10.68.3      10.10.68.1        udp         2123    2123
 "gtp"                  "GTP_c"
    2         326                1          51
      "05/19/2020 18:54:50" "05/19/2020 18:54:55"
7          yes   10.10.68.1      10.10.68.3        udp         63760   2152
 "gtp"                  "GTP_u"
    5         500                0          0
      "05/19/2020 18:54:50" "05/19/2020 18:54:50"
8          yes   10.10.68.3      10.10.68.1        udp         64784   2152
 "gtp"                  "GTP_u"
    5         500                0          0
      "05/19/2020 18:54:50" "05/19/2020 18:54:50"
11         yes   10.10.68.1      10.10.68.3        udp         2123    2123
 "gtp"                  "GTP_c"
    1         136                1          91
      "05/19/2020 18:54:50" "05/19/2020 18:54:50"
SEARCH COMPLETED.
Search End Time:     "05/19/2020 19:04:37" (UTC)
Total Records:       6
=====================================================
```

```
*A:7750-1>tools dump application-assurance group 1:1 admit-deny-stats
===============================================================================================
=============================================
Application-Assurance Group 1:1 Admit-Deny Statistics
===============================================================================================
=============================================
-----------------------------------------------------------------------------------------------
---------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
Packet Validation Statistics
(Packets)            (Packets)            (Packets)            (Packets)
-----------------------------------------------------------------------------------------------
---------------------------------------------
Error
0                    0                    0                    0
Fragments: Out-Of-Order
0                    0                    0                    0
Fragments: All
0                    0                    0                    0
Overload
N/A                  0                  N/A                    0
GTP Sanity
9                    0                    9                    0
-----------------------------------------------------------------------------------------------
---------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
GTP Filter Statistics
(Packets)            (Packets)            (Packets)            (Packets)
-----------------------------------------------------------------------------------------------
---------------------------------------------
GTP Filter: gtp-filter-partner1
 Entry: 1    echo-request
0                    0                    0                    0
 Entry: 2    echo-response
```

```
0                    0                    0                    0
 Entry: 3    error-indication
0                    0                    0                    0
 Entry: 4    supported-extension-headers-notification
0                    0                    0                    0
 Entry: 5    end-marker
0                    0                    0                    0
 Entry: 6    g-pdu
5                    0                    5                    0
 Message Type Default Action
0                    0                    0                    0
 GTPv2 Entry: 524  trace-session-activation
0                    1                    0                    0
 GTPv2 Entry: 525  trace-session-deactivation
0                    0                    0                    1
 GTPv2 Message Type Default Action
3                    0                    3                    0
 IMSI-APN Entry: 1031
N/A                  0                  N/A                    0
 IMSI-APN Filter Default Action
N/A                  0                  N/A                    0
 Max Payload Length
N/A                  0                  N/A                    0
 Message Type Header Sanity
N/A                  0                  N/A                    0
 Invalid TEID
N/A                  0                  N/A                    0
 Invalid Sequence Number
N/A                  0                  N/A                    0
 Invalid Source IP Address
N/A                  0                  N/A                    0
 Missing Mandatory IEs
N/A                  0                  N/A                    0
 GTP in GTP Action
N/A                  0                  N/A                    0
 GTP Tunnel DB Resource
N/A                  0                  N/A                    0
 Tunnel Endpoint Limit
N/A                  0                  N/A                    0
GTP Filter: gtp-filter-partner2
 Message Type Default Action
0                    0                    0                    0
 GTPv2 Message Type Default Action
0                    0                    0                    0
 IMSI-APN Entry: 1040
N/A                  0                  N/A                    0
 IMSI-APN Entry: 1041
N/A                  0                  N/A                    0
 IMSI-APN Filter Default Action
N/A                  0                  N/A                    0
 Max Payload Length
N/A                  0                  N/A                    0
 Message Type Header Sanity
N/A                  0                  N/A                    0
 Invalid TEID
N/A                  0                  N/A                    0
 Invalid Sequence Number
N/A                  0                  N/A                    0
 Invalid Source IP Address
N/A                  0                  N/A                    0
 Missing Mandatory IEs
N/A                  0                  N/A                    0
 GTP in GTP Action
N/A                  0                  N/A                    0
```

```
 GTP Tunnel DB Resource
N/A                  0                  N/A                    0
 Tunnel Endpoint Limit
N/A                  0                  N/A                    0
-----------------------------------------------------------------------------------------------
---------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
Session Filter Statistics
(Sessions)            (Packets)           (Sessions)            (Packets)
-----------------------------------------------------------------------------------------------
---------------------------------------------
Session Filter: restricted_access_partner1
 Entry: 10
1                    0                    0                    0
 Entry: 11
2                    0                    0                    0
 Entry: 20
0                    0                    0                    0
 Default Action
0                    0                    0                    0
Session Filter: restricted_access_partner2
 Entry: 10
0                    0                    0                    0
 Entry: 11
0                    0                    0                    0
 Entry: 20
0                    0                    0                    0
 Default Action
0                    0                    0                    0
Session Filter: restricted_downstream_partner1
 Entry: 10
0                    0                    1                    0
 Entry: 11
0                    0                    1                    0
 Entry: 20
0                    0                    0                    0
 Default Action
0                    0                    0                    0
Session Filter: restricted_downstream_partner2
 Entry: 10
0                    0                    0                    0
 Entry: 11
0                    0                    0                    0
 Default Action
0                    0                    0                    0
-----------------------------------------------------------------------------------------------
---------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
Flow Policer Statistics
(Flows)              (Flows)              (Flows)              (Flows)
-----------------------------------------------------------------------------------------------
---------------------------------------------
Subscriber Flow Count Policers
   limit_roamingSGW_Flows
0                    0                    0                    0
   limit_roamingSGWs_total_Flows
0                    0                    0                    0
-----------------------------------------------------------------------------------------------
---------------------------------------------
```

Conclusion

A 3GPP roaming interface using GTP presents a security risk to mobile access networks. The AA GTP stateful firewall protects the network infrastructure from untrusted roaming partner networks.

Application Assurance — HTTP and HTTPS Redirect

This chapter provides information about Application Assurance - HTTP and HTTPS Redirect.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 15.0.R5.

Overview

This chapter is intended for network architects and engineers working with Application Assurance AA). The user should have a basic understanding of AA policies (application service options, application filters, applications, application QoS policies), HTTP, and DNS.

Nokia recommends using the AppDB when configuring Application QoS Policies (AQPs) for traffic control policies. The AppDB is a default configuration file that can be obtained through Nokia's support organization, and contains information to classify all supported applications, including all relevant application ("app-") groups, applications, and app-filters.

Operators have scenarios where they need to allow some restricted web access to subscribers (while generally blocking all other Internet access). The allowed sites may be authentication sites, such as payment sites. Alternatively, the sites may be portals owned by the operators, which may offer app-store services or user account management.

There are several ways that HTTP redirection can be implemented in an AA policy:

- HTTP policy redirect with (or without) allowed URLs: redirects HTTP traffic to a web portal based on a policy decision
- HTTPS captive redirect with (or without) DNS IP cache allow-list: redirects all HTTP/HTTPS traffic to a web portal, except for traffic toward domains configured in a DNS IP cache
- URL filter triggered HTTP redirect or HTTPS captive redirect (SR OS Release 16.0): redirects all HTTP/HTTPS traffic, based on a URL filter policy decision, where on block decision, the request can be redirected to a message portal

AA is a preferred choice for the operator to configure HTTP redirect. As opposed to CPM-based redirect (offered as part of subscriber management), AA can provide high performance with no system (or ISA) limits on the number of HTTP sessions per second that can be redirected. ISA2 can support more than 5000 redirects per second. CPM-based redirect is quite limited and, in many customer deployments, not able to support the required redirect rates (CPM-based redirect can support a maximum of 250 simultaneous connections). Even if only a simple subscriber redirect is needed (without URL allow-lists or captive redirect), AA-based redirect should be preferred over CPM-based redirect.

The following sections describe the different methods of implementing HTTP and HTTPS redirection with URL allow-lists, why and when to use each one, provide the configuration, and describe how to troubleshoot and correct common errors.

HTTP redirection based on URL filtering is not included in this chapter. For URL filtering, see chapter Application Assurance — Local URL List Filtering.

Configuration

The following two use cases are described in this section:

- AA HTTP redirect with URL allow-lists
- HTTPS redirect using session filter captive redirect

AA HTTP redirect with URL allow-lists

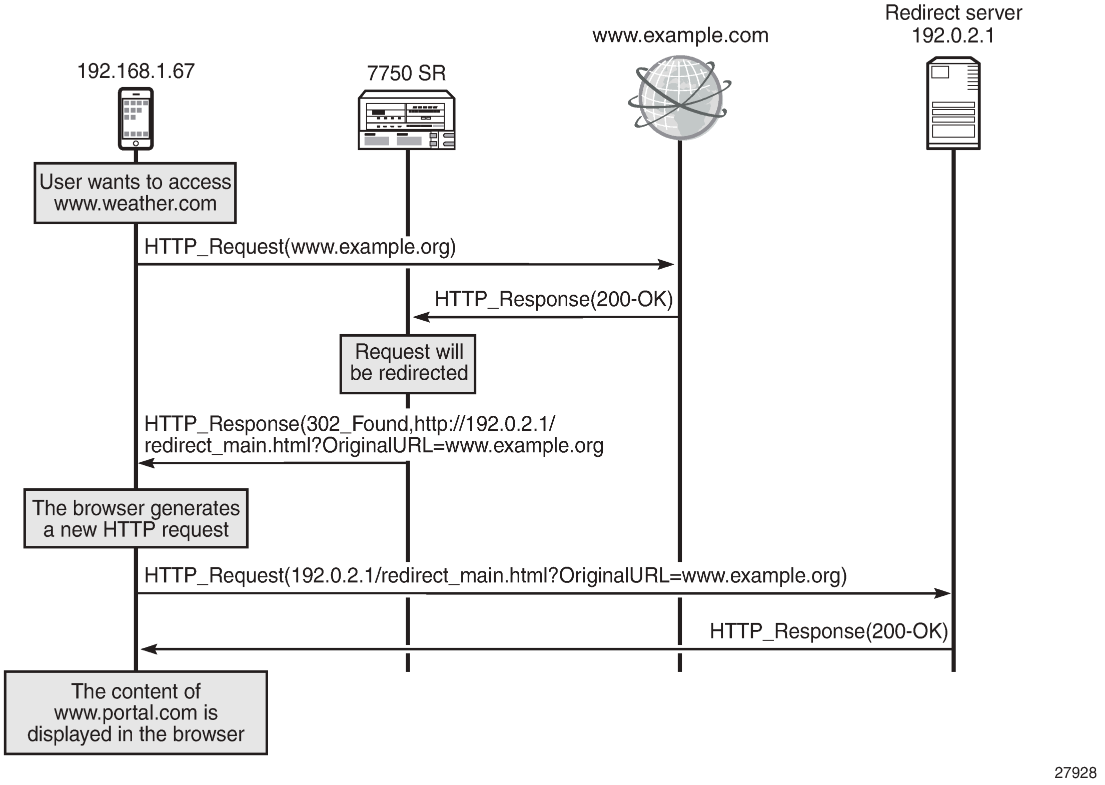
Operators have several use cases where they need to redirect a subscriber to a specific web page. Redirection may be the result of a policy (for example, the subscription does not allow certain content to be viewed) or for informative reasons (for example, the subscriber is not authenticated).

Operators may still need to allow the subscriber to access certain sites (for example, portals provided by the operator).

AA has the capability to perform HTTP redirection by using the "HTTP 302 Found" response code. This works as follows:

The subscriber generates an HTTP request to view a specific page. When AA receives the HTTP response, it classifies the flow and performs the redirect action by modifying that response. The web server replies with a "200 OK" response containing the page content, and AA modifies the HTTP response to a "302 Redirect" and sends it to the subscriber. The response contains the new URL that the subscriber should be redirected to. The subscriber's browser, upon receiving the response, automatically generates a new HTTP request for the URL received, as shown in Figure 20: HTTP redirect.

Figure 20: HTTP redirect


Redirect server
192.0.2.1
www.example.com
192.168.1.67
7750 SR
User wants to access www.weather.com
HTTP_Request(www.example.org)
HTTP_Response(200-OK)
Request will be redirected
HTTP_Response(302_Found,http://192.0.2.1/
redirect_main.html?OriginalURL=www.example.org
The browser generates a new HTTP request
HTTP_Request(192.0.2.1/redirect_main.html?OriginalURL=www.example.org)
HTTP_Response(200-OK)
The content of www.portal.com is displayed in the browser
27928

The redirect is shown in more detail in the following HTTP traces:

1. The browser sends an HTTP request for www.example.org . The HTTP packet contains the following information:

```
Hypertext Transfer Protocol
GET / HTTP/1.1\r\n
… …
Host: www.example.org\r\n
… …
```

- "GET" is followed by the path requested. In this example, the subscriber requested the root path (therefore "/").
- "Host" contains the hostname (domain) requested.

2. The web server replies with a "200 OK" message.
3. The SR OS node receives the server response and classifies the flow. The router determines that it will redirect the request. It replies with a message indicating the new URL along with the original URL. The HTTP packet contains the following information:

```
Hypertext Transfer Protocol
HTTP/1.1 302 Found\r\n
Location: http://192.0.2.1/redirect_main.html?OriginalURL=www.example.org/\r\n
… … …
```

- The HTTP response status code "302" indicates that the web browser should generate a new HTTP request.
- "Location" contains the URL that the subscriber should be redirected to.

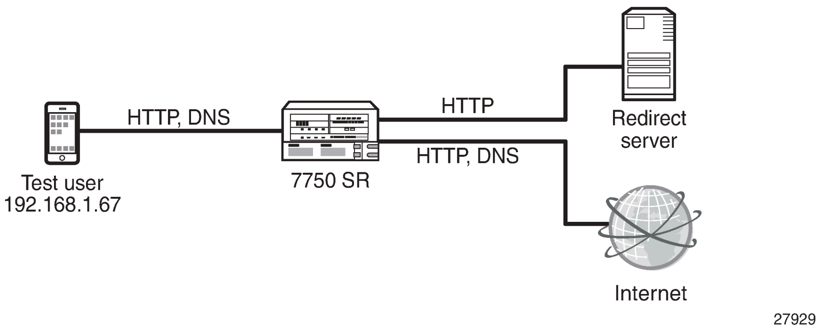
4. The subscriber's browser receives the redirect request and automatically generates a new HTTP request for the new URL. The HTTP packet sent by the browser is as follows:

```
Hypertext Transfer Protocol
GET /redirect_main.html?OriginalURL=www.example.org/ HTTP/1.1
Host: 192.0.2.1\r\n
…
```

- As in step 1, the GET request contains the path.
- Similarly, "`Host`" contains the hostname (domain) requested. "`Host`" can contain either a domain or an IP address.

Figure 21: Example setup shows the example setup containing a 7750 SR with ISA-AA. The setup requires a web client, a web server (which will be the redirect server), and Internet access through the router. A default configuration has been loaded from the AppDB.

Figure 21: Example setup



The following sections describe the configuration of the following objects:

- The redirect policy to configure the redirect URL, the information to be appended to the URL, plus how to handle non-HTTP traffic
- An application and app-filter for the redirect portal
- A charging group for the URLs in the allow-list
- An app-filter to define the URL in the allow-list
- An ASO that will be used as a policy decision to perform redirection
- An AQP to perform redirection

Creating the redirect policy

The redirect policy is configured as follows:

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
       http-redirect "policy-redirect" create
```

```
          description "redirects certain http traffic to a portal"
          template 5
          tcp-client-reset
          redirect-url "http://192.0.2.1/redirect_main.html?OriginalURL=$URL"
          no shutdown
       exit
----------------------------------------------
```

The preceding configuration defines:

- The template to be used ("template 5" means that macro substitution will be used). This allows the operator to append additional information to the redirect information, which can be used by the redirect portal to offer more details about why the subscriber was redirected or to offer more personalized services.
- Traffic that cannot be redirected will be TCP reset. This provides immediate feedback to the subscriber that the service requested is not available. The client application will not have to wait for the TCP session to expire.
- The redirect URL. The original URL will be appended.

Creating the application and app-filter for the redirect portal

The application "Captive Portal" is created and used in the app-filter that contains the redirect server address, as follows:

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
    policy
        begin
        application "Captive Portal" create
            description "HTTP Redirect portal"
            app-group "Web"
        exit

        app-filter
            entry 200 create
                server-address eq 192.0.2.1/32
                application "Captive Portal"
                no shutdown
            exit
        exit
        commit
    exit
----------------------------------------------
```

The preceding configuration defines:

- An application called "Captive Portal"
- An app-filter using destination address 192.0.2.1 (the address of the redirect server) and linking it to the application "Captive Portal"

Creating a charging group for the URLs in the allow-list

The following configuration defines a charging group and an application. The charging group created is assigned to the application.

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
  policy
    begin
    charging-group "allow-list" create
    description " Charging group for redirect portal and URLs in the allow-list"
       exit

  application "Captive Portal" create
    description "HTTP Redirect Portal"
    app-group "Web"
      charging-group "allow-list"
       exit
    commit
----------------------------------------------
```

Nokia recommends defining a charging group for the URLs in the allow-list. Even though an application (or several applications) can be configured, defining a charging group provides a simpler approach; AA application group mappings are not affected, and it is easier to configure AQPs. The redirect pages are usually free-of-charge, so configuring a charging group is the preferred way to aggregate these URLs.

As well as the redirect portal, DNS traffic should also be assigned to the "allow-list" charging group. Before performing an HTTP request, the subscriber's browser will first make a DNS request to resolve the hostname. Therefore, DNS traffic should be allowed.

This example assumes that the operator has loaded an AppDB configuration in the system. The AppDB contains the configuration (app-group, application, app-filter) for DNS traffic. The default application for DNS is as follows:

```
application "DNS" create
  description "Domain Name System"
  app-group "Network Infrastructure"
exit
```

This preceding configuration should be modified as follows:

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
policy
   begin
   application "DNS" create
      description "Domain Name System"
      app-group "Network Infrastructure"
      charging-group "allow-list"
   exit
   commit
----------------------------------------------
```

Creating an application for the URLs in the allow-list

The following configuration creates a new application that contains all the allowed URLs. The application is assigned to the "allow-list" charging group.

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
 policy
   begin
     application "URL-allow-list" create
       description "URLs to be allowed"
       app-group "Web"
       charging-group "allow-list"
     exit
   commit
----------------------------------------------
```

Creating app-filters for the allowed URLs

The following configuration creates an app-filter for the URLs in the allow-list. The configuration contains the URL and assigns it to an application.

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
 policy
   begin
   app-filter
   entry 61001 create
     description "dummy allowed domain"
     expression 1 http-host eq "*.domain.com$"
     application "URL-allow-list"
     no shutdown
   exit
   commit
----------------------------------------------
```

The operator can extend the preceding configuration and add more URLs to the allow-list.

Configuring the ASO and AQP

The following configuration defines:

- An ASO that will be used as a policy criterion to decide whether redirection will occur. The ASO is called "redirect" and has two values: "no" and "yes". The default value will be "no" (do not redirect).
- An AQP to perform redirection. The match criteria will be the preceding ASO and the value of the charging group. If the match criteria are met, the action will be to drop and redirect the flow using the "policy-redirect" defined.

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
       policy
           begin
           app-service-options
               characteristic "redirect" create
```

```
                   value "no"
                   value "yes"
                   default-value "no"
               exit
           exit

           app-qos-policy
               entry 400 create
                   match
                       charging-group neq "allow-list"
                       characteristic "redirect" eq "yes"
                   exit
                   action
                       drop
                       http-redirect "policy-redirect" flow-type dropped-flows
                   exit
                   no shutdown
               exit
            exit
            commit
----------------------------------------------
```

Having the charging group as match criteria is a future-proof solution: the configuration can easily be adapted in the future (if, for example, additional URLs should be allowed).

Note:
The use of ASOs is optional and shown here for completeness. It is used only in some policy-dependent redirect cases, where the operator may need to provide the redirect policy only to some subscribers. In most cases, all subscribers get the same redirect policy and the ASOs are not used. App-profiles may also be used to implement a match condition for a policy-based redirect.

Troubleshooting and debugging

The following show commands can be used to verify the configuration and correct possible errors.

The following command shows the redirect policy:

```
*A:Dut-C# show application-assurance group 1 http-redirect "policy-redirect" detail

===============================================================================
Application Assurance Group 1 HTTP Redirect policy-redirect
===============================================================================
Description        : redirects certain http traffic to a portal
Template           : 5
                   : Redirect supporting macro substitution using HTTP 302
Redirect URL       : http://10.99.99.40/redirect_main.html?OriginalURL=$URL
Captive Redirect   : No
    Redirect HTTPS : No
    VLAN ID        : N/A
Admin Status       : Up
AQP Ref            : Yes

-------------------------------------------------------------------------------
Group 1
-------------------------------------------------------------------------------
Redirects Sent        : 0
Client Resets Sent    : 0
Redirects Not Sent    :
    Out of Resources  : 0
    Config Errors     : 0
    Other Errors      : 0

-------------------------------------------------------------------------------
Total
-------------------------------------------------------------------------------
Redirects Sent        : 0
Client Resets Sent    : 0
Redirects Not Sent    :
    Out of Resources  : 0
    Config Errors     : 0
    Other Errors      : 0
===============================================================================
```

The preceding command provides the following information:

- The template used, along with a definition of that template
- The URL that the subscriber will be redirected to
- Whether this is a captive redirect and whether HTTPS will be redirected
- The admin status of the policy. An admin status of "Down" indicates that the policy is not in use.
- Whether this policy is referenced by an AQP. If it is not referenced, it is not used.
- "Redirects Sent" indicates the number of times that AA has sent a redirect message.
- "Client Resets Sent" indicates the number of times that TCP reset was sent.
- "Out of Resources" indicates the number of times that a redirect was not sent. The reasons relate to exhaustion of ISA resources (for example, memory).
- "Config Errors": for example, a command is in shutdown state
- "Other Errors": errors that are not in any of the preceding two categories

The following command shows the "Captive Portal" application:

```
*A:Dut-C# show application-assurance group 1 policy application "Captive Portal"

===============================================================================
Application Instance "Captive Portal"
===============================================================================
Description    : HTTP Redirect portal
App Group      : Web
Export Id      : none
Charging Group : allow-list

References
-------------------------------------------------------------------------------
AA Sub Stat    : none
Cflowd         : none
AQP Entries    :
App Filters    :
          200

Reference Counts
-----------------------------------
No. of cflowd references      : 0
No. of AQP entry references   : 0
No. of app filter references  : 1
===============================================================================
```

The preceding command provides the following information:

- The charging group assigned
- The app-filters linked (200)

The following command shows the app-filter for the redirect portal:

```
*A:Dut-C# show application-assurance group 1 policy app-filter 200

app-filter
    entry 200 create (0 flows, 0 B)
        server-address eq 10.99.99.40/32
        application "Captive Portal"
        no shutdown
    exit
exit
```

Using the preceding command, the operator can verify the number of flows and bytes matched and, therefore, determine if traffic toward the redirect portal was classified.

Show the AQP for redirection:

```
*A:Dut-C# show application-assurance group 1 policy app-qos-policy 400

================================================================================
Application QOS Policy Entry 400 (Application Based Policy)
================================================================================
Description : (Not Specified)
Admin State : in-service
Hits        : 0 flows
Conflicts   : 0 flows

Match :
    Charging Group            : neq allow-list
    ASO Characteristics       :
        redirect              : eq yes

Action :
    Drop                          : yes
    HTTP Redirect                 : policy-redirect flow-type dropped-flows
================================================================================
```

Using the preceding display, the operator can verify whether flows are matched to the AQP.

Operational considerations

- QUIC is a protocol originally developed by Google to reduce end-to-end latency. It is used by Google services and browsers (Chrome). Because QUIC cannot be redirected, traffic that would have been redirected may be dropped instead. This can be achieved in two ways:
  1. Allow UDP traffic on port 53 and block all other UDP traffic.

     QUIC is over UDP. Therefore, by allowing UDP on port 53 and blocking all other traffic, the operator can perform DNS requests, but all other UDP traffic (including QUIC) will be blocked.
  2. Create an application to block UDP traffic on port 443 at the first packet.

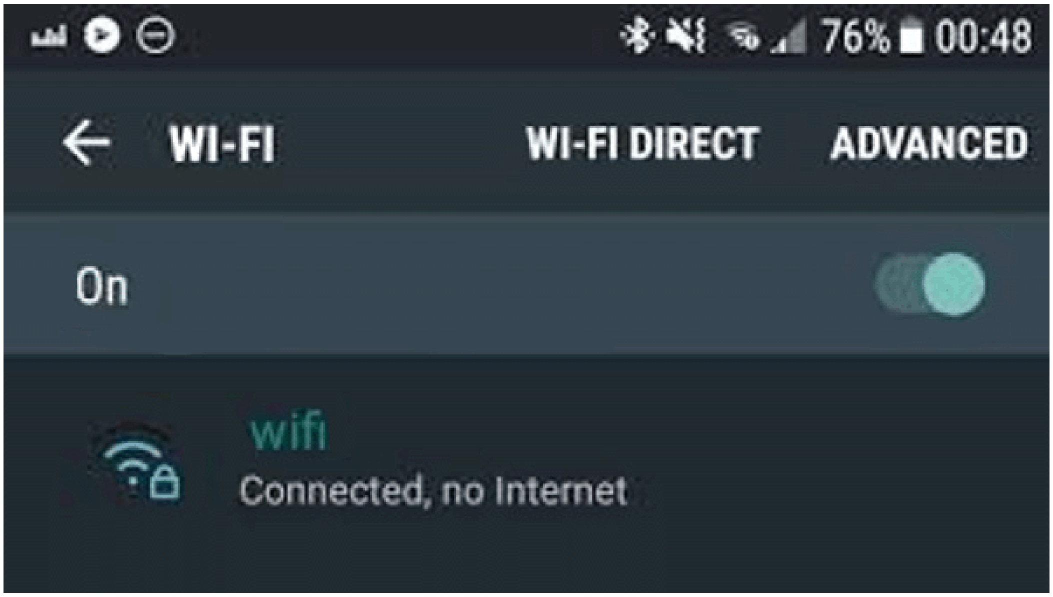
     QUIC is running on port 443. Therefore, the following app-filter will block QUIC traffic (UDP traffic on port 443). The "first-packet-validate" option is used for applications using a well-known TCP/UDP port and ensures that the policy will be applied from the first packet, while allowing AA to detect an unexpected application to the well-known port.

```
policy
   begin
   application "UDP 443 Validate" create
      description "First Packet Verify application for any UDP on port 443
                  such as QUIC"
      app-group "Web"
   exit
   app-filter
      entry 700 create
         ip-protocol-num eq udp
         server-port eq 443 first-packet-validate
         application "UDP 443 Validate"
         no shutdown
      exit
   exit
   commit
```

- Some mobile OSs (for example, Android) generate an HTTP request to a predefined "test URL" to verify that Internet access exists. The HTTP request is sent immediately after the device connects to a WiFi network. Even though the HTTP requests are generated in the background, redirecting these URLs will result in the redirect response to be presented in the web browser.

  Depending on the operator's use case, these URLs may be redirected. Redirecting these URLs may affect the DNS cache of the mobile device. The device thinks that it does not have full Internet access. On Android devices, the subscriber will see a message similar to the following:

  Figure 22: Connected device without full Internet access



The app-filter for the redirect server was configured using the server IP address (`server-address eq 192.0.2.1/32`). The app-filter for the redirect server can also be configured using a host (or any other HTTP-related) expression. An example configuration follows:

```
app-filter
```

```
   entry 200 create
       expression 1 http-host eq "*.redirect_domain.com$"
       application "Captive Portal"
       no shutdown
   exit
exit
```

The preceding configuration will classify HTTP traffic when the hostname is "*.redirect_domain.com$".

If this configuration approach is chosen, the operator must be certain that the app-filter will classify all the traffic toward the redirect portal. There may be cases where some content may be fetched from additional locations too (for example, "*.content_domain.com$"); therefore, additional app-filters should be configured.

To ensure that the configuration is correct and complete, a trace can be analyzed.

Table 8: AA redirect errors describes what each AA redirect error means.

Table 8: AA redirect errors

Redirect not sent	Description
Out of resources	No buffer resources
Config errors	Disabled redirect configuration, VLAN ID missing, session is not TCP
Other errors	Fragmented response packet

HTTPS redirect using session filter captive redirect

A second use case is HTTP redirect using session filter captive redirect. Captive redirect is used to redirect HTTP or HTTPS flows, without sending any traffic to the Internet. In the case of HTTP traffic, it is achieved by terminating the TCP sessions in AA. HTTP flows will then be redirected to a predefined URL, while non-HTTP flows will be TCP reset.

However, the majority of web traffic is HTTPS. Simply blocking (and TCP resetting) HTTPS traffic is not friendly to the end user: they are not given any explanation as to why access was not allowed, might think there was some different error, and will probably try again. AA can also redirect HTTPS traffic to a different page and provide information to the subscriber about why access to the site is not allowed.

Operators may still need to allow some HTTP traffic. Captive redirect can be combined with a DNS IP cache to allow (not redirect) certain HTTP traffic. The operator may configure a DNS IP cache for the allowed domains. This section will provide an example configuration to achieve this.

A DNS IP cache is used to ensure that traffic will be correctly classified and charged. After the feature has been enabled, AA will sniff subscribers' DNS responses and populate a cache containing the HTTP domain and the IP addresses that the DNS server provided. Traffic will only be classified into applications or app-groups if the hostname configured matches the one found, and the IP address of the packet matches an entry in the DNS cache.

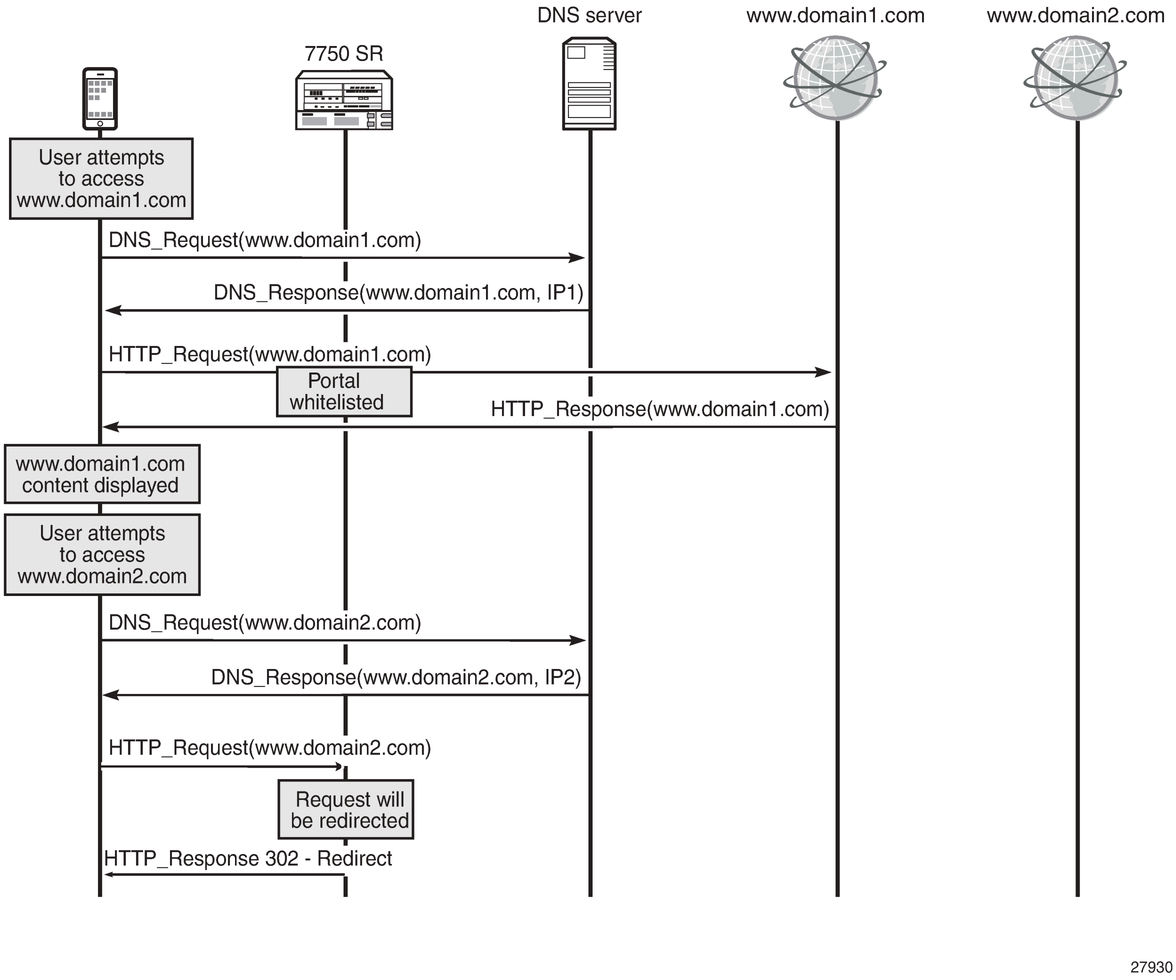
Additional information about DNS IP cache (including configuration instructions) is available in the Application Assurance — DNS IP Cache chapter.

Captive redirect is mostly used when the operator wants to redirect HTTPS traffic. Even though this chapter also provides the configuration for HTTP captive redirect, it is only shown for completeness.

Captive redirect of HTTP traffic is hardly ever used, adds more complexity, and provides no additional benefit over the HTTP redirect described in the previous section.

Figure 23: HTTP redirect using session filter captive redirect shows the high-level message flow of HTTP redirect using session filter captive redirect.

Figure 23: HTTP redirect using session filter captive redirect


DNS server
www.domain1.com
www.domain2.com
7750 SR
User attempts to access www.domain1.com
DNS_Request(www.domain1.com)
DNS_Response(www.domain1.com, IP1)
HTTP_Request(www.domain1.com)
Portal whitelisted
HTTP_Response(www.domain1.com)
www.domain1.com content displayed
User attempts to access www.domain2.com
DNS_Request(www.domain2.com)
DNS_Response(www.domain2.com, IP2)
HTTP_Request(www.domain2.com)
Request will be redirected
HTTP_Response 302 - Redirect
27930

Consider an operator that needs to allow traffic to www.domain1.com and redirect all other HTTP traffic, see Figure 23: HTTP redirect using session filter captive redirect. Therefore, the operator will configure the DNS IP cache to sniff DNS traffic and populate the cache with the IP addresses of that allowed domain.

- If the subscriber needs to access www.domain1.com , the DNS traffic will be analyzed by SR and the DNS IP cache will be updated. When the subscriber sends the HTTP request toward www.domain1.com , the traffic will be allowed because the destination IP address of the traffic will match the IP address contained in the cache.
- If the subscriber attempts to access www.domain2.com , the request will not be allowed because the destination IP address will not be present in the cache. This request will be redirected. The DNS IP cache will not be populated with the IP address of www.domain2.com .

Creating a DNS IP cache for the allowed domain

The following configuration will create and populate a DNS IP cache:

- DNS responses for the host "*.domain1.com$". Traffic to www.domain.com will be allowed.
- responses only from the DNS server with IP address 8.1.17.21 (which is considered a trusted DNS server)

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
      dns-ip-cache "domain1-cache" create
          description "dns cache for domain1"
          dns-match
              domain "domain1" expression "*.domain1.com$"
              server-address 8.1.17.21
          exit
          no shutdown
      exit
----------------------------------------------
```

Creating a default subscriber policy

The following configuration will create a default subscriber policy to snoop DNS responses. This AQP is used to populate the cache.

No match criteria are needed. AA will only analyze the DNS traffic to populate the cache.

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
policy
    begin
    app-qos-policy
        entry 50 create
            action
               dns-ip-cache "domain1-cache"
            exit
            no shutdown
        exit
    exit
    commit
exit
----------------------------------------------
```

Configuring captive redirect

The following will configure a captive redirect policy for the redirect portal. The configuration defines the redirect URL. Using macro substitution, the original URL that the subscriber requested will be appended in the redirect URL. Non-HTTP traffic will be TCP reset. The template specifies that the redirect template be used: "template 5" defines macro substitution using HTTP 302. HTTPS traffic will also be redirected.

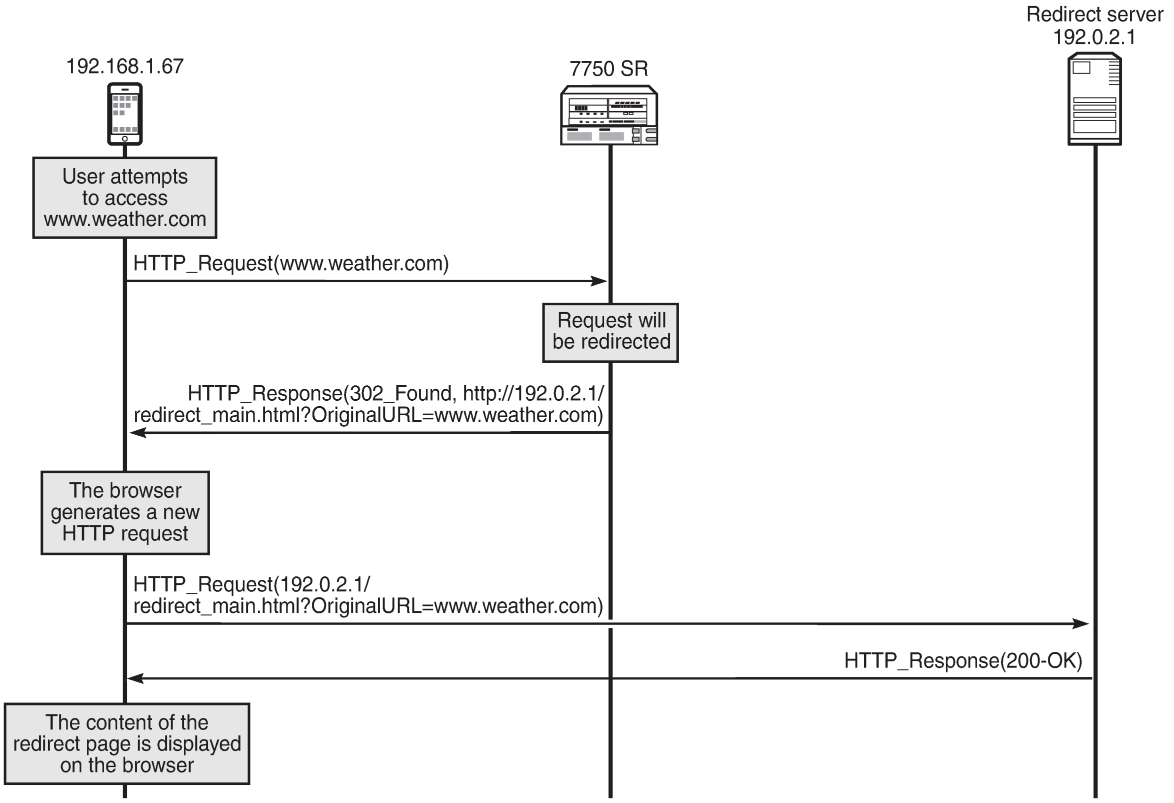
```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
http-redirect "capt-redirect" create
    template 5
    tcp-client-reset
    redirect-url "http:// 192.0.2.1/redirect_main.html?OriginalURL=$URL"
    captive-redirect
        redirect-https
    exit
    no shutdown
exit
---------------------------------------------
```

The preceding configuration did not define the VLAN ID. The VLAN ID refers to an AA IP aa-interface that the operators need to create within the subscriber service (IES or VPRN) for ESM subscribers, and is used by the ISA to respond to the subscriber. Operators need to create one interface per AA ISA card.

The VLAN ID does not need to be defined if the ISA group aa-sub-scale mode is configured for DSM.

Figure 24: HTTP redirection shows the high-level message flow of HTTP redirection.

Figure 24: HTTP redirection

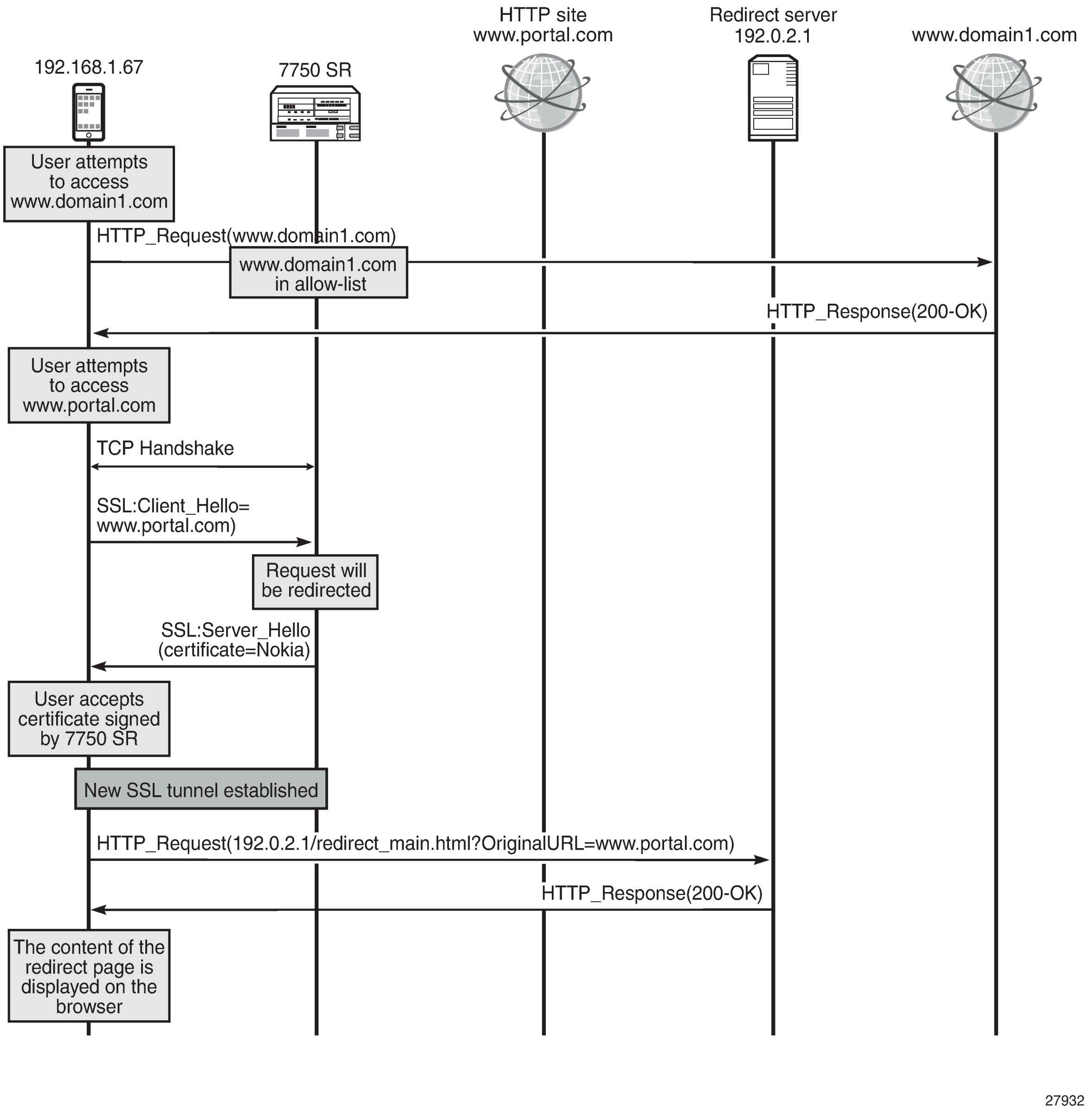

Redirect server
192.0.2.1
192.168.1.67
7750 SR
User attempts to access www.weather.com
HTTP_Request(www.weather.com)
Request will be redirected
HTTP_Response(302_Found, http://192.0.2.1/
redirect_main.html?OriginalURL=www.weather.com)
The browser generates a new HTTP request
HTTP_Request(192.0.2.1/
redirect_main.html?OriginalURL=www.weather.com)
HTTP_Response(200-OK)
The content of the redirect page is displayed on the browser
27931

As shown in Figure 24: HTTP redirection:

- The subscriber attempts to access a site.
- The SR OS node determines that traffic will be redirected. No traffic will be allowed to pass.
- The SR OS node returns an HTTP response message indicating the redirect.
- The browser automatically generates a new request toward the redirect portal.
- The content of the redirect page is shown to the subscriber.

Figure 25: Access to domain included in the allow-list versus HTTPS redirect shows how HTTP access to a domain included in the allow-list is granted, while access to an HTTPS site is redirected.

Figure 25: Access to domain included in the allow-list versus HTTPS redirect

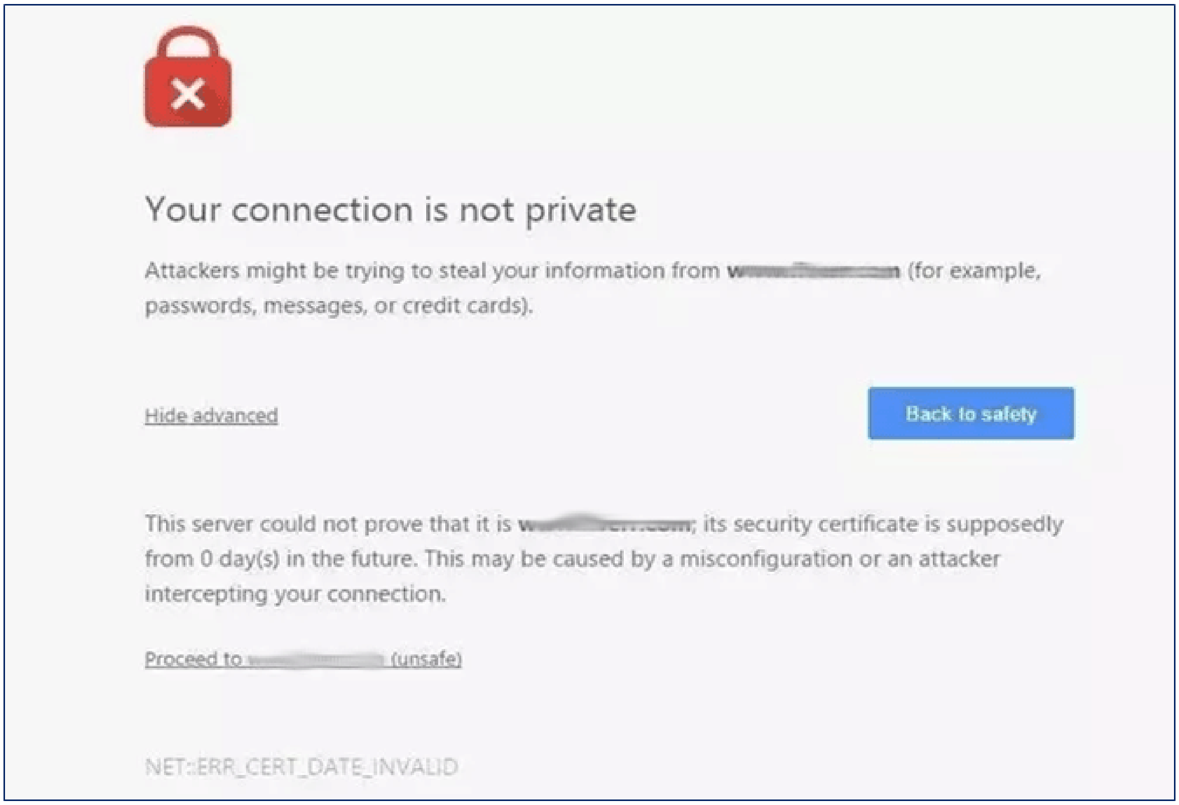

HTTP site
www.portal.com
Redirect server
192.0.2.1
www.domain1.com
192.168.1.67
7750 SR
User attempts to access www.domain1.com
HTTP_Request(www.domain1.com)
www.domain1.com in allow-list
HTTP_Response(200-OK)
User attempts to access www.portal.com
TCP Handshake
SSL:Client_Hello= www.portal.com)
Request will be redirected
SSL:Server_Hello (certificate=Nokia)
User accepts certificate signed by 7750 SR
New SSL tunnel established
HTTP_Request(192.0.2.1/redirect_main.html?OriginalURL=www.portal.com)
HTTP_Response(200-OK)
The content of the redirect page is displayed on the browser
27932

As shown in Figure 25: Access to domain included in the allow-list versus HTTPS redirect:

- The subscriber first attempts to access www.domain1.com. This site has been configured as a allowed site in SR OS.
- The subscriber next attempts to access the HTTPS site www.portal.com. This portal is not an allowed one. Therefore, the system will perform HTTPS redirection.

The system will return a self-signed Nokia certificate, similar to the following.

Figure 26: Subscriber receives warning



Your connection is not private

Attackers might be trying to steal your information from [redacted] (for example, passwords, messages, or credit cards).

This server could not prove that it is [redacted]; its security certificate is supposedly from 0 day(s) in the future. This may be caused by a misconfiguration or an attacker intercepting your connection.

NET::ERR_CERT_DATE_INVALID

- If the subscriber accepts the new certificate, a new SSL tunnel will be established and a new request toward the redirect server will be generated.
- Finally, the subscriber will see the contents of the redirect page.

Configuring a session filter

The following session filter is configured to:

- Allow DNS traffic (UDP traffic on port 53)
- Allow domains included in the DNS IP cache (traffic to www.domain1.com)
- Allow traffic to the redirect portal (portal IP address: 192.0.2.1)
- Redirect all other TCP traffic to the redirect portal

```
*A:Dut-C>config>app-assure>group# info
----------------------------------------------
session-filter "capt_redirect" create
    description "Session filter for captive redirect policy"
    default-action deny
    entry 5 create
        match
            ip-protocol-num udp
            dst-port eq 53
        exit
        action permit
    exit
    entry 10 create
        match
            dst-ip dns-ip-cache "domain1-cache"
         exit
         action permit
    exit
    entry 15 create
        description "Allow traffic to the redirect landing page server"
        match
            ip-protocol-num tcp
            dst-ip 192.0.2.1/32
            dst-port eq 80
        exit
        action permit
    exit
    entry 20 create
        match
            ip-protocol-num tcp
        exit
        action http-redirect "capt-redirect"
    exit
exit
----------------------------------------------
```

Troubleshooting and debugging: show commands

Operators can verify their configuration as follows.

The following command shows the session filter together with the number of flows matched:

```
*A:Dut-C# show application-assurance group 1 session-filter "capt_redirect"

===============================================================================
AA Session Filter Instance "capt_redirect"
===============================================================================
Description    : Session filter for captive redirect policy
Default Action : deny
    Event Log  : (Not Specified)
AQP Entries    :
-------------------------------------------------------------------------------
Filter Match Criteria
-------------------------------------------------------------------------------
Entry          : 5
Description    : (Not Specified)
IP Protocol    : udp
Dest Port      : eq 53
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 10
Description    : (Not Specified)
IP Protocol    : none
Dest IP Cache  : domain1-cache
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 15
```

```
Description    : Allow traffic to the redirect landing page server
IP Protocol    : tcp
Dest IP        : 192.0.2.1/32
Dest Port      : eq 80
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 20
Description    : (Not Specified)
IP Protocol    : tcp
Action         : http-redirect capt-redirect
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
No. of entries   : 4
===============================================================================
```

The preceding show output provides the following information:

- The "Default Action" specifies the action to take if a packet does not match any of the filters defined in the session filter (deny).
- The filters defined are:
  - Allowing DNS traffic (UDP traffic on port 53, entry 5)
  - Allowing traffic matching the DNS IP cache (entry 10)
  - Allowing traffic to the redirect portal (entry 15)
  - Redirecting all other traffic (entry 20)
- The "Hits" counter on each filter provides information about how many flows matched the corresponding filter.

The following command shows the DNS IP cache:

```
*A:Dut-C# show application-assurance group 1 dns-ip-cache "domain1-cache"

===============================================================================
Application Assurance Group 1 dns-ip-cache "domain1-cache"
===============================================================================
Admin Status                : Up
AQP Ref                     : Yes
Domain expressions          : 1 (out of 32)
Server addresses            : 1 (out of 64)
High watermark              : 90%
Low watermark               : 80%
Cache size                  : 10
-------------------------------------------------------------------------------
ISA                        Usage         (%)        Alarm            Hit Count
                                                    State
-------------------------------------------------------------------------------
1/2                            0        0.00        clear                    0
3/1                            0        0.00        clear                    0
===============================================================================
```

The preceding command shows:

- The administrative status of the DNS IP cache ("Up"). An administrative status of "Down" indicates that the DNS IP cache is not in use.
- Whether the DNS IP cache created is referenced by an AQP

- The number of configured domain expressions (in the example: "*.domain1.com$") and the maximum value (shown in brackets)
- The number of server addresses and the maximum value (shown in brackets)
- The number of times a cache lookup was successful ("Hit Count"). A hit count of 0 indicates a possible configuration error.

The following command shows information about the HTTP redirect:

```
*A:Dut-C# show application-assurance group 1 http-redirect "capt-redirect" detail

================================================================================
Application Assurance Group 1 HTTP Redirect capt-redirect
================================================================================
Description         : (Not Specified)
Template            : 5
                    : Redirect supporting macro substitution using HTTP 302
Redirect URL        : http://192.0.2.1/redirect_main.html?OriginalURL=$URL
Captive Redirect    : Yes
    Redirect HTTPS  : Yes
    VLAN ID         : N/A
Admin Status        : Up
AQP Ref             : No

--------------------------------------------------------------------------------
Group 1
--------------------------------------------------------------------------------
Redirects Sent         : 0
Client Resets Sent     : 0
Redirects Not Sent     :
    Out of Resources   : 0
    Config Errors      : 0
    Other Errors       : 0

--------------------------------------------------------------------------------
Total
--------------------------------------------------------------------------------
Redirects Sent         : 0
Client Resets Sent     : 0
Redirects Not Sent     :
    Out of Resources   : 0
    Config Errors      : 0
    Other Errors       : 0
================================================================================
```

The preceding output provides the following information:

- The template used and description about the template
- The redirect URL
- Whether this is a captive redirect and whether HTTPS will be redirected also
- The VLAN ID. Captive redirect uses the provisioned VLAN ID to send the HTTP response to the subscribers; therefore, this VLAN ID must be properly assigned in the same VPN as the subscriber (not needed for DSM).
- The admin status of the redirect. An admin status of "Down" indicates that the policy is not in use.
- "Redirects Sent" indicates the number of times that AA has sent a redirect message.
- "Client Resets Sent" indicates the number of times that TCP reset was sent.

- "Out of Resources" indicates the number of times that a redirect was not sent. The reasons relate to exhaustion of resources in ISA (such as memory).
- "Config Errors": for example, a command is in shutdown state
- "Other Errors": errors that are not in any of the preceding two categories

Troubleshooting and debugging: tools commands

Based on an FTP server with IP address 192.168.1.100, the operator can dump the contents of the DNS IP cache, as follows:

```
*A:Dut-C# tools dump application-assurance group 1 dns-ip-cache "domain1-cache" url ftp://
username:password@192.168.1.100/path/to/mylog.log
```

The preceding command will generate a text file, containing the following:

```
======================================================
Application-Assurance dns-ip-cache "domain1-cache"
Current Time:          "11/08/2017 12:45:32" (UTC)
  group:               1
  isa:                 3/2
  admin state:         no shutdown
  max-entries:         10
======================================================
ip-address  creationTime(UTC)  lastUpdated(sec)  numDNSResponses   lastMatchTime(UTC)  numTimes
Matched

Total entries in-use:   0

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

The log file shows the number of times that a cache entry was matched (numDNSResponses) and the number of times that traffic to this address was sent (numTimesMarches).

Operational considerations

HTTPS redirect is not available in Releases earlier than SR OS 14.0 (the session will simply be TCP reset).

Some browsers store certificates of popular sites. If the subscriber tries to access one of those sites, the browser will not accept the AA-signed certificate and the redirect page will never be displayed. The subscriber does not have an option to accept the certificate.

If the operator wants to configure an HTTPS redirect (either on VSR since Rel. 14.0 or in CMG since Rel.10.0), the FW ASL License is mandatory. It is needed to enable the use of the session filter by the captive redirect.

When using the standard HTTP redirect, there is no need for an aa-interface. However, when using the captive redirect, the operator needs one aa-interface per VPN service, not one interface per subscriber VLAN. To associate the redirect policy with the subscriber in the correct VPN for captive redirect, the operator needs to use ASO/ASO override, which will point to a redirect policy per VPN-this redirect policy will point to an aa-interface in the correct VPN service.

Captive redirect is special in the sense that no traffic is allowed from the Internet (apart from the traffic to the redirect portal).

HTTPS captive redirect cannot be used with ESM subscribers on L2-aware NAT. The forwarding model used for L2-aware NAT does not allow the use of HTTPS captive redirect. If configured, the session filter will be hit, but no TCP message will be sent to the subscriber. The AA statistics will show discarded packets.

Conclusion

This chapter shows how an operator can configure HTTP redirection. The chapter describes the different use cases, how HTTP redirection can be combined with DNS IP cache, and how HTTPS traffic can be redirected.

Application Assurance — HTTP In Browser Notification

This chapter provides information about Application Assurance HTTP in browser notification.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

The information and configuration in this chapter are based on SR OS Release 13.0.R2.

There are no specific prerequisites for this example.

Overview

Using the SR OS nodes and application assurance (AA), subscribers connected to an operator network can be sent fully customizable on-screen notification messages displayed in a non-disruptive and cost-effective manner through their web browser.

This chapter describes the different options for the operator to customize the notification messages returned to the subscriber using either different HTTP-notification policies or using the flexible HTTP-URL-PARAM VSA mechanism.

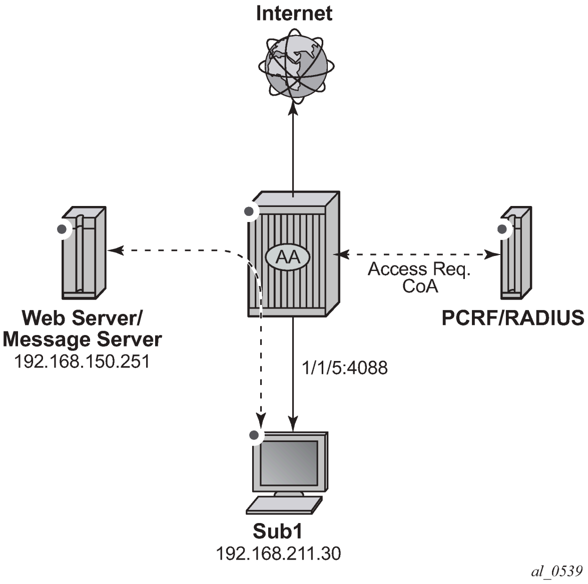
This chapter also describes the additional configuration required with the introduction in SR OS 13.0.R1 of the Notification status monitoring capability allowing the system to notify the subscriber at the next candidate flow instead of waiting for the next notification interval in case the previous notification did not result in a success.

Configuration

The setup comprises of the following elements, see Figure 27: HTTP notification – setup:

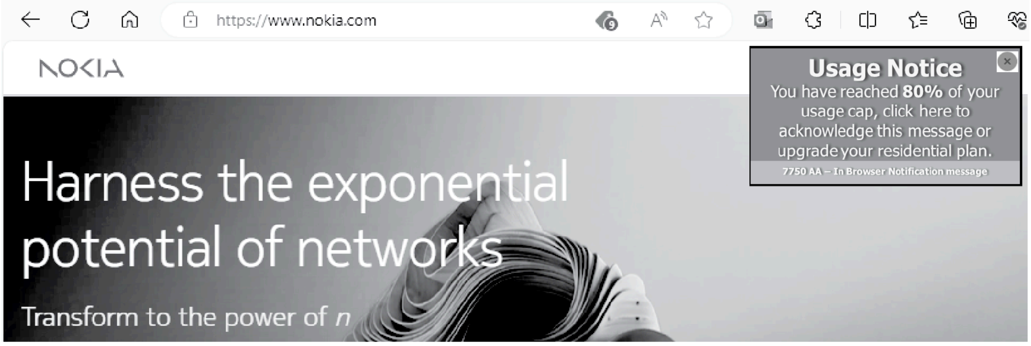
- 7750 SR node with ISA-AA.
- Apache web server (delivering notification Javascript and content).
- Subscriber (desktop/tablet/smartphone).
- Authentication, authorization and accounting (AAA) for subscriber authentication and policy modification.
- Internet access.

Figure 27: HTTP notification – setup



This example describes how to configure HTTP notification to display different notification messages. It demonstrates a simple example in the context of a residential deployment where a message is displayed when the subscriber reaches 80% or 100% of their maximum allowed volume (usage cap).

Figure 28: Notification Message Example – Quota 80%


https://www.nokia.com
NOKIA
Harness the exponential potential of networks
Transform to the power of *n*

Usage Notice
You have reached **80%** of your usage cap, click here to acknowledge this message or upgrade your residential plan.
7750 AA – In Browser Notification message

In this context the operator has two options:

- Use a dedicated http-notification policy per message type.
- Use a common http-notification policy for any message type together with the newly introduced Http-Url-Param RADIUS VSA.

This example provides configuration examples for both options.

HTTP Notification Policy per Message Type

In this option a dedicated http-notification policy for each notification message is required.

HTTP Notification Policy

Two dedicated HTTP notification policies are used to return a different message to the subscribers when reaching 80% and 100% of their usage cap, the interval in between notifications is set to 15 minutes.

```
configure
    application-assurance group 1
        http-notification "notification-quota-100" create
           description "100% Usage Cap Notification"
           script-url "http://192.168.150.251/In-Browser-Notification/script/quota-100.js"
           template 1
           interval 15
           no shutdown
        exit

configure
    application-assurance group 1
        http-notification "notification-quota-80" create
            description "80% Usage Cap Notification"
            script-url "http://192.168.150.251/In-Browser-Notification/script/quota-80.js"
            template 1
            interval 15
            no shutdown
        exit
```

Notification Status Monitoring

The operator then needs to enable the http-match-all-req feature for any HTTP request sent to the messaging server in order to monitor HTTP notification success and failures. This is done by creating a new application and enabling http-match-all-req within the app-filter.

Success and failure notifications include a specific HTTP encoded URI automatically interpreted as a success or a failure by Application Assurance on a per subscriber basis. If a failure is detected, the system will automatically attempt to notify a new candidate flow instead of waiting for the next notification interval.

```
configure
   application-assurance group 1:1 policy
      application "IBN Messaging Server" create
     app-group "Web"
   exit
   app-filter
     entry 100 create
        expression 1 http-host eq "^192.168.150.251$"
        http-match-all-req
        application "IBN Messaging Server"
        no shutdown
    exit
exit
```

App-Profiles and App-Service-Options

Event based HTTP notifications is enabled by a policy modification triggered via RADIUS or Gx by modifying the subscriber app-profile or using the Application Service Option (ASO) override.

In this implementation of the HTTP notification policy per message type, the following ASO configuration is used:

```
configure
    application-assurance group 1:1 policy
        app-service-option
            characteristic "quota-message-notification" create
                value "100"
                value "80"
                value "disabled"
                default-value "disabled"
            exit
        exit
        app-profile "1-1/Default" create
            divert
        exit
```

The ASO characteristic *quota-message-notification* values of 100 and 80 enable the App-Qos-Policy (AQP) *notification-quota-100* and *notification-quota-80* as defined below:

```
configure
    application-assurance group 1:1 policy app-qos-policy
        entry 1000 create
            match
                characteristic "quota-message-notification" eq "100"
                application neq "Advertising Statistics"
            exit
            action
                http-notification "notification-quota-100"
            exit
            no shutdown
        exit
        entry 1100 create
            match
                characteristic "quota-message-notification" eq "80"
                application neq "Advertising Statistics"
            exit
            action
                http-notification "notification-quota-80"
            exit
            no shutdown
        exit
```

RADIUS Policy

The following RADIUS CoA message is used to override the ASO characteristic of a residential subscriber so that a notification message can be returned to the subscriber when they reach 80% of their usage cap:

```
    NAS-Port-Id = "1/1/5:4088"
    Framed-IP-Address = 192.168.211.30
    Alc-AA-App-Service-Options = "quota-message-notification=80"
    Alc-App-Prof-Str = "1-1/Default"
```

Show Commands

Before the subscriber usage cap limit is reached, and before the RADIUS CoA message is received, the subscriber ASO parameter flag quota-message-notification is set to its default value *disabled* and therefore no App QoS Policy is triggered.

```
A:PE# show application-assurance group 1:1 aa-sub esm "sub1" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub1 (esm)
ISA assigned            : 1/2
App-Profile             : 1-1/Default
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : N/A
-------------------------------------------------------------------------------
Traffic                        Octets              Packets                Flows
-------------------------------------------------------------------------------
... ...
-------------------------------------------------------------------------------
Application Service Options (ASO)
-------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
-------------------------------------------------------------------------------
quota-message-notification       disabled                         default
===============================================================================
```

After the RADIUS CoA message is sent, the subscriber ASO characteristic *quota-message-notification* value is set to *80*, the subscriber-related App QoS Policy entry 1100 now matches for this subscriber:

```
A:PE# show application-assurance group 1:1 aa-sub esm "sub1" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub1 (esm)
ISA assigned            : 1/2
App-Profile             : 1-1/Default
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : N/A
-------------------------------------------------------------------------------
Traffic                        Octets              Packets                Flows
-------------------------------------------------------------------------------
... ...
-------------------------------------------------------------------------------
Application Service Options (ASO)
-------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
-------------------------------------------------------------------------------
quota-message-notification       80                               dyn-override
===============================================================================
```

The same command can be used to identify when the last successful subscriber notification occurred, see the Last HTTP Notified Time field:

```
A:PE# show application-assurance group 1:1 aa-sub esm "sub1" summary
================================================================================
Application-Assurance Subscriber Summary (realtime)
================================================================================
AA-Subscriber           : sub1 (esm)
ISA assigned            : 1/2
App-Profile             : 1-1/Default
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : 2014/06/24 15:35:49
--------------------------------------------------------------------------------
```

The operator can also identify how many notifications have been sent per http-notification policy per partition:

```
A:PE# show application-assurance group 1 http-notification "notification-quota-80"
================================================================================
Application Assurance Group 1 HTTP Notification "notification-quota-80"
================================================================================
Description  : 80% Usage Cap Notification
Template     : 1 - Javascript-url with subId and optional Http-Url-Param
Script URL   : http://192.168.150.251/In-Browser-Notification/script/quota-80.
               js
Admin Status : Up
AQP Ref      : Yes
Interval     : 15 minutes
--------------------------------------------------------------------------------
Group 1:1 Statistics
--------------------------------------------------------------------------------
Notified             : 2                     Succeeded : 2
Criteria Not Matched : 5                     Failed    : 0
================================================================================
```

The counter Criteria Not Matched is the number of HTTP flows which did not meet the AA ISA flow selection criteria for In Browser Notification. HTTP flow selection is constrained so that only HTTP web pages flows originating from a web browser are targeted, HTTP requests for content such as video or images are not candidate for notification.

HTTP Notification Customization using RADIUS VSA

Instead of using a dedicated HTTP notification policy for every single message type, the operator can return a RADIUS Http-Url-Param VSA at subscriber creation time or via CoA to customize the notification URL using a single policy. This VSA string is automatically appended to the end of the HTTP notification script-url by the SR OS which can then be used by the web server to decide which notification message to return to the subscriber.

SR OS Release supports 1 active HTTP Notification policy per subscriber, 8 different HTTP notification policies per AA ISA group and 1500 different values for the Http-Url-Param VSA. Therefore, using the Http-Url-Param VSA for the customization of the notification is the recommended model to scale the number of notification messages.

For example:

- RADIUS VSA (Alc-AA-Sub-Http-Url-Param): &message=quota80"
- 7750 SR HTTP Notification configured script-url: http://1.1.1.1/notification.js
- Subscriber HTTP request to the messaging server:

```
http://1.1.1.1/notification.js?SubId=sub1&var=&message=quota80
```

HTTP Notification Policy

A single HTTP notification policy is used to return different notification messages. The interval between notifications is set to 15 minutes.

```
configure
    application-assurance group 1
        http-notification "in-browser-notification" create
            description "Default HTTP Notification Policy"
            script-url "http://192.168.150.251/In-Browser-Notification/script/
                                                         notification-select.php"
            template 1
            interval 15
            no shutdown
        exit
```

Note:
This example does not describe the content of the *notification-select.php* file used to parse the URL parameters.

Notification Status Monitoring

The operator then needs to enable the http-match-all-req feature for any HTTP request sent to the messaging server in order to monitor HTTP notification success and failures. This is done by creating a new application and enabling http-match-all-req within the app-filter.

Success and failure notifications include a specific HTTP encoded URI automatically interpreted as a success or a failure by Application Assurance on a per subscriber basis. If a failure is detected, the system will automatically attempt to notify a new candidate flow, instead of waiting for the next notification interval.

```
configure
    application-assurance group 1:1 policy
       application "IBN Messaging Server" create
          app-group "Web"
       exit
    app-filter
       entry 100 create
          expression 1 http-host eq "^192.168.150.251$"
          http-match-all-req
          application "IBN Messaging Server"
          no shutdown
       exit
    exit
```

App-Profile and App-Service-Options

Similar to the previous example, HTTP notifications are enabled per subscriber using RADIUS or Gx by modifying the subscriber app-profile or using ASO override.

The following ASO configuration is used:

```
configure
    application-assurance group 1:1 policy
        app-service-option
            characteristic "in-browser-notification"
                value "enabled"
                value "disabled"
                default-value "disabled"
        exit
```

The ASO characteristic in-browser-notification value *enabled* is used to enable the app-qos-policy matching the http-notification policy in-browser-notification as shown below:

```
configure
    application-assurance group 1:1 policy app-qos-policy
        entry 1300 create
            match
                characteristic "in-browser-notification" eq "enabled"
                application neq "Advertising Statistics"
            exit
            action
                http-notification "in-browser-notification"
            exit
            no shutdown
```

RADIUS Policy

The following RADIUS CoA message is used to modify the ASO characteristic of a residential subscriber and assign a specific Http-Url-Param VSA. The *in-browser-notification* ASO characteristic with value *enabled* is dynamically assigned to the subscriber along with the **Http-Url-Param** *&message=quota80*:

```
    NAS-Port-Id = "1/1/5:4088"
    Framed-IP-Address = 192.168.211.30
    Alc-AA-App-Service-Options = "in-browser-notification=enabled"
    Alc-AA-Sub-Http-Url-Param = "&message=quota80"
    Alc-App-Prof-Str = "1-1/Default"
```

The subscriber HTTP request to the messaging server has the following format and includes the Http-Url-Param value as an argument of the URL:

```
http://192.168.150.251/In-Browser-Notification/script/notification-select.php?SubId=sub1&var=
&message=quota80
```

The web server can now use the parameter value to make a decision to return a suitable notification message related to the subscriber usage cap.

Show Commands

Both the *in-browser-notification* ASO characteristic with value *enabled* and the HTTP-Url-Param VSA can be shown as follows:

```
A:PE# show application-assurance group 1:1 aa-sub esm "sub1" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub1 (esm)
ISA assigned            : 1/2
App-Profile             : 1-1/Default
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : &message=quota80
Last HTTP Notified Time : N/A
-------------------------------------------------------------------------------
Traffic                        Octets              Packets                Flows
-------------------------------------------------------------------------------
... ...
-------------------------------------------------------------------------------
Application Service Options (ASO)
-------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
-------------------------------------------------------------------------------
in-browser-notification          enabled                          dyn-override
quota-message-notification       disabled                         default
===============================================================================
```

The operator can also display the HTTP URL parameters VSA currently in use, per AA ISA group:

```
A:PE## tools dump application-assurance group 1 http-url-param-list
-------------------------------------------------------------------------------
Application-Assurance Subscriber HTTP URL parameters for Group 1:
-------------------------------------------------------------------------------
==============================================
Http Url Parameter                  Sub Usage
----------------------------------------------
"&message=quota80"                  1
==============================================
Total entries displayed 1
```

Conclusion

This chapter, intended for Application Assurance (AA) network architects and engineers, provides two implementation options for configuring and deploying HTTP In Browser Notification. It also explains how to take advantage of the Http-Url-Param RADIUS VSA to flexibly define various messaging campaigns using a common AA notification policy.

Application Assurance — Local URL List Filtering

This chapter provides information about the Application Assurance local URL list filtering.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This chapter is applicable to all SR OS systems supporting Application Assurance and was initially based on SR OS Release 13.0.R2. The chapter has been updated with CLI and functionality changes introduced in SR OS Release 19.10.R1, but the CLI in the sections URL-list update and Cron job is based on SR OS Release 16.0.

There are no specific prerequisites for this feature.

Overview

The local URL-list filtering capability provided by Application Assurance offers the following:

- configuration of a list of URLs which are globally allowed (allow-list)
- configuration of a list of restricted URLs subscribers should be prevented from accessing (deny-list)

The lists are stored locally in the SR OS system compact flash (CF).

This chapter provides a guide for configuring deny-lists.

Deny-lists assist service providers to comply with regulatory requirements for network-wide URL filtering policies, such as:

- Court-ordered URL takedown
- Child protection
- Government-mandated URL takedown list

The SR OS uses the Application Assurance capabilities to extract the URL from a subscriber HTTP/HTTPS request and compare it to the list of URLs contained in the local file. If a match occurs, the subscriber request is redirected to a preconfigured web server landing page, typically describing why the access to this resource was denied.

The system supports both unencrypted and OpenSSL Triple Data Encryption Standard (3DES) encrypted file formats to protect the content of the list.

HTTP/HTTPS filtering

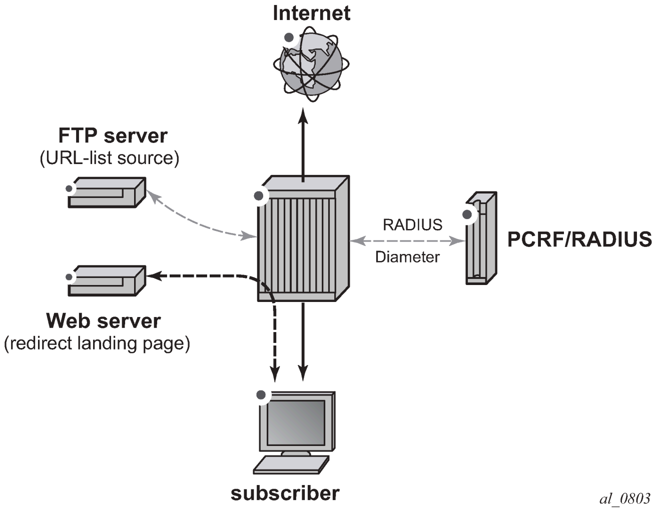
Each HTTP request within a TCP flow is analyzed and filtered. For HTTPS traffic, the system extracts the domain name information contained in the Transport Layer Security (TLS) Server Name Indication (SNI).

Setup details

The setup consists of the following elements, as shown in Figure 29: Local URL-list filtering setup:

- SR-series + ISA-AA
- Web server (redirect landing page)
- FTP server (source for the URL-list file)
- Subscriber (desktop/laptop/tablet/smart phone)
- Internet access
- Optional: AAA for subscriber authentication and policy modification

Figure 29: Local URL-list filtering setup



This chapter is written in the context of a residential or WiFi deployment. However, local URL-list filtering is also applicable to business VPN services.

Configuration

To configure the system for local URL-list filtering, the operator needs to:

- Create a URL-list policy referencing a valid URL-list file located on the system compact flash
- Create a URL-filter policy for local filtering by referencing the URL-list policy previously created

- Create an App-QoS-Policy (AQP) to apply this url-filter policy
- Optionally configure a cron job to automatically fetch a new list file and upgrade the URL list.

URL-list policy and URL-filter policy

In the following example, two dedicated URL-list and URL-filter policies show URL filtering based on a plain text file and an encrypted file:

```
configure
    application-assurance group 1
        url-list "denylist1-encrypted" create
            description "Demo URL Filtering List - Encrypted File"
            decrypt-key "ON3HU2GFPHmpOHwWbSGw/zdM4iuxzySpqS7pw/u3qIcuG4mABmrhc."
                                                                           hash2
            file "cf3:\aa-url-list\url-list1.encrypted"
            no shutdown
        exit
        url-filter "local-filter-list1-encrypted" create
            local-filtering
                deny-list "denylist1-encrypted"
                default-action allow
                http-redirect "redirect-denylist"
            exit
            no shutdown
        exit

configure
    application-assurance group 1
        url-list "denylist1-plaintext" create
            description "Demo URL Filtering List - Plaintext File"
            file "cf3:\aa-url-list\url-list1-plaintext.txt"
            no shutdown
        exit
        url-filter "local-filter-list1-txt" create
            local-filtering
                deny-list "denylist1-plaintext"
                default-action allow
                http-redirect "redirect-denylist"
            exit
            no shutdown
        exit
```

In the preceding example, both URL-filter policies are defined using **default-action allow**. The default action is used in case the file could not be loaded by the system, either at boot time or the first time the URL-list file was configured in the system. Possible causes are, for example:

- File corrupted, compact flash corrupted
- Incorrect file encryption format or password
- Wrong URL format in the file
- Too many URLs in the file

Operators should always use **default-action allow** when configuring the URL-filter policy associated with a URL-list file because the file or the CF may be corrupted, in which case the system logs an error and a trap is raised.

Note that if a valid URL-list file was previously in use, and an invalid file is uploaded and the URL-list policy upgraded using this file, then the system will continue using the previous list.

HTTP-redirect policy

Both URL-filter policies defined in the preceding example refer to the following http-redirect policy; subscribers accessing a URL from the URL-list file are redirected to the following landing page:

```
configure
    application-assurance group 1
        http-redirect "redirect-denylist" create
            description "Redirect for Local List URL Filtering"
            template 5
            tcp-client-reset
            redirect-url "http://172.16.70.100/Redirect/redirect-denylist.html?
                                                 Request edURL=$URL"
            no shutdown
        exit
```

URL-list file

File format

A URL-list file may contain either hostnames or URLs.

To create a URL-list containing hostnames, set **expression-match** in the url-list configuration:

```
config>app-assure>group# url-list url-list-name [create]
                description <description-string>
                no description
                decrypt-key key | hash-key | hash2-key [hash | hash2]
                no decrypt-key
                file file-url
                no file
                size url-list-size
                expression-match
                [no] shutdown
```

A URL-list with hostnames only (using **expression-match**) may contain the following wildcards:

- Head anchors character set [^ *]
- Tail anchors character set [$ *]
- Mid expression character set [\d \l \. **]
- Hex escaped characters [\x00 - \xFF]

Note that when **expression-match** is enabled, the list should contain hostnames only (with optional wildcards).

When configuring a URL-list with **expression-match** disabled (default), the system supports the following format for the URLs contained in the URL-list file:

- URLs without the HTTP keyword. For example:

```
        www.domain.com/path
```

- URLs with the HTTP keyword. For example:

```
http://www.domain.com/path
```

In all cases, the following is supported:

- Comment lines starting with the number sign character (#). For example:

```
# This is a comment line
```

- Printable ASCII characters. URLs using non-printable ASCII characters are percent-encoded by the web browser automatically and, therefore, need to be percent-encoded in the URL-list file.

File encryption

OpenSSL triple DES -nosalt is the supported encryption format. Files can be encrypted offline on a server using the following command:

```
openssl des3 -nosalt -in <input.txt> -out <output.enc>
```

List upgrade

The URL-list file can be upgraded using the **admin** command:

```
A:BNG# admin application-assurance group 1 url-list "denylist1-plaintext" upgrade
```

The upgrade result is logged in the system log-id 99:

```
A:BNG# show log log-id 99
===============================================================================
Event Log 99
===============================================================================
Description : Default System Log
Memory Log contents  [size=500   next event=72  (not wrapped)]
71 2015/07/07 13:09:25.01 EST MINOR: APPLICATION_ASSURANCE #4446 Base url-list success
"URL list "denylist1-plaintext" in ISA-AA group 1 has been updated. There are 3 entries in the
 URL list."
```

App-profiles and app-service-options

Application Assurance policies can be selectively applied to specific AA subscribers by modifying the app-profile assigned to the subscriber or using Application Service Option (ASO) override. See *7450 ESS, 7750 SR, and VSR Multiservice ISA and ESA Guide* for more information about modifying the app-profile or ASO assigned to AA subscribers (RADIUS, Gx, Override).

In this example, the following ASO configuration is used:

```
configure
    application-assurance group 1:1 policy
        app-service-options
            characteristic "local-list-filtering" create
                value "no"
                value "yes-encrypted"
                value "yes-plaintext"
                default-value "no"
            exit
        exit
        app-profile "1-1/Default" create
            divert
        exit
```

The ASO characteristic local-list-filtering value of yes-encrypted and yes-plaintext enable the AQP entry 210 and 220 in the example:

```
configure
    application-assurance group 1:1 policy app-qos-policy
        entry 210 create
            match
                characteristic "local-list-filtering" eq "yes-encrypted"
            exit
            action
                url-filter "local-filter-list1-encrypted"
            exit
            no shutdown
        exit
        entry 220 create
            match
                characteristic "local-list-filtering" eq "yes-plaintext"
            exit
            action
                url-filter "local-filter-list1-txt"
            exit
            no shutdown
        exit
```

If the url-filter policy needs to be applied to 100% of the subscribers in the network, it is also possible to remove the ASO match criteria.

URL-list update

The system supports a flexible mechanism to upgrade the URL list automatically, using either cron or the NSP, to comply with the regulatory requirements for list upgrade frequency.

To configure a crontab job to periodically upgrade the URL list, the operator needs to:

- Generate the file to be periodically executed and store it to compact flash. As an example, create a file with filename "fetch.txt" with the following content:

  ```
  file
  copy ftp://user:pwd@192.168.1.99/home/cmg/test/list.txt . force
  exit
  admin application-assurance group 1 url-list "test" upgrade
  ```

  The preceding commands will fetch a file from an ftp server and store it in compact flash. Assuming the operator has configured a local url-list "test" containing the file "list.txt", the url list will be upgraded.

- Configure the script policy:

  ```
  >config>system>script-control# info
  ----------------------------------------------
              script "bring"
                  location "cf3:/fetch.txt"
  ```

```
                no shutdown
            exit
            script-policy "test"
                results "cf3:/results.txt"
                script "bring"
                no shutdown
            exit
```

- Configure the crontab job:

```
>config>system>cron# info
----------------------------------------------
        schedule "bring_list"
            interval 60
            script-policy "test"
            no shutdown
        exit
----------------------------------------------
```

The preceding configuration will execute the commands stored in the file "fetch.txt" every 60 seconds. A value of 60 seconds was chosen for the test. In a real deployment, a list would be typically updated every 12 – 24 hours. A log file (results "results.txt") will also be created.

The end result will be that the system will automatically fetch from an ftp server a new list file, store it to compact flash and upgrade the URL list.

Show commands

url-list

The status of the URL list can be shown in the CLI. The url-list show command provides basic admin and operational status, as well as the number of URLs in the list. The command also provides reasons for any possible issue related to loading the list, as well as the last successfully deployed file and the last upgrade attempt. Therefore, the operator can determine whether the latest version of the file is currently in use or if an error occurred when trying to upgrade the list.

Show command output:

```
Label                    Description
Admin                    Status [Up | Down] - Administrative status of the url-list
Oper Status              [Up | Down] - Operational status of the url-list
Oper Flags               [admin-down | file-does-not-exist |invalid-file-format |
                         too-many-urls| switch-over-error]
File Deployed to ISA      [Yes | No] - This flag describes if the file located in
                          the compact flash is the one deployed in the ISA, in the
                          event the file is overwritten and before the admin upgrade
                          command is used this flag will display “No”.
Upgrade Statistics
Last Success              Last time the list was successfully upgraded
File Name                 File name for the last successful upgrade
URL Entries               Number of URLs loaded at the last success
Blank/Comment Lines       Number of blank or commented out lines
Last Attempt              Last time the operator tried to upgrade the list
Result                    Success | Failure. Result of the last upgrade
```

```
File Name                 File name for the last upgrade attempt
```

```
*A:Dut-C# show application-assurance group 1 url-list "Deny List1"
======================================================================
Application Assurance Group 1 url-list "Deny List1"
======================================================================
Description : (Not Specified)
Size : standard
Host Expressions : disabled
Admin Status : Up
Oper Status : Up
Oper Flags : <none>
File deployed to ISAs : Yes
-------------------------------------------------------------------------------
Upgrade Statistics
-------------------------------------------------------------------------------
Last Success : 11/02/2020 15:07:40
Deployed
File Name : cf1:/host.txt
URL Entries : 1 ( 0.01% full)
URL Characters : 6 (~0.00% full)
URL Host Expr Entries: 1 ( 0.01% full)
Blank/Comment Lines : 1
Last Attempt : 11/02/2020 15:07:40
Result : Success
File Name : cf1:/host.txt
======================================================================
```

url-filter

The **url-filter** show command provides its operational and admin status, as well as actions taken, such as the number of redirects. With URL list filtering, using a **default-action** set to **allow**, the only counters increasing are **allow**, **redirect**, and **default**.

```
*A:Dut-C>config>app-assure>group>url-filter# show application-assurance group 1 url-filter "Url
 Filter1"
===============================================================================
Application Assurance Group 1 URL Filter "Url Filter1"
===============================================================================
Description            : (Not Specified)
Admin Status           : Up
Oper Status            : Up
Oper Flags             : <none>
HTTP Request Filtering : all
AQP Referenced         : No
-------------------------------------------------------------------------------
URL Stats Summary
-------------------------------------------------------------------------------
Total Requests  : 0                    Default Action  : 0
Requests Allowed: 0                    Reqs Block/Redir: 0
-------------------------------------------------------------------------------
Local Filter
-------------------------------------------------------------------------------
deny-list              : Deny List1
  Admin Status         : Up
  Oper Status          : Up
  Oper Flags           : <none>
  Number of URLs       : 1
  Default Action       : block-all
  HTTP Redirect        : (Not Specified)
  URL-List Lookups     : 0
    Match              : 0
    Miss               : 0
    Default Action     : 0
===============================================================================
```

http-redirect

The **http-redirect** show command provides more information about how the traffic was blocked; for example, it differentiates TCP client reset used for HTTPS from regular redirect used for HTTP traffic.

```
A:BNG# show application-assurance group 1 http-redirect "redirect-denylist"
===============================================================================
Application Assurance Group 1 HTTP Redirect redirect-denylist
===============================================================================
Description        : Redirect for Local List URL Filtering
Template           : 5
                   : Redirect supporting macro substitution using HTTP 302
Redirect URL       : http://172.16.70.100/Redirect/redirect-denylist.
                     html?RequestedURL=$URL
Admin Status       : Up
AQP Ref            : No
-------------------------------------------------------------------------------
Summary Statistics
-------------------------------------------------------------------------------
Grp:Part             Redirects           Client Resets               Redirects
                          Sent                    Sent                Not Sent
-------------------------------------------------------------------------------
1:1                          2                       1                        0
-------------------------------------------------------------------------------
Total                        2                       1                        0
-------------------------------------------------------------------------------
===============================================================================
```

Cron job

This chapter provides details on how to verify that the cron job executes successfully. The operator can see the location and execution interval of the file which is executed periodically, the time it was last executed and any possible errors which may have occurred during execution using the following command:

```
 # show system cron schedule "bring_list"
===============================================================================
CRON Schedule Information
===============================================================================
Schedule                     : bring_list
Schedule owner               : TiMOS CLI
Description                  : none
Administrative status        : enabled
Operational status           : enabled
Script Policy                : test
Script Policy Owner          : TiMOS CLI
Script                       : bring
Script Owner                 : TiMOS CLI
Script source location       : cf3:/fetch.txt
Script results location      : cf3:/results.txt
```

```
Schedule type                : periodic
Interval                     : 0d 00:01:00 (60 seconds)
Repeat count                 : infinite
Next scheduled run           : 0d 00:00:51
End time                     : none
Weekday                      : none
Month                        : none
Day of month                 : none
Hour                         : none
Minute                       : none
Number of schedule runs      : 12
Last schedule run            : 2019/09/13 11:28:25 EEST
Number of schedule failures  : 0
Last schedule failure        : no error
Last failure time            : never

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

The following command provides more information on the execution cycle:

```
*A:4LS_CloudMG# show system script-control script-policy "test"

===============================================================================
Script-policy Information
===============================================================================
Script-policy                : test
Script-policy Owner          : TiMOS CLI
Administrative status        : enabled
Operational status           : enabled
Script                       : bring
Script owner                 : TiMOS CLI
Script source location       : cf3:/fetch.txt
Script results location      : cf3:/results.txt
Max running allowed          : 1
Max completed run histories  : 1
Max lifetime allowed         : 0d 01:00:00 (3600 seconds)
Completed run histories      : 1
Executing run histories      : 0
Initializing run histories   : 0
Max time run history saved   : 0d 01:00:00 (3600 seconds)
Script start error           : N/A
Last change                  : 2019/09/13 10:31:20 EEST
Max row expire time         : never
Last application             : cron
Last auth. user account      : not-specified

===============================================================================
Script Run History Status Information
-------------------------------------------------------------------------------
Script Run #25
-------------------------------------------------------------------------------
Start time    : 2019/09/13 11:29:25 EEST
End time      : 2019/09/13 11:29:26 EEST
Elapsed time  : 0d 00:00:01             Lifetime      : 0d 00:00:00
State         : terminated              Run exit code : noError
Result time   : 2019/09/13 11:29:26 EEST
Keep history  : 0d 00:59:41
Error time    : never
Results file  : cf3:/results.txt_20190913-082924-UTC.646420.out
Run exit      : Success
Error         : N/A
Application   : cron                    Auth. user ac*: not-specified
* indicates that the corresponding row element may have been truncated.
```

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

Every time a crontab job is executed, a log file is generated. The filename will be "results.txt_<timestamp>.out", where:

- results.txt: configured in the script-policy
- <timestamp>:file timestamp. As an example: 20190913-082324-UTC.646436.

An example log file will be: results.txt_20190913-082324-UTC.646436.out and the log file contents is as follows:

```
Pre-processing configuration file (V0v0)...
Completed processing 4 lines in 0.0 seconds
*A:4LS_CloudMG# file
*A:4LS_CloudMG# copy ftp://user:pwd@192.168.1.99/home/cmg/test/list.txt . force
Copying file ftp://user:pwd@192.168.1.99/home/cmg/test/list.txt ... OK
1 file copied.
*A:4LS_CloudMG# exit
*A:4LS_CloudMG# admin application-assurance group 1 url-list "test" upgrade
Executed 4 lines in 0.7 seconds from file cf3:\fetch.txt
```

Finally, using the following command, the operator may check when the list was last upgraded and verify that the cron job runs as intended:

```
# show application-assurance group 1 url-list "test"

===============================================================================
Application Assurance Group 1 url-list "test"
===============================================================================
Description              : (Not Specified)
Size                     : standard
Admin Status             : Up
Oper Status              : Up
Oper Flags               : <none>
File deployed to ISAs    : Yes

-------------------------------------------------------------------------------
Upgrade Statistics
-------------------------------------------------------------------------------
Last Success             : 09/13/2019 11:49:57
  Deployed
    File Name            : cf3:\list.txt
    URL Entries          : 8 ( 0.05% full)
    URL Characters       : 103 (~0.00% full)
    Blank/Comment Lines  : 0
Last Attempt             : 09/13/2019 11:49:57
  Result                 : Success
    File Name            : cf3:\list.txt
===============================================================================
```

Conclusion

This chapter, intended for Application Assurance (AA) network architects and engineers, provides two examples for deploying URL-list filtering, upgrading the list and displaying its statistics, as well as configuring a cron job so that the system will periodically fetch a new URL list file and upgrade the list automatically.

Application Assurance — Security Gateway Stateful Firewall

This chapter provides information about Application Assurance Security gateway stateful firewall.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This chapter is applicable to all 7750 SR/SR-c and 7450 ESS chassis supporting Application Assurance (AA).

The configuration was tested on Release 13.0.R2.

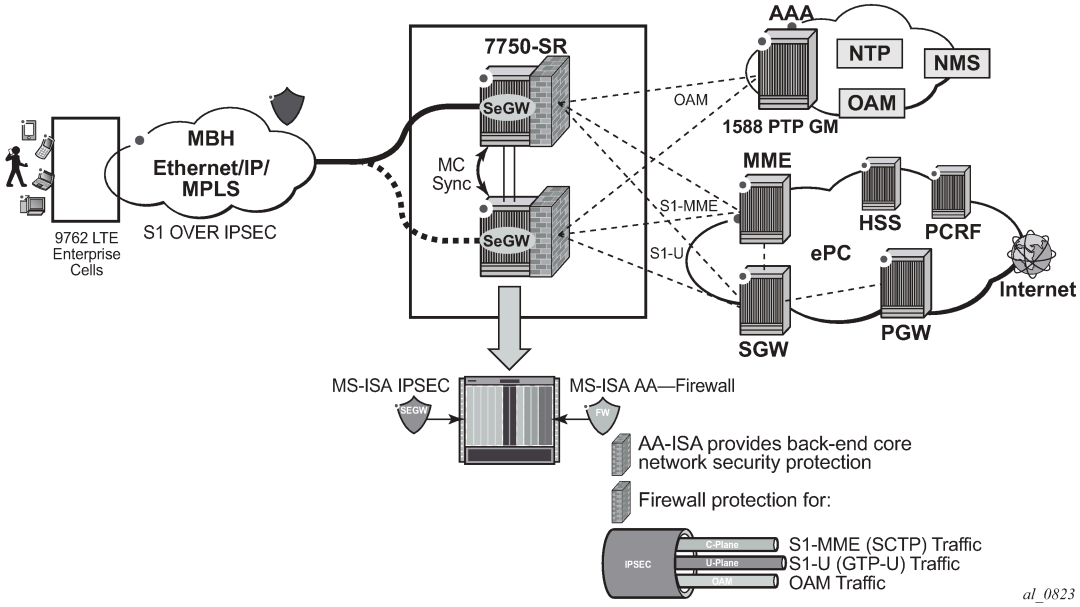
Overview

The SR OS 13.0.R1 AA stateful firewall feature runs on AA-ISA and extends application-level analysis to provide an in-line stateful service, integrated within the Security Gateway (SeGW). The feature provides protection for mobile infrastructure; Mobility Management Entities (MMEs), Serving Gateways (SGWs), and Network Management Systems (NMSs), against attacks from compromised base stations, evolved NodeBs (eNBs), or Femto Access Points (FAPs). AA stateful packet filtering, combined with AA L7 classification and control, provides advanced, next-generation firewall functionality. Using stateful packet filtering, the AA FW not only inspects packets at layers 3 to 7, but also monitors the connection state.

AA FW deployed within a SeGW in ultra-broadband access networks (3G/4G/Femto) provides back-end core network security protection, as per Figure 1. AA FW offers protection for the following 3rd Generation Partnership Project (3GPP) defined interfaces:

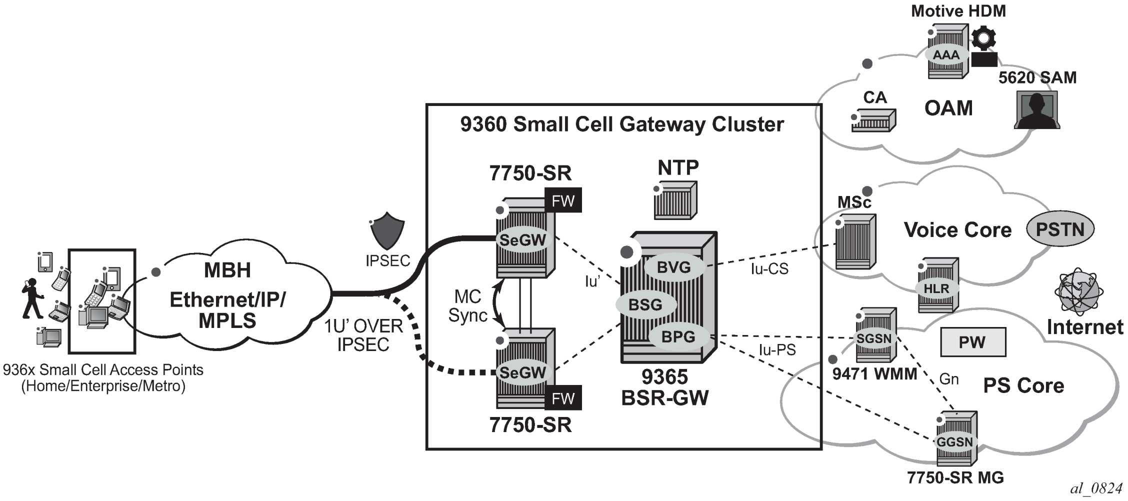
1. S1-MME
2. S1-U
3. Operations, Administration and Maintenance (OAM)

Figure 30: LTE SeGW Firewall Deployment



Similarly, the SeGW architecture for Femto deployment is based on two 7750 SR systems terminating the mobile backhaul side (front-end and connecting to, for example, a 9365 Base Station Router Gateway (BSR-GW) and other network elements of the packet core (back-end), as per Figure 31: SeGW in Small Cells Architecture:

Figure 31: SeGW in Small Cells Architecture



The two SeGWs run in stateful redundant mode: upon partial or total failure of the active SeGW for a set of IPSec tunnels, the other SeGW takes over without terminating the IPSec tunnels, providing hitless failover.

In addition to MS-ISA hardware dedicated to the IPSec function, each SeGW supports one or more additional MS-ISAs running AA to provide firewall capabilities. The firewall rules protect the BSR as well as the BSR-GW and packet core network elements (NEs) from malicious attacks or unauthorized traffic.

The objective of this chapter is to describe the required configuration within AA-ISA in order to enable AA FW and protection for S1-MME, S1-U, and OAM traffic. Basic knowledge of AA-ISA diversion configuration is assumed.

S1-MME Traffic Protection

The purpose of AA FW in this deployment is to protect the MME infrastructure against an attack from a compromised eNB or FAP. Network flooding attacks, malformed packets, and port scans are examples of denial of service (DoS) attacks that can be carried out using a compromised eNB or FAP.

AA FW provides inspection of the Stream Control Transmission Protocol (SCTP) used to communicate to the MME. Such inspection includes checking for SCTP payload protocol IDs (PPIDs), source /destination ports, SCTP chunk validation, and malformed SCTP packets (such as checksum validation). In addition, the operator can configure DoS flooding rules, such as policers to limit the bandwidth and/or flow counts of SCTP traffic.

S1-U Traffic Protection

The purpose of AA FW in this deployment is to protect the SGW infrastructure against an attack from a compromised eNB or FAP. AA FW supports protection against:

- malformed GPRS Tunneling Protocol User plane (GTP-U) packet attacks

  Checking packet sanity, which include GTP-U mandatory, optional, and extension header checks, as well as checks for invalid reserved information elements (IE) and missing IEs.
- unsupported GTP messages

  Filtering messages based on message type and/or message length.
- flooding attacks

  Shaping GTP traffic bandwidth, which limits the GTP-U bandwidth that a FAP can send to the core (SGW).

  Limiting GTP tunnels, which limits the number of concurrent GTP tunnels and/or setup rate of these tunnels from a FAP to the core network.

  To prevent the shared resources of bandwidth and the SGW processor from being consumed by an attacker, Nokia recommends the GTP flow rate limiting configuration.
- IP fragmentation-based attacks

  Applying various drop rules for IP fragmentation of GTP messages.

OAM Traffic Protection

The purpose of AA FW protection in this deployment is to protect against any abuse of OAM network resources, such as NMS.

Network flooding attacks, malformed packets, and port scans are examples of such attacks that can be carried out using a compromised eNB or FAP.

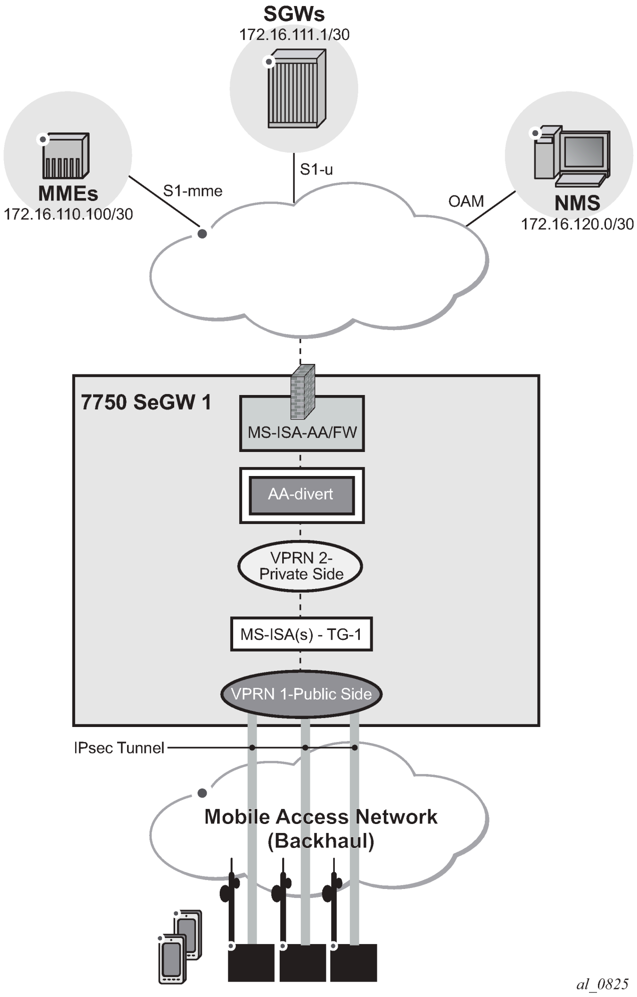
See the configuration described in the Application Assurance — Stateful Firewall chapter for this context of OAM protection in SeGW.

Configuration

AA-ISA Application QoS Policies (AQPs) are enhanced in Release 13.0.R1 with several new AQP actions that provide SCTP and GTP filtering functionality. As with all AQPs, these actions have partition-level scope, which allows different FW policies to be implemented by using AA partition concepts within the same AA-ISA.

The configuration topology in Figure 32: Configuration Topology shows the SeGW FW functionality of S1-U and S1-MME interfaces. Geo-redundancy, which is a very common deployment option, is not described in this here because it is described in the Multi-Chassis IPSec Redundancy chapter.

Figure 32: Configuration Topology



Pre-Setup Requirements

Configure tunnel ISAs with optional multi-chassis redundancy. See the Multi-Chassis IPSec Redundancy chapter for more information.

1. Divert AA traffic and apply basic firewall rules.

   Step 1.1. Divert private VPRN traffic into AA-ISA with AA multi-chassis redundancy.

   This step is required for any of the configurations in Steps 2, 3 or 4.

   There is no dependency between Steps 2, 3 or 4.

   In this example, one private VPRN is used for all traffic to/from eNBs. In some small cell deployments, eNB traffic is split into three different VPRNs: one for control (S1-MME), one for management (OAM), and one for bearer traffic (S1-U GTP-U). In that case, each of these VPRNs needs to be diverted into AA-ISA in order to provide firewall protection.

   First, define an application profile and transit IP policy, such as:

```
*A:7750-1>config>app-assure>group$ info
----------------------------------------------
            policy
                begin
                app-profile "default" create
                    description " App profile that applies to the whole SAP"
                    divert
                exit
                commit
            exit
            transit-ip-policy 1 create
                description "Per eNB-IP Sub policy"
                detect-seen-ip
                transit-auto-create
                    no shutdown
                exit
```

   Then, apply these policies to the SAP on the private side of the IPSec tunnel ISA:

```
*A:7750-1>config>service>vprn>if# info
----------------------------------------------
                sap tunnel-1.private:1 create
                    transit-policy ip 1
                    app-profile "default"
                exit
```

   This configuration achieves:

   - Traffic to/from the IPSec tunnel ISA private SAP is diverted to AA-ISA for the purpose of FW protection
   - Within AA-ISA, the diverted SAP is treated as a parent SAP. That is, instead of treating the whole SAP as a single subscriber, subscribers are auto-created within this SAP based on the IP address of the eNBs

   Step 1.2. Protect against malformed packets.

   In firewall deployments, it is recommended that overload-drop, error-drop, and fragment-drop (all) are enabled within the default sub-policy, as shown in the example below:

   - **overload-drop** ensures that AA-ISA, when overloaded, drops the excess traffic instead of allowing it through, without applying firewall rules.
   - **error-drop** ensures that AA-ISA drops malformed IP packets.

- **fragment-drop (all)** because many network DoS attacks use IP fragmentation to initiate attacks, the operator has the option to drop all fragmented traffic, drop out-of-order fragments only, or allow fragments through. Allowing fragments through is not recommended for firewall deployments.

```
*A:7750-1>config>app-assure>group>policy# app-qos-policy
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
                    entry 500 create
                        description "apply SeGW session filter rules"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            overload-drop
                            error-drop
                            fragment-drop all
                        exit
                        no shutdown
                    exit
                exit
----------------------------------------------
*A:7750-1>config>app-assure>group>policy#
```

Step 1.3. Limit total traffic from any eNB.

It is recommended that a total limit be placed on how much bandwidth and how many flows an eNB or FAP can generate toward the network, regardless of the type of traffic.

The limit values are a function of the number of end devices that are served by the eNB or FAP, plus some additional margin:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            policer "limit_eNBs_total_Flows" type flow-count-limit granularity subscriber
                create
                flow-count 1000
            exit
            policer "limit_eNBs_total_bw" type single-bucket-bandwidth granularity sub
                scriber create
                rate 500
                mbs 500
            exit
----------------------------------------------
*A:7750-1>config>app-assure>group#
```

Note:
If the traffic from eNB or FAP is separated into different private SAPs, based on traffic type (S1-AP, S1-U, or OAM), as with some deployment topologies, then the policing limit value is dependent on the SAP traffic type as well as the number of end devices. See policing limit settings in Steps 2 and 3.

Apply the configured policers as actions from within the default sub-policy AQP entry:

```
*A:7750-1>config>app-assure>group>policy# app-qos-policy entry 500
*A:7750-1>config>app-assure>group>policy>aqp>entry>action# flow-count-limit "limit_eNBs_
total_Flows"
*A:7750-1>config>app-assure>group>policy>aqp>entry>action# bandwidth-policer "limit_eNBs_
total_bw"
*A:7750-1>config>app-assure>group>policy>aqp>entry# info
```

```
----------------------------------------------
                        description "apply SeGW session filter rules"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            bandwidth-policer "limit_eNBs_total_bw"
                            flow-count-limit "limit_eNBs_total_Flows"
                            session-filter "SeGW_FW"
                            overload-drop
                            error-drop
                            fragment-drop all
                        exit
                        no shutdown
----------------------------------------------
*A:7750-1>config>app-assure>group>policy>aqp>entry#
```

Note:
All of the above listed actions use the traffic direction of subscriber-to-network. That is, they are not applied to traffic in the other direction (downstream) because the purpose of the firewall is to protect the network resources from upstream traffic coming from compromised eNBs or FAPs.

2. Configure AA-ISA to provide firewall protection to protect MMEs (S1-AP traffic).

Step 2.1. Create IP AA lists.

First, create an AA IP prefix list that contains eNB IP addresses or range of addresses:

```
*A:7750-1>config>app-assure# group 1:1
*A:7750-1>config>app-assure>group# ip-prefix-list "ALL_eNBs" create
*A:7750-1>config>app-assure>group>pfx>$ description "eNodeB subnet"
*A:7750-1>config>app-assure>group>pfx>$ prefix 172.16.100.0/24
*A:7750-1>config>app-assure>group>pfx>$ exit
```

Next, optionally create an AA IP list that contains MME IP addresses (in case there are more than one):

```
*A:7750-1>config>app-assure>group# ip-prefix-list "MMEs" create
*A:7750-1>config>app-assure>group>pfx>$ description "MME(s) subnet"
*A:7750-1>config>app-assure>group>pfx>$ prefix 172.16.110.100/30
*A:7750-1>config>app-assure>group>pfx>$ exit
```

After the above lists are created, they can be referenced and used in AA FW rules using session filters and AQPs.

Step 2.2. Allow only SCTP traffic towards MMEs — No port scanning.

A basic setup creates session-filter rules that will only allow SCTP traffic between eNBs and MMEs.

```
*A:7750-1>config>app-assure>group# session-filter "SeGW_FW" create
*A:7750-1>config>app-assure>group>sess-fltr$ default-action deny
*A:7750-1>config>app-assure>group>sess-fltr$ entry 1 create
*A:7750-1>config>app-assure>group>sess-fltr>entry$ description "allow SCTP to MM Es"
*A:7750-1>config>app-assure>group>sess-fltr>entry$ match
*A:7750-1>config>app-assure>group>sess-fltr>entry>match$ ip-protocol-num "sctp"
*A:7750-1>config>app-assure>group>sess-fltr>entry>match$ src-ip ip-prefix-list "ALL_eNBs"
*A:7750-1>config>app-assure>group>sess-fltr>entry>match$ dst-ip ip-prefix-list "MMEs"
*A:7750-1>config>app-assure>group>sess-fltr>entry>match$ dst-port eq 6005
*A:7750-1>config>app-assure>group>sess-fltr>entry>match$ exit
*A:7750-1>config>app-assure>group>sess-fltr>entry$ action permit
```

```
*A:7750-1>config>app-assure>group>sess-fltr>entry$ exit
```

Note:
In the above configuration, SCTP traffic on MMEs is assumed to be running on port 6005.

Next, the newly created session filter needs to be referenced from a default sub-policy AQP action, as follows:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
                    entry 500 create
                        description "apply SeGW session filter rules"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            session-filter "SeGW_FW"
                            overload-drop
                            error-drop
                            fragment-drop all
                        exit
                        no shutdown
                    exit
                exit
----------------------------------------------
*A:7750-1>config>app-assure>group>policy#
```

Using traffic direction **subscriber-to-network** in the above AQP entry achieves two objectives:

a. Protects MMEs by allowing only SCTP traffic to be initiated from eNB subnets toward MMEs. Port scanning toward MME is blocked.

b. Allows MMEs to have full access to eNBs.

Note:
It is important that an AQP, containing a session-filter action, does not contain any matching condition other than ASOs, traffic direction, or subscriber ID. Subscriber ID is not applicable in this deployment use-case.

Step 2.3. DoS protection: Limit the number of SCTP flows from eNBs.

In this step, the operator configures a flow count policer to limit the number of SCTP flows that an eNB can generate toward the MMEs. This protects the MMEs against a compromised eNB trying to set up many SCTP flows.

```
*A:7750-1# configure application-assurance group 1
*A:7750-1>config>app-assure>group# policer "sctp_flow_count" type flow-count-limit
 granularity subscriber create
*A:7750-1>config>app-assure>group>policer$ flow-count 2
*A:7750-1>config>app-assure>group>policer$ exit
```

In the above configuration, an eNB or FAP can have up to two flows at a time. In practice, there should only be one SCTP session, one flow in each direction, per eNB-MME pair. The above example uses two flows to leave a margin in case a second, backup, MME needs to communicate with the eNB, while still providing enough protection.

Add the defined policer as a **flow-count-limit** as an AQP action, as follows:

```
A:7750-1>config>app-assure>group>policy>aqp# entry 100
```

```
A:7750-1>config>app-assure>group>policy>aqp>entry$ info
----------------------------------------------
                        description "limit SCTP traffic"
                        match
                            traffic-direction subscriber-to-network
                            ip-protocol-num eq sctp
                        exit
                        action
                            flow-count-limit "sctp_flow_count"
                        exit
                        no shutdown
----------------------------------------------
A:7750-1>config>app-assure>group>policy>aqp>entry$
```

Step 2.3.1. Configure an AA FW events log.

It is sometimes advisable to configure a log that captures events related to various AA FW actions. Due to the limited size of the log and the large amount of traffic AA can handle, consider the usefulness of the information in the log when:

- debugging a configuration
- testing a configuration in a staged environment
- capturing infrequent actions

To configure a log:

```
*A:7750-1# configure application-assurance group 1:1
*A:7750-1>config>app-assure>group# event-log "FW_drops_log" create
*A:7750-1>config>app-assure>group>evt-log$ buffer-type circular
*A:7750-1>config>app-assure>group>evt-log$ max-entries 100000
*A:7750-1>config>app-assure>group>evt-log$ no shutdown
*A:7750-1>config>app-assure>group>evt-log$ exit
*A:7750-1>config>app-assure>group# info
----------------------------------------------
                ---snip---
            event-log "FW_drops_log" create
                buffer-type circular
                max-entries 100000
                no shutdown
            exit
```

To reference the configured log from within the deny action of the session filter:

```
*A:7750-1>config>app-assure>group>sess-fltr# info
----------------------------------------------
                default-action deny event-log "FW_drops_log"
                entry 1 create
                    description "allow SCTP to MMEs"
                    match
                        ip-protocol-num sctp
                        src-ip ip-prefix-list "ALL_eNBs"
                        dst-ip ip-prefix-list "MMEs"
                    exit
                    action permit
                exit
----------------------------------------------
*A:7750-1>config>app-assure>group>sess-fltr#
```

To view the log:

```
*A:7750-1# tools dump application-assurance group 1:1 event-log "FW_drops_log" isa 1/2
=====================================================
Application-Assurance event-log "FW_drops_log"
Current Time:          "06/10/2015 22:45:30" (UTC)
  group[:partition]:   1:1
  isa:                 1/2
  admin state:         no shutdown
  buffer-type:         circular
  max-entries:         100000
=====================================================
Event-source
                     Action        SubType     SubName                            Direction
 Src-ip                                  Dst-ip                                  Ip-protocol
 Src-port Dst-port Timestamp

Total Records:   0
=====================================================
*A:7750-1#
```

To clear all the entries within the specified log:

```
*A:7750-1# clear application-assurance group 1:1 event-log "FW_drops_log"
```

Step 2.4. DoS protection: Limit the SCTP bandwidth from eNB

Similar to the previous step, the operator configures a flow bandwidth policer to limit the amount of SCTP traffic that an eNB can generate toward the MMEs. This protects the MMEs against a compromised eNB trying to flood the MMEs.

```
*A:7750-1# configure application-assurance group 1
*A:7750-1>config>app-assure>group# info
----------------------------------------------
---snip---
            policer "sctp_bw_limit" type single-bucket-bandwidth granularity subscriber
 create
                rate 30
                mbs 10
            exit
---snip---
            exit
----------------------------------------------
*A:7750-1>config>app-assure>group#
```

In the above example, a single leaky-bucket policer is configured with a rate set to 30 kb/s and maximum burst size of 10 kbytes. This provides enough bandwidth to ensure normal operations, while still providing a ceiling limit of how much traffic any eNB can send toward the MMEs.

The value for this policer is a function of the amount of user equipment (UEs) served by the eNB/FAP. For example, in a small cell deployment, with 32 active users per 9962 FAP, the S1-MME bandwidth is estimated to be:

Uplink — toward MME : 2.7 kb/s

Downlink — from MME toward FAP : 28 kb/s

Add the defined policer as a subscriber policy, as follows:

```
A:7750-1>config>app-assure>group>policy>aqp# entry 100
```

```
A:7750-1>config>app-assure>group>policy>aqp>entry$ info
----------------------------------------------
                        description "limit SCTP traffic"
                        match
                            traffic-direction subscriber-to-network
                            ip-protocol-num eq sctp
                        exit
                        action
                            bandwidth-policer "sctp_bw_limit"
                            flow-count-limit "sctp_flow_count"
                        exit
                        no shutdown
----------------------------------------------
A:7750-1>config>app-assure>group>policy>aqp>entry$
```

Step 2.4.1 Configure additional limits for all traffic to MMEs.

To further protect the MMEs from a distributed attack, whereby a number of eNBs or FAPs are compromised, an AA FW can be configured to limit total traffic, not just from a single eNB as outlined in previous sections, but from all eNBs toward the MMEs.

It is recommended to configure the following three protection limits:

a. total bandwidth of SCTP toward MMEs

b. total number of flows toward MMEs

c. flow setup rate toward the MMEs

The configuration is shown below:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            policer "limit_total_sctp_bw" type single-bucket-bandwidth granularity system
 create
                rate 1200
                mbs 100
            exit
            policer "limit_total_sctp_flows" type flow-count-limit granularity system create
                flow-count 400
            exit
            policer "limit_total_sctp_flows_rate" type flow-rate-limit granularity system
 create
                rate 100
                mbs 100
            exit
----------------------------------------------
*A:7750-1>config>app-assure>group#
```

Note:

- The policers are of type **system** and not **subscriber** in order to be applied to all eNBs or FAPs, as is the case when auto-transit subscribers are created (see Step 1).
- The actual limits of these policers are a function of the total number of eNBs served by the SeGW. In the above configuration, it is assumed that there are 400 eNBs. Therefore, the total limit is 400 flows of SCTP traffic.
- A flow setup rate limit of 100 is set to protect MMEs from a storm of new SCTP flows.

The policers are then referenced from within the appropriate AQP entry that matches the MMEs traffic and SCTP:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
 ---snip---
                    entry 110 create
                        description " limit system traffic towards MMEs"
                        match
                            traffic-direction subscriber-to-network
                            src-ip eq ip-prefix-list "ALL_eNBs"
                            dst-ip eq ip-prefix-list "MMEs"
                        exit
                        action
                            bandwidth-policer "limit_total_sctp_bw"
                            flow-rate-limit "limit_total_sctp_flows_rate"
                            flow-count-limit "limit_total_sctp_flows"
                        exit
                        no shutdown
                    exit
*A:7750-1>config>app-assure>group>policy>aqp#
```

Note:
It is possible, but redundant, to add the **ip-protocol eq sctp** command as a match condition, because the configured session filter already ensures that only SCTP traffic can flow between eNBs and MMEs.

Step 2.5. Allow only specified SCTP PPIDs toward the MMEs.

In this step, the operator blocks all except the specified SCTP messages that contain configured PPIDs, using an AA SCTP filter configuration:

```
*A:7750-1>config>app-assure>group# sctp-filter
  - no sctp-filter <sctp-filter-name>
  - sctp-filter <sctp-filter-name> [create]

 <sctp-filter-name>   : [32 chars max]
 <create>             : keyword

 [no] description     - Configure a description of the SCTP filter
 [no] event-log       - Configure an event log for packets dropped by the SCTP
                        filter
      ppid            + Configure actions for specific or default PPIDs
                        (Payload Protocol Identifiers)
 [no] ppid-range      - Configure the range of allowable PPIDs for the SCTP
                        filter
```

The filter specifies either a range of PPIDs or individual PPIDs:

```
*A:7750-1>config>app-assure>group>sctp-fltr>ppid$ entry 1
  - entry <entry-id> value <ppid-value> action {permit|deny}
  - no entry <entry-id>

 <entry-id>           : [1..255]
 <ppid-value>         : [0..4294967295]D | [256 chars max]
 <permit|deny>        : permit|deny
```

The PPIDs can be specified either by their values or by names. Names are specified in RFC 4960. See Table 9: SCTP PPIDs .

Table 9: SCTP PPIDs

Value	SCTP PPID	Value	SCTP PPID
0	Reserved by SCTP	31	Service Area Broadcast Protocol (SABP)
1	IUA	32	Fractal Generator Protocol (FGP)
2	M2UA	33	Ping Pong Protocol (PPP)
3	M3UA	34	CalcApp Protocol (CALCAPP)
4	SUA	35	Scripting Service Protocol (SSP)
5	M2PA	36	NetPerfMeter Protocol Control Channel (NPMP-CONTROL)
6	V5UA	37	NetPerfMeter Protocol Data Channel (NPMP-DATA)
7	H.248	38	Echo (ECHO)
8	BICC/Q.2150.3	39	Discard (DISCARD)
9	TALI	40	Daytime (DAYTIME)
10	DUA	41	Character Generator (CHARGEN)
11	ASAP	42	3GPP RNA
12	ENRP	43	3GPP M2AP
13	H.323	44	3GPP M3AP
14	Q.IPC/Q.2150.3	45	SSH over SCTP
15	SIMCO <draft-kiesel-midcom-simco-sctp-00.txt>	46	Diameter in a SCTP DATA chunk
16	DDP Segment Chunk	47	Diameter in a DTLS/SCTP DATA chunk
17	DDP Stream Session Control	48	R14P. BER Encoded ASN.1 over SCTP
18	S1 Application Protocol (S1AP)	49	Unassigned
19	RUA	50	WebRTC DCEP
20	HNBAP	51	WebRTC String
21	ForCES-HP	52	WebRTC Binary Partial (deprecated)
22	ForCES-MP	53	WebRTC Binary
23	ForCES-LP	54	WebRTC String Partial (deprecated)

Value	SCTP PPID	Value	SCTP PPID
24	SBc-AP	55	3GPP PUA
25	NBAP	56	WebRTC String Empty
26	Unassigned	57	WebRTC Binary Empty
27	X2AP	58-4294967295	Unassigned
28	IRCP - Inter Router Capability Protocol		
29	LCS-AP		
30	MPICH2		

It is recommended to limit the SCTP traffic to only those packets with S1 AP PPID. The SCTP filter can be configured to deny all by default and only allow PPID S1 AP (by value = 18 or by name: *s1-application-protocol*) as follows:

```
*A:7750-1# configure application-assurance group 1:1
            sctp-filter "SCTP-PPID-Filter" create
                description "Allow only S1AP PPID"
                event-log "FW_drops_log"
                ppid
                    default-action deny
                    entry 1 value "s1-application-protocol" action permit
                exit
```

This configured SCTP filter is then referenced as an action from within an AQP entry:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
                    entry 100 create
                        description "limit SCTP traffic"
                        match
                            traffic-direction subscriber-to-network
                            ip-protocol-num eq sctp
                        exit
                        action
                            bandwidth-policer "sctp_bw_limit"
                            flow-count-limit "sctp_flow_count"
                            sctp-filter "SCTP-PPID-Filter"
                        exit
                        no shutdown
                    exit
----------------------------------------------
A:7750-1>config>app-assure>group>policy>aqp#
```

To view the packets allowed or denied by the configured SCTP filter:

```
*A:7750-1# show application-assurance group 1:1 sctp-filter "SCTP-PPID-Filter"
===============================================================================
Application Assurance Group 1:1 SCTP Filter "SCTP-PPID-Filter"
===============================================================================
Description           : Allow only S1AP PPID
Maximum PPID          : 4294967295
```

```
Minimum PPID          : 0
Default action        : deny
Configured PPIDs      : 1

Packets arrived       : 0
Packets denied
  Malformed packet    : 0
  PPID out of range   : 0
  PPID denied         : 0
Packets permitted     : 0
===============================================================================
*A:7750-1#
```

Note:
The SCTP malformed packet counter shown above increments when an AA SCTP filter encounters an SCTP packet that is malformed, such as:

- IP packet is too small to contain a common SCTP header
- SCTP chunk LEN < 4 bytes: each SCTP chunk header is 4 bytes, so the SCTP chunk cannot be smaller than this
- remaining space in the IP packet is too small to contain a chunk header (for example, your packet has 2 chunks and the 2nd chunk length goes beyond the IP length advertised)
- IP packet is too small to contain the chunk

Currently, the SCTP filter statistics cannot be reset on the fly without shutting down the SCTP filter.

Another way to view the effect of the configured SCTP filter is to check the firewall log, if configured:

```
*A:7750-1# tools dump application-assurance group 1:1 event-log "FW_drops_log" isa 1/2
```

3. ConfigureAA-ISA to protect SGW (GTP-U traffic).

The steps to configure the AA-ISA in an SeGW to protect against attacks toward the SGW are similar to the steps for SCTP traffic. While GTP filtering is very different from SCTP filtering, configuration to limit the flow counts, bandwidth, and session filter are similar.

Step 3.1. Create an AA IP list for SGWs.

In addition to the lists configured in step 2.1, the operator can optionally configure a list that contains the SGW IP addresses that are served by the SeGW, in case there is more than one.

```
*A:7750-1# configure application-assurance group 1:1 ip-prefix-list "SGWs" create
*A:7750-1>config>app-assure>group>pfx>$ description "Serving Gateways IPs"
*A:7750-1>config>app-assure>group>pfx>$ prefix 172.16.111.1/32
*A:7750-1>config>app-assure>group>pfx>$ prefix 172.16.111.2/32
*A:7750-1>config>app-assure>group>pfx>$ exit
```

Step 3.2. Allow only GTP-U traffic toward SGWs — No port scanning.

Similar to Step 2.2, create an GTP filter to allow only GTP traffic to/from eNBs to SGWs:

```
*A:7750-1>config>app-assure>group>sess-fltr# info
----------------------------------------------
                default-action deny event-log "FW_drops_log"
                ---snip---
                entry 2 create
                    description "allow GTP-u to SGWs"
                    match
                        ip-protocol-num udp
                        src-ip ip-prefix-list "ALL_eNBs"
                        dst-ip ip-prefix-list "SGWs"
                        dst-port eq 2152
                    exit
                    action permit
                exit
----------------------------------------------
*A:7750-1>config>app-assure>group>sess-fltr#
```

The following session filter needs to be added to the default sub-policy AQP, similar to Step 2.2:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
                    entry 500 create
                        description "apply SeGW session filter rules"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            session-filter "SeGW_FW"
                            overload-drop
                            error-drop
                            fragment-drop all
                        exit
                        no shutdown
                    exit
                exit
----------------------------------------------
```

For AA to recognize GTP traffic and perform sanity packet checking, configure a GTP filter at the group:partition level:

```
*A:7750-1# configure application-assurance group 1:1 gtp no shutdown
```

Step 3.3. DoS protection — Limit the number of GTP-U flows from eNBs.

AA can be configured to limit the number of GTP flows from an eNB. A GTP-U flow is defined by GTP-U packet destination IP + tunnel ID (TEID).

AA allows the operator to configure two limits: one that applies to the each eNB and one that applies for all GTP-U traffic from all eNBs:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            policer "GTPu-Flow-count-limit" type flow-count-limit granularity subscriber
 create
                flow-count 800
                gtp-traffic
            exit
```

The actual value of the flow count limit is a function of the number of UEs or devices served by an eNB or FAP. In the above case, it is assumed that there are 100 devices with a maximum of 8 GTP-U flows per device. For FAP, the number is typically around 32 devices per FAP. Note: By 3GPP standards, the maximum number of GTP-U tunnels per device is 16.

Assuming that there are 1000 eNBs or FAPs that are served by the SeGW, then to limit the total number of GTP-U flows, the operator can apply the following system policer:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
```

```
            policer "limit_total_GTPU_Flow_count" type flow-count-limit granularity system
 create
                flow-count 800000
                gtp-traffic
            exit
```

Configure AQPs to execute the policers:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
---snip---
                    entry 120 create
                        description "limit GTP-U traffic"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            flow-count-limit "GTPu-Flow-count-limit"
                        exit
                        no shutdown
                    exit
                    entry 130 create
                        description "limit TOTAL GTPU towards SGWs"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            flow-count-limit "limit_total_GTPU_Flow_count"
                        exit
                        no shutdown
                    exit
```

For GTP-U flow count policing, it is important that **aqp-initial-lockup** is enabled:

```
*A:7750-1# configure application-assurance group 1:1 aqp-initial-lookup
```

The above configured limits are applied only to upstream traffic, to protect the network. No limit is placed on the downstream traffic toward the eNBs.

Note:
For small cell deployments, the number of GTP-U tunnels per FAP is a function of:

a. deployment mode:

   i. residential = 32 (9962 MSEC-MS-MCI Enterprise) UEs

   ii. enterprise = 8 (9961 MSHC) UEs.

b. number of guaranteed bit rate (GBR) tunnels (max 8) and non-GBR tunnels (max 8) per UE.

Therefore, the GTP-U tunnel limit per FAP should be set to 32 x 8 = 256 for residential deployments or 8 x 8 = 64 for enterprise deployments.

The operator can view the effect of the configured policers on GTP traffic by running the following show routine:

```
*A:7750-1>show>app-assure>group# gtp
===============================================================================
Application Assurance Group 1:1 GTP
===============================================================================
Admin status : Up
Event log    : (Not Specified)
-------------------------------------------------------------------------------
GTP Statistics                                sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Incoming packets                                       0                      0
Packets denied
  UDP packet length                                    0                      0
  GTP message length                                   0                      0
  GTP version                                          0                      0
-------------------------------------------------------------------------------
Packets permitted                                      0                      0
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
GTP Policing Statistics                       sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Packets arrived                                        0                      0
Packets denied
  gtp-traffic flow-count policer                       0                      0
  Other                                                0                      0
-------------------------------------------------------------------------------
Packets permitted                                      0                      0
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
GTP Filter Statistics                         sub-to-net             net-to-sub
-------------------------------------------------------------------------------
Packets arrived                                        0                      0
Packets denied (gtp-filter)                            0                      0
Packets permitted
  gtp-filter                                           0                      0
  no gtp-filter                                        0                      0
-------------------------------------------------------------------------------
Total GTP packets permitted                            0                      0
===============================================================================
*A:7750-1>show>app-assure>group#
```

In the last section shown above, GTP filter statistics are related to GTP filters that are discussed and configured later in Step 3.5 of this chapter.

Step 3.4. DoS protection: Limit the GTP-U bandwidth from eNBs.

This step is similar to Step 3.3, but instead of configuring a flow count policer, the operator configures bandwidth policers:

```
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            policer "GTPU_bw_limit" type single-bucket-bandwidth granularity subscriber
 create
                rate 5000
                mbs 100
            exit

            policer "limit_total_GTPU_bw" type single-bucket-bandwidth granularity system
 create
                rate 2000000
                mbs 2000
            exit

*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
---snip---
                    entry 120 create
                        description "limit GTP-U traffic"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            bandwidth-policer "GTPU_bw_limit"
                            flow-count-limit "GTPu-Flow-count-limit"
                        exit
                        no shutdown
                    exit
                    entry 130 create
                        description "limit TOTAL GTPU towards SGWs"
                        match
                            traffic-direction subscriber-to-network
                        exit
                        action
                            bandwidth-policer "limit_total_GTPU_bw"
                            flow-count-limit "limit_total_GTPU_Flow_count"
                        exit
                        no shutdown
                    exit
```

The above configured limits are applied only to upstream traffic, to protect the network. No limit is placed on downstream traffic toward the eNB.

As a debugging tool, the operator can use the AA **flow-record-search** command to check the status of GTP flows through the ISA:

```
*A:7750-1# tools  dump application-assurance group 1:1 flow-record-search  isa 1/2 flow-
status active protocol "gtp"
=====================================================
Application-Assurance flow record search, Version 1.0
Search Start Time:     "06/16/2015 20:38:09" (UTC)
 Search Criteria:
  group[:partition]:   1:1
  isa:                 1/2
  protocol name:       "gtp"
  application name:    none specified
  app-group name:      none specified
  flow-status:         active
  start-flowId:        none specified
  classified:          none specified
  server-ip:           none specified
  server-port:         none specified
  client-ip:           none specified
  bytes-tx:            none specified
  flow-duration:       none specified
  max-count:           none specified
  search-type:         default
=====================================================
FlowId  Init  Src-ip                                  Dst-ip
     Ip-prot     Src-prt Dst-prt Protocol               Application            Pkts-tx
 Bytes-tx             Pkts-disc  Bytes-disc Time-ofp(UTC)         Time-olp(UTC)
SEARCH COMPLETED.
Search End Time:     "06/16/2015 20:38:09" (UTC)
Total Records:       0
=====================================================
*A:7750-1#
```

GTP flows that are to be denied by the previous AA configurations should not appear in the search results.

Step 3.5. Further GTP filtering and validation.

AA allows the operator to configure a GTP filter to enforce which GTP message types are allowed/ denied, as well as the maximum allowed GTP message length:

```
*A:7750-1>config>app-assure>group>gtp>gtp-fltr#
 [no] description     - Configure a description of the GTP filter
 [no] event-log       - Configure an event log for packets dropped by the GTP filter
 [no] max-payload-le* - Configure the maximum payload length of the GTP filter
      message-type    + Configure actions for specific or default messages

*A:7750-1>config>app-assure>group>gtp>gtp-fltr#
```

Note:
An AA GTP filter allows the operator to configure a maximum payload size for the GTP traffic. However, in this configuration example, no maximum payload size is configured.

The list of GTP message types are defined by 3GPP standard 3GPP TS 29.281 as per Table 10: GTP Messages .

Table 10: GTP Messages

Message Type Value (Decimal)	Message	Message Type Value (Decimal)	Message
1	“echo-request”	55	“”forward-relocation-complete
2	“echo-response”	56	“relocation-cancel-request”
3	“versi”n-not-supported:	57	“relocation-cancel-response”
4	“node-alive-request”	58	“forward-sms-context”
5	“ node-alive-response”	59	“forward-relocation-complete-acknowledge”
6	“redirection-request”	60	“forward-sms-context-acknowledge”
7	“redirection-response”	70	“ran-information-relay”
16	“create-pdp-context-request”	96	“mbms-notification-request”
17	“create-pdp-context-response”	97	“mbms-notification-response”
18	“update-pdp-context-request”	98	“mbms-notification-reject-request”
19	“update-pdp-context-response”	99	“mbms-notification-reject-response”
20	“delete-pdp-context-request”	100	“create-mbms-context-request”
21	“delete-pdp-context-response”	101	“create-mbms-context-response”
22	“initiate-pdp-context-activation-request”	102	“update-mbms-context-request”

Message Type Value (Decimal)	Message	Message Type Value (Decimal)	Message
23	“initiate-pdp-context-activation-response”	103	“update-mbms-context-response”
26	“error-indication”	104	“delete-mbms-context-request”
27	“pdu-notification-request”	105	“delete-mbms-context-response”
28	“pdu-notification-response”	112	“mbms-registration-request”
29	“pdu-notification-reject-request”	113	“mbms-registration-response”
30	“pdu-notification-reject-response”	114	“mbms-de-registration-request”
31	“supported-extension-headers-notification”	115	“mbms-de-registration-response”
32	“send-routing-information-for-gprs-request”	116	“mbms-session-start-request”
33	“send-routing-information-for-gprs-response”	117	“mbms-session-start-response”
34	“Failure-report-request”	118	“mbms-session-stop-request”
35	“failure-report-request”	119	“mbms-session-stop-response”
36	“note-ms-gprs-present-request”	120	“mbms-session-update-request”
37	“note-ms-gprs-present-response”	121	“mbms-session-update-response”
48	“identification-request”	128	“ms-info-change-notification-request”
49	“identification-response”	129	“ms-info-change-notification-response”
50	“sgsn-context-response”	240	“data-record-transfer-request”
51	“sgsn-context-request”	241	“data-record-transfer-response”
52	“sgsn-context-acknowledge”	254	“end-marker”
53	“forward-relocation-request”	255	“g-pdu”
54	“forward-relocation-response”		

Of the 67 GTP message types shown above, only 6 are allowed, by the standards, for GTP-U:

```
echo-request     echo-response      error-indication
g-pdu            end-marker         supported-extension-headers-notification
```

If these message types are permitted by the configured GTP filter, AA performs extensive GTP-U header checking on these six types.

Note:
If no GTP filter is configured, no extensive GTP-U header checks are performed. For example, if the operator wants to allow all GTP-U packets and perform all GTP header sanity checks, then a GTP filter that permits all message types needs to be configured, with the default action of permit and with no values, such as:

```
                 gtp-filter “allow-all” create
                      message-type
                         default-action permit
```

Because AA FW in an SeGW is protecting an S1-U interface running GTP-U, the GTP filter only needs to allow the six GTP messages that are permitted for GTP-U:

```
*A:7750-1>config>app-assure>group>gtp# info
----------------------------------------------
                gtp-filter "filter-gtp-msgs" create
                    description "allow only certain msg types"
                    message-type
                        default-action deny
                        entry 1 value "echo-request" action permit
                        entry 2 value "echo-response" action permit
                        entry 3 value "error-indication" action permit
                        entry 4 value "supported-extension-headers-notification" action
 permit
                        entry 5 value "end-marker" action permit
                        entry 6 value "g-pdu" action permit
                    exit
                exit
                no shutdown
----------------------------------------------
*A:7750-1>config>app-assure>group>gtp#
```

This GTP filter is then referenced from within an AQP entry action, as follows, in order for it to take effect:

```
*A:7750-1>config>app-assure>group>policy>aqp# info
----------------------------------------------
                    entry 120 create
                        description "limit GTP-U traffic"
                        match
                            traffic-direction subscriber-to-network
                            dst-ip eq ip-prefix-list "SGWs"
                        exit
                        action
                            bandwidth-policer "GTPU_bw_limit"
                            flow-count-limit "GTPu-Flow-count-limit"
                            gtp-filter "filter-gtp-msgs"
                        exit
                        no shutdown
                    exit
```

The operator can view the effect of the configured GTP filter on S1-U traffic using the following show routine:

```
*A:7750-1>show>app-assure>group# gtp gtp-filter "filter-gtp-msgs"
===============================================================================
Application Assurance Group 1:1 GTP Filter "filter-gtp-msgs"
===============================================================================
Description            : allow only certain msg types
Maximum payload length : (Not Specified)
Default action         : deny
Configured messages    : 6

Packets arrived        : 0
Packets denied
  Payload length       : 0
  Message type         : 0
  Mandatory header     : 0
  Extension header     : 0
  Information element  : 0
Packets permitted      : 0
===============================================================================
*A:7750-1>show>app-assure>group#
```

The above output is in addition to the information provided by the overall GTP show command:

```
*A:7750-1>show>app-assure>group# gtp
```

4. Configure AA-ISA to protect NMS (OAM Traffic).

Step 4.1. Create an IP AA list that contains the NMS server IPs.

```
*A:7750-1# configure application-assurance group 1:1
*A:7750-1>config>app-assure>group# info
----------------------------------------------
            ip-prefix-list "NMSs" create
                description "Network Management-OAM subnet"
                prefix 172.16.120.0/30
            exit
```

Note:
In the case of small cell deployments, different NMS servers need to be configured.

Step 4.2. Allow eNBs to initiate FTP- and ICMP-only traffic toward NMS, block port scanning.

```
*A:7750-1>config>app-assure>group>sess-fltr# info
----------------------------------------------
                default-action deny event-log "FW_drops_log"

                entry 3 create
                    description "allow FTP to NMS"
                    match
                        ip-protocol-num tcp
                        src-ip ip-prefix-list "ALL_eNBs"
                        dst-ip ip-prefix-list "NMSs"
                        dst-port eq 22
                    exit
                    action permit
                exit
                entry 4 create
                    description "allow ICMP to NMS"
                    match
                        ip-protocol-num icmp
                        src-ip ip-prefix-list "ALL_eNBs"
                        dst-ip ip-prefix-list "NMSs"
                    exit
                    action permit
                exit
----------------------------------------------
```

```
*A:7750-1>config>app-assure>group>sess-fltr#
```

The operator can view the effect of the session filter on traffic, in terms of how many times it is applied, using the following show routine:

```
*A:7750-1>show>app-assure>group# session-filter
===============================================================================
AA Session Filter Table
===============================================================================
Name                             Default Action    Referenced          Entries
-------------------------------------------------------------------------------
SeGW_FW                          deny              aqp                       4
-------------------------------------------------------------------------------
No. of session filters: 1
===============================================================================
*A:7750-1>show>app-assure>group#
*A:7750-1>show>app-assure>group# session-filter "SeGW_FW"
===============================================================================
AA Session Filter Instance "SeGW_FW"
===============================================================================
Description    : (Not Specified)
Default Action : deny
    Event Log  : FW_drops_log
AQP Entries    :
          500
-------------------------------------------------------------------------------
Filter Match Criteria
-------------------------------------------------------------------------------
Entry          : 1
Description    : allow SCTP to MMEs
IP Protocol    : sctp
Source IP List : ALL_eNBs
Dest IP List   : MMEs
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 2
Description    : allow GTP-u to SGWs
IP Protocol    : udp
Source IP List : ALL_eNBs
Dest IP List   : SGWs
Dest Port      : eq 2152
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 3
Description    : allow FTP to NMS
IP Protocol    : tcp
Source IP List : ALL_eNBs
Dest IP List   : NMSs
Dest Port      : eq 22
Action         : permit
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
Entry          : 4
Description    : allow ICMP to NMS
IP Protocol    : icmp
Source IP List : ALL_eNBs
Dest IP List   : NMSs
Action         : permit
```

```
    Event Log  : (Not Specified)
Hits           : 0 flows
-------------------------------------------------------------------------------
No. of entries   : 4
===============================================================================
*A:7750-1>show>app-assure>group#
```

Note:
The above configuration is generic and may need to be modified to suit the deployment requirements. For example, in the case of small cell SeGW deployment, traffic on other ports needs to be allowed to/from different NMS type servers, such as allowing TCP port 7003 and port 7013 to HDM servers. This can be accomplished by configuring additional entries in the above session filter.

Note:
By allowing port 22 for FTP, the AA FW automatically opens and closes the associated data channel ports. For more information about AA FW capabilities, with regard to OAM FW protection, see Application Assurance Stateful Firewall.

Conclusion

The SR OS AA stateful firewall feature runs on AA-ISA and extends application-level analysis to provide an in-line stateful service, integrated within the Security Gateway (SeGW).

AA stateful packet filtering, combined with AA Layer 7 classification and control, provides advanced, next-generation firewall functionality, protecting mobile network core infrastructure, such as MMEs, SGWs, and NMSs.

Application Assurance — Stateful Firewall

This chapter describes Application Assurance stateful firewall (FW) configurations for protecting residential and WiFi subscribers.

Topics in this chapter include:

Applicability

Initially, this chapter was written for SR OS Release 11.0.R1. The TCP validation section was added for SR OS Release 14.0.R4.

Overview

The AA SR OS 11.0.R1 stateful FW feature extends AA-ISA application level analysis to provide an in-line integrated stateful service that protects subscribers from malicious attacks. AA stateful packet filtering combined with AA L7 classification and control, empowers operators with advanced, next generation firewall functionality that is integrated within the Service Router. The AA stateful firewall (FW) and application firewall runs on AA-ISA. Using stateful inspection, the AA firewall not only inspects packets at layers 3-7, but also monitors and keeps track of the connection's state. If the operator configures a **deny** action within a session filter, then the matching packets (matching both the AA Application QoS policy (AQP) and associated session filter match conditions) are dropped and no flow session state/context is created.

AA FW can be used in all deployments of AA-ISA; mobile (MG OS) and fixed (SR OS), however the configurations examples here, while still very applicable (and almost 100% identical in mobile deployments) are focused on AA-ISA deployments in fixed networks.

The AA-ISA FW enabled solution provides:

- Stateful (and stateless) packet filtering and inspection with application-level gateway (ALG) support
- DoS attack protection

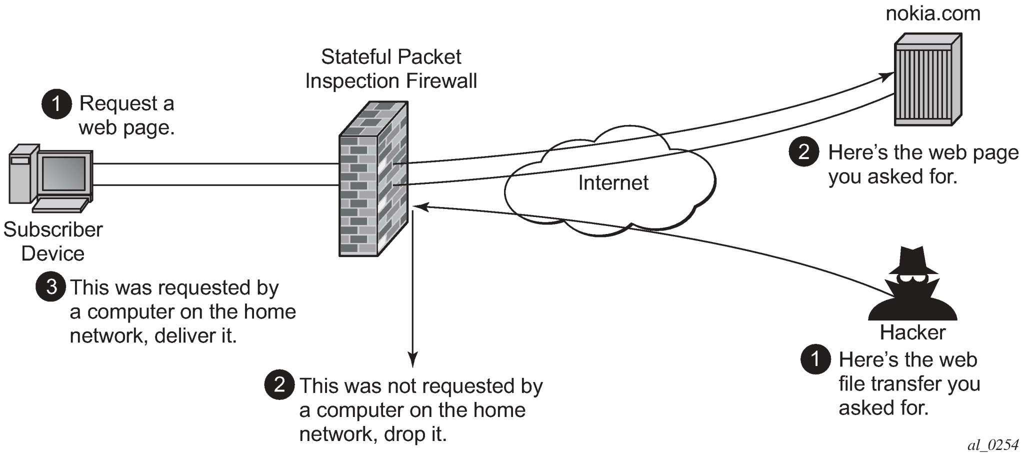
In SR OS Release 14.0, additional firewall functionalities were added, such as TCP-validation, threshold crossing alerts, syslog and statistics related to firewall actions.

The objective of this chapter is to describe the required configuration within AA-ISA (divert to AA-ISA basic knowledge is assumed) in order to enable AA FW and protect AA subscribers from attacks (Unsolicited attacks and DoS attacks), while still allowing pin-holing through the firewall, so that applications like peer to peer gaming and various ALGs (such as FTP) are not affected.

Stateful Filtering

By performing stateful inspection, AA-ISA takes into account which side initiated a session and acts accordingly. Stateful flow processing and inspection utilizes IP layers 3/ 4 header information to build a state of the flow within AA-ISA. Layer 7 inspection is used in order to provide ALG support. Stateful flow/ session processing takes note of the originator of the session and hence can allow traffic to be initiated from the subscriber, while denying (when configured) traffic originating from the network. Packets received from the network are inspected against the session filter and only those that are part of a subscriber initiated session are allowed.

Figure 33: Block Unsolicited Traffic


nokia.com
Stateful Packet Inspection Firewall
1 Request a web page.
Subscriber Device
Internet
2 Here's the web page you asked for.
3 This was requested by a computer on the home network, deliver it.
Hacker
1 Here's the web file transfer you asked for.
2 This was not requested by a computer on the home network, drop it.
al_0254

To support the example shown in Figure 33: Block Unsolicited Traffic, AA is configured with an action to block unsolicited traffic; traffic that is not originated/initiated from the subscriber. The direction field in match criteria of AQPs is utilized to enable this functionality.

Figure 34: SFW — Allow Gaming

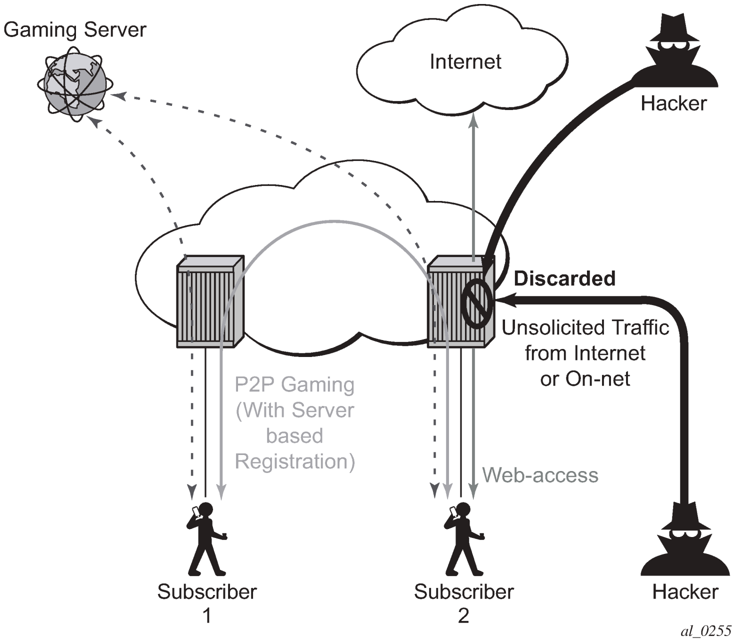
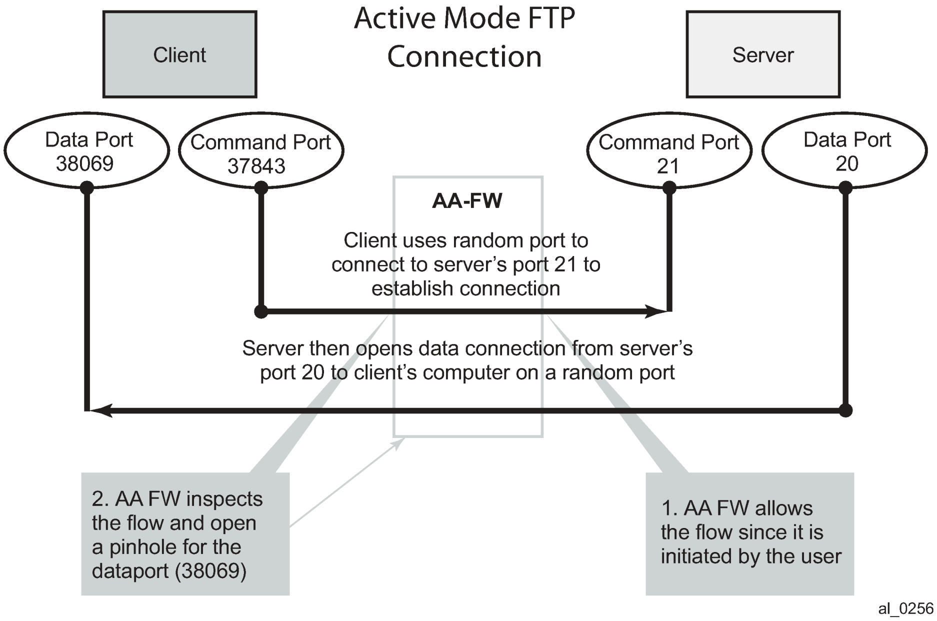

Gaming Server
Internet
Hacker
Discarded
Unsolicited Traffic from Internet or On-net
P2P Gaming (With Server based Registration)
Web-access
Subscriber 1
Subscriber 2
Hacker
al_0255

Figure 34: SFW — Allow Gaming shows a similar concept. It is used to allow UDP traffic for peer to peer applications (such as gaming). Once the traffic from one peer is seen by AA-ISA, a pin-hole is opened in the reverse direction to allow for the corresponding UDP traffic from the peer.

Stateless packet filtering on the other hand does not take note of the session initiator. It discards or allows packets independently of any previous packets. In addition to AA-ISA's support for stateless (and stateful) filtering, stateless packet filtering can be performed in the system using line card ACLs (and/or MGISM PCC rules in mobile gateway deployments).

Application Layer Gateway Filtering

Figure 35: ALG Support Example — FTP


Active Mode FTP Connection
Client
Server
Data Port 38069
Command Port 37843
Command Port 21
Data Port 20
AA-FW
Client uses random port to connect to server's port 21 to establish connection
Server then opens data connection from server's port 20 to client's computer on a random port
2. AA FW inspects the flow and open a pinhole for the dataport (38069)
1. AA FW allows the flow since it is initiated by the user
al_0256

AA FW inspection of packets at Layer 7 offers Application Layer Gateway functionality for a large list of applications (for example, FTP, SIP, RTSP, PPTP, IRC, etc.). These applications make use of control channels or flows that spawn other flows. AA FW inspects the payload of these control flows so it can open a pinhole in advance for unspawned data flows. Figure 35: ALG Support Example — FTP depicts an example of AA ALG support for FTP traffic.

Denial Of Service (DOS) Protection

DoS attacks work by consuming network and system resources, making them unavailable for legitimate network applications. Network flooding attacks, malformed packets and port scans are examples of such DoS attacks.

The aim of AA FW DOS protection is to protect subscribers and prevent any abuse of network resources.

Using AA FW stateful session filters, operators can protect their subscribers from any port scan scheme. This can be done by configuring the session filters to disallow any traffic that is initiated from the network.

Furthermore, AA ISA provides configurable flow policers. These policers, once configured, prevent a wide range of flooding attacks (such as ICMP PING flooding, UDP flooding, SYN Flood Attack...etc.). These policers provide protection at multiple levels; per system per application/application groups and per subscriber per applications/applications groups.

There are two types of AA ISA flow policers; flow setup rate policers and flow count policers. Flow setup rate policers limit the number of new flows, while flow count policers limit the total number of active flows.

In order to protect hosts and network resources, AA FW validates/checks different fields in the packet's header (checksum, TCP Flag, etc.) and if any fails it declares the packet to be invalid. This complements the SR OS subscriber management enhanced security features, such as IP (or MAC) anti-spoofing protection (such as protecting against LAND attacks) and network protocol DoS protections. The cut-through-drop AQP action must be configured in order to drop these types of invalid packets.

Virtual FW/Zone-Based FW

AA FW can provide up to 128 virtual FWs, each with its own FW policies. This is achieved through the use of AA-partitions.

In addition, AA subscribers within the same AA partition can have different application profiles with different Application Service Options (ASO) values. This provides a further control mechanism to enable/disable firewall rules.

For example, the operator may want to have some subscribers possess full firewall protection, while other subscribers not subscribed to this service to have a partial firewall protection that focuses on protecting network resources, rather than network and subscribers resources.

Configuration

AA-ISA AQPs were enhanced in R11.0.R1 with several AQP actions that provide session filtering functionality. As is the case of all AQPs, these have partition level scope, which allows different FW policies to be implemented by utilizing AA partitions concepts within the same AA-ISA group. Hence, multiple virtual AA FW instances can be realized, without the need for multiple physical instances of FWs to implement different FW policies.

The AA FW stateful session filter consists of multiple entries (similar to ACLs) with a match an action per entry. Actions are **deny** or **permit**. A **deny** action results in packets being discarded without creating a session/flow context. Match conditions include IP protocol types, source and destination IP addresses and ports. An overall default action is also configurable in case of no match to any session filter entry.

AQPs with session filter actions need to have — as a matching condition — traffic direction, ASOs, and/or a subscriber name. These AQP match rules cannot have any references to applications and/or application groups.

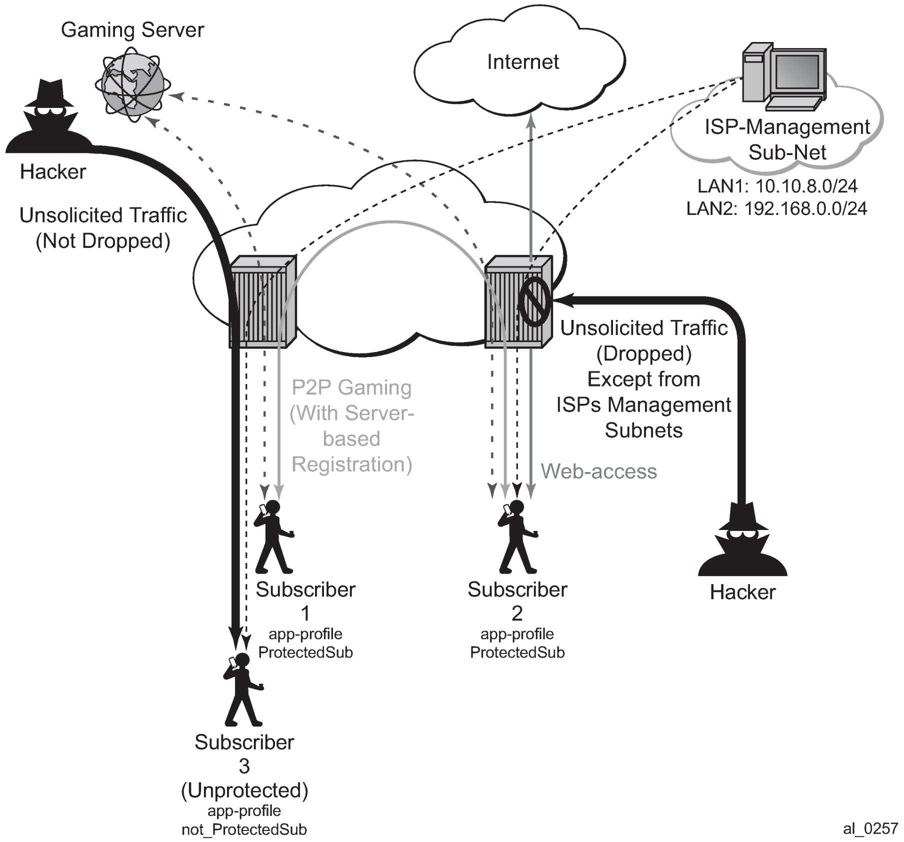
An AQP action to drop malformed/errored packets is also available.

Statistics are incremented when packets are dropped by a session filter. These are accounted against:

- protocol = denied by default policy
- application= unknown
- application group = unknown.

The configuration topology is shown in Figure 36: Configuration Topology.

Figure 36: Configuration Topology


Gaming Server
Internet
Hacker
ISP-Management Sub-Net
LAN1: 10.10.8.0/24
LAN2: 192.168.0.0/24
Unsolicited Traffic (Not Dropped)
Unsolicited Traffic (Dropped) Except from ISPs Management Subnets
P2P Gaming (With Server-based Registration)
Web-access
Subscriber 1
app-profile ProtectedSub
Subscriber 2
app-profile ProtectedSub
Hacker
Subscriber 3 (Unprotected)
app-profile not_ProtectedSub
al_0257

1. Application Profile configuration:

   There is nothing new introduced in application profiles in order to support FW. This section explains how to configure the application profile to allow differentiated FW services for different subscribers. In a nut shell, the AA common building construct/attribute for differentiated policy is ASO.

   To configure an ASO for FW protection:

```
configure application-assurance group 1:1 policy
  begin
  app-service-options
     characteristic "FW-Protection" create
            value "None"
            value "ON"
            default-value "None"
     exit
     characteristic "ISP-Protection" create
            value "None"
            value "ON"
            default-value "None"
     exit
     characteristic "DOS-Protection" create
            value "None"
            value "ON"
            default-value "None"
     exit
```

```
   exit
```

In the preceding example:

- ASO FW protection allows the operator to select if the subscriber is FW protected or not.
- ASO DOS protection refers to if the subscriber is protected from DOS attacks.
- ASO ISP protection is different from the preceding two as it protects the ISP resources by (in the example that follows) not allowing unsolicited traffic. This should be ON for all subscribers (it is then arguable if someone needs it to be defined in the ASO list, instead of merely configuring an AQP to protect ISP resources all the time).

These ASOs are referenced in appProfiles (and later in AQPs) as follows:

```
configure application-assurance group 1:1 policy
    begin
        app-profile "Protected" create
           divert
           characteristic "FW-Protection" value "ON"
           characteristic "ISP-Protection" value "ON"
           characteristic "DOS-Protection" value "ON"
    exit
```

The preceding application profile Protected is assigned to subscribers who opted/subscribed to the firewall protection service; for example sub 1 and sub 2 in the example shown in Figure 36: Configuration Topology.

Subscribers who are not protected (for example sub 3 in Figure 36: Configuration Topology) are assigned a different profile:

```
configure application-assurance group 1:1 policy
    begin
        app-profile "unProtected" create
           divert
           characteristic "FW-Protection" value "None"
           characteristic "ISP-Protection" value "ON"
           characteristic "DOS-Protection" value "None"
    exit
```

An alternative method to using application profiles/ASOs to provide differentiated services is to configure multiple partitions with different AQPs/ session filters. One partition for example will be for subscribers who are provided with firewall protection, while another is used for subscribers who are not protected. This configuration is simpler and provides statistics per partition. This example however covers the more complex case using ASOs/appProfiles.

2. Flow count policer configuration:

```

configure application-assurance group 1 policer Dos_police_Flow_count type flow-count-limit
 granularity subscriber create
   flow-count 500
exit
```

The preceding configuration limits the number of flows a subscriber can have at any time to 500. This is done to protect against DoS attacks. The value 500 is arbitrary and requires tuning for each deployment.

```
configure application-assurance group 1 policer Dos_Police_ICMPFlows type flow-count-limit
 granularity system create
   flow-count 5000
exit
```

This configuration limits the total number of flows that matches the configured AQP matching condition. It is used for ICMP applications to prevent mass port scanning.

3. TCP Protocol Validation configuration

```
configure application-assurance group 1:1 tcp-validate TCP_protect create
```

This simple configuration allows the operator to call TCP_protect policy from within an AQP action entry.

The operator can also configure the policy to be "strict", in which case the AA checks for valid sequence and acknowledgements numbers. In this example, the "strict" option is not used.

4. Application configuration

The following configuration is standard with AppDB. It is shown here for reference.

```
configure application-assurance group 1:1 policy begin
  application ICMP create
  exit
  app-filter
   entry 1540 create
    protocol eq "non_tcp_udp"
    ip-protocol-num eq icmp
    application "ICMP"
    no shutdown
   exit
   entry 35500 create
    protocol eq "non_tcp_udp"
    ip-protocol-num eq ipv6-icmp
    application "ICMP"
    no shutdown
   exit
```

5. Session-Filter

The following displays session-filter configuration commands to be used in Step 6 later.

```
configure application-assurance group 1:1 session-filter <name> create
description  <description>
    default-action permit|deny    # default=deny
    entry n create
        description <entry-description>
        match
            ip-protocol-num <ip-protocol-number>
            no src-ip  <ip4_or_v6-address/mask>
            no dst-ip  <ip4_or_v6-address/mask>
            no src-port {eq|gt|lt} <port-num> #or
                    range <start-port-num> <end-port-num>
            no dst-port {eq|gt|lt} <port-num> #or
                    range <start-port-num> <end-port-num>
        exit
        action permit|deny
    exit
    entry m create
    ---snip---
```

Parameters

- **entry** *n* — A session filter can have multiple match-action rules, each of these match-action rules represent an entry within the session-filter. The entries are executed in order. If a match is found, within one entry, the subsequent entries within the session-filter are skipped (not evaluated).
- **default-action [permit | deny]** — This action is performed if no match is found for any of the configured entries within the session-filter. Default is deny.
  - A **deny** action will drop the packet and will not allow a flow record to be allocated for that flow. A **drop** action within AA AQP will drop the packet but it will still create flow record.
  - A **permit** action will allow the packet to flow through the system. A flow record is also allocated. The packet may get dropped by other configured AQP actions (due to header check failures).
- **description** *description-string*

  This configures a text string, up to 80 characters, which can be used to describe the use of the session-filter.
- **match** — Keywords to perform the action specified under the **action** keyword only if the conditions in the match section are met.
  - **ip-protocol** *ip-protocol-number*

    ip-protocol-number — 1..255

    - Decimal, hexadecimal or binary representation

    - Supported IANA IP protocol names:

    - crtp, crudp, egp, eigrp, encap, ether-ip, gre, icmp, idrp, igmp, igp, ip, ipv6, ipv6-frag, ipv6-icmp, ipv6-no-nxt, ipv6-opts, ipv6-route, isis, iso-ip, l2tp, ospf-igp, pim, pnni, ptp, rdp, rsvp, sctp, stp, tcp, udp, vrrp
  - **src-ip>/dst-ip** *ipv4-address/mask* **src-ip/dst-ip** *iv6-address/mask*

    - Source/destination IP address within the packet header.

    - IPv4 or IPv6 formats are allowed, with prefixes masks.
  - **src-port** *src-port-numbers*

    src-port {eq | gt | lt} *port-num*

    eq — equal, exact match

    gt — match port numbers that are greater than the one specified.

    lt — match port numbers that are smaller than the one specified.

    port-num — 0..65535 (Applicable to TCP, UDP and SCTP protocols only.)
  - **src-port range** *start-port-num end-port-num*

    range — Keyword- that match port numbers within the specified range:

    start-port-num — 0..65535

    end-port-num — 0..65535
  - **dst-port** *dst-port-number*

    - Same as source port number explained above, but applied against destination port number.

- **action deny | permit**
  - **deny** or **permit** action is only executed if a match is found.
  - **deny** action will drop the packet and will not create a flow record.
  - **permit** action will allow the packet to go through (unless another different action is found that causes it to be dropped).
- **no entry** *entry-id*
  - Causes the entry to be deleted.
- **no session-filter** *session-filter-name*
  - Causes the session filter to be deleted.

```
config application-assurance group 1:1
       session-filter " denyUnsolictedwMgntCntrl" create
       description “S-FW opted-in sub – allow ISP access"
       default-action deny
       entry 10 create
           description "allow ICMP access from ISP LAN1"
           match
               ip-protocol-num icmp
               src-ip 10.10.8.0/24
           exit
           action permit
        exit
        entry 20 create
            description "allow ICMP access from ISP LAN2"
           match
               ip-protocol-num icmp
               src-ip 192.168.0.0/24
           exit
           action permit
       exit
       entry 30 create
           description "allow all TCP (e.g. FTP/telnet)access from ISP LAN2"
           match
               ip-protocol-num tcp
               src-ip 192.168.0.0/24
           exit
           action permit
       entry 40 create
           description "allow TCP on port 80 /HTTP access from ISP LAN1"
           match
               ip-protocol-num tcp
               src-ip 10.10.8.0/24
               dst-port eq 80
           exit
           action permit
       exit
```

This session filter is used to protect systems located in LAN2. It drops all unsolicited traffic except for FTP coming from LAN1.

```
configure application-assurance group 1:1
   session-filter "protectISPLan2" create
       description "S-FW to deny all unsolicited requests to LAN2"
       default-action deny
       entry 10 create
```

```
            description "allow ftp access from ISP LAN1"
            match
                ip-protocol-num tcp
                src-ip 10.10.8.0/24
                dst-port eq 21
            exit
            action permit
        exit
    exit
```

6. AQP configuration:

```
configure application-assurance group 1:1  policy
  begin
 app-qos-policy

  entry 100 create
     description "Protecting ISP1 from DoS attacks from subs"
     match
        traffic-direction subscriber-to-network
        characteristic "ISP-Protection" eq "ON"
        dst-ip eq 10.10.8.0/24
     exit
     action
        flow-count-limit Dos_police_Flow_count
    exit
    no shutdown
  exit

  entry 105 create
    description "Protecting ISP2 from DoS attacks from subs"
    match
        traffic-direction subscriber-to-network
        characteristic "ISP-Protection" eq "ON"
        dst-ip eq 192.168.0.0/24
    exit
    action
        flow-count-limit Dos_police_Flow_count
    exit
    no shutdown
  exit
```

These AQPs protect the ISP network by limiting the number of concurrent flows. Dropping malformed packets is done by entry 130 (later).

To guard against ICMP flooding attacks, a flow count policer (defined earlier) is used as follows:

```
configure application-assurance group 1:1 policy
  begin
 app-qos-policy entry 107 create
    match
       application eq "ICMP"
       traffic-direction subscriber-to-network
    exit
    action
       flow-count-limit Dos_Police_ICMPFlows
    exit
    no shutdown
  exit
```

To guard against attacks exploiting TCP handshake mechanisms, TCP validate policy (defined earlier) is used as follows:

```
configure application-assurance group 1:1 policy
 begin
  app-qos-policy
     entry 108 create
        match
           characteristic "ISP-Protection" eq "ON"
        exit
        action
           tcp-validate "TCP_protect"
        exit
        no shutdown
     exit
     entry 109 create
        match
           characteristic "FW-Protection" eq "ON"
        exit
        action
           tcp-validate "TCP_protect"
        exit
       no shutdown
    exit
```

TCP validation works on both direction and needs to be called in from within a sub-default AQP entry. Therefore, this AQP action cannot be restricted to one ISP versus another because no destination IP can be specified. The configuration shown runs TCP validation policy when ISP-Protection or FW-protection ASOs are enabled.

The preceding configuration will ensure, for example, that no TCP session starts without the proper handshake message exchanges.

In order to protect ISP LAN2 from all incoming traffic (unsolicited), the operator configures entry 120.

```
   entry 120 create
      match
         traffic-direction subscriber-to-network
         characteristic "ISP-Protection" eq "ON"
      exit
      action
         session-filter "protectISPLan2"
      exit
      no shutdown
   exit
```

ProtectISPLan2 session filter drops all unsolicited traffic to LAN2 (highly secure) except for access to FTP services coming from ISP LAN1. Details of these configurations are shown in Session-Filter (step 5).

To enable stateful protection for opted-in subscribers:

```
configure application-assurance group 1:1 policy
 begin
  app-qos-policy
     entry 110 create
        description "FW for managed opted-in subs"
        match
           traffic-direction network-to-subscriber
           characteristic "FW-Protection" eq "ON"
        exit
        action
           session-filter "denyUnsolictedwMgntCntrl"
        exit
        no shutdown
     exit
```

The preceding AQP protects opt-in subscribers from unsolicited traffic but still allows unsolicited traffic from ISP subnets to manage the subscriber’s network.

Dropping malformed/illegal packets and protecting against DOS attacks is done via the following entry 130 and 131.

```
   entry 130 create
      match
         traffic-direction subscriber-to-network

        characteristic "DOS-Protection" eq "ON"
    exit
    action
        flow-count-limit Dos_police_Flow_count
    exit
    no shutdown
  exit
  entry 131 create
    match
       characteristic "DOS-Protection" eq "ON"
    exit
    action
       error-drop
       overload-drop
       fragment-drop all
    exit
    no shutdown
  exit
```

7. Configuration of Threshold Crossing Alerts (TCA).

   Operators can configure AA to generate TCAs for various firewall related parameters, such as error-drop, session-filter hits, TCP-validate, fragment-drop-all etc. as well as flow count policers. An example of a TCA used for TCP_validation policy is as follows:

```
configure application-assurance group 1:1 statistics  threshold-crossing-alert
     tcp-validate "TCP_protect" direction from-sub create
     high-wmark 50 low-wmark 40
   exit
```

   Unlike the other TCAs, in order to configure TCAs for flow count policers, operators need first to configure AA admit-deny to allocate ISA resources to record, such as:

```
configure application-assurance group 1:1 statistics aa-admit-deny policer-stats-resources
```

   Then, a TCA can be configured for any flow based policer in the system, such as:

```
configure application-assurance group 1:1 statistics  threshold-crossing-alert
     policer "Dos_police_Flow_count" direction from-sub create
     high-wmark 300 low-wmark 199
   exit
```

The system allows the various AA-admit-deny statistics to be exported via XML according to the configured accounting policy on the system. SAM-A can then use these statistics to generate the right reports / alerts.

As a prerequisite, an accounting policy is configured for aa-admit-deny statistics:

```
configure log accounting-policy 5 record aa-admit-deny
```

Then, the operator can configure AA to export the statistics related to various firewall functions configured in the system, such as:.

```
configure application-assurance group 1:1 statistics aa-admit-deny
                    accounting-policy 5
                    collect-stats
                    session-filter-stats
                    policer-stats-resources
                    tcp-validate-stats
                exit
```

GTP and STCP admit deny stats are related to firewall deployment within a SeGW, which is not covered within the scope of this chapter.

Show Routine — AQP:

```
*A:PE-1# show application-assurance group 1:1 policy app-qos-policy 110

================================================================================
Application QOS Policy Entry 110 (Default Subscriber Policy)
================================================================================
Description : FW for managed opted-in subs
Admin State : in-service
Hits:       : 0 flows
Conflicts   : 0 flows

Match :
    Traffic Direction        : network-to-subscriber
    ASO Characteristics      :
        FW-Protection                  : eq ON
Action :
    Session Filter               : denyUnsolictedwMgntCntrl
=================================================================
```

Show Routines — Session Filter:

```
*A:PE-1# show application-assurance group 1:1 session-filter
             "denyUnsolictedwMgntCntrl"
================================================================================
AA Session Filter Instance "denyUnsolictedwMgntCntrl"
================================================================================
Description    : (Not Specified)
Default Action : deny
    Event Log  : (Not Specified)
AQP Entries:    :
          110
------------------------------------------------------------

--------------------------------------------------------------------------------
Filter Match Criteria
------------------------------------------------------------ -----------------------------------
---------------------------------------------
```

```
Entry          : 10
Description    : allow ICMP access from ISP LAN1
IP Protocol    : icmp
Source IP      : 10.10.8.0/24
Action         : permit
    Event Log  : (Not Specified)
Hits:          : 0 flows
----------------------------------------------------------- -----------------------------------
--------------------------------------------
Entry          : 20
Description    : allow ICMP access from ISP LAN2
IP Protocol    : icmp
Source IP      : 192.168.0.0/24
Action         : permit
    Event Log  : (Not Specified)
Hits:          : 0 flows
----------------------------------------------------------- -----------------------------------
--------------------------------------------
Entry          : 30
Description    : allow all TCP (e.g. FTP/telnet)access from ISP LAN2
IP Protocol    : tcp
Source IP      : 192.168.0.113/320/24
Action         : permit
    Event Log  : (Not Specified)
Hits:          : 0 flows
-----------------------------------------------------------
-------------------------------------------------------------------------------
Entry          : 40
Description    : allow TCP on port 80 /HTTP access from ISP LAN1
IP Protocol    : tcp
Source IP      : 10.10.8.0/24
SourceDest Port      : eq 80
Action         : permit
    Event Log  : (Not Specified)
Hits:          : 0 flows
----------------------------------------------------------- -----------------------------------
--------------------------------------------
No. of entries   : 4
==========================================================
```

Show Routines — TCP Validation:

```
*A:PE-1# show application-ass group 1:1 tcp-validate "TCP_protect"
===============================================================================
Application Assurance Group 1:1 tcp-validate "TCP_protect"
===============================================================================
Description      : (Not Specified)
Event log        : (Not Specified)
Strict Validation: No
AQP referenced   : Yes

-------------------------------------------------------------------------------
Decision Statistics                         sub-to-net               net-to-sub
-------------------------------------------------------------------------------
Total
-------------------------------------------------------------------------------
Allowed
  Octets                                             0                        0
  Packets                                            0                        0
Dropped
  Octets                                             0                        0
  Packets                                            0                        0
```

```
Dropped Reason
-------------------------------------------------------------------------------
Bad Flags
  Octets                                             0                        0
  Packets                                            0                        0
Bad Options
  Octets                                             0                        0
  Packets                                            0                        0
Bad Sequence Number
  Octets                                             0                        0
  Packets                                            0                        0
Bad Acknowledgment Number
  Octets                                             0                        0
  Packets                                            0                        0
No Establishment
    Octets                                           0                        0
  Packets                                            0                        0
SYN After Conn Establishment
  Octets                                             0                        0
  Packets                                            0                        0
Asymmetric Traffic
  Octets                                             0                        0
  Packets                                            0                        0
Traffic After Reset (RST)
  Octets                                             0                        0
  Packets                                            0                        0
Fragmented
  Octets                                             0                        0
  Packets                                            0                        0
```

```

*A:PE-1# show application-assurance threshold-crossing-alert detail
===============================================================================
Application Assurance Threshold Crossing Alerts
===============================================================================
-------------------------------------------------------------------------------
policer "Dos_police_Flow_count" from-sub
-------------------------------------------------------------------------------
Group:Part     : 1:1                    Trigger on     : denied-traffic
High watermark : 300                    Low watermark  : 199
Last raised    : N/A                    Last cleared   : N/A
State          : cleared
-------------------------------------------------------------------------------
tcp-validate "TCP_protect" from-sub
-------------------------------------------------------------------------------
Group:Part     : 1:1                    Trigger on     : denied-traffic
High watermark : 50                     Low watermark  : 40
Last raised    : N/A                    Last cleared   : N/A
State          : cleared
No. of TCAs : 2
===============================================================================
*A:PE-1#

*A:PE-1>tools>dump>app-assure>group# admit-deny-stats
===============================================================================
Application-Assurance Group 1:1 Admit-Deny Statistics
===============================================================================
----------------------------------------------------------------------------------
Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
Packet Validation Statistics
         (Packets)            (Packets)            (Packets)            (Packets)
----------------------------------------------------------------------------------
Error
```

```
                0                    0                    0                    0
Fragments: Out-Of-Order
                0                    0                    0                    0
Fragments: All
                0                    0                    0                    0
Overload
              N/A                    0                  N/A                    0
----------------------------------------------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
Session Filter Statistics
        (Sessions)            (Packets)           (Sessions)            (Packets)
----------------------------------------------------------------------------------
Session Filter: test
 Entry: 1                                                                   0
  0                    0                    0
 Default Action                                                             0
  0                    0                    0
----------------------------------------------------------------------------------

Admitted Sub-To-Net    Denied Sub-To-Net  Admitted Net-To-Sub    Denied Net-To-Sub
TCP Validation Statistics
         (Packets)            (Packets)            (Packets)            (Packets)
----------------------------------------------------------------------------------
test                                                                        0
  0                    0                    0
TCP_protect                                                                 0
  0                    0                    0
TCP_protect_ISP1                                                            0
  0                    0                    0
----------------------------------------------------------------------------------
*A:PE-1>tools>dump>app-assure>group#
```

Conclusion

The AA stateful packet filtering feature combined with AA Layer 7 classification and control empowers operators with an advanced, next generation firewall functionality that is integrated within SR OS. This chapter focused on traditional stateful and stateless session firewall functionality.

Application Assurance — Usage Monitoring and Policy Control via Diameter Gx Protocol

This chapter provides information about the diameter (Gx) control feature.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

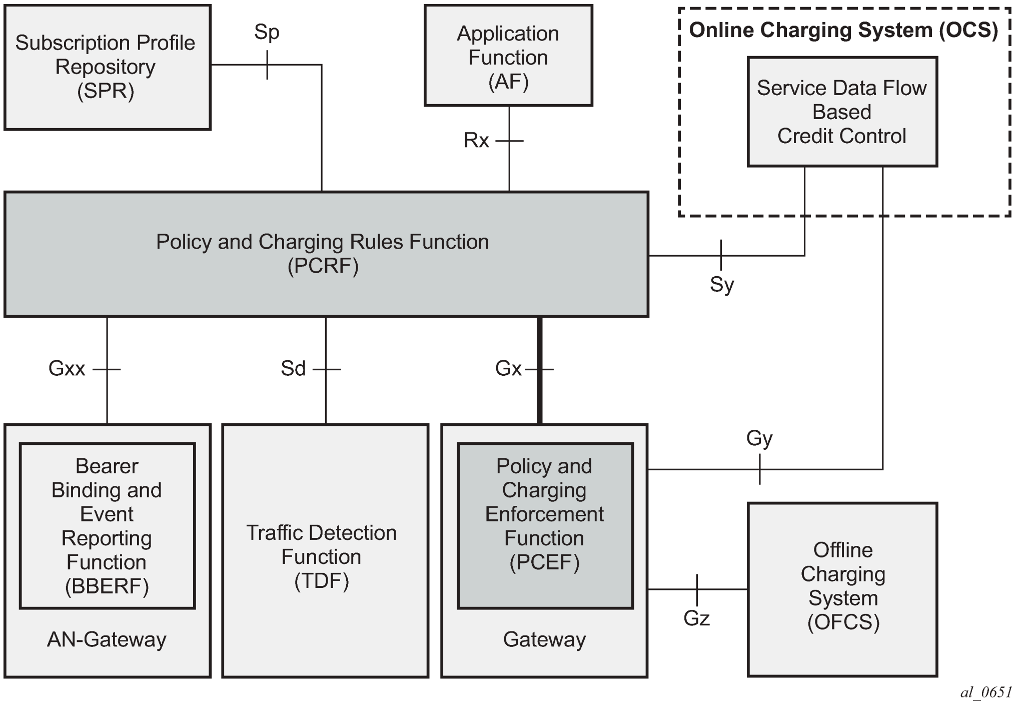
This chapter was initially based on SR OS Release 13.0.R1, but the Diameter Base configuration in the current edition is based on SR OS Release 19.10.R1.

Overview

The Gx reference point is defined in the Policy and Charging Control (PCC) architecture within the 3rd Generation Partnership Project (3GPP) standardization body. The Gx reference point is used for policy and charging control. The PCC architecture is defined in the 23.203 3GPP technical specification, while the Gx functionality is defined in the 29.212 3GPP technical specification. The SR OS implementation of Gx supports both Release 11 and Release 12 of the specification. Gx is an application of the Diameter Base Protocol (RFC 6733).

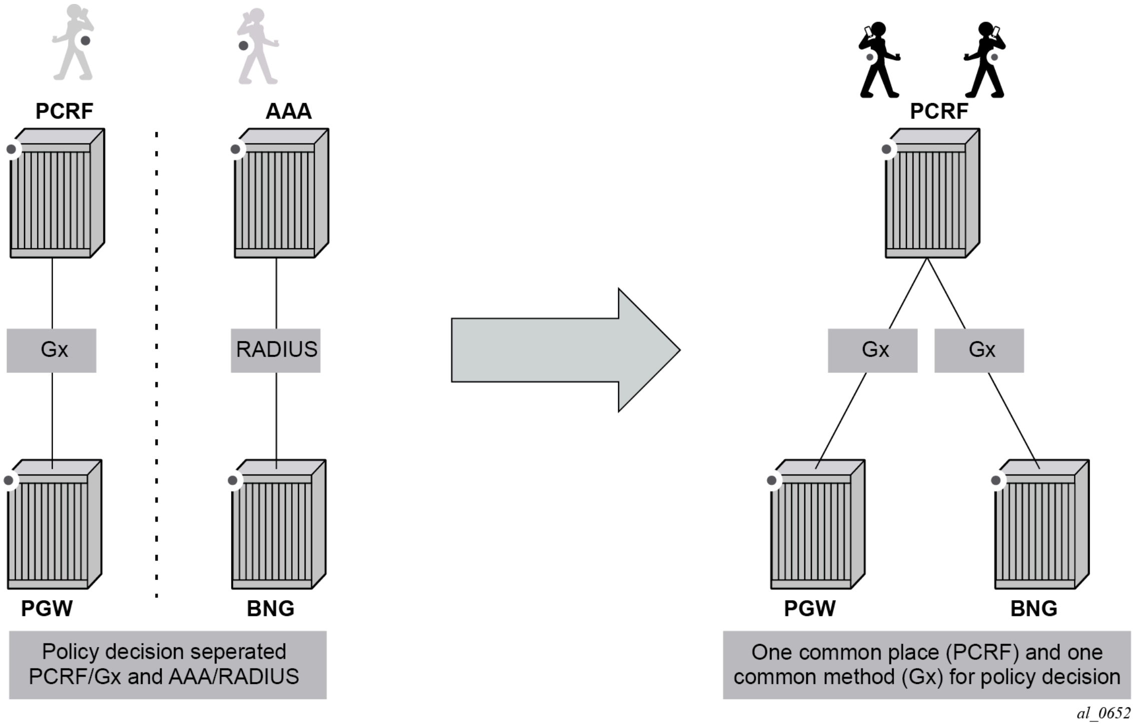
As shown in Figure 37: Gx reference point, Gx is placed between a policy server Policy and Charging Rule Function (PCRF) and a traffic forwarding node Policy and Charging Enforcement Function (PCEF) that enforces rules set by the policy server.

Figure 37: Gx reference point


Subscription Profile Repository (SPR)
Sp
Application Function (AF)
Online Charging System (OCS)
Service Data Flow Based Credit Control
Rx
Policy and Charging Rules Function (PCRF)
Sy
Gxx
Sd
Gx
Gy
Bearer Binding and Event Reporting Function (BBERF)
AN-Gateway
Traffic Detection Function (TDF)
Policy and Charging Enforcement Function (PCEF)
Gateway
Gz
Offline Charging System (OFCS)
al_0651

Although the Gx reference point is defined within the 3GPP standardization body, its applicability has also spread to wire-line operations to achieve mobile–fixed convergence gains by streamlining policy management functions into a single Gx based infrastructure, see Figure 38: Convergence.

Figure 38: Convergence

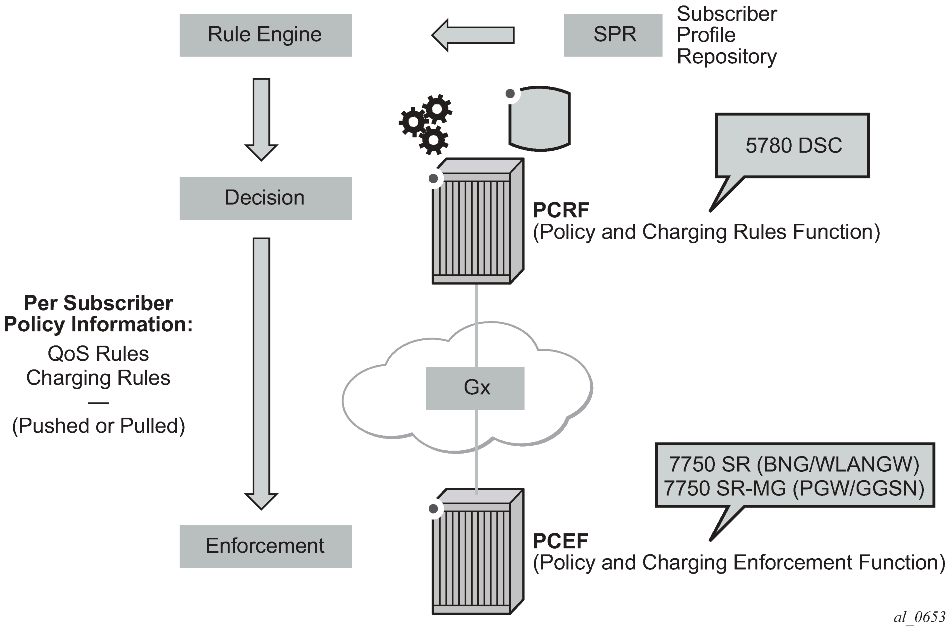

PCRF
AAA
Gx
RADIUS
PGW
BNG
Policy decision seperated
PCRF/Gx and AAA/RADIUS
PCRF
Gx
Gx
PGW
BNG
One common place (PCRF) and one
common method (Gx) for policy decision
al_0652

Gx support on SR OS is applicable to Enhanced Subscriber Management (ESM) functions, including the Application Assurance (AA) functions. The focus of this chapter is on the AA aspects of Gx.

The SR OS based Gx interface offers the following functionalities:

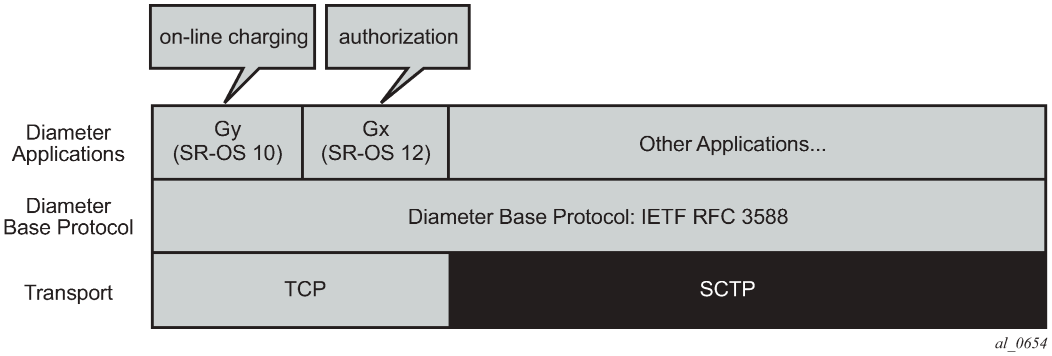
- ESM subscriber-based policy decision providing
  - QoS attributes
  - charging attributes
  - subscriber identification
- Usage management
  - usage reporting from PCEF to PCRF

Figure 39: Gx reference point



Note that Gx does not provide subscriber authentication or subscriber IP address assignment.

Figure 40: Diameter protocol stack



Policy assignment use case

The SR OS accepts the following policy information from PCRF using Gx:

- Subscriber profile strings and SLA profile strings.
- Subscriber-QoS-overrides.
- Application profile strings.
- Application subscriber options (ASOs) related to AA.

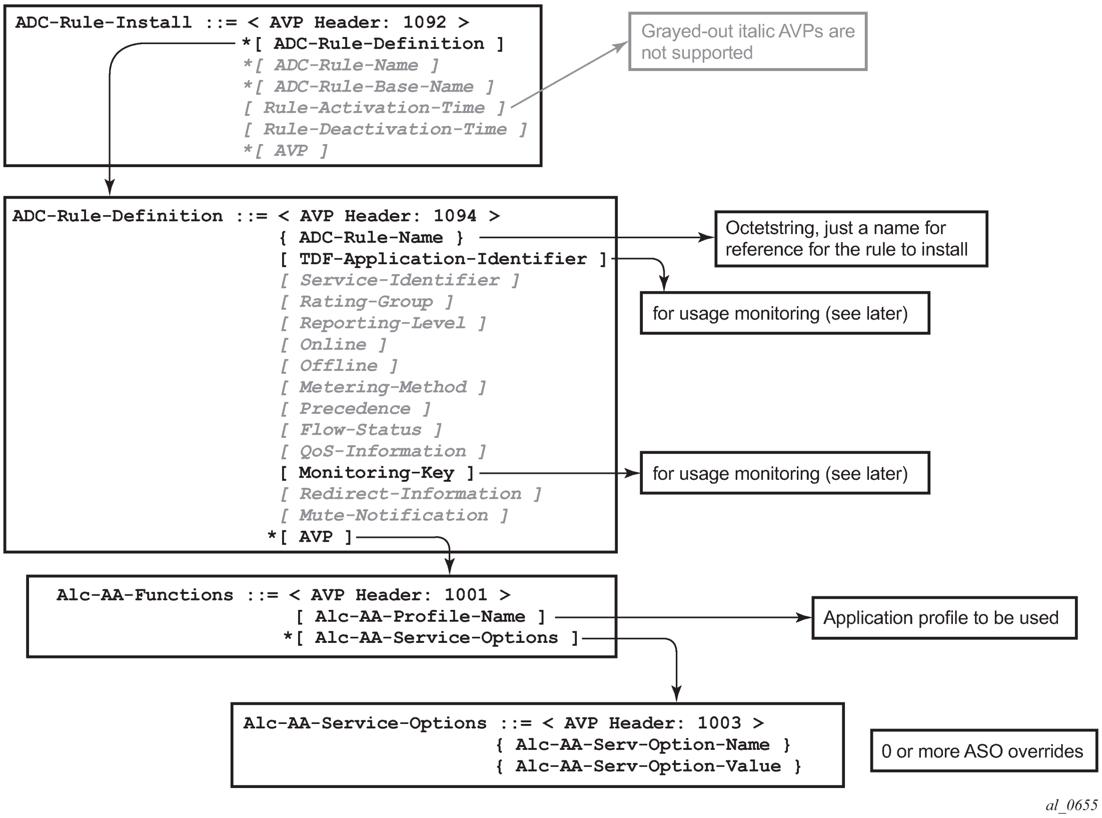
Gx operates at subscriber host level and creates an "IP-CAN Session" (IP Connectivity Access Network) for every subscriber host. However, as AA operates at the subscriber level, AA related policies apply to all the hosts belonging to that subscriber.

This chapter covers AA related functionalities, namely: application profile and ASO assignments and override. These functionalities are defined in either:

1. Application Detection and Control (ADC) rules—per 3GPP Release 11— **or**
2. Policy and Charging Control (PCC) rules—per 3GPP Release 12—.
   - **Application Profile** Alc-AA-Profile-Name Attribute-Value-Pair (AVP)
     - RADIUS equivalent is Alc-App-Prof-Str Vendor-Specific-Attribute (VSA)
   - **ASO overrides** Alc-AA-Service-Options AVP
     - RADIUS equivalent is Alc-AA-App-Service-Options VSA

Details of the ADC rules and related Nokia defined AVPs defined for use by AA are shown in Figure 41: ADC rules and related Nokia-defined AVPs defined for use by AA.

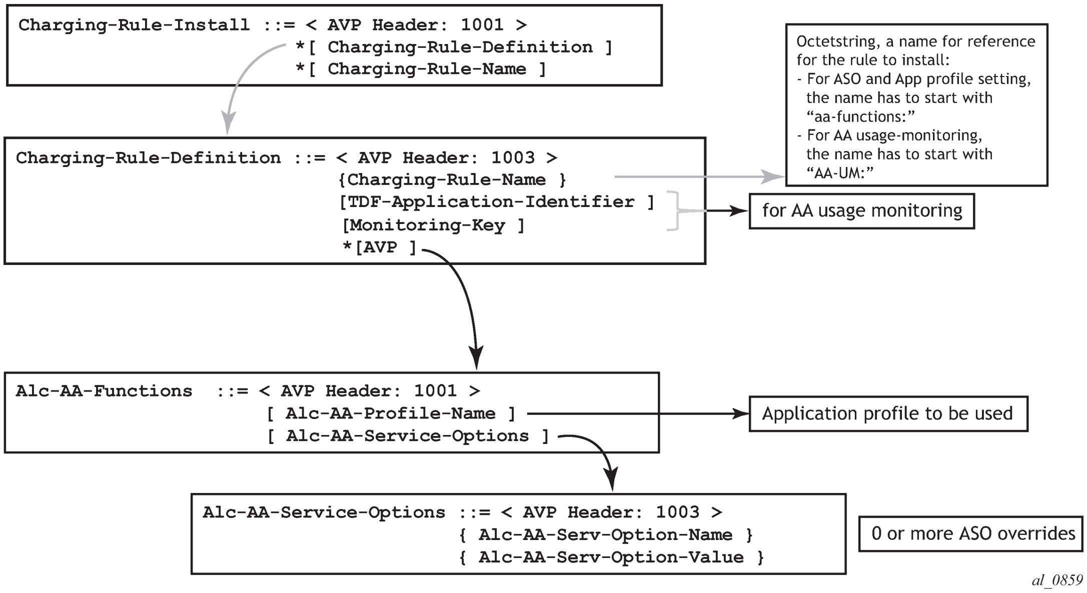
Figure 41: ADC rules and related Nokia-defined AVPs defined for use by AA


ADC-Rule-Install ::= < AVP Header: 1092 >
*[ADC-Rule-Definition]
*[ADC-Rule-Name]
*[ADC-Rule-Base-Name]
[Rule-Activation-Time]
[Rule-Deactivation-Time]
*[AVP]
Grayed-out italic AVPs are not supported
ADC-Rule-Definition ::= < AVP Header: 1094 >
{ ADC-Rule-Name }
[TDF-Application-Identifier]
[Service-Identifier]
[Rating-Group]
[Reporting-Level]
[Online]
[Offline]
[Metering-Method]
[Precedence]
[Flow-Status]
[QoS-Information]
[Monitoring-Key]
[Redirect-Information]
[Mute-Notification]
*[AVP]
Octetstring, just a name for reference for the rule to install
for usage monitoring (see later)
for usage monitoring (see later)
Alc-AA-Functions ::= < AVP Header: 1001 >
[Alc-AA-Profile-Name]
*[Alc-AA-Service-Options]
Application profile to be used
Alc-AA-Service-Options ::= < AVP Header: 1003 >
{ Alc-AA-Serv-Option-Name }
{ Alc-AA-Serv-Option-Value }
0 or more ASO overrides
al_0655

The ADC-Rule-Install is at the root level of the GX message.

As for 3GPP Release 12, the details of the PCC rules and related Nokia-defined AVPs defined for use by AA are shown in Figure 42: PCC rules and related Nokia-defined AVPs defined for use by AA.

Figure 42: PCC rules and related Nokia-defined AVPs defined for use by AA



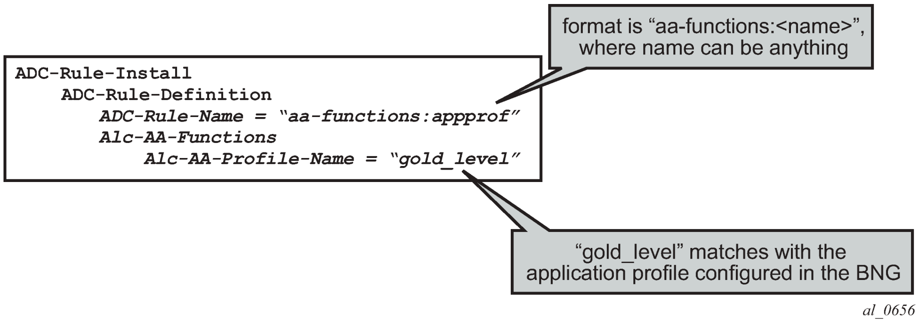
```
Charging-Rule-Install ::= < AVP Header: 1001 >
                         *[ Charging-Rule-Definition ]
                         *[ Charging-Rule-Name ]

Charging-Rule-Definition ::= < AVP Header: 1003 >
                             {Charging-Rule-Name }
                             [TDF-Application-Identifier ]
                             [Monitoring-Key ]
                             *[AVP ]

Alc-AA-Functions  ::= < AVP Header: 1001 >
                      [ Alc-AA-Profile-Name ]
                      [ Alc-AA-Service-Options ]

Alc-AA-Service-Options ::= < AVP Header: 1003 >
                           { Alc-AA-Serv-Option-Name }
                           { Alc-AA-Serv-Option-Value }
```

Octetstring, a name for reference for the rule to install:
- For ASO and App profile setting, the name has to start with "aa-functions:"
- For AA usage-monitoring, the name has to start with "AA-UM:"
for AA usage monitoring
Application profile to be used
0 or more ASO overrides
al_0859

The PCC-Rule-Install, as in the case of ADC-Rule-Install, is at the root level of the GX message.

An example of the AVPs to install the application profile "gold_level" using 3GPP Release 11 (/ADC rules) is shown in Figure 42: PCC rules and related Nokia-defined AVPs defined for use by AA.

Figure 43: ADC rule example of AVPs to install the application profile "gold_level"

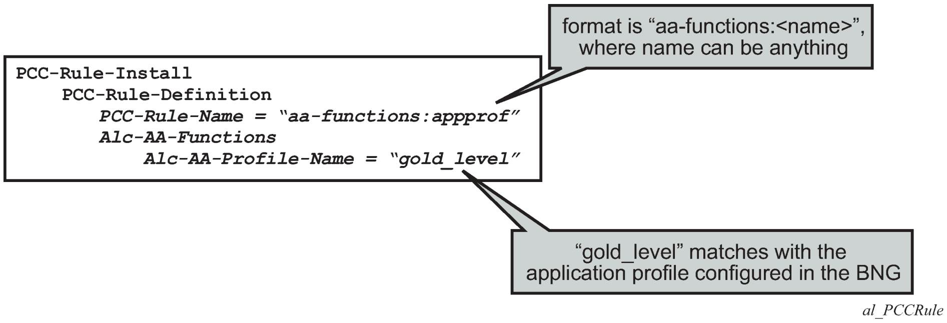


```
ADC-Rule-Install
    ADC-Rule-Definition
       ADC-Rule-Name = "aa-functions:appprof"
       Alc-AA-Functions
           Alc-AA-Profile-Name = "gold_level"
```

format is "aa-functions:<name>", where name can be anything
"gold_level" matches with the application profile configured in the BNG
al_0656

An example of the AVPs to install the "gold_level" application profile using 3GPP Release 12 (/PCC rules) is shown in Figure 45: Capture of the ADC rule assignment of the "gold_level" appProfile.

Figure 44: PCC rule example of AVPs to install the application profile “gold_level”



```
PCC-Rule-Install
    PCC-Rule-Definition
        PCC-Rule-Name = “aa-functions:appprof”
        Alc-AA-Functions
            Alc-AA-Profile-Name = “gold_level”
```

format is “aa-functions:<name>”, where name can be anything

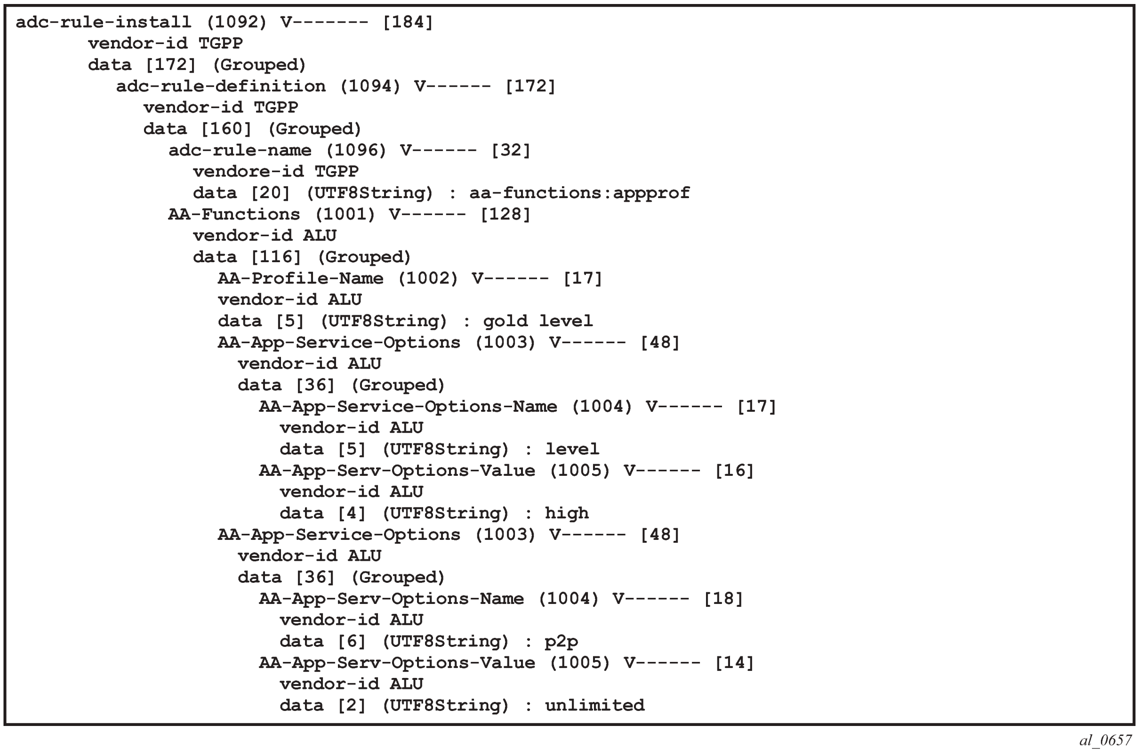
“gold_level” matches with the application profile configured in the BNG

al_PCCRule

Note:
ADC-Rule-Names and PCC-Rule-Names have to start with *aa-functions* when they contain an Alc-AA-Functions AVP.

The assignment of the *gold_level* appProfile is shown in another format in Figure 45: Capture of the ADC rule assignment of the “gold_level” appProfile.

Figure 45: Capture of the ADC rule assignment of the “gold_level” appProfile



```
adc-rule-install (1092) V------- [184]
      vendor-id TGPP
      data [172] (Grouped)
        adc-rule-definition (1094) V------ [172]
          vendor-id TGPP
          data [160] (Grouped)
            adc-rule-name (1096) V------ [32]
              vendore-id TGPP
              data [20] (UTF8String) : aa-functions:appprof
            AA-Functions (1001) V------ [128]
              vendor-id ALU
              data [116] (Grouped)
                AA-Profile-Name (1002) V------ [17]
                vendor-id ALU
                data [5] (UTF8String) : gold level
                AA-App-Service-Options (1003) V------ [48]
                  vendor-id ALU
                  data [36] (Grouped)
                    AA-App-Service-Options-Name (1004) V------ [17]
                      vendor-id ALU
                      data [5] (UTF8String) : level
                    AA-App-Serv-Options-Value (1005) V------ [16]
                      vendor-id ALU
                      data [4] (UTF8String) : high
                AA-App-Service-Options (1003) V------ [48]
                  vendor-id ALU
                  data [36] (Grouped)
                    AA-App-Serv-Options-Name (1004) V------ [18]
                      vendor-id ALU
                      data [6] (UTF8String) : p2p
                    AA-App-Serv-Options-Value (1005) V------ [14]
                      vendor-id ALU
                      data [2] (UTF8String) : unlimited
```

al_0657

Application profiles and ASO overrides can be changed on-the-fly with a Re-Authentication-Request (RAR) message according to these rules:

- If an Application profile is present in the Gx message it is applied first. Then ASO AVPs are applied when present (in the Gx message). In other words:

- If a RAR message only contains the same application profile and no ASO overrides, then all previous ASO overrides are removed.
- When a RAR message contains the same application profile and new ASO overrides, then the new ASO overrides are applied, and the previous ASO overrides are removed.
- When a RAR message contains a new application profile, all previous ASO overrides are removed and replaced with the ASOs in the RAR if present.
- When a RAR message does not contain an application profile but only ASO overrides, then the new ASO overrides are added to the existing ASO overrides.

Note that a single Gx ADC (or PCC) rule cannot contain both AA subscriber policies (appProfile/ASO) and AA Usage monitoring (as outlined later). These have to be in separate ADC (or PCC) rules.

Usage management/monitoring use-case

The AA-ISA can monitor application usage at the subscriber level and report back to the PCRF whenever the usage exceeds the threshold(s) set by the PCRF when receiving requests from the PCRF over the Gx interface.

Usage monitoring can be used by operators to report to PCRF when:

- The AA-ISA detects the start of a subscriber application by setting the usage threshold to a very low value.
- A pre-set usage volume per subscriber application is exceeded.

AA can monitor subscriber’s traffic for any defined:

- Application
- Application group, and/or
- Charging group

The AA-ISA reports the accumulated usage when:

- A usage threshold is reached.
- The PCRF explicitly disables the usage monitoring.
- The PCRF requests a report.
- The ADC (or PCC) rule associated with the monitoring instance is removed or deactivated.
- A session is terminated.

An AA defined application, application group and/or charging group is automatically allowed to be referenced by an ADC (or PCC) rule for the purpose of usage monitoring only if:

{It is already selected for either XML or RADIUS per subscriber accounting

OR

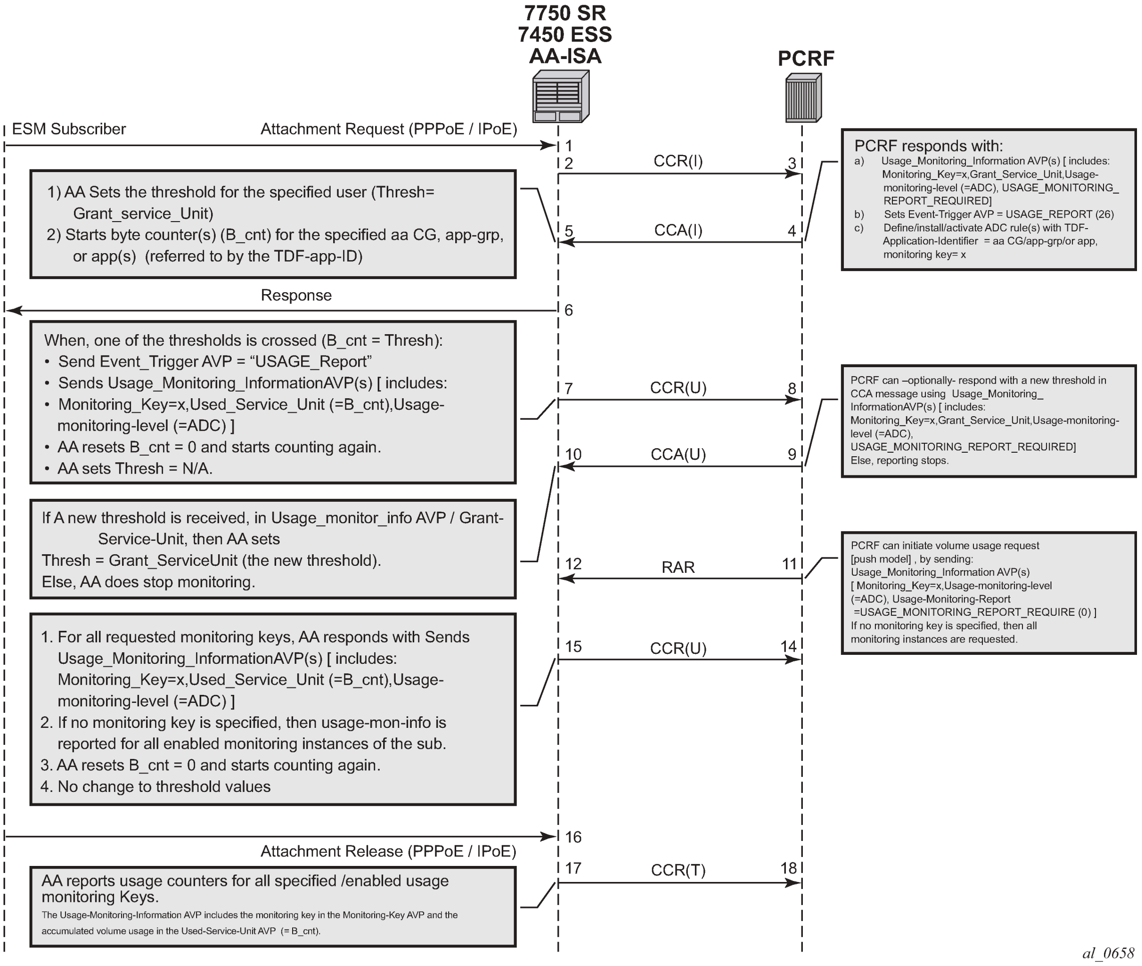
It is explicitly enabled by the operator for per subscriber statistics collection}

AND

Usage monitoring is enabled for the given AA group:partition

Figure 46: Call flow diagram illustrates the different messaging/call flows involved in application level usage monitoring. Details of the different supported AVPs used in these messages are illustrated later.

Figure 46: Call flow diagram


7750 SR
7450 ESS
AA-ISA
PCRF
ESM Subscriber
Attachment Request (PPPoE / IPoE)
1
2
CCR(I)
3
PCRF responds with:
a) Usage_Monitoring_Information AVP(s) [includes: Monitoring_Key=x,Grant_Service_Unit,Usage-monitoring-level (=ADC), USAGE_MONITORING_REPORT_REQUIRED]
b) Sets Event-Trigger AVP = USAGE_REPORT (26)
c) Define/install/activate ADC rule(s) with TDF-Application-Identifier = aa CG/app-grp/or app, monitoring key= x
1) AA Sets the threshold for the specified user (Thresh= Grant_service_Unit)
2) Starts byte counter(s) (B_cnt) for the specified aa CG, app-grp, or app(s) (referred to by the TDF-app-ID)
5
CCA(I)
4
Response
6
When, one of the thresholds is crossed (B_cnt = Thresh):
• Send Event_Trigger AVP = “USAGE_Report”
• Sends Usage_Monitoring_InformationAVP(s) [includes:
• Monitoring_Key=x,Used_Service_Unit (=B_cnt),Usage-monitoring-level (=ADC)]
• AA resets B_cnt = 0 and starts counting again.
• AA sets Thresh = N/A.
7
CCR(U)
8
PCRF can –optionally- respond with a new threshold in CCA message using Usage_Monitoring_InformationAVP(s) [includes: Monitoring_Key=x,Grant_Service_Unit,Usage-monitoring-level (=ADC), USAGE_MONITORING_REPORT_REQUIRED]
Else, reporting stops.
10
CCA(U)
9
If A new threshold is received, in Usage_monitor_info AVP / Grant-Service-Unit, then AA sets
Thresh = Grant_ServiceUnit (the new threshold).
Else, AA does stop monitoring.
12
RAR
11
PCRF can initiate volume usage request [push model] , by sending: Usage_Monitoring_Information AVP(s) [Monitoring_Key=x,Usage-monitoring-level (=ADC), Usage-Monitoring-Report =USAGE_MONITORING_REPORT_REQUIRE (0)]
If no monitoring key is specified, then all monitoring instances are requested.
1. For all requested monitoring keys, AA responds with Sends Usage_Monitoring_InformationAVP(s) [includes: Monitoring_Key=x,Used_Service_Unit (=B_cnt),Usage-monitoring-level (=ADC)]
2. If no monitoring key is specified, then usage-mon-info is reported for all enabled monitoring instances of the sub.
3. AA resets B_cnt = 0 and starts counting again.
4. No change to threshold values
15
CCR(U)
14
16
Attachment Release (PPPoE / IPoE)
17
CCR(T)
18
AA reports usage counters for all specified /enabled usage monitoring Keys.
The Usage-Monitoring-Information AVP includes the monitoring key in the Monitoring-Key AVP and the accumulated volume usage in the Used-Service-Unit AVP (= B_cnt).
al_0658

The AA-ISA/PCEF supports Usage-Thresholds AVPs that refer to the thresholds (in bytes) at which point an event needs to be sent back to the PCRF, (see Figure 44: PCC rule example of AVPs to install the application profile “gold_level”).

Time based thresholds are not supported.

AA supports the “grant-service-unit” AVP using the following possible values (AVP):

- CC-Input-Octets AVP (code 412): from subscriber total byte count threshold.
- CC-Output-Octet AVP (code 414): to subscriber total byte count threshold.
- CC-Total-octets AVP (code 421): threshold of aggregate traffic (input and output byte counters).

As shown in Figure 46: Call flow diagram, (T=7), AA sends a Credit Control Request (CCR_ message) with a "USAGE_REPORT" Event-Trigger AVP to the PCRF when the usage counter reaches the configured usage monitoring threshold for a given subscriber (and given application group). AA counters are reset (to zero) when the monitoring threshold is reached (and an event is sent back to the PCRF). The counter(s) however does not stop counting newly arriving traffic. AA counters only include “admitted” packets. Any packets that were discarded by AA due to, for example, policing actions are not counted for usage-monitoring purposes.

The TDF-Application-Identifier AVP (within the ADC or PCC rule) refers to an AA Charging group, an AA application group or to an AA application. TDF-Application-Identifiers (for example, charging-groups) have to be manually entered at the PCRF to match the AA charging groups defined in the AA. If the TDF-Application-Identifier refers to a name that is used for both a charging group and an application (or an application group), AA monitors the charging group. In other words, the AA charging group has a higher precedence than the AA application group.

Gx usage monitoring AVP summary

For 3GPP Release 11 (using ADC rules), the following AVPs are used for AA-Usage monitoring:

```
ADC-Rule-Install ::= < AVP Header: 1092 >
                            *[ ADC-Rule-Definition ]
                            *[ ADC-Rule-Name ]


    ADC-Rule-Definition ::= < AVP Header: 1094 >
                            { ADC-Rule-Name }
                            [ TDF-Application-Identifier ]; AA app/app-grp/chrg-grp
                            [ Monitoring-Key  ];

    Usage-Monitoring-Information::= < AVP Header: 1067 >
                                   [ Monitoring-Key ]
                                0,2[ Granted-Service-Unit ]
                                        Granted-Service-Unit ::= < AVP Header: 431 >
                                   [ CC-Total-Octets ]
                                   [ CC-Input-Octets ]
                                   [ CC-Output-Octets ]

                            0,2[ Used-Service-Unit ]
    Used-Service-Unit ::= < AVP Header: 446 >
                            [ CC-Total-Octets ] ;
                            [ CC-Input-Octets ]
                            [ CC-Output-Octets ]

                                 [ Usage-Monitoring-Level ]
; ADC_RULE_LEVEL (2)

                                 [ Usage-Monitoring-Report ]
; immidiate report -- USAGE_MONITORING_REPORT_REQUIRED (0)

                                 [ Usage-Monitoring-Support ]
; to disable :  USAGE_MONITORING_DISABLED (0)
```

For 3GPP Release 12 (using PCC rules), the following AVPs are used for AA-Usage monitoring:

```
Charging-Rule-Install ::= < AVP Header: 1001 >
                               *[ Charging-Rule-Definition ]
                               *[ Charging-Rule-Name ]


    Charging-Rule-Definition ::= < AVP Header: 1003 >
                                  { Charging-Rule-Name } ;/ starts with “UM-AA:”
                        [ TDF-Application-Identifier ]; AA app/app-grp/chrg-grp
                        [ Monitoring-Key ];


    Usage-Monitoring-Information::= < AVP Header: 1067 >
```

```
                             [ Monitoring-Key ]
                          0,2[ Granted-Service-Unit ]
                                Granted-Service-Unit ::= < AVP Header: 431 >
                                 [ CC-Total-Octets ]
                                 [ CC-Input-Octets ]
                                 [ CC-Output-Octets ]

                          0,2[ Used-Service-Unit ]
    Used-Service-Unit ::= < AVP Header: 446 >
                            [ CC-Total-Octets ] ;
                            [ CC-Input-Octets ]
                            [ CC-Output-Octets ]

                             [ Usage-Monitoring-Level ]
; PCC_RULE_LEVEL (1)

                             [ Usage-Monitoring-Report ]
; immidiate report -- USAGE_MONITORING_REPORT_REQUIRED (0)

                             [ Usage-Monitoring-Support ]
; to disable :  USAGE_MONITORING_DISABLED (0)
```

Configuration

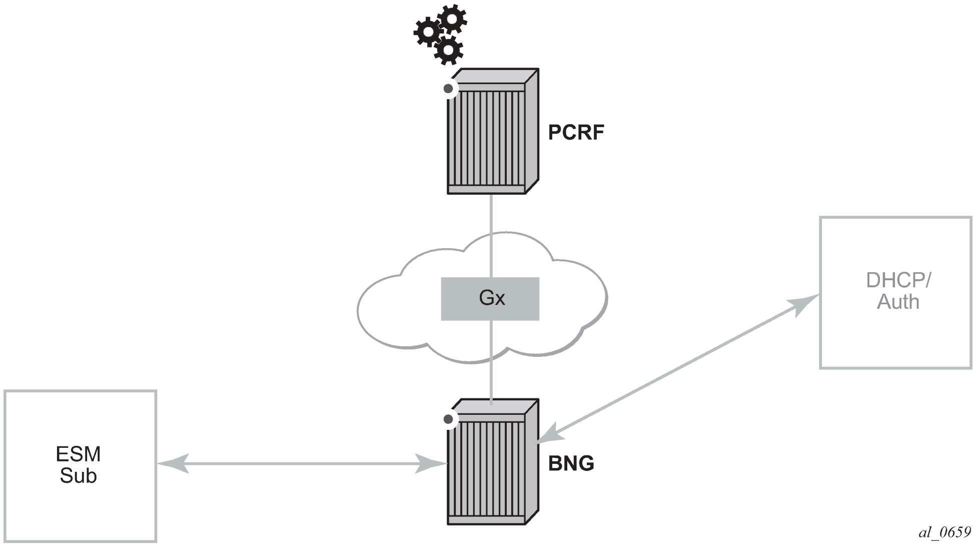
This configuration example highlights the commands illustrating how Gx can be used to:

- Override AppProfile and ASO characteristics.
- Set and retrieve AA level usage monitoring metrics.

While the configuration associated with setting up the Gx interface toward the PCRF is shown for the sake of completeness, that aspect of the configuration is not explored in detail, and only a 3GPP Release 11 model is shown Similarly, the Gx policies and usage monitoring associated with ESM host policies (non-AA aspects) are out of the scope of this chapter.

The configuration on the 7750 node is the same, independent of whether the PCRF supports 3GPP Release 11 (ADC) or Release 12(PCC) to provide AA policy control function.

Figure 47: Example configuration setup



The BNG is set up with at least one IOM and one MS-ISA MDA configured as ISA-AA.

```
configure
    card 1
        card-type iom3-xp
        mda 1
            mda-type m20-1gb-xp-sfp
            no shutdown
        exit
        mda 2
            mda-type isa-aa
            no shutdown
        exit
        no shutdown
    exit
    card 3
        card-type iom3-xp
        mda 1
            mda-type isa-aa
            no shutdown
        exit
        mda 2
            mda-type isa-aa
            no shutdown
        exit
        no shutdown
    exit
```

The configurations in this example are broken down into four main steps:

1. Configuring the Gx interface (high-level)
2. Configuring AA application profiles and ASOs (high-level)
3. Configuring AA applications filters (high-level)
4. Configuring AA usage-monitoring

The focus of this configuration example is on Step 4, and the updated show routines related to AA ESM subscriber state are shown at the end of Step 2.

1. Configuring the Gx interface (high-level).

These commands bring up the Gx diameter control channel between the Gx Controller(/Server), also known as PCRF, and the PCEF(/BNG).

```
configure
    aaa
        diameter
                node "bng-gx.realm-1.com" create
                    description "Authentication and Policy Management"
                    source-address 192.0.2.2
                    peer index 1 "dra-1.realm-1.com" create
                        address 10.1.0.10
                        no shutdown
                    exit
                exit
            exit
        exit
    exit
```

The diameter node "*bng-gx.realm-1.com*" is then referenced under subscriber management.

```
configure
    subscriber-mgmt
        diameter-application-policy "diamAppPlcy" create
            application gx
            diameter-node "bng-gx.realm-1.com" destination-realm "realm-1.com"
        exit
    exit
```

Then the created subscriber management policy "*diamAppPlcy*" is applied to the subscriber interface.

```
configure
    service
        customer 1 create
            description "Default customer"
        exit
        ies 1 customer 1 vpn 1 create
            description "Default Ies description for service id 1"
            subscriber-interface "ies-1-172.16.0.0" create
                address 172.16.0.0/12
                group-interface "grp-1-35782656-1" create
                    dhcp
                        server 172.16.200.200
                        trusted
                        lease-populate 2000
                        gi-address 172.16.0.0
                        no shutdown
                    exit
                    diameter-application-policy "diamAppPlcy"
                    sap 1/1/4:1 create
                        description "sap-grp-1"
                        sub-sla-mgmt
                            def-sub-profile "sub_prof"
                            def-sla-profile "sla_prof"
                            def-app-profile "app_prof_1"
                            sub-ident-policy "sub_ident_A_1"
                            multi-sub-sap 2
                            no shutdown
                        exit
                    exit
                exit
            exit
            service-name "ACG Ies 1"
            no shutdown
        exit
```

Now verify the configuration and connectivity towards the PCRF by running the following command:

```
# show aaa diameter-node "bng-gx.realm-1.com" peers

===============================================================================
Peers
===============================================================================
Host identity                          Status        Default Preference Active
-------------------------------------------------------------------------------
dra-1.realm-1.com                      I-Open        No      50         Yes
-------------------------------------------------------------------------------
No. of peers: 1
===============================================================================
```

The Peer-State-Machine State (PSM), as per RFC 6733, has the value **I**-OPEN indicating that the peer is operational. The "**I**-" stands for **I**nitiator state, in this case the BNG is the initiator.

A detailed look into the traffic statistics between the PCEF and the PCRF (Gx controller) can be viewed using a show statistics command (see below). These statistics provide a breakdown of the messages exchanged:

```
# show aaa diameter-node "bng-gx.realm-1.com" peer "dra-1.realm-1.com" statistics

===============================================================================
Peer "dra-1.realm-1.com"
===============================================================================
Message                              Sent                 Received
-------------------------------------------------------------------------------
Capabilities-Exchange-Request        7                    0
Capabilities-Exchange-Answer         0                    7
Disconnect-Peer-Request              0                    0
Disconnect-Peer-Answer               0                    0
Device-Watchdog-Request              1217                 778
Device-Watchdog-Answer               778                  1217
Application message request          0                    0
Application message answer           0                    0

Last cleared time: N/A
===============================================================================

# show subscriber-mgmt diameter-application-policy "diamAppPlcy" statistics

===============================================================================
Diameter node statistics for policy "diamAppPlcy"
===============================================================================
Message                              Requests             Answers
-------------------------------------------------------------------------------
Initial Credit-Control               2                    2
Update Credit-Control                14                   14
Termination Credit-Control           1                    1
Re-Auth                              2                    2
Abort-Session                        0                    0
```

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

Request message transmission failur* 0
Request message retransmissions      0

Result code                          Sent                 Received
-------------------------------------------------------------------------------
(1xxx) Informational                 0                    0
(2xxx) Success                       0                    14
(3xxx) Protocol Errors               0                    0
(4xxx) Transient Failures            0                    0
(5xxx) Permanent Failures            0                    0
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

2. Configuring AA application profiles and ASOs (high-level)

   To illustrate the use of application profiles and ASO overrides using Gx RAR messages, four ASOs and 2 appProfiles are defined, as follows.

   “app_prof_1” is the default app-profile used when a subscriber is created on AA.

```
configure
    application-assurance
        group 129:34883 create
            policy
                begin
                app-service-options
                    characteristic "permitDNS" persist-id 1 create
                        value "no"
                        value "yes"
                        default-value "yes"
                    exit
                    characteristic "permitRDP" persist-id 2 create
                        value "no"
                        value "yes"
                        default-value "yes"
                    exit
                    characteristic "permitHTTP" persist-id 3 create
                        value "no"
                        value "yes"
                        default-value "yes"
                    exit
                exit
                app-profile "app_prof_1" create
                    description "Application Profile Id app_prof_1"
                    divert
                exit
                app-profile "app_prof_2" create
                    description "Application Profile Id app_prof_2"
                    divert
                exit
```

3. Configuring AA applications filters (high-level)

   First create the application group, as follows.

```
configure
    isa
        application-assurance-group 129 create
            primary 3/2
            backup 1/2
            partitions
            divert-fc be
            no shutdown
        exit
```

Then create the partition and associated charging groups, application groups, applications, etc.

```
configure
    application-assurance
        group 129:34883 create
            policy
                begin
                charging-group "0_rated" create
                    export-id 1
                exit
                charging-group "default_charge_group" create
                    export-id 255
                exit
                default-charging-group "default_charge_group"
                app-group "Other" create
                    export-id 8
                exit
                app-group "Peer to Peer" create
                    export-id 3
                exit
                app-group "Remote Connectivity" create
                    export-id 4
                exit
                app-group "Unknown"
                    charging-group "0_rated"
                    export-id 1
                exit
                app-group "Web" create
                    export-id 10
                exit
                application "DNS" create
                    description "default-description for application DNS"
                    app-group "Other"
                    export-id 12
                exit
                application "BitTorrent" create
                    app-group "Peer to Peer"
                    export-id 3
                exit
                application "HTTP" create
                    description "default-description for application HTTP"
                    app-group "Web"
                    export-id 26
                exit
                application "RDP" create
                    description "default-description for application RDP"
                    app-group "Remote Connectivity"
                    export-id 61
                exit
                application "Unknown"
                    charging-group "0_rated"
                    export-id 1
                exit
            exit
            commit
        exit
    exit
exit
```

Example app-filter definitions defining HTTP, DNS, Bittorrent and RDP applications are as follows.

```
configure
    application-assurance
        group 129:34883
            policy
                begin
                app-filter
                    entry 6 create
                        description "default-description for AppFilter entry 6"
                        protocol eq "rdp"
                        ip-protocol-num eq tcp
                        application "RDP"
                        no shutdown
                    exit
                    entry 9 create
                        description "default-description for AppFilter entry 9"
                        protocol eq "dns"
                        ip-protocol-num eq udp
                        server-port eq range 53 55
                        application "DNS"
                        no shutdown
                    exit
                    entry 20 create
                        description "default-description for AppFilter entry 20"
                        protocol eq "bittorrent"
                        ip-protocol-num eq tcp
                        application "BitTorrent"
                        no shutdown
                    exit
                    entry 38 create
                        description "default-description for AppFilter entry 38"
                        protocol eq "http"
                        ip-protocol-num eq tcp
                        server-port gt 8738
                        application "HTTP"
                        no shutdown
                    exit
                exit
                commit
            exit
        exit
    exit
```

Note:
The focus of this example is on the definition of app-filters and/or AQPs. These are listed above (and below) for illustration purposes. The "sample" AQP configurations and app-filters shown here should not be used in a real-life configuration. Their configuration should follow the information in Application Assurance — Application Identification and User-Defined Applications.

Example AQP configurations for blocking DNS, RDP and HTTP traffic are as follows.

```
configure
    application-assurance
        group 129:34883
            policy
                begin
                app-qos-policy
                    entry 2 create
```

```
                        match
                            application eq "DNS"
                            characteristic "permitDNS" eq "no"
                        exit
                        action
                            drop
                        exit
                        no shutdown
                    exit
                    entry 3 create
                        match
                            application eq "HTTP"
                            characteristic "permitHTTP" eq "no"
                            ip-protocol-num neq 0
                        exit
                        action
                            drop
                        exit
                        no shutdown
                    exit
                    entry 4 create
                        match
                            application eq "RDP"
                            app-group eq "Remote Connectivity"
                            characteristic "permitRDP" eq "no"
                            ip-protocol-num neq udp
                        exit
                        action
                            drop
                        exit
                        no shutdown
                    exit
                exit
                commit
            exit
        exit
    exit
```

When an ESM subscriber is created, it is associated with the default AA app-profile, as seen using the show command below.

```
*A:BNG-1>show>app-assure>group# aa-sub esm "sub_172.16.0.2" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub_172.16.0.2 (esm)
ISA assigned            : 3/2
App-Profile             : app_prof_1
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : N/A
-------------------------------------------------------------------------------
Traffic                        Octets              Packets                Flows
-------------------------------------------------------------------------------
From subscriber:
  Admitted                          0                    0                    0
  Denied                            0                    0                    0
  Active flows                                                                0
To subscriber:
  Admitted                          0                    0                    0
  Denied                            0                    0                    0
  Active flows                                                                0
Flow counts:
  Terminated                                                                  0
  Short duration                                                              0
  Med duration                                                                0
  Long duration                                                               0
Total flow duration :  0 seconds
-------------------------------------------------------------------------------
Top App-Groups                              Octets        Packets        Flows
-------------------------------------------------------------------------------
None
-------------------------------------------------------------------------------
Application Service Options (ASO)
-------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
-------------------------------------------------------------------------------
permitDNS                        yes                              default
permitRDP                        yes                              default
permitHTTP                       yes                              default
===============================================================================
*A:BNG-1>show>app-assure>group#
```

After the PCRF sends out AppProfile and ASO override AVPs, using either PCC or ADC rules, in RAR messages (as shown below) it can be seen that the new parameters (new profile and new values for permitDNS and permitHTTP ASOs) are updated for that ESM subscriber.

Figure 48: PCRF AVPs override call flow diagram

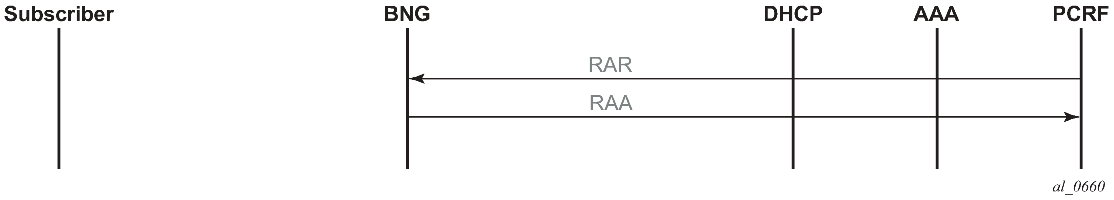
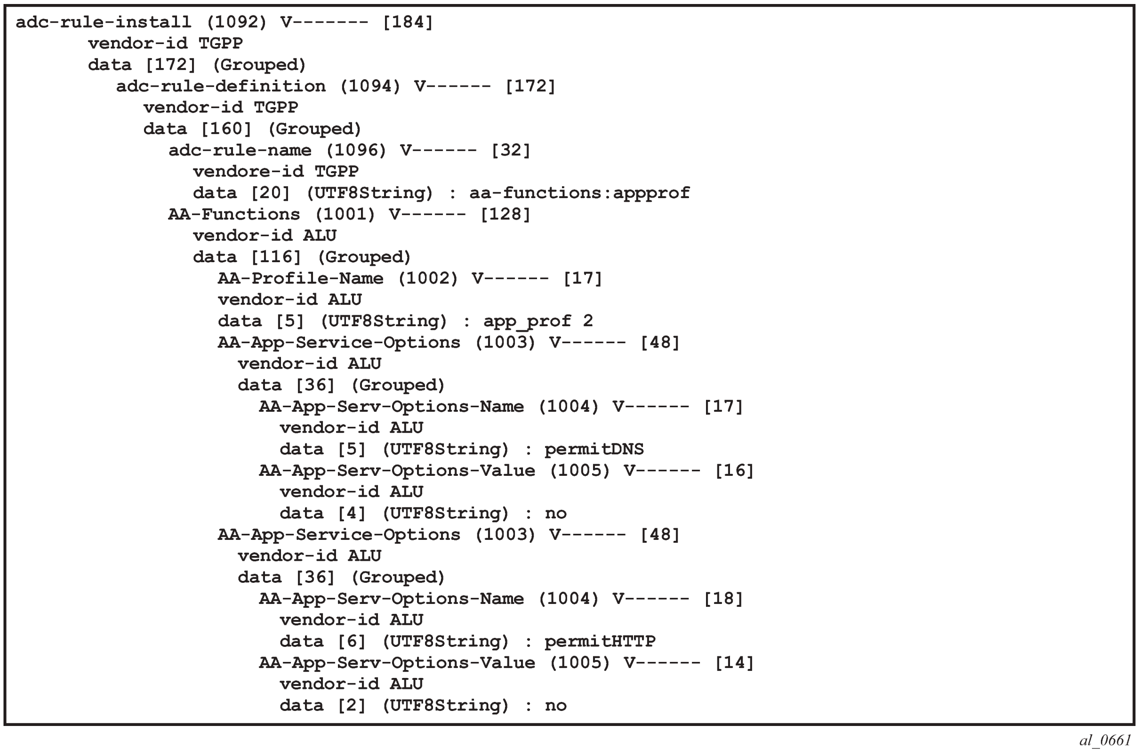


Figure 49: RAR containing ASOs and AppProfile override AVPs example



```
adc-rule-install (1092) V------- [184]
      vendor-id TGPP
      data [172] (Grouped)
        adc-rule-definition (1094) V------ [172]
          vendor-id TGPP
          data [160] (Grouped)
            adc-rule-name (1096) V------ [32]
              vendore-id TGPP
              data [20] (UTF8String) : aa-functions:appprof
            AA-Functions (1001) V------ [128]
              vendor-id ALU
              data [116] (Grouped)
                AA-Profile-Name (1002) V------ [17]
                vendor-id ALU
                data [5] (UTF8String) : app_prof 2
                AA-App-Service-Options (1003) V------ [48]
                  vendor-id ALU
                  data [36] (Grouped)
                    AA-App-Serv-Options-Name (1004) V------ [17]
                      vendor-id ALU
                      data [5] (UTF8String) : permitDNS
                    AA-App-Serv-Options-Value (1005) V------ [16]
                      vendor-id ALU
                      data [4] (UTF8String) : no
                AA-App-Service-Options (1003) V------ [48]
                  vendor-id ALU
                  data [36] (Grouped)
                    AA-App-Serv-Options-Name (1004) V------ [18]
                      vendor-id ALU
                      data [6] (UTF8String) : permitHTTP
                    AA-App-Serv-Options-Value (1005) V------ [14]
                      vendor-id ALU
                      data [2] (UTF8String) : no
```

```
*A:BNG-1>show>app-assure>group# aa-sub esm "sub_172.16.0.2" summary
===============================================================================
Application-Assurance Subscriber Summary (realtime)
===============================================================================
AA-Subscriber           : sub_172.16.0.2 (esm)
ISA assigned            : 3/2
App-Profile             : app_prof_2
App-Profile divert      : Yes
Capacity cost           : 1
Aarp Instance Id        : N/A
HTTP URL Parameters     : (Not Specified)
Last HTTP Notified Time : N/A
-------------------------------------------------------------------------------
Traffic                        Octets             Packets               Flows
-------------------------------------------------------------------------------
From subscriber:
  Admitted                          0                   0                   0
  Denied                            0                   0                   0
  Active flows                                                              0
To subscriber:
  Admitted                          0                   0                   0
  Denied                            0                   0                   0
  Active flows                                                              0
Flow counts:
  Terminated                                                                0
  Short duration                                                            0
  Med duration                                                              0
  Long duration                                                             0
Total flow duration :  0 seconds
```

```
-------------------------------------------------------------------------------
Top App-Groups                              Octets        Packets        Flows
-------------------------------------------------------------------------------
None

-------------------------------------------------------------------------------
Application Service Options (ASO)
-------------------------------------------------------------------------------
Characteristic                   Value                            Derived from
-------------------------------------------------------------------------------
permitDNS                        no                               dyn-override
permitRDP                        yes                              default
permitHTTP                       no                               dyn-override
===============================================================================
```

4. Configuring AA usage monitoring

Once the applications, application groups and/or charging groups are defined and configured (see previous steps), the operator needs:

- to enable the collection of per-subscriber statistics so they can be used for Gx based usage-monitoring. This step is not needed for any app/appgrp or charging group that is already enabled for per-subscriber statistics. In other words, if XML or RADIUS accounting is enabled for a given app/appgrp or charging group, then Gx usage-monitoring is also automatically enabled.
- to enable usage-monitoring for the given AA group:partition.

```
configure
    application-assurance
        group 129:34883
            statistics
                aa-sub
                    usage-monitoring
                    app-group "Unknown" export-using accounting-policy
                                                   radius-accounting-policy
                    charging-group "0_rated" export-using accounting-policy
                                                   radius-accounting-policy
                    charging-group "default_charge_group" export-using
                                                          accounting-policy
                    radius-accounting-policy
                    application "BitTorrent" no-export
                exit
```

In the preceding example:

- The usage-monitoring command is used to enable Gx usage monitoring for the specified AA partition.
- The aa-group and charging-group commands specify which charging groups and AA groups are selected for export. In this case *0-rated*, *Unknown*, and *default_charge_group* are selected for RADIUS accounting and they automatically qualify for Gx-usage monitoring.
- The BitTorrent application however needs to be explicitly configured as "no-export" as it needs to be enabled for Gx-usage monitoring.

The operator can display the number of usage monitoring rules for a given subscriber. This is shown below after the ESM subscriber is created, but before any ADC rules are installed for usage-monitoring by PCRF, so AA reports that no rules apply ("0").

```
*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor status
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
```

```
Usage Monitor Status
===============================================================================
Type             Name                            Oper Status
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
No. of rules: 0
===============================================================================
*A:BNG-1>show>app-assure>group#
```

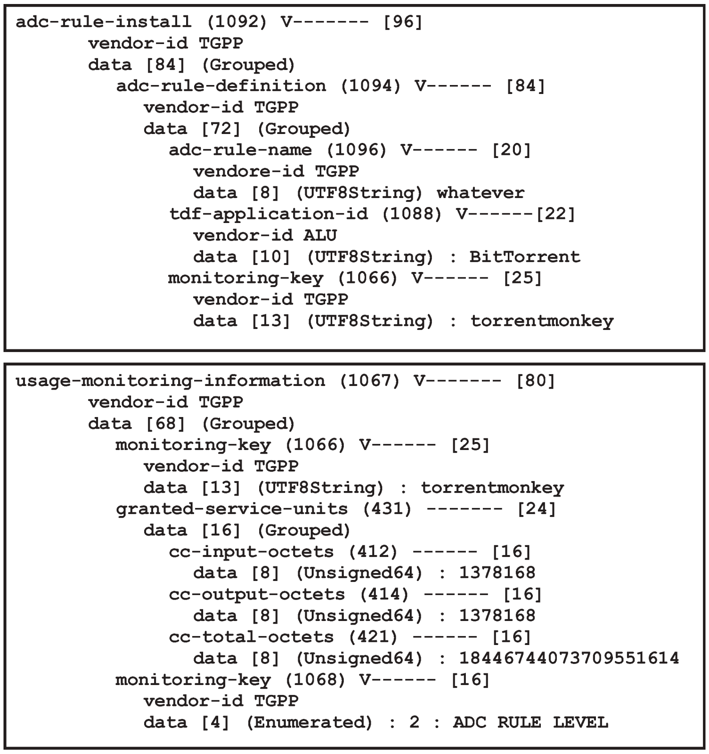
The PCRF then sends a RAR message with a usage monitoring ADC or PCC rule for the BitTorrent application to set the usage thresholds for BitTorrent for the ESM subscriber "alcatel_A_1" to (in bytes):

Input (from sub) 1378168

Output (to sub) 1381148

Total traffic (up and down) 18446744073709551614

Figure 50: RAR containing usage monitoring ADC rules example



```
adc-rule-install (1092) V------- [96]
      vendor-id TGPP
      data [84] (Grouped)
        adc-rule-definition (1094) V------ [84]
          vendor-id TGPP
          data [72] (Grouped)
            adc-rule-name (1096) V------ [20]
              vendore-id TGPP
              data [8] (UTF8String) whatever
            tdf-application-id (1088) V------[22]
              vendor-id ALU
              data [10] (UTF8String) : BitTorrent
            monitoring-key (1066) V------ [25]
              vendor-id TGPP
              data [13] (UTF8String) : torrentmonkey
```

```
usage-monitoring-information (1067) V------- [80]
      vendor-id TGPP
      data [68] (Grouped)
        monitoring-key (1066) V------ [25]
          vendor-id TGPP
          data [13] (UTF8String) : torrentmonkey
        granted-service-units (431) ------- [24]
          data [16] (Grouped)
            cc-input-octets (412) ------ [16]
              data [8] (Unsigned64) : 1378168
            cc-output-octets (414) ------ [16]
              data [8] (Unsigned64) : 1378168
            cc-total-octets (421) ------ [16]
              data [8] (Unsigned64) : 18446744073709551614
        monitoring-key (1068) V------ [16]
          vendor-id TGPP
          data [4] (Enumerated) : 2 : ADC RULE LEVEL
```

al_0662

This is then reflected on the AA-ISA:

```
*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor status
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
Usage Monitor Status
===============================================================================
Type             Name                            Oper Status
-------------------------------------------------------------------------------
application      BitTorrent                      active
-------------------------------------------------------------------------------
No. of rules: 1
```

```
===============================================================================
*A:BNG-1>show>app-assure>group#
```

Note the “active” oper status is set since there is at least one usage monitoring threshold associated with this application.

Given that there is no traffic flowing yet to or from the subscriber the counters currently are “0”:

```

*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor count
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
Usage Monitor Credit Statistics
===============================================================================
Application: "BitTorrent"
Direction     Status                    Granted                 Used     % Used
-------------------------------------------------------------------------------
to sub        valid                     1378168                    0         0%
from sub      valid                     1381148                    0         0%
both          valid        18446744073709551614                    0         0%
===============================================================================
*A:BNG-1>show>app-assure>group#
```

The status is set to “valid” since a threshold (or Grant) is received.

When, at a later stage, traffic starts flowing again usage-monitor subscriber statistics are updated as shown below.

```
*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor count
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
Usage Monitor Credit Statistics
===============================================================================
Application: "BitTorrent"
Direction     Status                    Granted                 Used     % Used
-------------------------------------------------------------------------------
to sub        valid                     1378168               137816        10%
from sub      valid                     1381148                13781         1%
both          valid        18446744073709551614               151597         5%
===============================================================================
*A:BNG-1>show>app-assure>group#
```

The PCRF can also at the same time set ADC or PCC rules for other applications (such as the *0_rated* and the default_charging_group charging groups).

In the following case, the PCRF installs an ADC usage monitoring rule for:

- Charging group: “0-rated”, but without usage thresholds
- Charging group: “default_charge_group”, and sets only a threshold for “to sub” traffic.

This results in having a usage policy for the “0-rated” charging group installed but this is not active since there are no grants associated with it:

```
*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor status
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
Usage Monitor Status
===============================================================================
Type             Name                             Oper Status
-------------------------------------------------------------------------------
application      BitTorrent                       active
```

```
charging-group   0_rated                          inactive
charging-group   default_charge_group             active
-------------------------------------------------------------------------------
No. of rules: 3
===============================================================================
*A:BNG-1>show>app-assure>group#
```

Note that the "inactive" status for the "0-rated" charging group is due to no grants being received.

Moreover, detailed counters show:

```
*A:BNG-1>show>app-assure>group# aa-sub esm "alcatel_A_1" usage-monitor count
===============================================================================
Application-Assurance Subscriber "alcatel_A_1" (esm)
Usage Monitor Credit Statistics
===============================================================================
Application: "BitTorrent"
Direction     Status                    Granted                 Used     % Used
-------------------------------------------------------------------------------
to sub        valid                     1378168               137816        10%
from sub      valid                     1381148                13781         1%
both          valid        18446744073709551614               151597         5%
-------------------------------------------------------------------------------
Charging-Group: "0_rated"
Direction     Status                    Granted                 Used     % Used
-------------------------------------------------------------------------------
to sub        invalid                       n/a                    0        n/a
from sub      invalid                       n/a                    0        n/a
both          invalid                       n/a                    0        n/a
-------------------------------------------------------------------------------
Charging-Group: "default_charge_group"
Direction     Status                    Granted                 Used     % Used
-------------------------------------------------------------------------------
to sub        valid                     1000000              1378084       100%
from sub      invalid                       n/a                 1574        n/a
both          invalid                       n/a              1379658        n/a
===============================================================================
*A:BNG-1>show>app-assure>group#
```

Again, the "invalid" status above reflects the fact that no grants have been received.

Conclusion

The introduction of the diameter (/Gx) control feature on the SR-series BNG enables operators to consolidate policy management systems used in wire-line and wireless environments into a single system. This provides an increase in operational efficiency as mobile and fixed networks convergence gains more traction.

This example illustrates how policy control and usage monitoring of the SR-series BNG Application Assurance services can be achieved over standard 3GPP Diameter Gx protocol.

Deterministic Large Scale NAT44

This chapter provides information about deterministic large scale NAT44 configurations.

Topics in this chapter include:

Applicability

The information and configuration in this chapter was initially based on SR OS Release 11.0.R3, and is updated to SR OS Release 14.0.R4.

Overview

Deterministic Network Address Translation (NAT) is a mode of operation where mappings between the NAT subscriber and the outside IP address and port range are allocated at the time of configuration.

In deterministic NAT for Large Scale NAT IPv4-to-IPv4 (LSN44) subscribers, each LSN44 subscriber is permanently mapped to an outside IP address and a dedicated (deterministic) port-block based on a specific algorithm.

Logging is not needed in this case because the reverse mapping can be obtained using the reverse of the preceding algorithm.

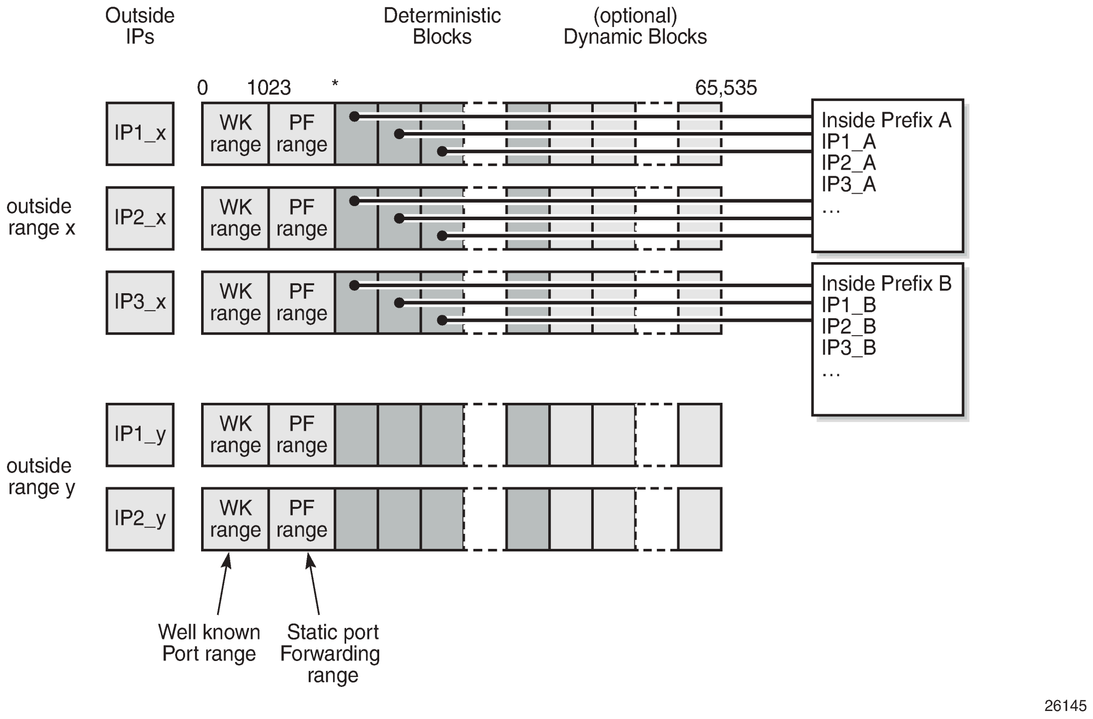
A deterministic LSN44 subscriber can have only one deterministic port-block that can (optionally) be extended by one or multiple dynamic port-blocks in case all ports in deterministic port-block are exhausted.

In case an LSN44 subscriber has been assigned both deterministic and dynamic port blocks, logging for the dynamic port-block allocation/de-allocation is required.

A scalable logging solution for dynamic port-blocks is achievable using RADIUS or IPFIX.

Logging for dynamic port-blocks is out of the scope of this chapter.

Figure 51: Deterministic NAT Mapping



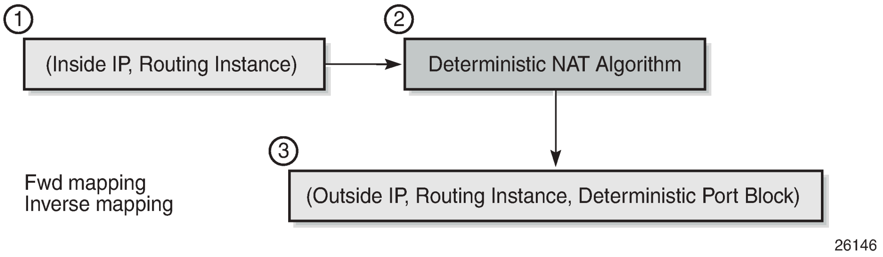
Algorithm

The deterministic NAT algorithm makes a predictable mapping between the (inside IP, routing instance) and the (outside IP, routing instance, deterministic port block).

The algorithm is revertive, meaning that a given (outside IP, routing instance, deterministic port block) will derive a given (inside IP, routing Instance).

The algorithm is loosely based on draft-donley-behave-deterministic-cgn-00.txt, which allows for the dynamic expansion of the port-blocks once the ports in the original deterministic port-block are exhausted.

Figure 52: Deterministic NAT Algorithm

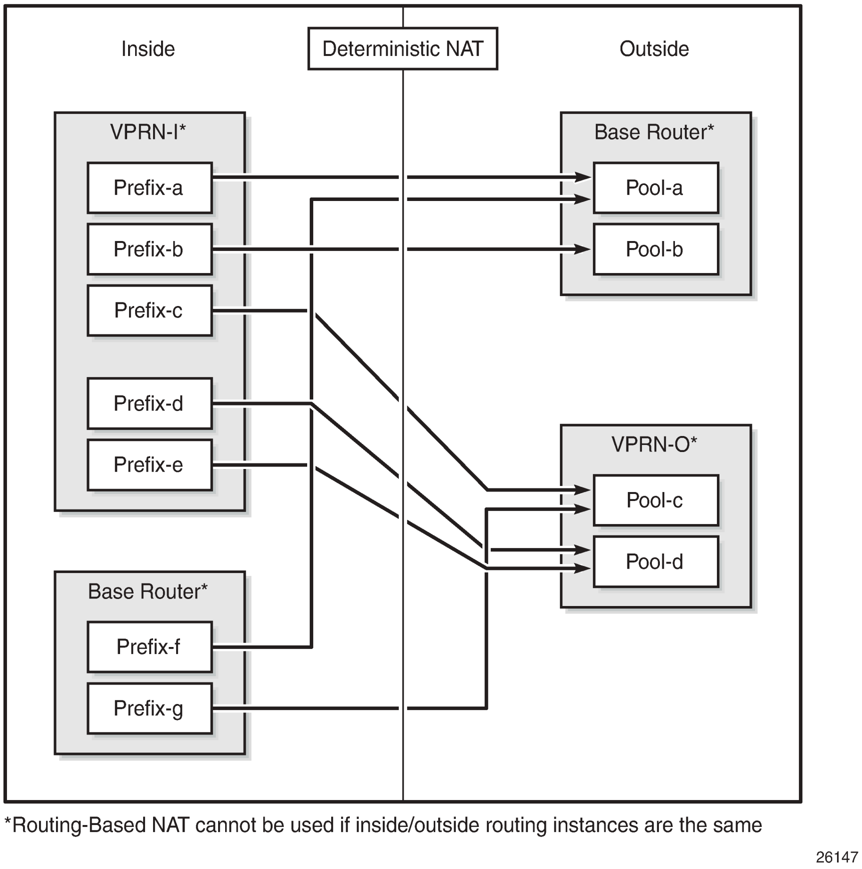

1
(Inside IP, Routing Instance)
2
Deterministic NAT Algorithm
3
(Outside IP, Routing Instance, Deterministic Port Block)
Fwd mapping
Inverse mapping
26146

Deterministic mapping

Any inside prefix in any routing instance can be mapped to any pool in any routing instance.

In deterministic NAT, prefixes from multiple routing instances can be mapped to the same outside pool, also prefixes from a single inside routing instance can be selectively mapped to different outside pools.

Figure 53: Deterministic Mapping: Inside -> Outside Routing Instances



Mapping rules

A deterministic LSN44 subscriber is mapped to only one deterministic block which can further be extended to multiple dynamic blocks if ports within the deterministic block are exhausted.

The subscriber-limit is the number of subscribers that can be deterministically mapped to one outside IP address (i.e. compression ratio) and **must** be a power of 2.

The total number of deterministic ports (DetP) per outside IP address is determined by this subscriber-limit and the number of deterministic ports per subscriber.

The remaining ports (DynP) beyond the deterministic port range up to 65535 will be dedicated for dynamic use when a deterministic block is exhausted.

Every host using an inside prefix is guaranteed one dedicate block in the deterministic port ranges.

If the inside prefix length is m < 32-n, where 2^n=subscriber-limit, then the prefix must be broken into pieces so that all hosts (subscriber-limit) in each piece maps exactly to one outside IP address.

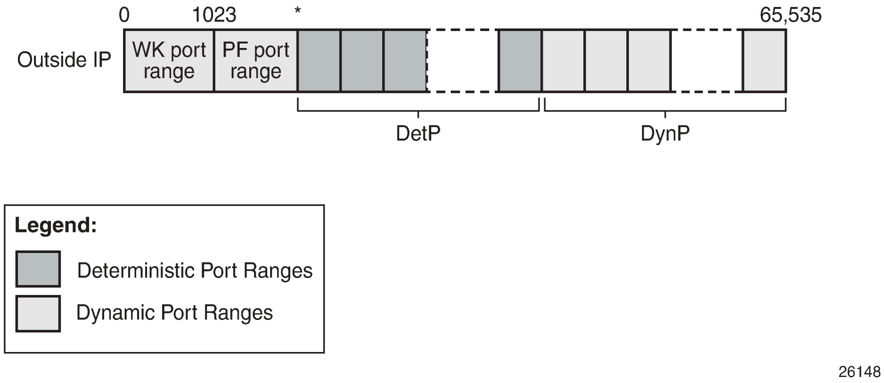
For example, if there is an inside prefix 192.168.0.0/23, with m=23 and a maximum number of 256 hosts; and the subscriber-limit set to 256, then n=8. This results in 23 < 24 (32-8) and so this inside prefix has to be broken into 2 pieces, in other words, this inside prefix will fit into 2 outside IP addresses, each of 256 port-blocks.

In case that the prefix length is m ≥ 32-n, where 2^n=subscriber-limit, then all hosts from the configured prefix are mapped to the same outside IP.

For example, if there is an inside prefix 192.168.1.0/25, with m=25 and a maximum number of 128 hosts, and the subscriber-limit set to 256, then n=8. This results in 25 > 24 (32-8), so definitely 128 hosts can fit in one outside IP because there are 256 available port-blocks, in other words, this inside prefix will fit into one outside IP where 128 blocks have been used out of the 256 port-blocks available, and the rest (256-128) are wasted.

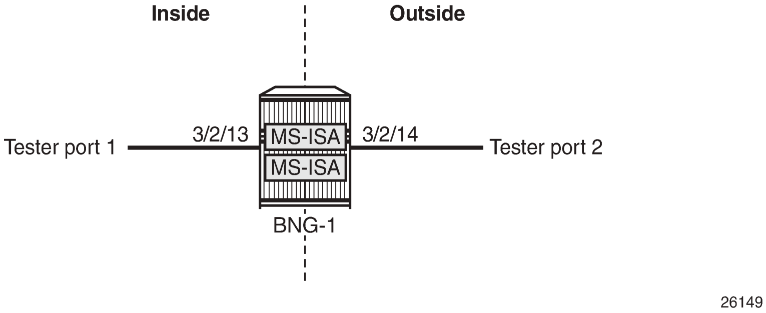
Overbooking of the outside address pool is not supported in deterministic NAT.

Figure 54: Deterministic mapping: outside IP port-blocks/ranges



Configuration

Figure 55: Example topology



Configuration prerequisites

Card and MDA configuration.

```
configure
    card 2
        card-type iom3-xp
        mda 1
            mda-type isa-bb
            no shutdown
        exit
        mda 2
            mda-type isa-bb
            no shutdown
        exit
        no shutdown
    exit
exit
```

Note:
Private address ranges are used in outside pools within this chapter but normally public address ranges would be used.

Create the NAT group, and add the MS-ISAs created above to the NAT group; up to 10 MS-ISAs of type isa-bb can be configured under the NAT group.

```
configure
    isa
        nat-group 1 create
            mda 2/1
            mda 2/2
            active-mda-limit 1
            no shutdown
        exit
    exit
exit
```

Configuration commands

A NAT **outside pool** is configured using the following command:

```
configure  {router | service vprn <service-id>}
  nat
    outside
      pool <nat-pool-name> [nat-group <nat-group-id> type <pool-type> create]
        port-reservation {blocks <num-blocks> | ports <num-ports>}
        port-forwarding-range <range-end>
        subscriber-limit <subscriber-limit>
        deterministic
          port-reservation <det-num-ports>
        exit
          address-range <start-ip-address> <end-ip-address> create
        exit
      exit
    exit
  exit
```

where:

nat-pool-name — Specifies the name of the NAT pool up to 32 characters max.

nat-group-id — Specifies the NAT group ID. The values are 1 — 4.

pool-type — Species the pool type (**large-scale**).

num-blocks — Specifies the number of dynamic port-blocks per outside IP address. The values are 1 — 64512

num-ports — Specifies the number of ports per dynamic block. The values are 1 — 32256

range-end — Specifies the upper limit of the port range available for static port forwarding. The values are 1023 — 65535

subscriber-limit — Specifies the maximum number of subscribers per outside IP address.

A power of 2 (2^n) number for deterministic NAT

[1,2,4,8,16,32,64,128,256,512,1024,2048, 4096, 8192,16348, 32768]

1..65535 for non-deterministic NAT

default: 65535 for non-deterministic

det-num-ports — Specifies the number of ports in a deterministic port block that is allocated and dedicated to a single subscriber during the configuration phase. The values are 1..65535

start-ip-address — Specifies the first outside IP address in the a.b.c.d format.

end-ip-address — Specifies the last outside IP address in the a.b.c.d format.

Note:

- When the subscriber limit equals 1, each subscriber is mapped to a single outside IP address, though the NAPT (port translation) function is still performed.
- 1:1 NAT mode in combination with deterministic NAT is not supported.

A NAT **policy** is configured using the following command:

```
configure service nat
 nat-policy <nat-policy-name> [create]
    block-limit <[1..40]>
    pool <nat-pool-name> {router <router-instance> | service-name <service-name>}
 exit
```

where:

nat-policy-name — Specifies the NAT policy name up to 32 characters max.

block-limit —The maximum number of deterministic plus dynamic port blocks that can be assigned to a single inside IP address. In other words, the maximum number of dynamic port blocks that can be assigned to an inside IP address when the deterministic port block is exhausted equals (block-limit - 1).

nat-pool-name — Specifies the NAT pool name up to 32 characters max.

router-instance — Specifies the router instance the pool belongs to, either by router name or service ID.

<router-name> | *<service-id>*

The router name values are *Base* or *service-id* [1..2147483647]

service-name — Specifies the name of the service up to 64 characters max.

A NAT **inside prefix** is configured using the following command:

```
configure [router| service vprn <service-id>]
    nat
        inside
            classic-lsn-max-subscriber-limit <max>
            deterministic
                prefix <ip-prefix/length> subscriber-type <nat-sub-type>
                nat-policy <nat-policy-name> create
                    map start <lsn-sub-address> end <lsn-sub-address> to <outside-ip-address>
                    no shutdown
                exit
            exit
        exit
    exit
```

where:

max — The power of 2 (2^n) number that must match the largest subscriber limit number in a deterministic pool referenced from this inside routing instance. The range for this command is the same as the subscriber-limit command under the pool hierarchy. The values are 1,2,4,8 — 32768

ip-prefix/length — A prefix on the inside encompassing subscribers that will be deterministically mapped to an outside IP address and port block in the corresponding pool.

<ip-prefix/ip-pref*>	<ipv4-prefix>/<ipv4-prefix-length> |
	<ipv6-prefix>/<ipv6-prefix-length>
<*ipv4-prefix*>	a.b.c.d (host bits must be 0)
<*ipv4-prefix-length*>	[0..32]
<*ipv6-prefix*>	x:x:x:x:x:x:x:x (eight 16-bit pieces)
	x:x:x:x:x:x:d.d.d.d
	x - [0..FFFF]H
	d - [0..255]D
<*ipv6-prefix-length*>	[0..128]
<*nat-sub-type*>:	classic-lsn-sub
<*nat-policy-name*>	Specifies a NAT policy name up to 32 characters in length.

Following rules apply to the **classic-lsn-max-subscriber-limit**:

- Should be greater than or equal to the largest subscriber-limit of all pools referenced by the NAT policies within the corresponding inside routing instance.
- Must be configured before any inside prefix configuration.
- Must be 2^n and affects the ingress hashing of deterministic subscribers and also non-deterministic subscribers in case both are configured under the same inside router instance.

Three cases are now configured to demonstrate the use of deterministic and dynamic port-block usage:

- Case 1: Mapping multiple prefixes from the same VRF (VPRN 15001) into the same outside pool, routing instance “Base”.
- Case 2: Mapping multiple prefixes from the same VRF (VPRN 15001) into different outside pools, routing instance VPRN 15002
- Case 3: Mapping overlapping prefixes from different VRFs (VPRN 15001 and VPRN 15002) into the same outside pool, routing instance “Base”.

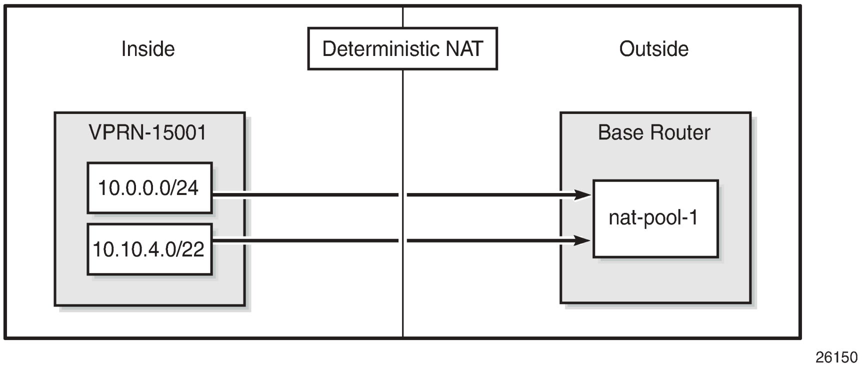
In each case all of the traffic is NATed.

Case 1

Configured with:

- Mapping multiple prefixes of the same VRF into the same outside pool.
- NAT all traffic.

Figure 56: Case 1



The NAT **outside pool** is configured as follows:

```
configure
    router
        nat
            outside
                pool "nat-pool-1" nat-group 1 type large-scale create
                    port-reservation ports 180
                    port-forwarding-range 4023
                    subscriber-limit 128
                    deterministic
                        port-reservation 300
                    exit
                    address-range 192.168.0.1 192.168.0.100 create
                    exit
                    no shutdown
                exit
            exit
        exit
    exit
exit
```

The NAT **policy** is configured as follows:

```
configure
    service
        nat
            nat-policy "nat-policy-1" create
                block-limit 4
                pool "nat-pool-1" router Base
            exit
        exit
    exit
exit
```

The NAT **inside prefixes** are configured as follows:

```
configure
    service
        vprn 15001 customer 1 create
            nat
                inside
                    destination-prefix 0.0.0.0/0
                    classic-lsn-max-subscriber-limit 256
                    deterministic
                        prefix 10.0.0.0/24 subscriber-type classic-lsn-sub
                        nat-policy "nat-policy-1" create
                            map start 10.0.0.0 end 10.0.0.255 to 192.168.0.1
                            no shutdown
                        exit
                        prefix 10.10.4.0/22 subscriber-type classic-lsn-sub
                        nat-policy "nat-policy-1" create
                            map start 10.10.4.0 end 10.10.7.255 to 192.168.0.3
                            no shutdown
                        exit
                    exit
                exit
            exit
            no shutdown
        exit
    exit
exit
```

map statements are automatically created when the prefix is created and it is **no shutdown**.

Show commands

The subscriber-limit is set to 128 for the 10.0.0.0/24 prefix, so it is broken into two smaller /25 prefixes each. Each of these smaller prefixes are mapped into a specific outside IP address.

To show the first Large Scale NAT (LSN) subscriber of the first /25 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-router 15001 inside-ip 10.0.0.0

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.0.0.0]
NAT policy                  : nat-policy-1
Subscriber ID               : 276824064
-------------------------------------------------------------------------------
```

```
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.0.0.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.0.1

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last subscriber mapping to the same 192.168.0.1 outside IP address has inside address 10.0.0.127.

To show the first LSN subscriber of the second /25 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-router 15001 inside-ip 10.0.0.128

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.0.0.128]
NAT policy                  : nat-policy-1
Subscriber ID               : 276824192
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.0.0.128/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.0.2

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last subscriber mapping to the same 192.168.0.2 outside IP address has inside address 10.0.0.255.

To show the base router LSN blocks corresponding to the first inside IP address within the 10.0.0.0/24 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.0

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.1 [4024..4323]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.0

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

To show the base router LSN blocks corresponding to the last inside IP address within the 10.0.0.0/24 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.255

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.2 [42124..42423]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.255

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

The subscriber-limit is 128 for the 10.10.4.0/22 prefix, so it is broken into eight /25 prefixes. Each of these smaller prefixes are mapped into a specific outside IP address.

To show the first LSN subscriber of the first /25 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.10.4.0

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.10.4.0]
NAT policy                  : nat-policy-1
Subscriber ID               : 276824320
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.10.4.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.0.3

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last subscriber mapping to the same 192.168.0.3 outside IP address has inside address 10.10.4.127.

To show the first LSN subscriber of the last /25 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.10.7.128

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.10.7.128]
NAT policy                  : nat-policy-1
Subscriber ID               : 276825216
-------------------------------------------------------------------------------
```

```
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.10.7.128/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.0.10

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

To show the base router LSN blocks corresponding to the first inside IP within 10.10.4.0/24 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.10.4.0

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.3 [4024..4323]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.10.4.0

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

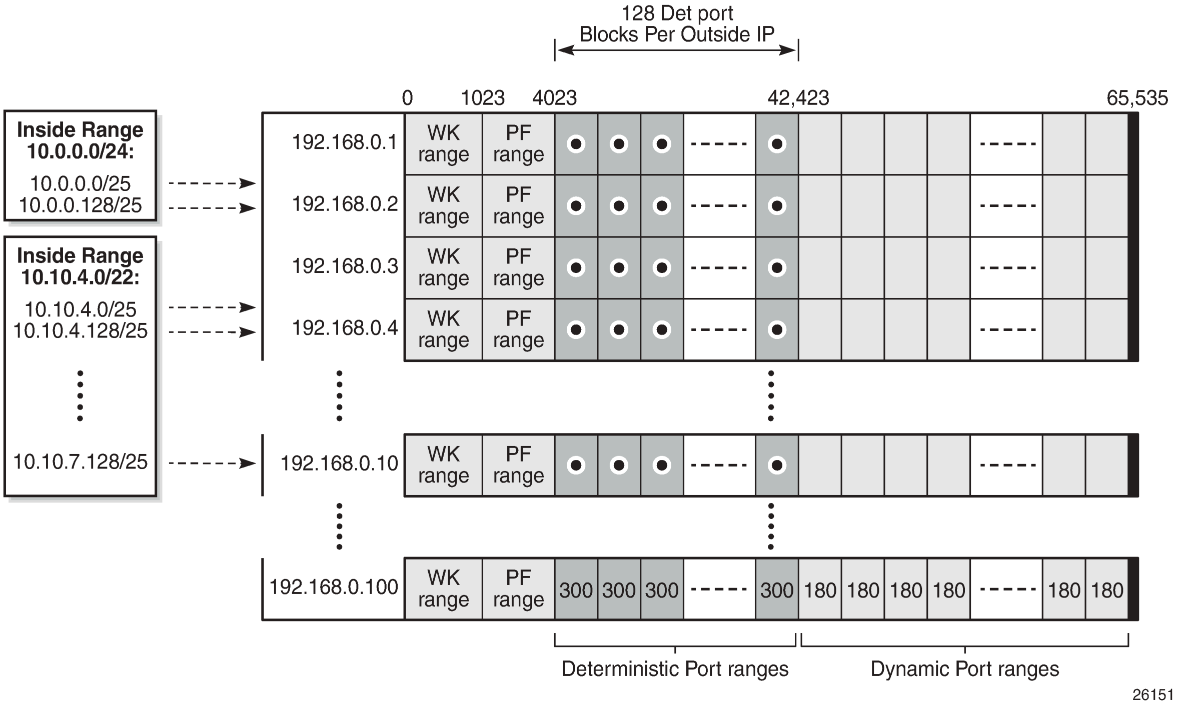
To show the base router LSN blocks corresponding to the last inside IP within 10.10.4.0/24 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.10.7.255

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.10 [42124..42423]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.10.7.255

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

Mapping results

According to this configuration, each inside IP address has one deterministic block of 300 ports and can have up to three dynamic blocks (block-limit = 4) each of 180 ports, allowing a maximum of 300+3*180 = 840 flows.

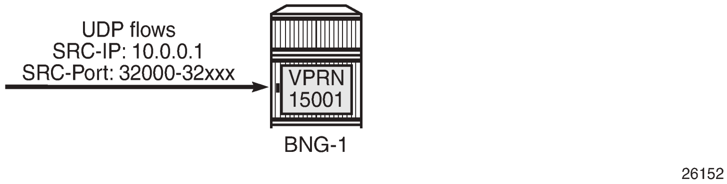
Figure 57: Case 1 results



Sending flows

For the inside IP 10.0.0.1, several UDP flows will be sent and both the deterministic and dynamic blocks mappings will be verified.

Figure 58: Case 1 flows



When sending 300 UDP flows or less, all flows are mapped to a single deterministic block because the number of ports in a deterministic block is 300. There is no logging; because no dynamic blocks are used, only the deterministic block is used.

To show LSN blocks on the outside routing instance *Base* and the outside ports allocated for the inside IP 10.0.0.1, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.1

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.1 [4324..4623]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

When increasing the number of flows such that: 301 < number of flows < 480

- In addition to the deterministic block (300 ports), there will be an extension by 1 dynamic block of 180 ports (port-reservation=180).
- Logging occurs for the dynamic port-block.

To show the base router LSN blocks and the outside ports allocated to the inside IP address 10.0.0.1, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.1

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.1 [4324..4623]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

192.168.0.1 [44044..44223]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:40:41
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

-------------------------------------------------------------------------------
Number of blocks: 2
===============================================================================
*A:PE1#
```

Logging is verified using Log 99 (in case event-control *nat* events are generated) which shows the mapping details to the new dynamic block as follows:

```
2 2016/10/28 12:40:41.51 UTC MINOR: NAT #2012 Base NAT
"{12} Map  192.168.0.1 [44044-44223] MDA 2/1 -- 276824065 classic-lsn-sub %1 vprn15001 10.0.0.1
 at 2016/10/28 12:40:41"
```

When increasing the number of flows such that: 481 < number of flows < 660

- In addition to the deterministic block (300 ports), there will be an extension by 2 dynamic blocks of 180 ports each.
- Logging occurs for the dynamic port-blocks.

To show LSN blocks on the outside routing instance *Base* and the outside ports allocated for the inside IP 10.0.0.1, the following command is used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.1

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.1 [4324..4623]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

192.168.0.1 [44044..44223]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:40:41
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

192.168.0.1 [44224..44403]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:41:52
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

-------------------------------------------------------------------------------
Number of blocks: 3
===============================================================================
*A:PE1#
```

Logging is verified using Log 99 (in case event-control *nat* events are generated) which shows the mapping details to the new dynamic block as follows:

```
3 2016/10/28 12:41:52.66 UTC MINOR: NAT #2012 Base NAT
"{13} Map  192.168.0.1 [44224-44403] MDA 2/1 -- 276824065 classic-lsn-sub %1 vprn15001 10.0.0.1
 at 2016/10/28 12:41:52"
```

When increasing the number of flows such that :661 < number of flows < 840

- In addition to the deterministic block (300 ports), there will be an extension by 3 dynamic blocks of 180 ports each.
- Logging occurs for the dynamic port-blocks.

To show LSN blocks on the outside routing instance “Base” and the outside ports allocated for the inside IP 10.0.0.1, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.0.0.1

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.0.1 [4324..4623]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/27 11:18:59
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1
```

```
192.168.0.1 [44044..44223]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:40:41
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

192.168.0.1 [44224..44403]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:41:52
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

192.168.0.1 [44404..44583]
Pool                                  : nat-pool-1
Policy                                : nat-policy-1
Started                               : 2016/10/28 12:43:46
Inside router                         : vprn15001
Inside IP address                     : 10.0.0.1

-------------------------------------------------------------------------------
Number of blocks: 4
===============================================================================
*A:PE1#
```

Logging is verified using Log 99 (in case event-control *nat* events are generated) which shows the mapping details to the new dynamic block as follows:

```
4 2016/10/28 12:43:46.71 UTC MINOR: NAT #2012 Base NAT
"{14} Map  192.168.0.1 [44404-44583] MDA 2/1 -- 276824065 classic-lsn-sub %1 vprn15001 10.0.0.1
 at 2016/10/28 12:43:46"
```

When increasing number of flows such that the number of flows > 840

- No more extension by dynamic blocks (block-limit = 4) allowed.
- Any flows more than 840 will be dropped and the relevant NAT statistics incremented.

To verify NAT statistics, first check the NAT group/member and MS-ISA associated with the outside IP 192.168.0.1/32:

```
*A:PE1# show router route-table 192.168.0.1/32

===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
192.168.0.1/32                                Remote  NAT       01d01h26m  0
       NAT outside to mda 2/1                                       0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
*A:PE1#
```

To check which group/member does this MS-ISA belong to, the following command can be used:

```
*A:PE1# show isa nat-group 1 members

===============================================================================
ISA Group 1 members
===============================================================================
Group Member     State          Mda  Addresses  Blocks     Se-% Hi Se-Prio
-------------------------------------------------------------------------------
1     1          active         2/1  175        23648      < 1  N  0
-------------------------------------------------------------------------------
No. of members: 1
===============================================================================
*A:PE1#
```

To verify relevant statistics for this NAT group/member, the following command can be used:

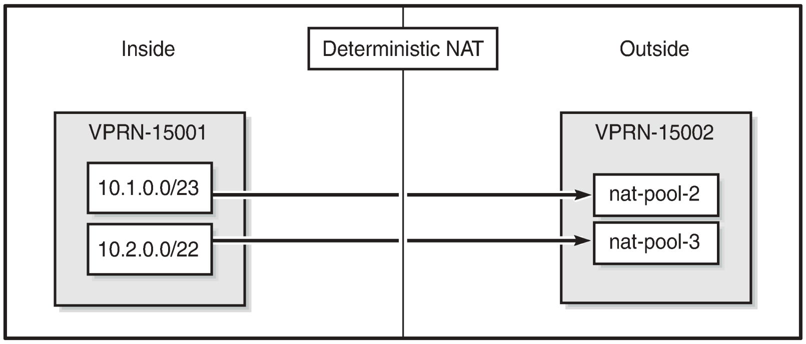
```
*A:PE1# show isa nat-group 1 member 1 statistics | match flow
no matching flow                                        : 56818
max flow exceeded                                       : 0
TCP no flow for RST                                     : 0
TCP no flow for FIN                                     : 0
TCP no flow                                             : 128094
flow log failed                                         : 0
new flow                                                : 1470768
found flow                                              : 39661850
flow create logged                                      : 0
flow delete logged                                      : 0
flow log pkt tx                                         : 0
flow create failed, key ambiguous                       : 0
flow create failed, conflicting policies                : 0
*A:PE1#
```

Case 2

Configured with:

- Mapping multiple prefixes from the same VRF into different outside pools.
- NAT all traffic.

Figure 59: Case 2



26153

The NAT **outside pool** are configured as follows:

```
configure
    service
        vprn 15002 customer 1 create
            nat
                outside
                    pool "nat-pool-2" nat-group 1 type large-scale create
                        port-reservation ports 80
                        subscriber-limit 256
                        deterministic
                            port-reservation 180
                        exit
                        address-range 192.168.2.1 192.168.2.200 create
                        exit
                        no shutdown
                    exit
                    pool "nat-pool-3" nat-group 1 type large-scale create
                        port-reservation ports 120
                        port-forwarding-range 4023
                        subscriber-limit 64
                        deterministic
                            port-reservation 840
                        exit
                        address-range 192.168.3.1 192.168.3.200 create
                        exit
                        no shutdown
                    exit
                exit
            exit
        exit
    exit
exit
```

The NAT **policies** are configured as follows:

```
configure
    service
        nat
            nat-policy "nat-policy-2" create
                block-limit 4
                pool "nat-pool-2" router 15002
            exit
            nat-policy "nat-policy-3" create
                block-limit 2
                pool "nat-pool-3" router 15002
            exit
        exit
    exit
exit
```

The NAT **inside prefix** is configured as follows:

```
configure
    service
        vprn 15001 customer 1 create
            nat
                inside
                    destination-prefix 0.0.0.0/0
                    classic-lsn-max-subscriber-limit 256
                    deterministic
                        prefix 10.1.0.0/23 subscriber-type classic-lsn-sub
```

```
                                                nat-policy "nat-policy-2" create
                            map start 10.1.0.0 end 10.1.1.255 to 192.168.2.1
                            no shutdown
                        exit
                        prefix 10.2.0.0/22 subscriber-type classic-lsn-sub
                                                nat-policy "nat-policy-3" create
                            map start 10.2.0.0 end 10.2.3.255 to 192.168.3.1
                            no shutdown
                        exit
                    exit
                exit
            exit
        exit
    exit
exit
```

Show commands

The subscriber-limit corresponding to the 10.1.0.0/23 prefix is 256, so the 10.1.0.0/23 prefix is broken into two /24 prefixes. Each of these smaller prefixes are mapped into a specific outside IP address.

To show the first LSN subscriber of the first /24 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.1.0.0

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.1.0.0]
NAT policy                  : nat-policy-2
Subscriber ID               : 276829472
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.1.0.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : 15002
Outside IP address          : 192.168.2.1

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last subscriber mapping to the same 192.168.2.1 outside IP address has inside address 10.1.0.255.

To show the first LSN subscriber of the second /24 prefix for inside routing instance 15001, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.1.1.0

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.1.1.0]
NAT policy                  : nat-policy-2
Subscriber ID               : 276829728
-------------------------------------------------------------------------------
```

```
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.1.1.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : 15002
Outside IP address          : 192.168.2.2

--------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
================================================================================
*A:PE1#
```

The last subscriber mapping to the same 192.168.2.2 outside IP address has inside address 10.1.1.255.

To show the VPRN-15002 LSN blocks corresponding to the first inside IP address within 10.1.0.0/23 prefix, the following command can be used:

```
*A:PE1# show router 15002 nat lsn-blocks inside-ip 10.1.0.0

================================================================================
Large-Scale NAT blocks for vprn15002
================================================================================
192.168.2.1 [1024..1203]
Pool                                  : nat-pool-2
Policy                                : nat-policy-2
Started                               : 2016/10/28 12:53:23
Inside router                         : vprn15001
Inside IP address                     : 10.1.0.0

--------------------------------------------------------------------------------
Number of blocks: 1
================================================================================
*A:PE1#
```

To show the VPRN-15002 LSN blocks corresponding to the last inside IP address within 10.1.0.0/23 prefix, the following command can be used:

```
*A:PE1# show router 15002 nat lsn-blocks inside-ip 10.1.1.255

================================================================================
Large-Scale NAT blocks for vprn15002
================================================================================
192.168.2.2 [46924..47103]
Pool                                  : nat-pool-2
Policy                                : nat-policy-2
Started                               : 2016/10/28 12:53:23
Inside router                         : vprn15001
Inside IP address                     : 10.1.1.255

--------------------------------------------------------------------------------
Number of blocks: 1
================================================================================
*A:PE1#
```

The subscriber-limit corresponding to the 10.2.0.0/22 prefix is 64,so the 10.2.0.0/22 prefix is broken into sixteen /26 prefixes. Each of these /26 prefixes is mapped to a specific outside IP address.

To show the first LSN subscriber for the inside routing instance 15001 for the first /26 prefix, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.2.0.0

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                    : [LSN-Host@10.2.0.0]
NAT policy                    : nat-policy-3
Subscriber ID                 : 276829984
-------------------------------------------------------------------------------
Type                          : classic-lsn-sub
Inside router                 : 15001
Inside IP address prefix      : 10.2.0.0/32
ISA NAT group                 : 1
ISA NAT group member          : 1
Outside router                : 15002
Outside IP address            : 192.168.3.1

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last inside address mapping to the 192.168.3.1 outside address is 10.2.0.63.

To show the first LSN subscriber for the inside routing instance 15001 for the last /26 prefix, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.2.3.192

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                    : [LSN-Host@10.2.3.192]
NAT policy                    : nat-policy-3
Subscriber ID                 : 276830944
-------------------------------------------------------------------------------
Type                          : classic-lsn-sub
Inside router                 : 15001
Inside IP address prefix      : 10.2.3.192/32
ISA NAT group                 : 1
ISA NAT group member          : 1
Outside router                : 15002
Outside IP address            : 192.168.3.16

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last inside address mapping to the 192.168.3.16 outside address is 10.2.3.255.

To show the VPRN-15002 LSN blocks corresponding to the first inside IP address within the 10.2.0.0/22 prefix, the following command can be used:

```
*A:PE1# show router 15002 nat lsn-blocks inside-ip 10.2.0.0

===============================================================================
Large-Scale NAT blocks for vprn15002
===============================================================================
192.168.3.1 [4024..4863]
Pool                                  : nat-pool-3
Policy                                : nat-policy-3
Started                               : 2016/10/28 12:53:23
Inside router                         : vprn15001
Inside IP address                     : 10.2.0.0

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

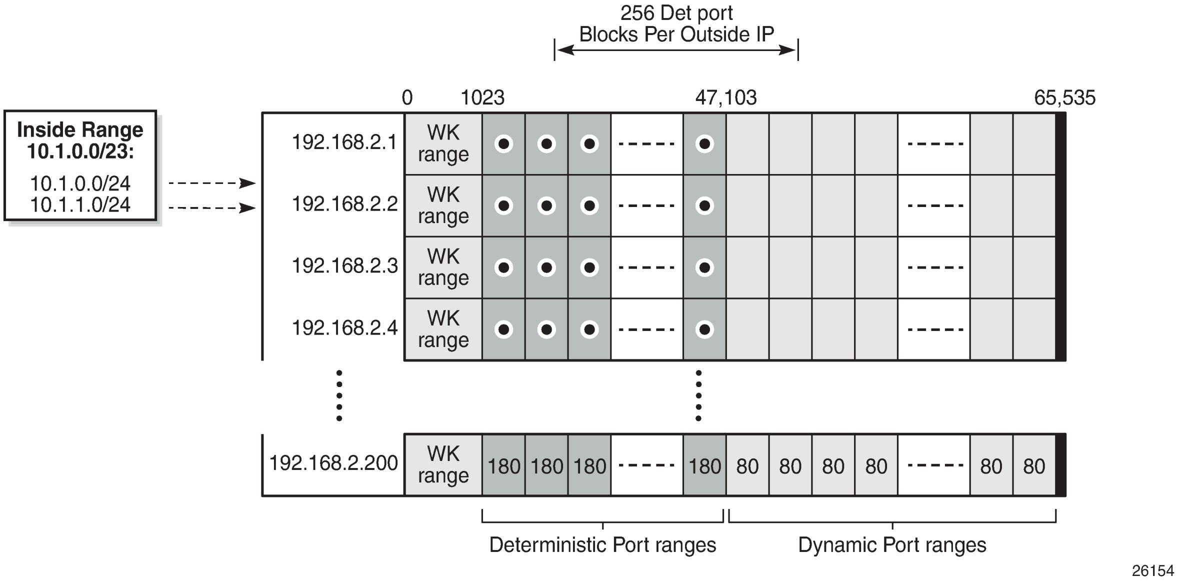
To show the VPRN-15002 LSN blocks corresponding to the last inside IP within 10.2.0.0/22 prefix, the following command can be used:

```
*A:PE1# show router 15002 nat lsn-blocks inside-ip 10.2.3.255

===============================================================================
Large-Scale NAT blocks for vprn15002
===============================================================================
192.168.3.16 [56944..57783]
Pool                                  : nat-pool-3
Policy                                : nat-policy-3
Started                               : 2016/10/28 12:53:23
Inside router                         : vprn15001
Inside IP address                     : 10.2.3.255

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

Mapping results

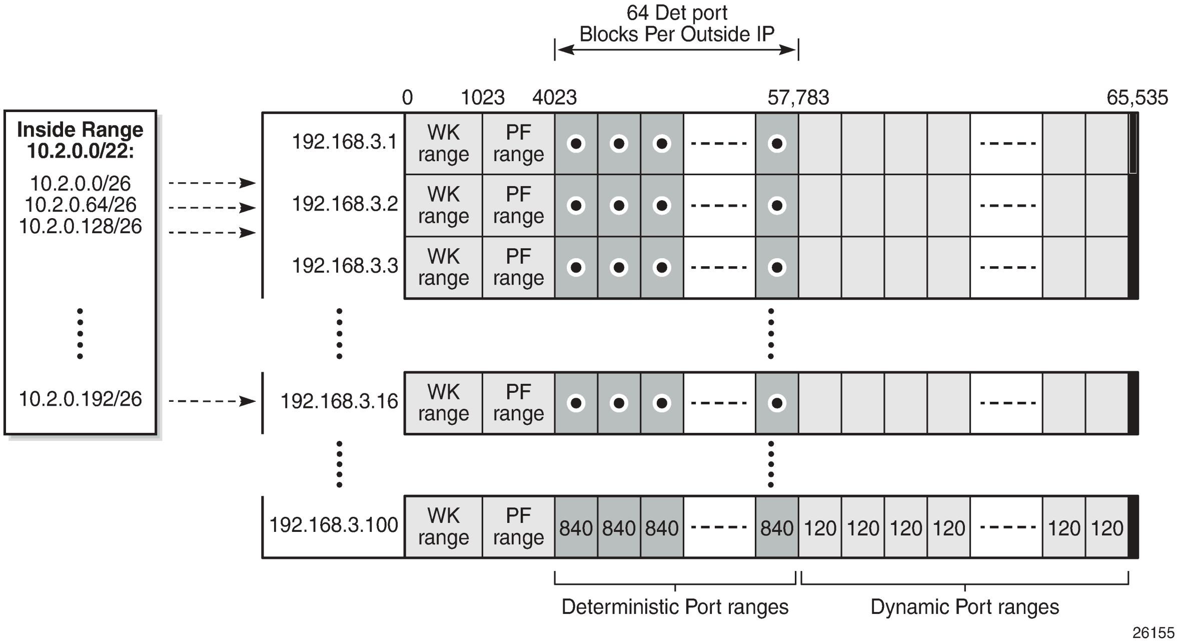
According to this configuration, for the 10.1.0.0/23 prefix, each inside IP address has one deterministic block of 180 ports and can have up to three dynamic blocks (block-limit =4) each of 80 ports, allowing for a maximum of 180+3*80 = 420 flows.

Figure 60: Case 2: Prefix 10.1.0.0/23 results



According to this configuration, for the 10.2.0.0/22 prefix, each inside IP address has one deterministic block of 840 ports, and can have up to one dynamic block (block-limit =2) of 120 ports, allowing for a maximum of 840+120 = 960 flows.

Figure 61: Case 2: Prefix 10.2.0.0/22 results

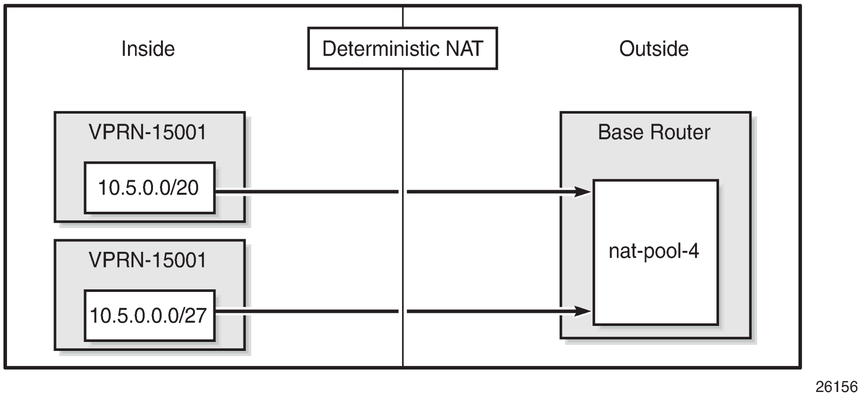


Case 3

Configured with:

- Mapping overlapping prefixes from different VRFs into the same outside pool.
- NAT all traffic.

Figure 62: Case 3



The NAT **outside pool** is configured as follows:

```
configure
    router "Base"
        nat
            outside
                pool "nat-pool-4" nat-group 1 type large-scale create
                    port-reservation ports 461
                    port-forwarding-range 4023
                    subscriber-limit 64
                    deterministic
                        port-reservation 500
                    exit
                    address-range 192.168.4.1 192.168.4.100 create
                    exit
                    no shutdown
                exit
            exit
        exit
    exit
exit
```

The NAT **policy** is configured as follows:

```
configure
    service
        nat
            nat-policy "nat-policy-4" create
                block-limit 4
                pool "nat-pool-4" router Base
            exit
        exit
    exit
```

```
exit
```

The NAT **inside prefix** is configured as follows:

```
configure
    service
        vprn 15001 customer 1 create
            nat
                inside
                    destination-prefix 0.0.0.0/0
                    classic-lsn-max-subscriber-limit 256
                    deterministic
                        prefix 10.5.0.0/20 subscriber-type classic-lsn-sub
                                                nat-policy "nat-policy-4" create
                            map start 10.5.0.0 end 10.5.15.255 to 192.168.4.1
                            no shutdown
                        exit
                    exit
                exit
            exit
        exit
    exit
exit

configure
    service
        vprn 15002 customer 1 create
            nat
                inside
                    destination-prefix 0.0.0.0/0
                    classic-lsn-max-subscriber-limit 128
                    deterministic
                        prefix 10.5.0.0/27 subscriber-type classic-lsn-sub
                                                nat-policy "nat-policy-4" create
                            map start 10.5.0.0 end 10.5.0.31 to 192.168.4.65
                            no shutdown
                        exit
                    exit
                exit
            exit
        exit
    exit
exit
```

Show commands

For the 10.5.0.0/20 prefix on VPRN 15001, the subscriber-limit is 64.The 10.5.0.0/20 prefix will be broken into 64 smaller /26 prefixes, each will be mapped into a specific outside IP address.

To show the first LSN subscriber for the inside routing instance 15001 of the first /26 prefix, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.5.0.0 inside-router 15001

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.5.0.0]
NAT policy                  : nat-policy-4
Subscriber ID               : 276825344
```

```
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.5.0.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.4.1

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last inside address mapping to the 192.168.4.1 outside address is 10.5.0.63.

To show the first Large Scale NAT (LSN) subscriber for the inside routing instance 15001 of the last /26 prefix, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.5.15.192 inside-router 15001

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.5.15.192]
NAT policy                  : nat-policy-4
Subscriber ID               : 276829376
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15001
Inside IP address prefix    : 10.5.15.192/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.4.64

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

The last inside address mapping to the 192.168.4.64 outside address is 10.5.15.255.

To show the base router LSN blocks corresponding to the first inside IP address within the 10.5.0.0/20 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.5.0.0 inside-router 15001

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.4.1 [4024..4523]
Pool                                  : nat-pool-4
Policy                                : nat-policy-4
Started                               : 2016/10/27 13:11:38
Inside router                         : vprn15001
Inside IP address                     : 10.5.0.0

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

To show the base router LSN blocks corresponding to the last inside IP address within the 10.5.0.0/20 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.5.15.255 inside-router 15001

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.4.64 [35524..36023]
Pool                                  : nat-pool-4
Policy                                : nat-policy-4
Started                               : 2016/10/27 13:11:38
Inside router                         : vprn15001
Inside IP address                     : 10.5.15.255

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

For the 10.5.0.0/27 prefix in VPRN 15002, the subscriber-limit is 64.The 10.5.0.0/27 prefix will be mapped into one outside IP address.

To show the first LSN subscriber for the inside routing instance 15002 of the 10.5.0.0/27 prefix, the following command can be used:

```
*A:PE1# show service nat lsn-subscribers inside-ip 10.5.0.0 inside-router 15002

===============================================================================
NAT LSN subscribers
===============================================================================
Subscriber                  : [LSN-Host@10.5.0.0]
NAT policy                  : nat-policy-4
Subscriber ID               : 276829440
-------------------------------------------------------------------------------
Type                        : classic-lsn-sub
Inside router               : 15002
Inside IP address prefix    : 10.5.0.0/32
ISA NAT group               : 1
ISA NAT group member        : 1
Outside router              : "Base"
Outside IP address          : 192.168.4.65

-------------------------------------------------------------------------------
No. of LSN subscriber instances: 1
===============================================================================
*A:PE1#
```

To show the LSN blocks corresponding to the first inside IP address within the 10.5.0.0/27 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.5.0.0 inside-router 15002

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.4.65 [4024..4523]
Pool                                  : nat-pool-4
Policy                                : nat-policy-4
Started                               : 2016/10/27 13:12:02
Inside router                         : vprn15002
Inside IP address                      : 10.5.0.0

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

To show the LSN blocks for the last inside IP address within the 10.5.0.0/27 prefix, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.5.0.31 inside-router 15002

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.4.65 [19524..20023]
Pool                                   : nat-pool-4
Policy                                 : nat-policy-4
Started                                : 2016/10/27 13:12:02
Inside router                          : vprn15002
Inside IP address                      : 10.5.0.31

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

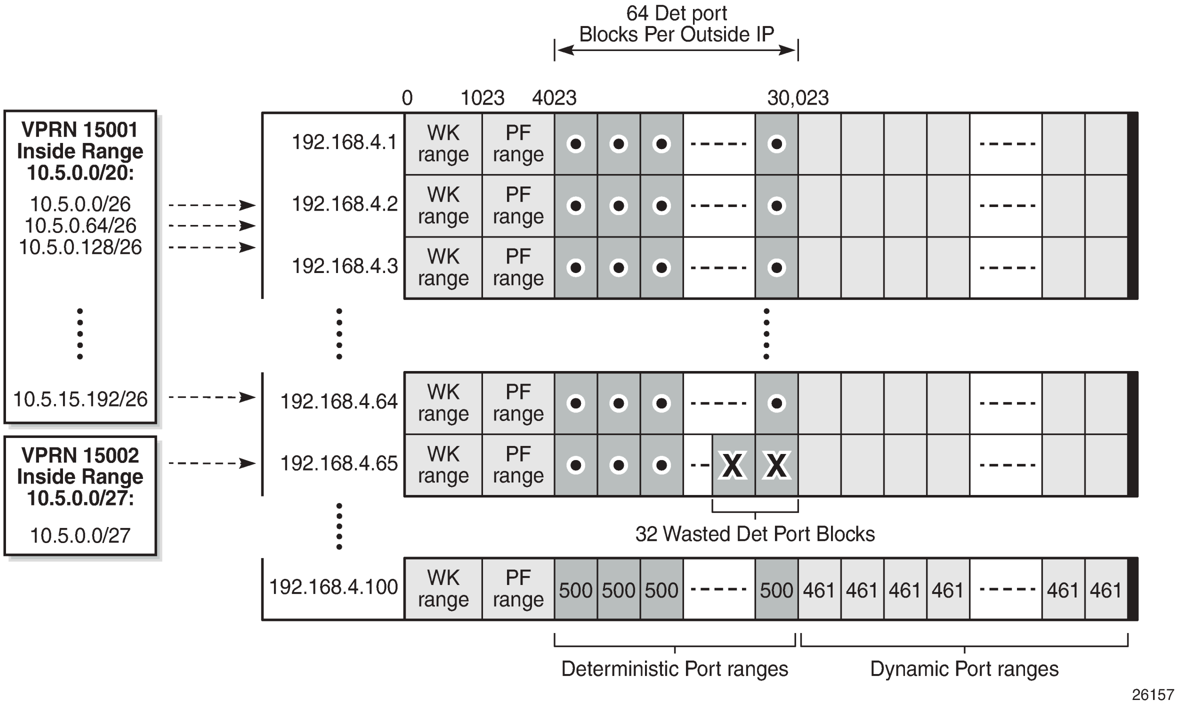
Mapping results

According to this configuration, each inside IP address within VPRN 15001 has one deterministic block of 500 ports and can have up to three dynamic blocks (block-limit =4) of 461 ports each, allowing a maximum of 500+3*461 = 1883 flows.

According to this configuration each inside IP address within VPRN 15002 has one deterministic block of 500 ports and can have up to three dynamic blocks (block-limit =4) of 461 ports each, allowing a maximum of 500+3*461 = 1883 flows.

For VPRN 15002, since the number of LSN subscribers (32) is less than the number of deterministic blocks (64), then 32 deterministic blocks will be wasted, specifically 32*500 = 16,000 ports will be wasted which is not good in terms of capacity planning.

Figure 63: Case 3 results



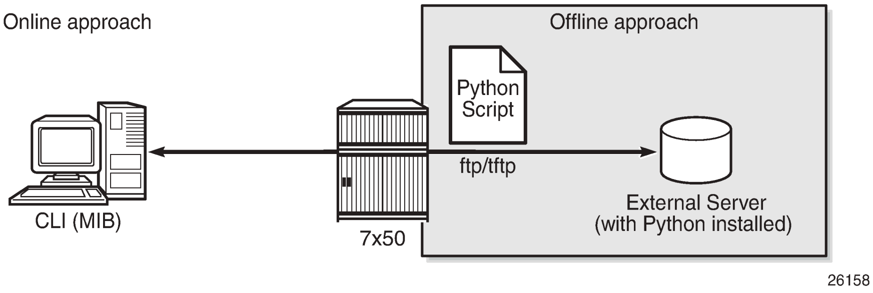
Inverse mapping

In deterministic LSN44, the inside IP addresses are mapped to outside IP addresses and corresponding port blocks based on a deterministic algorithm. The inverse mapping that reveals the subscriber identity behind the NAT is based on the reversal of this algorithm.

Inverse mappings can be done either online or offline:

- Online — Locally on the SR-series node, via CLI (MIB)
- Offline — Externally, via a Python script. The purpose of such an offline approach is to provide fast queries without accessing the SR-series node.

Figure 64: Inverse mapping approach



Online approach

A **tools** command is available which shows the reverse mapping (outside to inside) for deterministic NAT instead of using logging.

```
tools dump nat deterministic-mapping outside-ip <ipv4-address> router <router-instance>
 outside-port <[1..65535]>

 <ipv4-address>       : a.b.c.d
 <router-instance>    : <router-name>|<service-id>
                        router-name    - "Base"
                        service-id     - [1..2147483647]
```

Using Case 3 as an example, to obtain (inside IP, inside routing instance), the inverse mapping for a specific (outside IP, outside routing instance, outside port) is done as follows:

```
*A:PE1# tools dump nat deterministic-mapping outside-ip 192.168.4.1 router "Base" outside-port
 4024
classic-lsn-sub inside router 15001 ip 10.5.0.0 -- outside router Base ip 192.168.4.1 port 4024
 at Fri Oct 28 13:04:22 UTC 2016
*A:PE1#
```

```
*A:PE1# tools dump nat deterministic-mapping outside-ip 192.168.4.65 router "Base" outside-port
 4024
classic-lsn-sub inside router 15002 ip 10.5.0.0 -- outside router Base ip 192.168.4.65 port
 4024 at Fri Oct 28 13:04:45 UTC 2016
*A:PE1#
```

Offline approach

The purpose of such an offline approach is to provide fast queries without the need to directly query the SR-series node.

This is achieved by generating and exporting a Python script for reverse querying, which is a manual operation that needs to be repeated every time there is configuration change in deterministic NAT.

The script is exported (manually) to the external system.

To configure remotely the location for the Python script, the following command is used:

```
configure service nat deterministic-script location <remote-url>
```

remote-url — A remote location where the script is stored:

[{ftp://|tftp://}<login>:<pswd>@ <remote-locn>/][<file-path>]

Maximum length is 180 characters.

Once the script location is specified, the script can be exported to that location using the following command:

```
admin nat save-deterministic-script
```

Using the following command the status of the script can be checked, and whether it is necessary to re-save (export) the script or not:

```
*A:PE1# show service nat deterministic-script

===============================================================================
Deterministic NAT script data
===============================================================================
Location                    : ftp://*:*@123.123.123.123/pub/python/detnat.py
Save needed                 : no
Last save result            : success
Last save time              : 2016/10/28 13:05:41
===============================================================================
*A:PE1#
```

The external system must have the Python scripting language installed with the following modules: getopt, math, os, socket, and sys.

The Python script can then be run on the external server; the parameters are as follows:

```
[user@123.123.123.123 ~]$ ./detnat.py
Error: need exactly one of --forward or --backward arguments

Usage: detnat.py DIRECTION PARAMETERS
Perform forward or backard NAPT according to the configured deterministic rules.

DIRECTION:
  -f, --forward             Translate from inside to outside address/port
  -b, --backward            Translate from outside to inside address/port

PARAMETERS:
  -a, --address=IP-ADDRESS  The address to translate. IPv6 addresses can be
                            specified in shorthand or full notation.
  -p, --port=PORT           The outside port in case of backward translation.
  -s, --service=SERVICE-ID  The service where the IP-ADDRESS originates from.
                            This is the inside service in case of forward
                            translation and the outside service in case of
                            backward translation.
                            To specify the base router, this option must be
                            omitted.

  -h, --help                Show this help message
[user@123.123.123.123 ~]$
```

where deterministic-nat.py is the name of the python script previously exported.

As an example of a forward query:

```
[user@123.123.123.123 ~]$ ./detnat.py -f -s 15001 -a 10.0.0.1
classic-lsn-sub has public ip address 192.168.0.1 from base router and is using ports [4324 -
 4623]
[user@123.123.123.123 ~]$
```

As an example of a reverse query:

```
[user@123.123.123.123 ~]$ ./detnat.y -b -s 0 -a 192.168.0.1 -p 4325
classic-lsn-sub has private ip address 10.0.0.1 from service 15001
[user@123.123.123.123 ~]$
```

Simultaneous support of deterministic and non-deterministic NAT

Deterministic NAT can be used simultaneously with non-deterministic NAT within the same inside routing instance. However, they cannot share the same pool.

An outside pool can be only deterministic (although expandable by dynamic ports blocks) or non-deterministic at any given time (a non-deterministic pool is a pool that contains dynamic port-blocks only).

The following show a configuration using deterministic NAT simultaneously with non-deterministic NAT.

The NAT **outside pool** are configured as follows:

```
configure
    router
        nat
            outside
                pool "nat-pool-1" nat-group 1 type large-scale create
                    port-reservation ports 180
                    port-forwarding-range 4023
                    subscriber-limit 128
                    deterministic
                        port-reservation 300
                    exit
                    address-range 192.168.0.1 192.168.0.100 create
                    exit
                    no shutdown
                exit
                pool "nat-pool-Non-Deterministic" nat-group 1 type large-scale create
                    address-range 192.168.7.1 192.168.7.100 create
                    exit
                    no shutdown
                exit
            exit
        exit
    exit
exit
```

The NAT **policies** are configured as follows:

```
configure
    service
        nat
            nat-policy "nat-policy-1" create
                block-limit 4
                pool "nat-pool-1" router Base
            exit
            nat-policy "nat-policy-Non-Deterministic" create
                pool "nat-pool-Non-Deterministic" router Base
            exit
        exit
    exit
exit
```

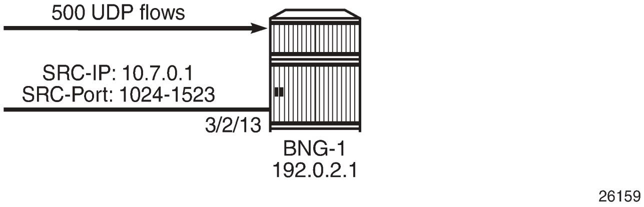
The NAT **inside prefixes** are configured as follows:

```
configure
    service
        vprn 15001 customer 1 create
            nat
                inside
```

```
                    destination-prefix 0.0.0.0/0
                    classic-lsn-max-subscriber-limit 256
                    deterministic
                        prefix 10.0.0.0/24 subscriber-type classic-lsn-sub
                        nat-policy "nat-policy-1" create
                            map start 10.0.0.0 end 10.0.0.255 to 192.168.0.1
                            no shutdown
                        exit
                    exit
                    nat-policy "nat-policy-Non-Deterministic"
                exit
            exit
            no shutdown
        exit
    exit
exit
```

In this example, the inside IP prefixes that do not match any of the deterministic prefixes will be NATed using a non-deterministic pool.

Figure 65: Sending flows: deterministic + non-deterministic NAT



To check which NAT pool/NAT policy is used for NATing the inside IP 10.7.0.1, the following command can be used:

```
*A:PE1# show router nat lsn-blocks inside-ip 10.7.0.1

===============================================================================
Large-Scale NAT blocks for Base
===============================================================================
192.168.7.100 [1024..1527]
Pool                                  : nat-pool-Non-Deterministic
Policy                                : nat-policy-Non-Deterministic
Started                               : 2016/10/28 13:24:56
Inside router                         : vprn15001
Inside IP address                     : 10.7.0.1

-------------------------------------------------------------------------------
Number of blocks: 1
===============================================================================
*A:PE1#
```

Conclusion

This example provides the commands required for configuring deterministic LSN44 NAT. Both deterministic as well as non-deterministic NAT are supported, with simultaneous operation being possible.

Inverse query can be done online or offline to retrieve the NAT mappings. Logging is not needed as long as there are no dynamic blocks assigned to LSN44 subscribers.

IP/GRE Termination

This chapter provides configuration and troubleshooting commands for IP/GRE termination.

Topics in this chapter include:

Applicability

The chapter was initially written for SR OS Release 9.0.R8. The CLI in the current edition corresponds to SR OS Release 22.2.R2.

Initially, the 7750 SR only supported GRE SDP tunnels which use pseudowire encapsulation. In SR OS Release 8.0.R5 and later, the 7750 SR supports tunneling IPv4 packets in an IPv4 Generic Routing Encapsulation (GRE) tunnel.

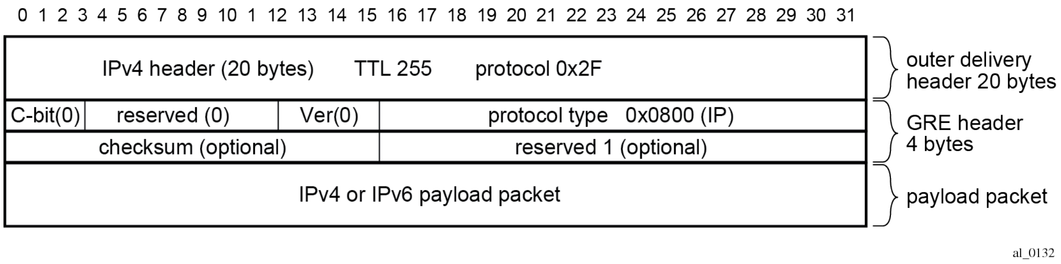
The IP GRE tunnel termination configuration described in this chapter requires an MS-ISA. IP GRE tunnels without ISA are beyond the scope of this chapter.

Overview

A common use case for IP/GRE tunneling is remote access to a VPRN over a public IP network because IP/GRE tunneling allows encapsulated packets to follow a path based on the outer IP header which is useful when the inner IP packet cannot or should not be forwarded natively over this path.

GRE allows packets of one protocol, the payload protocol, to be encapsulated by packets of another protocol, called the delivery protocol. Figure 66: GRE packet format shows the GRE packet format with an outer delivery header, GRE header, and payload packet:

Figure 66: GRE packet format



The outer delivery and GRE header for outgoing traffic is as follows.

- Outer delivery header

  - The source address in the IPv4 delivery header is the configured source address.
  - The destination address in the IPv4 delivery header is the configured remote IP (or the backup remote IP) address.
  - The IP protocol value in the IPv4 delivery header is 0x2F or 47 (GRE).
  - The DSCP in the IPv4 outer delivery header is:
    - set to the value configured for the tunnel;
    - otherwise, the DSCP value from the payload packet is copied into the outer delivery header.
  - The TTL in the IPv4 outer delivery header is set to 255.
- GRE header
  - The checksum (C) bit in the GRE header is set to 0 (no checksum present).
  - The version in the GRE header is 0.
  - The protocol type in the GRE header is 0x0800 for IPv4.

The outer delivery and GRE header for incoming traffic is as follows:

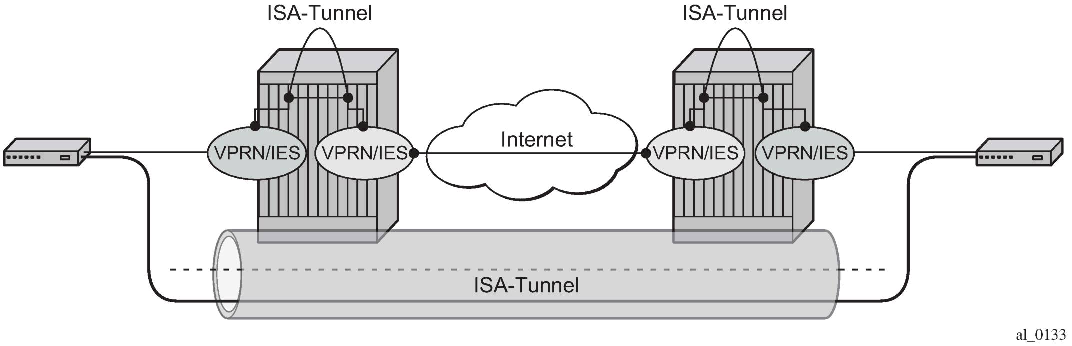
- Outer delivery header
  - If the packet is a fragment (more fragments=1, non-zero fragment offset), it is dropped.
  - If the checksum (C) bit in the GRE header is set, then the included checksum is validated; if the checksum is incorrect, the packet is discarded.
  - If the version in the GRE header is not 0, the packet is discarded.
  - If the source/destination address pair in the IPv4 delivery header is any other combination, the packet is dropped.
- GRE header
  - If the checksum (C) bit in the GRE header is set, then the included checksum is validated; if the checksum is incorrect, the packet is discarded.
  - If the version in the GRE header is not 0, the packet is discarded.

Implementation

Encapsulation, de-encapsulation and other datapath operations related to IP/GRE are handled by the ISA-tunnel MDA.

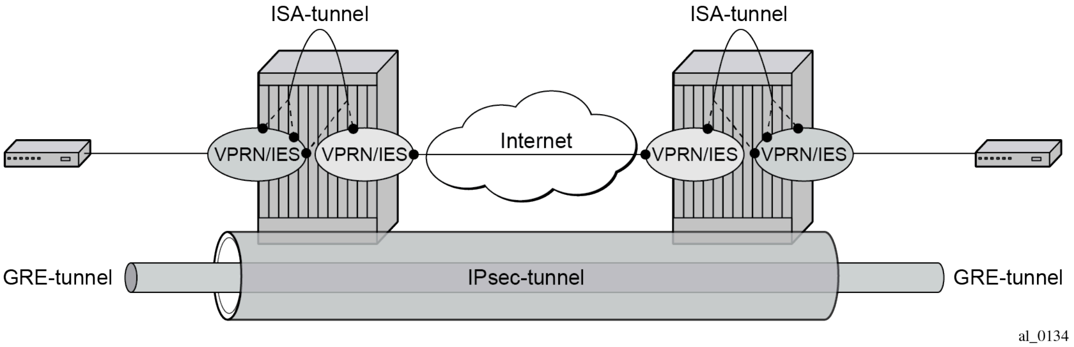
For GRE tunnels configured as SDPs (which are not covered by this section), no ISA-tunnel MDA is required.

Figure 67: Implementation



SR OS nodes initially supported the IP/GRE tunnels with static routes and BGP. IP/GRE tunnels have been enhanced by adding support for OSPF and BFD on private tunnel interfaces (used with static routes, OSPF, or BGP) and GRE protection by tunneling into an IPSec tunnel.

Figure 68: IP/GRE over IPSec tunnel



Configuration

ISA-tunnel MDA

The ISA-tunnel MDA supports IP/GRE and IPSec tunnels and is configured as follows:

```
# on PE-1:
configure
    card 1
        mda 2
            mda-type isa2-tunnel
            no shutdown
        exit
    exit
    card 2
        mda 2
            mda-type isa2-tunnel
            no shutdown
        exit
    exit
```

The following command checks the MDA configuration:

```
*A:PE-1# show mda

===============================================================================
MDA Summary
===============================================================================
Slot  Mda   Provisioned Type                            Admin     Operational
                Equipped Type (if different)            State     State
-------------------------------------------------------------------------------
1     1     p10-10g-sfp                                 up        up
      2     isa2-tunnel                                 up        up
                p-isa2-ms
2     1     me40-1gb-csfp                               up        up
      2     isa2-tunnel                                 up        up
                me-isa2-ms
===============================================================================
```

Tunnel groups and tunnel group restrictions

The first step of the GRE tunnel configuration is to configure a tunnel group.

A tunnel group can have one tunnel ISA designated primary and optionally one tunnel-ISA designated backup. When a GRE tunnel is created, it is assigned to the primary tunnel-ISA in its tunnel group. If the primary tunnel-ISA fails, the backup tunnel-ISA (if not already claimed by another tunnel group) takes over for the failed card.

```
*A:PE-1>config>isa# tunnel-group 1 ?
  - tunnel-group <tunnel-group-id> [create]
  - tunnel-group <tunnel-group-id> isa-scale-mode <isa-scale-mode> [create]
  - no tunnel-group <tunnel-group-id>

 <tunnel-group-id>    : [1..16]
 <isa-scale-mode>     : tunnel-limit-2k
                        k=1024
 <create>             : keyword - mandatory while creating an entry.


 [no] active-mda-num* - Configure number of active MDAs
 [no] backup          - Configure ISA-Tunnel-Group backup ISA
 [no] description     - Configure the ISA group description
 [no] esa-vm          - Configure the esa-vm
 [no] ipsec-responde* - Enable/Disable responder-only for IPsec Ikev2 tunnels only
 [no] mda             - Configure MDA to associate with
 [no] multi-active    - Configure multi-active status of tunnel-group
 [no] primary         - Configure ISA-Tunnel-Group primary ISA
 [no] reassembly      - Configure reassembly wait time
 [no] shutdown        - Administratively enable/disable an ISA-Tunnel-Group
      stats-collecti* + Configure ISA statistics collection parameters

# on PE-1:
configure
    isa
        tunnel-group 1 create
            primary 1/2
            backup 2/2
            no shutdown
        exit
```

The failed tunnels are re-established using a cold-standby on the backup tunnel-ISA. Cold-standby means the backup tunnel-ISA has no state or configuration information about the tunnels prior to the failure.

A tunnel ISA cannot be primary for more than one tunnel group:

```
*A:PE-1>config>isa# tunnel-group 2 create
*A:PE-1>config>isa>tunnel-grp$ primary 1/2
MINOR: IPSECGRPMGR #1003 The specified MDA is primary in another Tunnel Group
```

A tunnel ISA cannot be primary in one tunnel group and backup in another tunnel group:

```
*A:PE-1>config>isa# tunnel-group 2 create
*A:PE-1>config>isa>tunnel-grp# backup 1/2
MINOR: IPSECGRPMGR #1003 The specified MDA is primary in another Tunnel Group
```

The following commands shows the ISA tunnel group (after tunnel group 2 has been removed):

```
*A:PE-1# show isa tunnel-group

===============================================================================
ISA Tunnel Groups
===============================================================================
Tunnel    PrimaryIsa             BackupIsa    ActiveIsa    Admin     Oper
GroupId                                                    State     State
-------------------------------------------------------------------------------
1         1/2                    2/2          1/2          Up        Up
-------------------------------------------------------------------------------
No. of ISA Tunnel Groups: 1
===============================================================================
```

The following command shows the number of the IP (GRE) tunnels, after configuring IES and VPRN services with tunnel interfaces:

```
*A:PE-1# show ip tunnel count
--------------------------------------------------------------------------------
IP Tunnels: 2
--------------------------------------------------------------------------------
```

The following command shows all IP tunnels:

```
*A:PE-1# show ip tunnel

===============================================================================
IP Tunnels
===============================================================================
TunnelName                       SapId                          SvcId      Admn
 Local Address                                                  DlvrySvcId Oper
  OperRemoteAddress
-------------------------------------------------------------------------------
gre-tunnel-1                     tunnel-1.private:1             1          Up
 192.168.1.1                                                    2          Up
  192.168.2.1
protected-gre-tunnel             tunnel-1.private:5             3          Up
 192.168.11.1                                                   3          Up
  192.168.22.1
-------------------------------------------------------------------------------
IP Tunnels: 2
```

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

The detailed tunnel information is as follows:

```
*A:PE-1# show ip tunnel "gre-tunnel-1"

===============================================================================
IP Tunnel Configuration Detail
===============================================================================
Service Id       : 1                       Sap Id           : tunnel-1.private:1
Tunnel Name      : gre-tunnel-1
Description      : None
GRE Header       : Yes
Delivery Service : 2
GRE Keys Set     : False
GRE Send Key     : N/A                     GRE Receive Key  : N/A
Admin State      : Up                      Oper State       : Up
Source Address   : 192.168.1.1
Remote Address   : 192.168.2.1
Backup Address   : (Not Specified)
Oper Remote Addr : 192.168.2.1
DSCP             : None
Reassembly       : inherit
Clear DF Bit     : false                   IP MTU           : max
Encap IP MTU     : max
Pkt Too Big      : true
Pkt Too Big Num  : 100                     Pkt Too Big Intvl: 10 secs
Frag Required    : true
Frag Req Count   : 100                     Frag Req Interval: 10 secs
Propagate IPv6 P*: true
Propagate IPv4 P*: true
Oper Flags       : None
Transport Profile: (Not Specified)
Last Oper Changed: 05/12/2022 08:40:02
Host ISA         : 1/2
TCP MSS Adjust
    Public       : Disabled
    Private      : Disabled

-------------------------------------------------------------------------------
Target Address Table
-------------------------------------------------------------------------------
Destination IP                          IP Resolved Status
-------------------------------------------------------------------------------
10.0.0.2                                Yes
-------------------------------------------------------------------------------

===============================================================================
IP Tunnel Statistics: gre-tunnel-1
===============================================================================
Errors Rx        : 0                       Errors Tx        : 0
Pkts Rx          : 51                      Pkts Tx          : 49
Bytes Rx         : 3575                    Bytes Tx         : 3483
Key Ignored Rx   : 0                       Too Big Tx       : 0
Seq Ignored Rx   : 0
Vers Unsup. Rx   : 0
Invalid Chksum Rx: 0
Key Mismatch Rx  : 0
===============================================================================

===============================================================================
Fragmentation Statistics
===============================================================================
Encapsulation Overhead                 : 24
Temporary Private MTU                  : max
Pre-Encapsulation
    Fragmentation Count                : 0
    Last Fragmented Packet Size        : 0
Post-Encapsulation
    Fragmentation Count                : 0
    Last Fragmented Packet Size        : 0
===============================================================================
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

Interfaces

The interface toward the Internet (or WAN):

- can be a network interface or VPRN/IES interface.
- provides IP reachability.

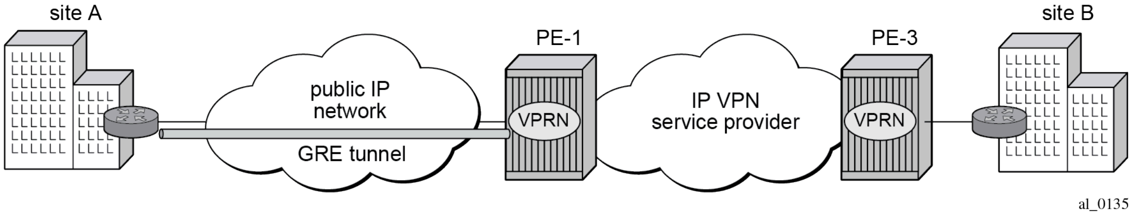
The tunnel public interface:

- can be an IES or VPRN interface.
- represents the public side of the tunnel-ISA.

The delivery VPRN/IES service (the service connected to the Internet) must have at least one IP interface associated with a public tunnel SAP in order to receive and process GRE encapsulated packets.

The public tunnel SAP type has the format **tunnel-***id***.private|public**:*tag* (where the *id* corresponds to the tunnel group). Figure 70: GRE for remote access to a VPRN service shows the example topology, where CE-2 in customer site A is connected to PE-1.

Figure 69: GRE for remote access to a VPRN service



The IES service with public tunnel SAP is configured on PE-1 as follows:

```
*A:PE-1>config>service>ies>if# sap ?
  - no sap <sap-id>
  - sap <sap-id> [create]

 <sap-id>
              ---snip---
                        tunnel-id      - tunnel-<id>.<private|public>:<tag>
                          tunnel         - keyword
                          id             - [1..16]
                          tag            - [0..4094]
              ---snip---

# on PE-1:
configure
    service
```

```
        ies 2 name "IES 2" customer 1 create
            interface "int-tunnel-public" create
                address 192.168.1.2/30
                tos-marking-state untrusted
                sap tunnel-1.public:1 create
                exit
            exit
            interface "int-PE-1-CE-2" create
                address 192.168.12.1/24
                sap 1/1/2:2 create
                exit
            exit
            no shutdown
        exit
```

PE-1 has address 192.168.1.2/30 assigned to the interface "int-tunnel-public" in IES 2. In a similar way, CE-2 has address 192.168.2.2/30 assigned to the interface "int-tunnel-public" in IES 2.

In order to reach 192.168.2.0/30 on CE-2, a static route is configured on PE-1, as follows:

```
# on PE-1:
configure
    router Base
        static-route-entry 192.168.2.0/30
            next-hop 192.168.12.2
                no shutdown
```

In a similar way, a static route is configured on CE-2 to reach 192.168.1.0/30 on PE-1.

Mask /32 is not supported on the public tunnel. When address 192.168.1.2/32 is configured on the interface "int-tunnel-public", the public tunnel cannot be created, as follows:

```
*A:PE-1>config>service>ies>if# address 192.168.1.2/32
*A:PE-1>config>service>ies>if# sap tunnel-1.public:1 create
INFO: PIP #1288 Cannot bind when there are /32 or /128 addresses configured
```

Therefore, the address configured on the interface will have mask /30 instead of /32, as shown earlier.

The tunnel private interface:

- can be an IES or VPRN interface.
- represents the private side of the tunnel ISA.

The private tunnel SAP has the format **tunnel-***id***.private|public**:*tag* (where the *id* corresponds to the tunnel group) as shown in the following example where an unprotected GRE tunnel is configured in the SAP context.

```
*A:PE-1>config>service>vprn>if# sap ?
  - no sap <sap-id>
  - sap <sap-id> [create]

 <sap-id>
              ---snip---
                        tunnel-id      - tunnel-<id>.<private|public>:<tag>
                          tunnel         - keyword
                          id             - [1..16]
                          tag            - [0..4094]
              ---snip---

# on PE-1:
configure
```

```
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                    ---snip---
```

It is not mandatory to have the same tag (internal dot1q) in private and public GRE tunnels.

```
    sap tunnel-1.private:1 <=> sap tunnel-1.public:2
```

Unprotected GRE tunnel configuration

To associate an unprotected GRE tunnel with a private tunnel SAP, the **ip-tunnel** command is configured in the SAP context.

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
        ---snip---
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        ---snip---
```

The **dest-ip** keyword followed by the private IP address of the remote tunnel endpoint is mandatory.

If this remote IP address is not within the subnet of the local private endpoint, then the tunnel will not come up.

The following parameters are configured in the **ip-tunnel** context:

- The source address of the GRE tunnel. This is the source IPv4 address of GRE encapsulated packets sent by the delivery service. It must be an address in the subnet of the associated public tunnel SAP interface.
- The remote IP address. If this address is reachable in the delivery service (there is a route), then this is the destination IPv4 address of GRE encapsulated packets sent by the delivery service.
- The backup remote IP address. If the remote IP address of the tunnel is not reachable, then the backup remote IP address is the destination IPv4 address of GRE encapsulated packets sent by the delivery service.
- The delivery service. This is the identifier or name of the IES or VPRN service where GRE encapsulated packets are injected and terminated. The delivery service can be the same service where the private tunnel SAP interface resides.
- The DSCP marking in the outer IP header of GRE encapsulated packets. If this is not configured, then the default copies the DSCP from the inner IP header to the outer IP header.

  ```
  # on PE-1:
  ```

```
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        dscp af22
                        no shutdown
                    exit
                    ---snip---
```

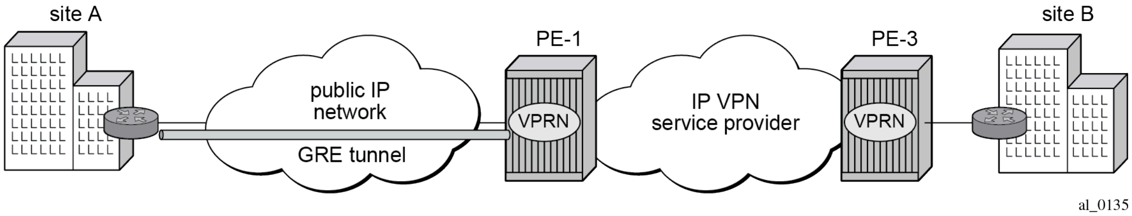
- A private tunnel SAP can have only one IP/GRE tunnel (per SAP).

```
*A:PE-1>config>service>vprn>if# sap tunnel-1.private:1 ip-tunnel "gre-tunnel-2" create
MINOR: SVCMGR #5120 Only one IP tunnel allowed per SAP
```

IP/GRE tunneling via static route

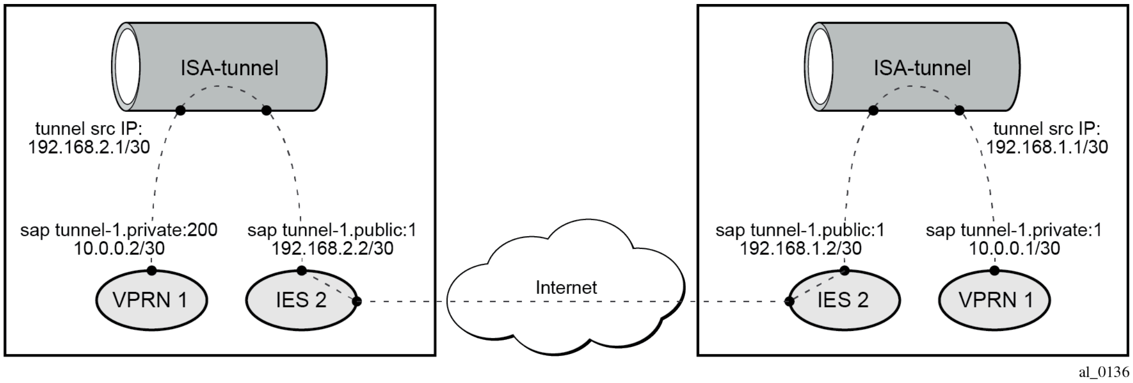
A static route can reference the GRE tunnel directly (by next-hop IP address) or the GRE tunnel can be the resolved next-hop for an indirect static route (Figure 70: GRE for remote access to a VPRN service).

Figure 70: GRE for remote access to a VPRN service



The details of both ends on the GRE tunnel, at site A and PE-1, are shown in Figure 71: IP/GRE tunneling via static route. The node at left hand side is CE-2 at site A.

Figure 71: IP/GRE tunneling via static route



The following shows the configuration of VPRN 1 on PE-1.

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                loopback
            exit
            static-route-entry 172.16.2.1/32
                next-hop 10.0.0.2
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            no shutdown
        exit
```

The configuration of the VPRN on CE-2 is similar.

To check the static route status:

```
*A:PE-1# show router 1 static-route

===============================================================================
Static Route Table (Service: 1)  Family: IPv4
===============================================================================
Prefix                                        Tag         Met    Pref Type Act
   Next Hop                                    Interface
-------------------------------------------------------------------------------
172.16.2.1/32                                 0           1      5    NH   Y
   10.0.0.2                                    int-gre-tunnel
-------------------------------------------------------------------------------
No. of Static Routes: 1
===============================================================================
```

IP/GRE tunneling via BGP peering

In this section, the configuration has BGP running inside the GRE tunnel.

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            autonomous-system 64496
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                loopback
            exit
            interface "loopback2" create
                address 172.31.1.1/24
                loopback
            exit
            static-route-entry 172.16.2.1/32
                next-hop 10.0.0.2
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            bgp
                group "group-1"
                    type internal
                    export "export-bgp-172.31"
                    local-address 172.16.1.1
                    neighbor 172.16.2.1
                    exit
                exit
                no shutdown
            exit
            no shutdown
```

It is mandatory to configure the autonomous system in the **vprn** context, otherwise the BGP session will not be established.

The configuration of the VPRN on CE-2 is similar.

The following command on PE-1 shows the summary of the BGP sessions. The BGP session between peers 172.16.1.1 in VPRN 1 on PE-1 and 172.16.2.1 in VPRN 1 on CE-2 is established for address family IPv4.

```
*A:PE-1# show router 1 bgp summary all

===============================================================================
BGP Summary
===============================================================================
Legend : D - Dynamic Neighbor
===============================================================================
Neighbor
Description
ServiceId          AS PktRcvd InQ  Up/Down   State|Rcv/Act/Sent (Addr Family)
                      PktSent OutQ
-------------------------------------------------------------------------------
172.16.2.1
1               64496       7    0 00h01m12s 1/1/1 (IPv4)
                            7    0
-------------------------------------------------------------------------------
```

In this example, PE-1 exports BGP route 172.31.1.0/24 and CE-2 exports BGP route 172.31.2.0/24. The route table for VPRN 1 on PE-1 includes the following BGP route:

```
*A:PE-1# show router 1 route-table protocol bgp

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                  Metric
-------------------------------------------------------------------------------
172.31.2.0/24                                 Remote  BGP       00h00m45s  170
       10.0.0.2                                                   1
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

IP/GRE tunneling via OSPFv2 peering

OSPF can be run on IES and VPRN IP interfaces associated with private IP/GRE tunnel SAPs.

All OSPF features are supported, including area 0 and non-area 0 support, virtual links, authentication, BFD, configurable protocol timers.

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                loopback
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            ospf
                area 0.0.0.0
                    interface "int-gre-tunnel"
                    exit
                    interface "loopback1"
                    exit
                exit
                no shutdown
            exit
            no shutdown
        exit
```

The configuration on CE-2 is similar.

The following command shows the OSPF neighbors for VPRN 1:

```
*A:PE-1# show router 1 ospf neighbor

===============================================================================
Rtr vprn1 OSPFv2 Instance 0 Neighbors
===============================================================================
Interface-Name                   Rtr Id          State      Pri  RetxQ   TTL
   Area-Id
-------------------------------------------------------------------------------
int-gre-tunnel                   192.0.2.2       Two Way    1    0       32
   0.0.0.0
-------------------------------------------------------------------------------
No. of Neighbors: 1
===============================================================================
```

The OSPF routes in the routing table of VPRN 1 are as follows:

```
*A:PE-1# show router 1 route-table protocol ospf

===============================================================================
Route Table (Service: 1)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
172.16.2.1/32                                 Remote  OSPF      00h00m22s  10
       10.0.0.2                                                     2
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
```

```
        B = BGP backup route available
        L = LFA nexthop available
        S = Sticky ECMP requested
===============================================================================
```

IP/GRE tunneling protection using IPSec tunnel mode

To provide protection against potential threats such as spoofing, the GRE packets can be encrypted and authenticated using IPSec.

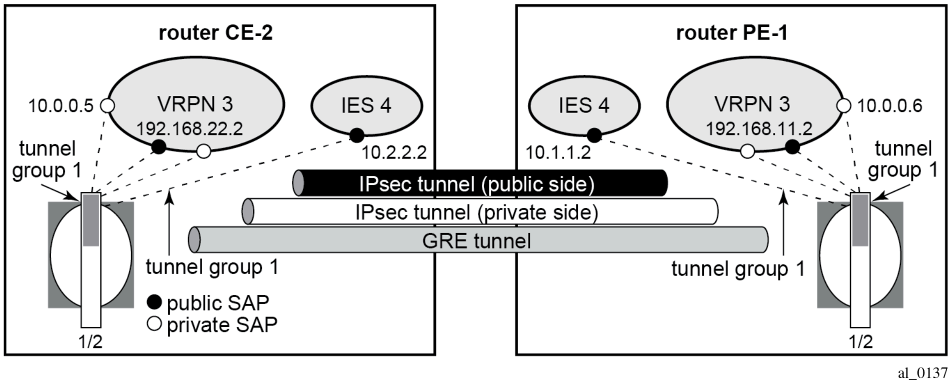
GRE packets receive IPSec protection by forwarding them, after encapsulation by a tunnel-ISA, into an IPSec tunnel supported by another (or the same) tunnel ISA.

Note that when configuring GRE protection by an IPSec tunnel:

- A GRE tunnel and its protecting IPSec tunnel may belong to the same or different tunnel groups (the same tunnel group is assumed in the following example).
- A GRE tunnel and its protecting IPSec tunnel may be assigned to the same tunnel ISA (if they belong to the same tunnel group) or different tunnel ISAs.
- A single IPSec tunnel can protect one or more GRE tunnels in addition to other IP traffic that meets the IPSec security policy.
- The private IPSec tunnel SAP interface and public GRE tunnel SAP interface are always part of the same VPRN. The private GRE tunnel SAP interface can be part of this same VPRN or a different VPRN.

In the following example, the GRE tunnel and its protecting IPSec tunnel belong to the same tunnel group.

Figure 72: Example GRE over IPSec tunnel



IPSec configuration

An **ike-policy** and **ipsec-transform** must be configured on PE-1 and CE-2, as follows:

```
# on PE-1, CE-2:
configure
    ipsec
        ike-transform 1 create
            dh-group 5
        exit
        ike-policy 1 create
            ike-transform 1
        exit
        ipsec-transform 1 create
            esp-encryption-algorithm aes256
        exit
```

The public/private side of the GRE tunnel and the private side of the IPSec tunnel are in the same VPRN, as shown in the following configuration example:

```
# on PE-1:
configure
    service
        vprn 3 name "VPRN 3" customer 1 create
            ipsec
                security-policy 1 create
                    entry 1 create
                        local-ip 192.168.11.0/24
                        remote-ip 192.168.22.0/24
                    exit
                exit
            exit
            interface "int-private-ipsec-1" tunnel create
                sap tunnel-1.private:3 create
                    ipsec-tunnel "ipsec-tunnel-for-gre-tunnel" create
                        security-policy 1
                        local-gateway-address 10.1.1.1 peer 10.2.2.1 delivery-service 4
                        dynamic-keying
                            ike-policy 1
                            pre-shared-key "pass"
                            transform 1
                        exit
                        no shutdown
                    exit
                exit
            exit
            interface "int-public-gre-1" create
                address 192.168.11.2/24
                sap tunnel-1.public:4 create
                exit
            exit
            interface "int-private-gre-1" tunnel create
                address 10.0.0.6/30
                sap tunnel-1.private:5 create
                    ip-tunnel "protected-gre-tunnel" create
                        dest-ip 10.0.0.5
                        gre-header
                        source 192.168.11.1
                        remote-ip 192.168.22.1
                        delivery-service 3
                        no shutdown
                    exit
                exit
            exit
            static-route-entry 192.168.22.0/24
                ipsec-tunnel "ipsec-tunnel-for-gre-tunnel"
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:3
                    vrf-target target:64496:3
                    no shutdown
                exit
            exit
            no shutdown
        exit
```

The following displays a configuration example of the public side of the IPSec tunnel:

```
# on PE-1:
configure
    service
        ies 4 name "IES 4" customer 1 create
            interface "public-ipsec-1" create
                address 10.1.1.2/24
                tos-marking-state untrusted
                sap tunnel-1.public:3 create
                exit
            exit
            interface "int2-PE-1-CE-2" create
                address 192.168.112.1/30
                sap 1/1/2:4 create
                exit
            exit
            no shutdown
        exit
```

The following static route is configured in the base router on PE-1:

```
# on PE-1:
configure
    router Base
        static-route-entry 10.2.2.0/24
            next-hop 192.168.112.2
                no shutdown
```

The configuration is similar on CE-2.

The following command shows that the tunnel "protected-gre-tunnel" with SAP tunnel-1.private:5 is up:

```
*A:PE-1# show ip tunnel

===============================================================================
IP Tunnels
===============================================================================
TunnelName                       SapId                          SvcId      Admn
 Local Address                                                  DlvrySvcId Oper
  OperRemoteAddress
-------------------------------------------------------------------------------
gre-tunnel-1                     tunnel-1.private:1             1          Up
 192.168.1.1                                                    2          Up
  192.168.2.1
protected-gre-tunnel             tunnel-1.private:5             3          Up
 192.168.11.1                                                   3          Up
  192.168.22.1
-------------------------------------------------------------------------------
IP Tunnels: 2
===============================================================================
```

The following command shows the IP/GRE tunnel information for this IPSec tunnel:

```
*A:PE-1# show ip tunnel "protected-gre-tunnel"
```

```
===============================================================================
IP Tunnel Configuration Detail
===============================================================================
Service Id       : 3                     Sap Id           : tunnel-1.private:5
Tunnel Name      : protected-gre-tunnel
Description      : None
GRE Header       : Yes
Delivery Service : 3
GRE Keys Set     : False
GRE Send Key     : N/A                   GRE Receive Key  : N/A
Admin State      : Up                    Oper State       : Up
Source Address   : 192.168.11.1
Remote Address   : 192.168.22.1
Backup Address   : (Not Specified)
Oper Remote Addr : 192.168.22.1
DSCP             : None
Reassembly       : inherit
Clear DF Bit     : false                 IP MTU           : max
Encap IP MTU     : max
Pkt Too Big      : true
Pkt Too Big Num  : 100                   Pkt Too Big Intvl: 10 secs
Frag Required    : true
Frag Req Count   : 100                   Frag Req Interval: 10 secs
Propagate IPv6 P*: true
Propagate IPv4 P*: true
Oper Flags       : None
Transport Profile: (Not Specified)
Last Oper Changed: 05/12/2022 08:46:26
Host ISA         : 1/2
TCP MSS Adjust
    Public       : Disabled
    Private      : Disabled

-------------------------------------------------------------------------------
Target Address Table
-------------------------------------------------------------------------------
Destination IP                          IP Resolved Status
-------------------------------------------------------------------------------
10.0.0.5                                Yes
-------------------------------------------------------------------------------

===============================================================================
IP Tunnel Statistics: protected-gre-tunnel
===============================================================================
Errors Rx        : 0                     Errors Tx        : 0
Pkts Rx          : 0                     Pkts Tx          : 0
Bytes Rx         : 0                     Bytes Tx         : 0
Key Ignored Rx   : 0                     Too Big Tx       : 0
Seq Ignored Rx   : 0
Vers Unsup. Rx   : 0
Invalid Chksum Rx: 0
Key Mismatch Rx  : 0
===============================================================================

===============================================================================
Fragmentation Statistics
===============================================================================
Encapsulation Overhead                 : 24
Temporary Private MTU                  : max
Pre-Encapsulation
    Fragmentation Count                : 0
    Last Fragmented Packet Size        : 0
Post-Encapsulation
```

```
    Fragmentation Count                : 0
    Last Fragmented Packet Size        : 0
===============================================================================
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

By default, the IPSec tunnel is down if it is not used by any traffic, as follows:

```
*A:PE-1# show ipsec tunnel

===============================================================================
IPsec Tunnels
===============================================================================
TunnelName                       LocalAddress      SvcId        Admn   Keying
  SapId                            RemoteAddress     DlvrySvcId   Oper   Sec
                                                                         Plcy
-------------------------------------------------------------------------------
ipsec-tunnel-for-gre-tunnel      10.1.1.1          3            Up     Dynamic
  tunnel-1.private:3               10.2.2.1          4            Down   1
-------------------------------------------------------------------------------
IPsec Tunnels: 1
===============================================================================
```

After it is used by traffic, the status will be changed to be up.

```
*A:PE-1# ping router 3 10.0.0.5
PING 10.0.0.5 56 data bytes
64 bytes from 10.0.0.5: icmp_seq=1 ttl=64 time=1.42ms.
64 bytes from 10.0.0.5: icmp_seq=2 ttl=64 time=1.35ms.
64 bytes from 10.0.0.5: icmp_seq=3 ttl=64 time=1.26ms.
64 bytes from 10.0.0.5: icmp_seq=4 ttl=64 time=1.34ms.
64 bytes from 10.0.0.5: icmp_seq=5 ttl=64 time=1.28ms.

---- 10.0.0.5 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 1.26ms, avg = 1.33ms, max = 1.42ms, stddev = 0.058ms
```

The IPSec tunnel is now up, as follows:

```
*A:PE-1# show ipsec tunnel

===============================================================================
IPsec Tunnels
===============================================================================
TunnelName                       LocalAddress      SvcId        Admn   Keying
  SapId                            RemoteAddress     DlvrySvcId   Oper   Sec
                                                                         Plcy
-------------------------------------------------------------------------------
ipsec-tunnel-for-gre-tunnel      10.1.1.1          3            Up     Dynamic
  tunnel-1.private:3               10.2.2.1          4            Up     1
-------------------------------------------------------------------------------
IPsec Tunnels: 1
===============================================================================
```

BFD support on private tunnel interfaces

BFD is supported on IP interfaces associated with private IP/GRE tunnel SAPs. The BFD state of the interface can be used by static routes, OSPFv2, or BGP configured on the interface. It is not used to trigger a switchover to the backup remote IP address of the GRE tunnel.

The following displays a static route example:

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                bfd 100 receive 100 multiplier 3
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no dscp
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                loopback
            exit
            static-route-entry 172.16.2.1/32
                next-hop 10.0.0.2
                    bfd-enable
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            no shutdown
```

The following command shows that the BFD session on interface "int-gre-tunnel" is up for protocol static:

```
*A:PE-1# show router 1 bfd session

===============================================================================
Legend:
  Session Id = Interface Name | LSP Name | Prefix | RSVP Sess Name | Service Id
  wp = Working path   pp = Protecting path
===============================================================================
BFD Session
===============================================================================
Session Id                                        State      Tx Pkts    Rx Pkts
  Rem Addr/Info/SdpId:VcId                      Multipl     Tx Intvl   Rx Intvl
  Protocols                                        Type     LAG Port     LAG ID
  Loc Addr                                                             LAG name
```

```
-------------------------------------------------------------------------------
int-gre-tunnel                                       Up          N/A        N/A
  10.0.0.2                                            3         1000       1000
  static                                         cpm-np          N/A        N/A
  10.0.0.1
-------------------------------------------------------------------------------
No. of BFD sessions: 1
===============================================================================
```

When no static routes are configured and OSPF is configured instead, the configuration of VPRN 1 on PE-1 is as follows:

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                bfd 100 receive 100 multiplier 3
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                loopback
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            ospf
                area 0.0.0.0
                    interface "int-gre-tunnel"
                        bfd-enable
                        no shutdown
                    exit
                    interface "loopback1"
                        no shutdown
                    exit
                exit
                no shutdown
            exit
            no shutdown
```

The following shows that the BFD session is up for protocol OSPF on interface "int-gre-tunnel":

```
*A:PE-1# show router 1 bfd session

===============================================================================
Legend:
  Session Id = Interface Name | LSP Name | Prefix | RSVP Sess Name | Service Id
  wp = Working path   pp = Protecting path
===============================================================================
BFD Session
===============================================================================
Session Id                                        State      Tx Pkts    Rx Pkts
  Rem Addr/Info/SdpId:VcId                      Multipl     Tx Intvl   Rx Intvl
  Protocols                                        Type     LAG Port     LAG ID
  Loc Addr                                                             LAG name
-------------------------------------------------------------------------------
int-gre-tunnel                                       Up          N/A        N/A
  10.0.0.2                                            3         1000       1000
  ospf2                                          cpm-np          N/A        N/A
  10.0.0.1
-------------------------------------------------------------------------------
No. of BFD sessions: 1
===============================================================================
```

When BGP is configured instead of OSPF, the configuration of VPRN 1 on PE-1 is as follows:

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            autonomous-system 64496
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                bfd 100 receive 100 multiplier 3
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        no shutdown
                    exit
                exit
            exit
            interface "loopback1" create
                address 172.16.1.1/32
                bfd 100 receive 100 multiplier 3
                loopback
            exit
            static-route-entry 172.16.2.1/32
                next-hop 10.0.0.2
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:1
                    vrf-target target:64496:1
                    no shutdown
                exit
            exit
            bgp
                group "group-1"
                    type internal
                    local-address 172.16.1.1
                    neighbor 172.16.2.1
                        bfd-enable
                    exit
                exit
                no shutdown
            exit
            no shutdown
```

The following command shows that the BFD session is up for protocol BGP on interface "loopback1":

```
*A:PE-1# show router 1 bfd session

===============================================================================
Legend:
  Session Id = Interface Name | LSP Name | Prefix | RSVP Sess Name | Service Id
  wp = Working path   pp = Protecting path
===============================================================================
BFD Session
===============================================================================
Session Id                                        State      Tx Pkts    Rx Pkts
  Rem Addr/Info/SdpId:VcId                      Multipl     Tx Intvl   Rx Intvl
  Protocols                                        Type     LAG Port     LAG ID
  Loc Addr                                                             LAG name
-------------------------------------------------------------------------------
loopback1                                            Up          N/A        N/A
  172.16.2.1                                          3         1000       1000
  bgp                                            cpm-np          N/A        N/A
  172.16.1.1
-------------------------------------------------------------------------------
No. of BFD sessions: 1
===============================================================================
```

IP/GRE termination – Advanced topics

DSCP value of outer delivery header

- Default behavior — The DSCP value from the payload header is copied into the outer GRE header. This is a one to one copy and no QoS classifications are required. It is performed when no DSCP value is configured under the private GRE tunnel.
- Non default behavior — DSCP is configured under the private SAP (following example using DSCP af41).

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel" tunnel create
                address 10.0.0.1/30
                sap tunnel-1.private:1 create
                    ip-tunnel "gre-tunnel-1" create
                        dest-ip 10.0.0.2
                        gre-header
                        source 192.168.1.1
                        remote-ip 192.168.2.1
                        delivery-service 2
                        dscp af41
                        no shutdown
                    exit
                exit
---snip---
```

The log filter output showsTOS=88 (DSCP=af41) in the public network.

```
*A:PE-1# show filter log 102

===============================================================================
Filter Log
===============================================================================
Admin state : Enabled
Description : (Not Specified)
Destination : Memory
Wrap        : Enabled
-------------------------------------------------------------------------------
Maximum entries configured : 1000
Number of entries logged   : 5
-------------------------------------------------------------------------------
2022/05/12 08:54:19  Ip Filter: 2:10  Desc:
SAP: tunnel-1.private:1  Direction: Egress
Src MAC: 02-0f-ff-00-02-c9  Dst MAC: 00-00-00-07-a0-bd  EtherType: 0800
Src IP: 10.0.0.1  Dst IP: 10.0.0.2  Flags: 0  TOS: 88  TTL: 64 Len: 84
Protocol: ICMP  Type: Echo Request  Code: 0
---snip---
```

IP MTU

It is possible to configure the IP MTU of a private tunnel SAP interface. This sets the maximum IP packet size payload (including IP header) that can be sent into the tunnel (it applies to the packet size before the tunnel encapsulation is added).

```
# on PE-1:
configure
    service
        vprn 1 name "VPRN 1" customer 1 create
            interface "int-gre-tunnel"
                ip-mtu 1476
            ---snip---
```

When an IPv4 packet needs to be forwarded to the tunnel and is larger than IP MTU bytes:

- If the DF bit is clear, the payload packet is IP fragmented to the MTU size prior to tunnel encapsulation.
- If the DF bit is set, the payload packet is discarded.

The IP MTU range supported is from 512 to 9000 bytes.

The following command shows the configured IP MTU and the operational IP MTU for the GRE tunnel:

```
*A:PE-1# show router 1 interface "int-gre-tunnel" detail | match MTU
IP MTU           : 1476
IP Oper MTU      : 1476
```

Statistics and accounting

Collect-stats can be configured under public and private SAPs.

For public SAPs:

```
# on PE-1:
```

```
configure
    service
        ies "IES 2"
            interface "int-tunnel-public"
                sap tunnel-1.public:1
                    collect-stats
                exit
            exit
```

For private SAPs:

```
# on PE-1:
configure
    service
        vprn "VPRN 1"
            interface "int-gre-tunnel"
                sap tunnel-1.private:1
                    collect-stats
                exit
            exit
```

Filtering, policing, and QoS

An IP filter and QoS policy can be applied to the ingress and egress traffic of the private and public SAPs.

Public SAPs:

```
# on PE-1:
configure
    service
        ies "IES 2"
            interface "int-tunnel-public"
                sap tunnel-1.public:1
                    ingress
                        qos 10
                        filter ip 1
                    exit
                    egress
                        qos 20
                        filter ip 2
                    exit
                exit
            exit
```

Private SAPs:

```
# on PE-1:
configure
    service
        vprn "VPRN 1"
            interface "int-gre-tunnel"
                sap tunnel-1.private:1
                    ingress
                        qos 10
                        filter ip 1
                    exit
                    egress
                        qos 20
                        filter ip 2
                    exit
                exit
                ---snip---
```

Mirroring

The public and private SAPs can be mirrored.

```
# on PE-1:
debug
    mirror-source 99
        sap tunnel-1.private:3 egress ingress
        sap tunnel-1.public:1 egress ingress
        no shutdown
    exit
exit
```

Conclusion

This chapter provides configuration and show commands for IP/GRE termination.

L2-aware NAT (with dNAT and MNPs)

This chapter provides information about Network Address Translation (NAT) in combination with Enhanced Subscriber Management (ESM).

Topics in this chapter include:

Applicability

The information and configuration in this chapter are applicable to all SR OS nodes that support Broadband (BB) MS-ISAs, including the virtual version of MS-ISA in VSR, and was tested on SR OS 15.0.R4.

Overview

L2-aware network address translation (NAT) is an enhanced NAT functionality that offers better IPv4 address conservation ratios than traditional NAT44. L2-aware NAT relies on tight integration between enhanced subscriber management (ESM) and NAT, and in the context of ESM, it maps (or binds) an ESM subscriber to an outside IP address and a single port block (PB). In this fashion, the private (inside) IPv4 address in L2-aware NAT is abstracted from the binding.

Note:
Binding is distinct from the flow/session concept. Flow or session state is maintained for each specific conversation between the two end nodes (this includes IP addresses, ports, and protocol), whereas binding is mapping between an inside entity (the subscriber ID in L2-aware NAT or the IPv4 source IP address in large scale NAT44 (LSN44)) and the outside IP address and the PB. Flows/session can then consume ports within the PB.

On the public side (outside), L2-aware NAT allows multiple hosts (inside IPv4 addresses) under the same subscriber to share the same outside IPv4 address and a PB. This contrasts with traditional NAT44 where each inside IPv4 address (host) is mapped to a unique IPv4 address and PB combination.

On the private side (inside), the abstraction of the host IPv4 address from the binding allows assignment of the same inside IPv4 address to multiple hosts that belong to different subscribers. This sharing of an inside IPv4 address between hosts that belong to different subscribers means that each host still has its own instance of the shared IPv4 address. For example, host 1 of subscriber 1 and host 2 of subscriber 2 can be both assigned the same inside IP address (for instance 10.10.1.1).

A binding in L2-aware NAT is defined as the following:

```
<subId, nat-policy> → <outside routing context, outside srcIP address, outside
PB>
```

Where:

Inside

`Sub-id`

Subscriber ID in the ESM context.

`Nat-policy`

NAT policy associated with the ESM subscriber. This association is performed within the sub-profile. A subscriber can have multiple NAT policies (MNPs), in which case each subscriber will have one binding per NAT policy.

Outside

`Outside routing context`

Outside routing context that contains an L2-aware pool. NAT traffic is sent out of the node via this routing context.

`Outside SrcIP address`

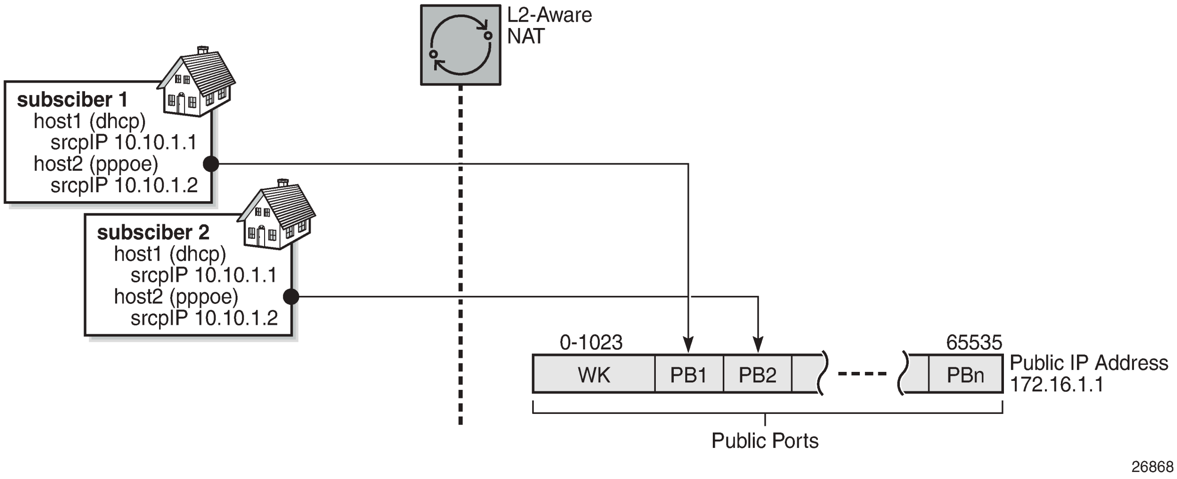
A subscriber in conjunction with the NAT policy is assigned this outside IP address. The source IP address of the subscriber traffic will be replaced by this IP address.

`Outside PB`

A subscriber in conjunction with the NAT policy is assigned this outside PB. A source (TCP/UDP) port of the subscriber traffic will be replaced by one of the source ports from this PB. In the ICMP case, the query ID is selected from this PB.

A simplified example of an L2-aware NAT binding is shown in Figure 73: NAT Binding. Subscriber 1 is mapped to the outside IP address 172.16.1.1 and PB 1. This means that the source IP addresses (10.10.1.1 and 10.10.1.2) for all hosts under Subscriber 1 will be translated to 172.16.1.1 and the source port will be translated to one of the ports from PB 1. Similar logic can be applied to Subscriber 2 mappings, where Subscriber 2 hosts will use PB 2 from the same outside IP address.

Figure 73: NAT Binding



NAT binding in MS-ISA will be used to translate traffic arriving from the subscriber. When the first packet of the flow from the subscriber is received, the translation is performed and a flow state is created. This flow state is maintained for the duration of the flow. All consecutive packets of the flow will rely on the flow state lookup for translation.

The flow mapping in L2-aware NAT is characterized as:

```
<subId, inside srcIP address, inside srcPort, foreign ip address, foreign port, protocol> →
 <outside routing context, outside IP address, outside port, destination ip address, foreign
 port, protocol>
```

Flow mapping is more specific than a binding (which is a partial mapping), and the fields in flow mapping are self-explanatory. The flow mapping contains the original source IP address, which is needed to determine where to send the return (downstream) traffic.

The subId field needs to be unique in the mappings at the subscriber level, allowing all the other fields (including the source IP address) to overlap between the subscribers. This is the basis for sharing of the source IP addresses between the subscribers.

In flow mapping, there is a naming distinction between the foreign IP address and the destination IP address. A foreign IP address is the original destination IP address in the packet sourced by the host. This foreign IP address can be translated by destination-based NAT (dNAT), just like source IP is translated by NAT. In this case, the translated foreign IP address on the outside in the SR OS CLI syntax becomes a destination IP address. This naming distinction helps to differentiate two IP addresses of the same field in the IP header, one before dNAT is performed and the other after dNAT is performed.

The reason why there is only a foreign port (and no destination port) is that in SR OS the original destination port is not translated by dNAT, while the original destination IP is.

L2-aware NAT requires that ESM and NAT are collocated in the same SR OS node. Because of the tight integration between ESM and NAT, logging of L2-aware NAT resources can be integrated into ESM accounting via AAA.

This chapter does not discuss L2-aware NAT intra-chassis redundancy. Intra-chassis redundancy for L2-aware NAT offers protection against MS-ISA failure. The supported modes are:

- Active/Standby MS-ISA

- Active/Active MS-ISA
- L2-aware Bypass

Each L2-aware NAT intra-chassis redundancy mode is described in the *7450 ESS, 7750 SR, and VSR Multiservice ISA and ESA Guide*.

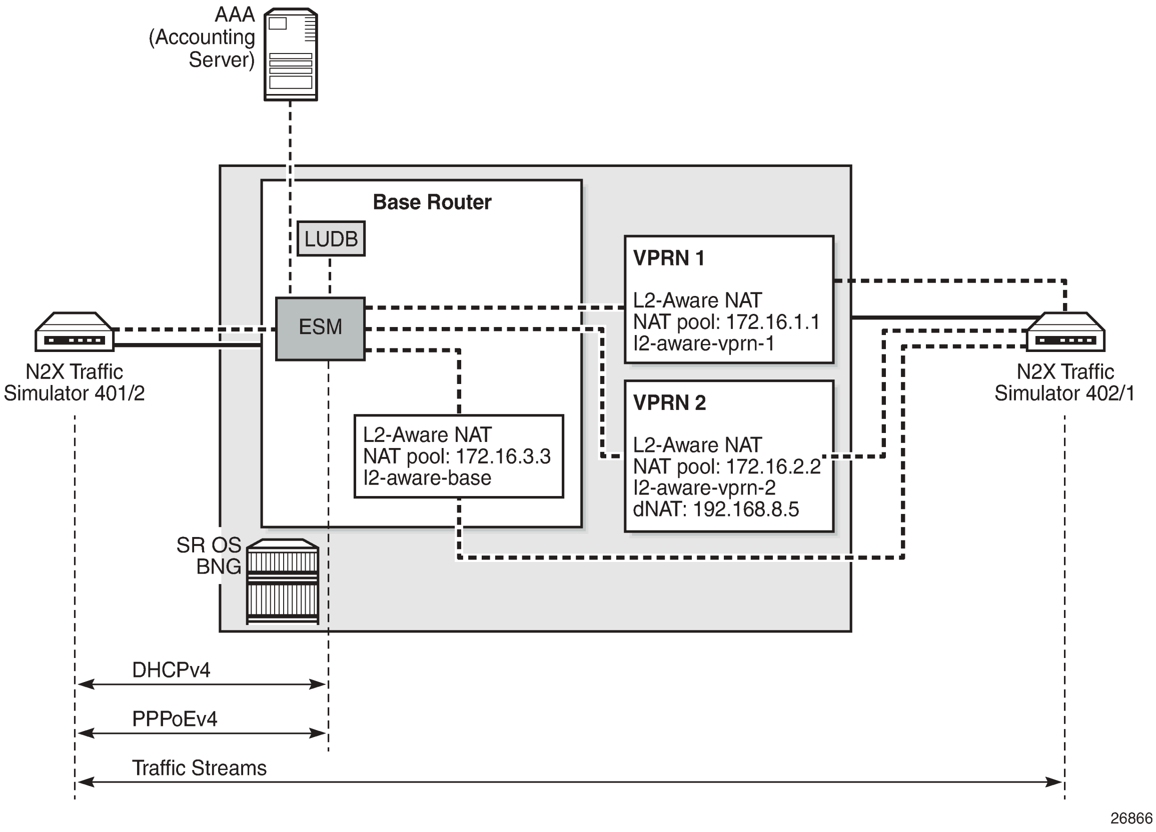
Configuration

The test setup in this configuration example will accommodate:

- Two L2-aware subscribers.
- Two hosts per L2-aware subscriber – one DHCPv4 host and one PPPoEv4 host per L2-aware subscriber.
- IP addresses overlap between the subscribers.
- Traffic split from a single host to three different NAT pools and outside routing contexts. This is achieved by selecting a NAT policy based on the destination IP address of the traffic.
- Destination-based NAT (dNAT) – some of the traffic is subjected to dNAT, where the destination IP address in the outgoing packet is translated (in addition to the source IP address and source port).

The test setup is shown in Figure 74: Test Setup.

Figure 74: Test Setup



The N2X traffic simulator on the left is used to simulate two homes (subscribers) with DHCPv4 and PPPoE hosts. Those hosts are terminated on an IES interface in a Base router in the BNG and they are further associated with the L2-aware NAT pools in VPRN1, VPRN2, and the Base router. ESM host authentication and IP address assignment is performed via local user database (LUDB). Accounting for the subscribers is sent to an AAA server and the accounting records contain NAT logging information.

Once the subscriber hosts are instantiated, traffic is sent through L2-aware NAT. This traffic is then used to further examine operation of L2-aware NAT in SR OS BNG.

The L2-aware NAT configuration in SR OS is split into two major parts:

- NAT-specific configuration – this is the subject of this document.
- ESM-specific configuration focuses on ESM subscribers and is described in other chapters in this volume. There are some key parts of ESM configuration that are essential for understanding L2-aware NAT operation and those parts will be described in this section.

  For the sake of completeness, the remaining, more generic part of ESM configuration is described in Appendix A – Generic ESM Configuration.

NAT-specific Configuration

L2-aware NAT requires the system IP address to be configured in the node. Without the system IP address, L2-aware NAT will not be operational. Lack of the system IP address within the L2-aware node is a common problem in most troubleshooting scenarios. The system IP is used within the system to pass traffic between the carrier IOM (where MS-ISA is attached) and the MS-ISA itself. The following is the command to configure the system IP address:

```
configure
    router
        interface "system"
            address 192.0.2.2/32
            no shutdown
        exit
```

The **inside** NAT context in the following code identifies the private (inside) side of NAT, which in this example is configured in the Base routing context.

The address in the **l2-aware** context (in this example, address 10.10.1.254/24) represents the default gateway and an L2-aware subnet. All ESM hosts with IP addresses within the configured L2-aware subnet (in this example, 10.10.2.0/24) are subjected to L2-aware NAT.

ESM hosts with IP addresses out of the L2-aware subnet can be instantiated as regular (not L2-aware) hosts, effectively bypassing L2-aware NAT within the L2-aware enabled subscriber.

The **outside** NAT context contains the NAT pool configuration that is used for address translation. In this example, traffic is mapped to three outside contexts: Base, VPRN 1, and VPRN 2. IP address 172.16.3.3 is used for translation in the Base outside routing context, IP address 172.16.1.1 in the VPRN 1 outside routing context, and IP address 172.16.2.2 in the VPRN 2 outside routing context.

VPRN 1 and VPRN 2 also contain NAT pools with their own IP addresses.

```
configure
    router
        nat
            inside
                l2-aware
                    address 10.10.1.254/24
                exit
            exit
            outside
                pool "l2-aware-base" nat-group 1 type l2-aware create
                    port-reservation ports 2000
                    address-range 172.16.3.3 172.16.3.3 create
                    exit
                    no shutdown
                exit
            exit
        exit
```

```
configure
    service
        vprn 1 name "1" customer 1 create
            nat
                outside
                    pool "l2-aware-vprn-1" nat-group 1 type l2-aware create
                        port-reservation ports 2000
                        address-range 172.16.1.1 172.16.1.1 create
                        exit
                        no shutdown
                    exit
                exit
            exit
         exit
         vprn 2 name "2" customer 1 create
             nat
                 outside
                     port-reservation ports 2000
                     pool "l2-aware-vprn-2" nat-group 1 type l2-aware create
                         address-range 172.16.2.2 172.16.2.2 create
                         exit
                         no shutdown
                     exit
                 exit
             exit
         exit
```

NAT pool selection (and with this, the outside routing context) is performed based on the foreign IP address (traffic destination) in the packet. This is configured in the **nat-prefix-list**, which is then applied to the L2-aware subscriber via **sub-profile**.

Packets arriving from the inside and destined to networks 192.168.7.0/24 and 192.168.8.0/24 will be respectively routed to the VPRN 1 and VPRN 2 routing context, which are implied through NAT policies “l2-aware-vrf1” and “l2-aware-vrf2” referenced in the **nat-prefix-list**.

All other traffic will use the default NAT policy “l2-aware-base”, which is directly referenced in the **sub-profile**. This NAT policy is pointing to the Base outside routing context. Also, traffic identified by NAT classifier “vrf2” will be subject to dNAT.

In the example, NAT classifier “vrf2” identifies UDP traffic destined to port 5001. This implies that destination IP address for traffic that is mapped to VPRN 2 with UDP destination port 5001 will be translated. The new destination IP address (192.168.8.5) is configured in the NAT classifier “vrf2”, which is then applied via NAT policy “l2-aware-vrf2”.

All other traffic (destined to a destination UDP port other than 5001) will pass transparently through the classifier without the destination IP address being modified. Traffic that is subject to dNAT can be identified (or classified) based on destination-port only.

The **block-limit** command in L2-aware pool is automatically set to “1”. That is, each L2-aware subscriber is assigned exactly one PB. Additional PB allocation is not supported. That is why it is important that this single PB allocated to each L2-aware subscriber is initially configured with a large enough number of ports – this is performed via pool configuration (in this example, port-block size is set to 2000 ports). Once all the ports in this PB are exhausted, no new ports can be allocated.

```
configure
    subscriber-mgmt
        sub-profile "sub-prof-1" create
            nat-policy "l2-aware-base"
            nat-prefix-list "l2aware-mnp"
        exit
```

```
configure
    service
        nat
            nat-classifier "vrf2" create
                entry 1 create
                    action dnat ip-address 192.168.8.5
                    match protocol udp
                        dst-port-range start 5001 end 5001
                    exit
                exit
            exit
            nat-policy "l2-aware-base" create
                block-limit 1
                pool "l2-aware-base" router Base
            exit
            nat-policy "l2-aware-vrf1" create
                block-limit 1
                pool "l2-aware-vprn-1" router 1
            exit
            nat-policy "l2-aware-vrf2" create
                dnat
                    nat-classifier "vrf2"
                exit
            exit
            nat-prefix-list "l2aware-mnp" application l2-aware-dest-to-policy create
                prefix 192.168.7.0/24 nat-policy "l2-aware-vrf1"
                prefix 192.168.8.0/24 nat-policy "l2-aware-vrf2"
            exit
        exit
```

On the ESM side, **anti-spoof** configuration is of particular importance in L2-aware NAT.

- L2-aware deployments with bridged homes (multiple ESM hosts within a subscriber) must have the anti-spoofing option set to **ip-mac**. This will allow each host within the home to be distinctly identified in ESM and consequently receive proper ESM treatment (filters, QoS, accounting, and so on).
- L2-aware deployments with routed CPEs must have the anti-spoofing option set to **nh-mac**. In this case, only one host per ESM subscriber is allowed and this single host is the routed CPE. There can be multiple hosts behind this routed CPE and they will be all seen and properly treated (using NAT) by the MS-ISA. However, ESM will not be aware of them and all downstream traffic will be sent to the routed CPE (ESM host), which will then further route traffic to the proper destination.

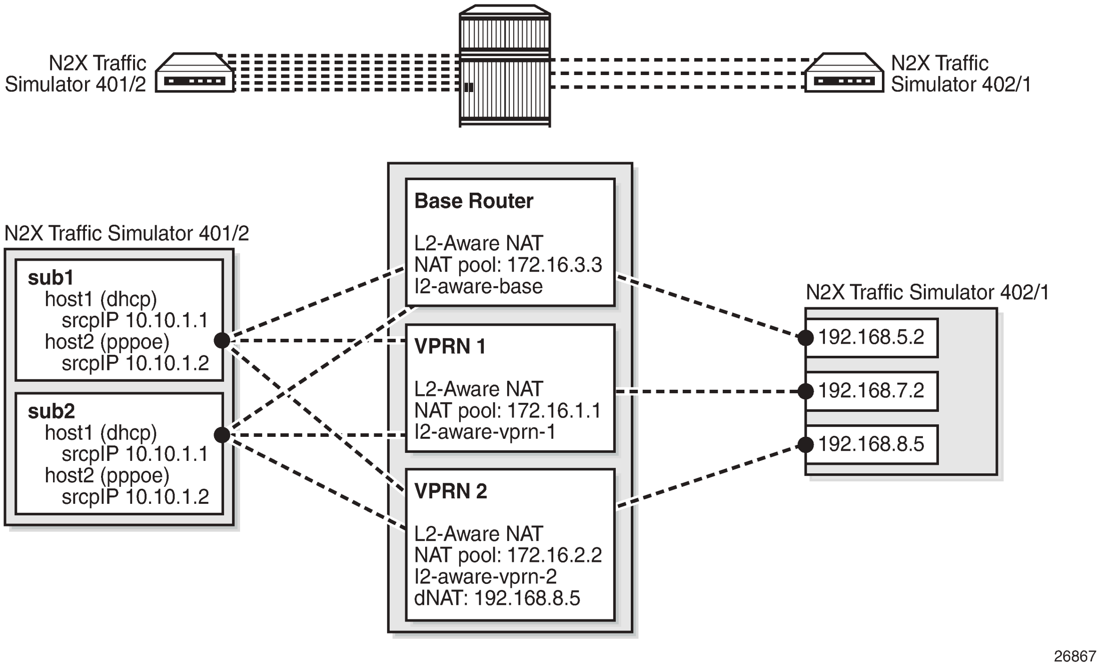
  From an ESM perspective, this approach is different from that for non-L2-aware ESM deployments. Non-L2-aware ESM deployments allow multiple ESM hosts per subscriber with the **nh-mac** anti-spoofing option enabled, where the ESM hosts have different MACs.

```
configure
```

```
    subscriber-mgmt
        msap-policy "msaps" create
            sub-sla-mgmt
                sub-ident-policy "sub_ident_pol"
                multi-sub-sap limit 10
            exit
            ies-vprn-only-sap-parameters
                anti-spoof ip-mac
            exit
        exit
```

This preceding described configuration is shown in Figure 75: Logical Mapping of Subscribers to L2-aware Pool.

Figure 75: Logical Mapping of Subscribers to L2-aware Pool



Logging

In this example, NAT logging is integrated with ESM accounting. This is configuration dependent and L2-aware NAT also supports syslog to convey information, about a subscriber that is hidden behind NAT, to the operator.

With RADIUS logging, NAT-specific information is carried in the **Alc-Nat-Port-Range** VSA, which must be explicitly enabled in **accounting-policy**. The ESM accounting model used in this example is per host-accounting with interim updates. For brevity, only the accounting records for host 1 of the subscriber 1 are shown following:

- Acct *Start* is generated when the ESM host is created. The first host for the subscriber will allocate NAT resources in the MS-ISA.

- Acct *Interim-Update* message is generated periodically (configuration dependent) and additionally carries information about consumed bytes flowing through queues/policers associated with the subscriber host.
- Acct *Stop* is generated when the ESM host is terminated. NAT resources are then released. The significance of this is that freeing the NAT resources coincides with the termination of the subscriber, without having to wait for NAT flows in the MS-ISA to time out, long after the user has ceased.

Appendix B – Logging lists a more comprehensive stream of logs for all four hosts that are established in this exercise.

```
root@linux:/var/log/freeradius/radacct/192.168.114.2# more detail-20170814
Mon Aug 14 09:14:45 2017
        Acct-Status-Type = Start
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:29:02 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502727285
        Request-Authenticator = Verified

Mon Aug 14 09:20:11 2017
        Acct-Status-Type = Interim-Update
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Session-Time = 326
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:34:28 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Alc-Acct-Triggered-Reason = regular
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x0001000000000000b220
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000001c8
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502727611
        Request-Authenticator = Verified

Mon Aug 14 09:30:34 2017
        Acct-Status-Type = Stop
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Session-Time = 949
        Acct-Terminate-Cause = User-Request
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:44:51 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x00010000000000037974
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000008e5
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502728234
        Request-Authenticator = Verified

root@linux:/var/log/freeradius/radacct/192.168.114.2#
```

Operational Commands

In this example, three streams of traffic are run from host 1 of each subscriber (six streams in total). Traffic flows in the upstream direction (from the inside routing context to the outside routing context) and it will cause L2-aware NAT to allocate resources and create mappings that can be observed with the show and debug commands.

Host 1 of each subscriber is initiating three streams to three different destinations. Accordingly, traffic from the same host will be (based on the destination) routed to three different L2-aware NAT pools and outside routing contexts. One of the streams (stream 3) will, in addition to its source IP address/port, have its destination IP address translated as well.

The six traffic stream definitions are shown in Table 11: Traffic Streams:

- All six streams have the same source IP address, 10.10.1.1. The significance of this is that both subscribers have hosts that are assigned the same inside (private) IP address.
- All streams carry UDP protocol.
- The source UDP port is 5000 for all streams.

- The UDP foreign port is 5000 except for the stream 3, which is assigned UDP foreign port 5001. Traffic with the foreign port 5001 is identified through the NAT classifier and is selected for dNAT (foreign IP address will be translated).
- Each host sends traffic to three different destinations and this is signified by three different foreign IP addresses:
  - The stream with the foreign IP address 192.168.5.2 is sent to the Base routing context (`pool l2-aware-base`).
  - The stream with the foreign IP address 192.168.7.2 is sent to the VPRN 1 routing context (`pool l2-aware-vprn-1`).
  - The stream with the foreign IP address 192.168.8.2 is sent to the VPRN 2 routing context (`pool l2-aware-vprn-2`).
- The column labelled Out IP (Outside IP address) represents a newly translated source (private) IP address by L2-aware NAT. Each of the three pools (one per routing context in Base, VPRN 1, and VPRN 2) has its own single outside IP address configured. In this example, each pool has a single outside IP address configured but, usually, multiple address ranges per NAT pool are supported.
- Each subscriber is allocated a single PB per pool with 2000 ports (configuration dependent) in each pool – this is reflected in the PB column. Subscriber 1 is allocated the same PB range in all three pools (one per routing context). A similar setup is shown for subscriber 2.

  In this example, the PBs allocated per subscriber in three pools coincidentally match due to a low number of subscribers in the system (only two subscribers in our example). Usually, with an increased number of subscribers in the system, the probability for matching PB allocations per subscriber across pools would decrease. That is, each pool allocates PB independently of any other pool.
- The Out Port column indicates a new (translated) source UDP source port, after L2-aware NAT is performed.
- The Dest IP column indicates that the foreign IP address is modified only for stream 3, which is subject to dNAT (identified by UDP foreign port 5001 via a NAT classifier). The remaining streams do not have foreign IP addresses translated.
- The Destination ports on the outside in all three streams remain unchanged. This applies even for stream 3, which is subject to dNAT (only the foreign IP address is translated by dNAT functionality).

Dynamically allocated parameters by L2-aware NAT are in the following traffic streams table (fields in bold typeface - **Out IP, PB, Out Port, Dest IP**) populated based on the observation in the system once the traffic is run. The values are collected based on the output provided by debug commands and a **tools dump nat sessions** command (both of them are shown further in the text).

Table 11: Traffic Streams

Host	Strm	Packet Fields Before NAT (inside routing context – Base)					Packet Fields After NAT (outside routing context)					
		Src IP	**Prot**	**Src Port**	**Forgn Port**	**Forgn IP**	**Outside Rtr**	**Out IP**	**PB**	**Out Port**	**Dest IP**	**Dest Port**
Sub1 Host1	1	10.10.1.1	UDP	5000	5000	192.168.5.2	Base	**172.16.3.3**	1024-3023	**2186**	192.168.5.2	5000
	2					192.168.7.2	VRF1	**172.16.1.1**	1024-3023	**1124**	192.168.7.2	
	3				5001	192.168.8.2	VRF2	**172.16.2.2**	1024-3023	**3690**	**192.168.8.5**	5001
Sub2 Host1	4	10.10.1.1	UDP	5000	5000	192.168.5.2	Base	**172.16.3.3**	3024-5023	**4470**	192.168.5.2	5000
	5					192.168.7.2	VRF1	**172.16.1.1**	3024-5023	**4666**	192.168.7.2	
	6					192.168.8.2	VRF2	**172.16.2.2**	3024-5023	**4698**	192.168.8.2	

Show Commands

A good starting point for troubleshooting is a generic NAT command, **show service nat overview**, with the output that would normally point to misconfiguration:

```
*A:BNG1# show service nat overview

===============================================================================
NAT overview
===============================================================================
Inside/        Policy/                                    Type
Outside        Pool
-------------------------------------------------------------------------------
*              l2-aware-base                              l2aware
```

```
Base           l2-aware-base

*              l2-aware-vrf1                             l2aware
vprn1          l2-aware-vprn-1

*              l2-aware-vrf2                             l2aware
vprn2          l2-aware-vprn-2
===============================================================================

No firewall policies found.
*A:BNG1#
```

The output of the show port statistics command provides information relevant to packet flow between the carrier IOM (to which the MS-ISA is attached) and the MS-ISA itself.

In the following output, "1/2/nat-in-l2" is the port on the carrier IOM that represents the private (inside) side of NAT and "1/2/nat-out-ip" is the port on the carrier IOM that represents the public (outside) side of NAT. In this example, traffic is sent only in one direction, from private side to public side, and therefore:

- Traffic is egressing port "1/2/nat-in-l2" on the IOM and entering MS-ISA on the private side (11789 packets).
- Traffic is leaving MS-ISA on the public side and ingressing port "1/2/nat-out-ip" on the IOM (11789 packets).

```
*A:BNG1# show port statistics

===============================================================================
Port Statistics on Slot 1
===============================================================================
Port                    Ingress         Ingress          Egress          Egress
Id                      Packets          Octets         Packets          Octets
-------------------------------------------------------------------------------
1/1/2                    223162        25147835               0               0
1/1/3                    815114        78685140          614895        60630323
1/1/4                     15484         1531339          230226        25728265
1/1/5                     84116         8375228            1054           69854
1/1/10                     5115          835288             345          110442
1/2/nat-out-ip            11789         1131744               0               0
1/2/nat-in-l2                 0               0           11789         1532570

===============================================================================
Port Statistics on Slot A
===============================================================================
Port                    Ingress         Ingress          Egress          Egress
Id                      Packets          Octets         Packets          Octets
-------------------------------------------------------------------------------
A/1                    81952168               0           28663         2721059
===============================================================================
```

NAT resources are allocated during the subscriber instantiation phase (when the first host for an L2-aware subscriber is created), before any data traffic from the subscriber side is even initiated.

The following command is run to verify that the two subscribers, each with two hosts (one DHCP and one PPPoE) are online. In addition, this command lists limited input related to L2-aware NAT:

- NAT policies associated with the subscriber
- Outside IP address allocated to the subscriber
- Port range in a PB

As shown in the output of this command, both L2-aware subscribers have hosts with overlapping IP addresses, which is a unique characteristic in L2-aware NAT functionality.

```
*A:BNG1# show service active-subscribers

===============================================================================
Active Subscribers
===============================================================================
-------------------------------------------------------------------------------
Subscriber sub-1 (sub-prof-1)
-------------------------------------------------------------------------------
NAT Policy: l2-aware-base
Outside IP: 172.16.3.3
Ports     : 1024-3023

NAT Policy: l2-aware-vrf1
Outside IP: 172.16.1.1 (vprn1)
Ports     : 1024-3023

NAT Policy: l2-aware-vrf2
Outside IP: 172.16.2.2 (vprn2)
Ports     : 1024-3023

-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
(1) SLA Profile Instance sap:[1/1/5:1.1] - sla:sla-1
-------------------------------------------------------------------------------
IP Address
                MAC Address          Session        Origin       Svc        Fwd
-------------------------------------------------------------------------------
10.10.1.2
                00:00:65:01:01:02    PPP 1          IPCP         3          Y
10.10.1.1
                00:00:65:01:01:01    N/A            DHCP         3          Y
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
Subscriber sub-2 (sub-prof-1)
-------------------------------------------------------------------------------
NAT Policy: l2-aware-base
Outside IP: 172.16.3.3
Ports     : 3024-5023

NAT Policy: l2-aware-vrf1
Outside IP: 172.16.1.1 (vprn1)
Ports     : 3024-5023

NAT Policy: l2-aware-vrf2
Outside IP: 172.16.2.2 (vprn2)
Ports     : 3024-5023

-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
(1) SLA Profile Instance sap:[1/1/5:1.2] - sla:sla-1
-------------------------------------------------------------------------------
IP Address
                MAC Address          Session        Origin       Svc        Fwd
-------------------------------------------------------------------------------
10.10.1.2
                00:00:65:01:02:02    PPP 1          IPCP         3          Y
10.10.1.1
                00:00:65:01:02:01    N/A            DHCP         3          Y
-------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------
Number of active subscribers : 2
===============================================================================
*A:BNG1#
```

NAT-specific **show** commands can provide more information about the L2-aware subscribers. One such command is **show service nat l2-aware-subscribers**. In this example, this command is run in a simplest form, but it can be expanded with additional filtering capabilities:

```
show service nat l2-aware-subscribers [nat-policy <policy-name>] [nat-group <nat-group-id>]
 [member <[1..255]>] [outside-router <router-instance>] [outside-ip <outside-ip-address>]
 [firewall-policy <policy-name>]
```

```
*A:BNG1# show service nat l2-aware-subscribers

===============================================================================
Layer-2-Aware NAT subscribers
===============================================================================

Subscriber                            : sub-1
-------------------------------------------------------------------------------
ISA NAT group                         : 1
ISA NAT group member                  : 1
UPnP policy                           : (None)
Default NAT policy                    : l2-aware-base
Per-host port block size              : N/A
Firewall policy                       : (None)

Policy                                : l2-aware-base
Purpose                               : nat
Outside router                        : Base
Outside IP                            : 172.16.3.3
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 1024-3023

Policy                                : l2-aware-vrf1
Purpose                               : nat
Outside router                        : vprn1
Outside IP                            : 172.16.1.1
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 1024-3023

Policy                                : l2-aware-vrf2
Purpose                               : nat
Outside router                        : vprn2
Outside IP                            : 172.16.2.2
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 1024-3023


Subscriber                            : sub-2
-------------------------------------------------------------------------------
ISA NAT group                         : 1
ISA NAT group member                  : 1
UPnP policy                           : (None)
Default NAT policy                    : l2-aware-base
Per-host port block size              : N/A
Firewall policy                       : (None)
```

```
Policy                                : l2-aware-base
Purpose                               : nat
Outside router                        : Base
Outside IP                            : 172.16.3.3
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 3024-5023

Policy                                : l2-aware-vrf1
Purpose                               : nat
Outside router                        : vprn1
Outside IP                            : 172.16.1.1
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 3024-5023

Policy                                : l2-aware-vrf2
Purpose                               : nat
Outside router                        : vprn2
Outside IP                            : 172.16.2.2
DNAT default IP address override      : (Not Specified)
DNAT disabled by override             : false
Ports                                 : 3024-5023

-------------------------------------------------------------------------------
No. of subscribers: 2

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

Host-level L2-aware NAT information can be obtained via the **show service nat l2-aware-hosts** command. This command also provides additional filters for more targeted results:

```
show service nat l2-aware-hosts [outside-router <router-instance>] [outside-ip <outside-ip-
address>] [inside-ip-prefix <ip-prefix/mask>]
```

In this case, the most generic version of this command is run where all four hosts and their L2-aware NAT specific information is shown:

```
*A:BNG1# show service nat l2-aware-hosts

===============================================================================
Layer-2-Aware NAT hosts
===============================================================================
Subscriber                    : sub-1
Inside IP address             : 10.10.1.1
-------------------------------------------------------------------------------
Policy                        : l2-aware-base
Bypassing                     : false
Outside router                : "Base"
Outside IP address            : 172.16.3.3
Port block                    : N/A

Policy                        : l2-aware-vrf1
Bypassing                     : false
Outside router                : 1
Outside IP address            : 172.16.1.1
Port block                    : N/A

Policy                        : l2-aware-vrf2
Bypassing                     : false
```

```
Outside router              : 2
Outside IP address          : 172.16.2.2
Port block                  : N/A

Subscriber                  : sub-1
Inside IP address           : 10.10.1.2
-------------------------------------------------------------------------------
Policy                      : l2-aware-base
Bypassing                   : false
Outside router              : "Base"
Outside IP address          : 172.16.3.3
Port block                  : N/A

Policy                      : l2-aware-vrf1
Bypassing                   : false
Outside router              : 1
Outside IP address          : 172.16.1.1
Port block                  : N/A

Policy                      : l2-aware-vrf2
Bypassing                   : false
Outside router              : 2
Outside IP address          : 172.16.2.2
Port block                  : N/A

Subscriber                  : sub-2
Inside IP address           : 10.10.1.1
-------------------------------------------------------------------------------
Policy                      : l2-aware-base
Bypassing                   : false
Outside router              : "Base"
Outside IP address          : 172.16.3.3
Port block                  : N/A

Policy                      : l2-aware-vrf1
Bypassing                   : false
Outside router              : 1
Outside IP address          : 172.16.1.1
Port block                  : N/A

Policy                      : l2-aware-vrf2
Bypassing                   : false
Outside router              : 2
Outside IP address          : 172.16.2.2
Port block                  : N/A

Subscriber                  : sub-2
Inside IP address           : 10.10.1.2
-------------------------------------------------------------------------------
Policy                      : l2-aware-base
Bypassing                   : false
Outside router              : "Base"
Outside IP address          : 172.16.3.3
Port block                  : N/A

Policy                      : l2-aware-vrf1
Bypassing                   : false
Outside router              : 1
Outside IP address          : 172.16.1.1
Port block                  : N/A

Policy                      : l2-aware-vrf2
Bypassing                   : false
Outside router              : 2
Outside IP address          : 172.16.2.2
Port block                  : N/A

-------------------------------------------------------------------------------
No. of hosts: 4
===============================================================================
```

The following two commands are more generic and their output displays packet and event statistics related to NAT operation in two contexts:

- Per MS-ISA
- Per member within the ISA

  A member within an MS-ISA is a concept related to intra-chassis redundancy. CPM maintains a copy of the member configuration so that it can download it to a rescuing MS-ISA during the switchover (when an MS-ISA fails). An MS-ISA can contain one or more (in A/A redundancy model) members.

```
show isa nat-group 1 mda 1/2 statistics
show isa nat-group 1 member 1 statistics
```

The output is the following:

```
*A:BNG1# show isa nat-group 1 mda 1/2 statistics
===============================================================================
ISA NAT Group 1 MDA 1/1
===============================================================================
no resource                                             : 0
pkt rx on wrong port                                    : 0
unsupported protocol                                    : 0
no host or host group                                   : 0
no ip or port                                           : 0
no matching flow                                        : 0
max flow exceeded                                       : 0
TCP no flow for RST                                     : 0
TCP no flow for FIN                                     : 0
TCP no flow                                             : 0
addr. dep. filtering                                    : 0
ICMP type unsupported                                   : 0
ICMP local unsupported                                  : 0
ICMP/ICMPv6 checksum error                              : 0
ICMP embedded checksum error                            : 0
ICMP/ICMPv6 unsupported embedded L4                     : 0
ICMP/ICMPv6 too short                                   : 0
ICMP/ICMPv6 length error                                : 0
Pkt not IPv4 or IPv6                                    : 0
Pkt rcv error                                           : 0
Pkt error                                               : 0
IPv4 header checksum violation                          : 0
IP header malformed                                     : 0
IP malformed packet                                     : 0
IP ttl zero                                             : 0
IPv4 opt /IPv6 ext headers                              : 0
IP undefined error                                      : 0
IPv6 fragments unsupported                              : 0
TCP/UDP malformed                                       : 0
TCP/UDP checksum failure                                : 0
Pkt send error                                          : 0
no buf to copy pkt                                      : 0
no policy                                               : 0
policy not in use by subscriber                         : 0
locked by mgmt core                                     : 0
port range log failed                                   : 0
```

```
MTU exceeded                                            : 0
DS Lite unrecognized next hdr                           : 0
DS Lite unknown AFTR                                    : 0
too many fragments for IP packet                        : 0
too many fragmented packets                             : 0
too many fragment holes                                 : 0
too many frags buffered                                 : 0
fragment list expired                                   : 0
fragment rate too high                                  : 0
flow log failed                                         : 0
no multiple host or subscr. IPs allowed                 : 0
NAT64 disabled                                          : 0
NAT64 invalid src addr                                  : 0
NAT64 frag has zero checksum                            : 0
NAT64 v4 has zero checksum                              : 0
NAT64 ICMP frag unsupported                             : 0
Reassembly Failures                                     : 0
subscriber-id unknown                                   : 0
packet hashed to wrong MDA                              : 0
Nptv6 map failed                                        : 0
Nptv6 IID all 1's                                       : 0
UPnP rate too high                                      : 0
Dest. IP is unroutable                                  : 0
no TCP/UDP checksum                                     : 0
to local                                                : 0
to local ignored                                        : 0
new flow                                                : 0
TCP closed                                              : 0
TCP expired                                             : 0
UDP expired                                             : 0
ICMP/ICMPv6 expired                                     : 0
GRE expired                                             : 0
ICMP local                                              : 0
found flow                                              : 0
ARPs ignored                                            : 10
Fragments RX L2A                                        : 0
Fragments RX LSN                                        : 0
Fragments RX DSL                                        : 0
Fragments RX DORMANT                                    : 0
Fragments RX OUT                                        : 0
Fragments TX L2A                                        : 0
Fragments TX LSN                                        : 0
Fragments TX DSL                                        : 0
Fragments TX DORMANT                                    : 0
Fragments TX NAT64                                      : 0
Fragments TX OUT                                        : 0
flow create logged                                      : 0
flow delete logged                                      : 0
flow log pkt tx                                         : 0
Reassembled Pkts                                        : 0
subscriber-id cached                                    : 0
subscriber-id delayed                                   : 0
subscriber-id timeout                                   : 0
forwarded DS-Lite V6 pkts                               : 0
V6 pkts fragmented                                      : 0
dropped multicast                                       : 0
SSDP rx                                                 : 0
SSDP tx                                                 : 0
SSDP dropped                                            : 0
UPnP rx                                                 : 0
UPnP tx                                                 : 0
UPnP dropped                                            : 0
no radius resources                                     : 0
no radius connection resources                          : 0
```

```
Dest. NAT dest. IP mismatch                             : 0
Dest. NAT foreign IP mismatch                           : 0
temp. no policy                                         : 0
no SNat enabled                                         : 0
no default policy                                       : 0
flow create failed, key ambiguous                       : 0
flow create failed, conflicting policies                : 0
NAT64 unrecognized next hdr                             : 0
ICMP/ICMPv6 unsupported embedded L3                     : 0
TCP reset waiting                                       : 0
IPv6 downstream prohibited                              : 0
IPv6 ext. hdr parse error                               : 0
any V6 L4 expired                                       : 0
ICMP embedded IPv6 ext. hdr parse error                 : 0
ICMPv6 type unsupported                                 : 0
ICMP/ICMPv6 fragmented error                            : 0
Service-chaining rx                                     : 0
Service-chaining tx                                     : 0
Service-chaining encapsulation error                    : 0
Service-chaining decapsulation error                    : 0
Service-chaining filter drop                            : 0
PPPoE uplink down                                       : 0
ICMP6 PTB dropped (mtu < 1280)                          : 0
firewall addr. dep. filtering                           : 0
unresolved L2-aware V6 host                             : 0
===============================================================================
*A:BNG1#
```

Resources monitoring is performed via the following commands:

```
tools dump nat isa resources nat-group <id> member <id>
tools dump nat isa resources mda <card-id/mds-id>
show isa nat-system-resources nat-group <id> member <id>
```

The NAT scale depends on the hardware (vSIMs, MS-ISA1, MS-ISA2, VSR-NAT).

Tools Commands

To display the complete session state, the following command can be used:

```
tools dump nat sessions [nat-group <nat-group-id>] [mda <mda-id>] [protocol     {gre|icmp|
icmp6|tcp|udp|unknown}] [inside-ip <ip-prefix[/ip-prefix-    length]>] [inside-router <router-
instance>] [inside-port <port-number>]     [outside-ip <ipv4-address>] [outside-port <port-
number>] [foreign-ip
    <ip-address>] [foreign-port <port-number>] [dslite-address
    <ipv6-address>] [wlan-gw-ue <ieee-address>] [next index <index>] [upnp]     [member
 <member-id>] [nat-policy <policy-name>] [dest-ip <ip-address>]     [firewall-policy <policy-
name>] [address-type <addr-type>] [vas-filter
    <filter-name>] [vas-filter-entry <vas-filter-entry-id>]
    [l2-aware-subscriber <sub-ident>]
```

The output of this command is used to populate dynamically allocated fields (in bold) in Table 11: Traffic Streams. The command, in its basic form, is run as:

```
*A:BNG1# tools dump nat sessions

===============================================================================
Matched 6 sessions on Slot #1 MDA #2
===============================================================================
Owner               : L2-aware Subscr sub-1
Policy              : l2-aware-vrf1
FlowType            : UDP               Timeout (sec)       : 299
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.1.1
Outside Port        : 1124
Foreign IP Addr     : 192.168.7.2
Foreign Port        : 5000
Dest IP Addr        : 192.168.7.2
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
Owner               : L2-aware Subscr sub-2
Policy              : l2-aware-vrf1
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.1.1
Outside Port        : 4666
Foreign IP Addr     : 192.168.7.2
Foreign Port        : 5000
Dest IP Addr        : 192.168.7.2
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
Owner               : L2-aware Subscr sub-2
Policy              : l2-aware-base
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.3.3
Outside Port        : 4470
Foreign IP Addr     : 192.168.5.2
Foreign Port        : 5000
Dest IP Addr        : 192.168.5.2
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
Owner               : L2-aware Subscr sub-1
Policy              : l2-aware-vrf2
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.2.2
Outside Port        : 2068
Foreign IP Addr     : 192.168.8.2
Foreign Port        : 5001
Dest IP Addr        : 192.168.8.5
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
Owner               : L2-aware Subscr sub-1
Policy              : l2-aware-base
FlowType            : UDP               Timeout (sec)       : 299
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.3.3
Outside Port        : 2186
Foreign IP Addr     : 192.168.5.2
Foreign Port        : 5000
Dest IP Addr        : 192.168.5.2
Nat Group           : 1
Nat Group Member    : 1
```

```
-------------------------------------------------------------------------------
Owner               : L2-aware Subscr sub-2
Policy              : l2-aware-vrf2
FlowType            : UDP               Timeout (sec)       : 299
Inside IP Addr      : 10.10.1.1
Inside Port         : 5000
Outside IP Addr     : 172.16.2.2
Outside Port        : 3690
Foreign IP Addr     : 192.168.8.2
Foreign Port        : 5000
Dest IP Addr        : 192.168.8.2
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
```

Clear Commands

The following command clears L2-aware subscribers:

```
clear nat isa nat-group 1 member 1 l2-aware-
```

Note:
ESM subscribers are also deleted with this command. Therefore, this command should be used with caution because DHCP hosts in ESM do not send termination messages back to the client. This leaves the DHCP client in a state where it is not aware that its DHCP state in the L2-aware node has been deleted.

Debug Commands

Troubleshooting ESM subscribers is described in other configuration guides focusing on ESM. This chapter focuses only on the NAT aspect of ESM in L2-aware NAT.

L2-aware NAT debug output shows session (or flow) initiation information. As traffic is run, translations of the source IP address/port are performed and this information is displayed for debugging purposes. Traffic from both subscribers in this example is debugged with these two commands:

```
*A:BNG1# show debug
debug
    nat
        l2-aware-sub id 1 subscriber "sub-1"
        l2-aware-sub id 2 subscriber "sub-2"
    exit
exit
```

The following configuration sends debug output to the Telnet/SSH sessions screen:

```
A:BNG1>config>log# info
----------------------------------------------
        log-id 50
            from debug-trace
            to session
            no shutdown
        exit
```

Debug output shows translations that are being performed on the incoming traffic, which consists of six streams in this example:

```
1 2017/08/31 08:37:39.223 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-1
Initiated flow from 172.16.3.3(port 2186) to 192.168.5.2(port 5000) protocol UDP"

2 2017/08/31 08:37:39.223 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-1
Initiated flow from 172.16.1.1(port 1124) to 192.168.7.2(port 5000) protocol UDP"

3 2017/08/31 08:37:39.483 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-1
Initiated flow from 172.16.2.2(port 3690) to 192.168.8.2(port 5001) protocol UDP"

4 2017/08/31 08:37:39.743 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-2
Initiated flow from 172.16.3.3(port 4470) to 192.168.5.2(port 5000) protocol UDP"

5 2017/08/31 08:37:39.743 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-2
Initiated flow from 172.16.1.1(port 4666) to 192.168.7.2(port 5000) protocol UDP"

6 2017/08/31 08:37:40.003 WEST MINOR: DEBUG #2001 Base NAT "NAT: L2-Aware@sub-2
Initiated flow from 172.16.2.2(port 4698) to 192.168.8.2(port 5000) protocol UDP”
```

Note that dNAT information is not displayed in the debug output.

Conclusion

L2-aware NAT integrates NAT functionality into ESM. NAT resources (outside IP address and a PB) are allocated per ESM subscriber and, consequently, all the hosts within the subscriber share those NAT resources. Each subscriber is initially assigned a single PB that will be used during the lifetime of the subscriber. The number of ports in this PB must be large enough to accommodate the needs of the subscriber during its lifetime.

The integration between ESM and NAT in L2-aware NAT introduces the following:

- Using the subscriber-id field (instead of source IPv4 address) to identify the NAT binding allows the source IPv4 address to be abstracted from the binding. This allows hosts of different subscribers to have overlapping IP addresses .

  IPv4 addresses cannot overlap between the hosts of the same subscriber, but they can overlap between the hosts of different subscribers.
- NAT logging can be integrated into ESM RADIUS accounting, simplifying operations and removing the need for a separate network resource dedicated to the collection of NAT logs.
- NAT resources are released when the ESM subscriber is disconnected. This prevents lingering of NAT resources in the system waiting for the timer of the last session to expire long after the entity that initially requested those resources ceased to exist. This improves resource utilization in a system with a high number of subscribers.

L2-aware NAT should be considered for deployment because it allows coupling between ESM and NAT.

Appendix A – Generic ESM Configuration

Appendix A describes the generic part of the ESM configuration.

ESM configuration starts with a **subscriber-interface** configured in an inside routing context. In this example, subscriber hosts are instantiated in the IES 3 service, under the **group-interface “group-int-1”**, which is created under the **subscriber-interface “sub-int-1”**. Authentication and address assignment of the subscriber hosts is performed via local user database (LUDB) user-db “user-db-1”. The IP addresses that are assigned to the hosts are statically configured in LUDB (no DHCP server is used in this setup).

```
configure
    service
        ies 3 customer 1 create
            subscriber-interface "sub-int-1" create
                address 10.10.1.254/24
                group-interface "group-int-1" create
                    dhcp
                        proxy-server
                            emulated-server 10.10.1.254
                            no shutdown
                        exit
                        option
                            action keep
                            circuit-id
                            remote-id
                        exit
                        trusted
                        lease-populate 100
                        gi-address 10.10.1.254
                        user-db "user-db-1"
                        no shutdown
                    exit
                    pppoe
                        anti-spoof mac-sid-ip
                        policy "pppoe_pol"
                        session-limit 100
                        sap-session-limit 100
                        user-db "user-db-1"
                            no shutdown
                        exit
                        no shutdown
                    exit
                exit
            exit
            service-name "ies-3"
            no shutdown
        exit
    exit
```

Subscriber SAPs are automatically created based on the VLAN tags carried in the initial control packets of the subscriber hosts. This VLAN auto-detection and SAP auto-creation is configured under the capture SAP hierarchy. The capture SAP is configured to support LUDB authentication for dynamic DHCPv4/PPPoEv4 host instantiation, as follows:

```
configure
    service
        vpls 10 name "10" customer 1 create
            sap 1/1/5:1.* capture-sap create
                trigger-packet dhcp pppoe
                dhcp-user-db "user-db-1"
                pppoe-policy "pppoe_pol"
                pppoe-user-db "user-db-1"
            exit
        exit
```

The first interaction between ESM and L2-aware NAT is performed through a **sub-profile**. Specifically, a **nat-policy** within the **sub-profile** is used to associate the ESM subscriber with NAT.

A **nat-prefix-list** within the sub-profile provides the place to configure multiple nat-policies per subscriber. Selection of **nat-policy** (and with this, the NAT pool and outside routing context) is based on the destination IP address in the packet sent by the subscriber.

RADIUS accounting records passed to the accounting server will contain subscriber-host specific information, including NAT logs. This is configured via a RADIUS accounting policy, which is referenced in the **sub-profile**.

```
configure
    subscriber-mgmt
        sub-profile "sub-prof-1" create
            nat-policy "l2-aware-base"
            nat-prefix-list "l2aware-mnp"
            radius-accounting
               policy "acct"
          exit
      exit
```

Although an **sla-profile** is a mandatory configuration for a subscriber, the **sla-profile** content is not relevant for understanding L2-aware NAT concepts. A basic **sla-profile** is chosen with default settings: a default **qos-policy 1** and no **ip-filters**.

The **sub-profile** and **sla-profile** are associated with the subscriber(-host) during the authentication phase and, in this case, this is achieved through LUDB.

```
configure
    subscriber-mgmt
        sla-profile "sla-1" create
        exit
```

The MSAP policy is a mandatory configuration for dynamically created SAPs (Managed SAP or MSAPs). It is used to determine SAP parameters during the MSAP creation process based on control traffic (DHCP or PPPoE) of the first host (MSAP trigger traffic). Parameters defined in **msap-policy** that are of relevance to our example are:

- **Sub-ident-policy** → this is a mandatory parameter in ESM that governs mapping of subscriber-related profiles (`sub-profile`, `SLA-profile`) to the subscriber host during the authentication phase.
- **Anti-spoof filter** → this is of particular importance in L2-aware NAT and is already described in more detail in the Configuration section.

```
configure
    subscriber-mgmt
        msap-policy "msaps" create
            sub-sla-mgmt
                sub-ident-policy "sub_ident_pol"
                multi-sub-sap limit 10
            exit
            ies-vprn-only-sap-parameters
                anti-spoof ip-mac
            exit
        exit
```

Sub-ident-policy is a mandatory configuration in ESM. It determines the mapping method between the sub/SLA profiles and the corresponding strings obtained during the authentication phase for the subscriber.

Subscriber strings obtained during the authentication phase point, in some form (determined by **sub-ident-policy**), to the configured sub/SLA profiles (in the SR OS node) that will be associated with the subscriber.

```
configure
    subscriber-mgmt
        sub-ident-policy "sub_ident_pol" create
            sub-profile-map
                use-direct-map-as-default
            exit
            sla-profile-map
                use-direct-map-as-default
            exit
        exit
```

In this example, authentication of the subscriber hosts and IP address assignment is performed through LUDB. The hosts are identified based on **circuit-id** and **remote-id** fields in DHCP and PPPoE control packets. **Sla-sub-profile** strings in the LUDB are directly mapped to the configured sub/sla-profiles in SR OS node. This direct mapping is implied by the preceding **use-direct-map-as-default** command within **sub-ident-policy**. The LUDB carries only ESM specific configuration. There is no NAT relevant configuration present in the LUDB.

```
configure
    subscriber-mgmt
        local-user-db "user-db-1" create
            ipoe
                match-list circuit-id remote-id
                host "sub-1-host-1" create
                    host-identification
                        circuit-id string "sub-1"
                        remote-id string "host-1"
                    exit
                    address 10.10.1.1
                    identification-strings 254 create
                        subscriber-id "sub-1"
                        sla-profile-string "sla-1"
                        sub-profile-string "sub-prof-1"
                    exit
                    msap-defaults
                        group-interface "group-int-1"
                        policy "msaps"
                        service 3
                    exit
                    options
                        subnet-mask 255.255.255.0
                    exit
                    no shutdown
                exit
                host "sub-2-host-1" create
                    host-identification
                        circuit-id string "sub-2"
                        remote-id string "host-1"
                    exit
                    address 10.10.1.1
                    identification-strings 254 create
                        subscriber-id "sub-2"
                        sla-profile-string "sla-1"
                        sub-profile-string "sub-prof-1"
                    exit
                    msap-defaults
                        group-interface "group-int-1"
                        policy "msaps"
                        service 3
                    exit
                    options
                        subnet-mask 255.255.255.0
                    exit
                    no shutdown
                exit
            exit
            ppp
                match-list circuit-id remote-id
                host "sub-1-host-2" create
                    host-identification
                        circuit-id string "sub-1"
                        remote-id string "host-2"
                    exit
                    address 10.10.1.2/24
                    identification-strings 254 create
                        subscriber-id "sub-1"
                        sla-profile-string "sla-1"
                        sub-profile-string "sub-prof-1"
                    exit
                    msap-defaults
                        group-interface "group-int-1"
                        policy "msaps"
                        service 3
                    exit
                    no shutdown
                exit
                host "sub-2-host-2" create
                    host-identification
                        circuit-id string "sub-2"
                        remote-id string "host-2"
                    exit
                    address 10.10.1.2/24
                    identification-strings 254 create
                        subscriber-id "sub-2"
                        sla-profile-string "sla-1"
                        sub-profile-string "sub-prof-1"
                    exit
                    msap-defaults
                        group-interface "group-int-1"
                        policy "msaps"
                        service 3
                    exit
                    no shutdown
                exit
            exit
            no shutdown
        exit
```

The accounting policy identifies the type of accounting and the attributes that will be carried in an accounting message. Among the ESM-specific attributes, **nat-port-range** is the only NAT-related attribute that will carry NAT logging information for the subscriber.

```
configure
    subscriber-mgmt
        radius-accounting-policy "acct" create
            no queue-instance-accounting
            host-accounting interim-update
            update-interval 5
            include-radius-attribute
                circuit-id
                framed-ip-addr
                mac-address
                nas-identifier
                nas-port-id
                nat-port-range
                remote-id
                sla-profile
                sub-profile
                subscriber-id
                user-name
                alc-acct-triggered-reason
            exit
            session-id-format number
            radius-server-policy "aaa"
        exit
```

Appendix B – Logging

Appendix B shows the RADIUS logging stream for all four hosts used in this chapter:

```
root@ linux:/var/log/freeradius/radacct/192.168.114.2# more detail-20170814
Mon Aug 14 09:14:45 2017
        Acct-Status-Type = Start
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:29:02 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502727285
        Request-Authenticator = Verified

Mon Aug 14 09:14:53 2017
        Acct-Status-Type = Start
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003159916D66"
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:29:10 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:02"
        Acct-Unique-Session-Id = "cc1701cf565fcdf5"
        Timestamp = 1502727293
        Request-Authenticator = Verified

Mon Aug 14 09:15:02 2017
        Acct-Status-Type = Start
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003259916D70"
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:29:20 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:01"
        Acct-Unique-Session-Id = "57659d99038f7c0d"
        Timestamp = 1502727302
        Request-Authenticator = Verified

Mon Aug 14 09:15:09 2017
        Acct-Status-Type = Start
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003559916D76"
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:29:26 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:02"
        Acct-Unique-Session-Id = "433924b3304e2d9e"
        Timestamp = 1502727309
        Request-Authenticator = Verified

Mon Aug 14 09:19:42 2017
        Acct-Status-Type = Interim-Update
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003159916D66"
        Acct-Session-Time = 289
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:33:59 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:02"
        Alc-Acct-Triggered-Reason = regular
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x00010000000000009024
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x00010000000000000171
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "cc1701cf565fcdf5"
        Timestamp = 1502727582
        Request-Authenticator = Verified

Mon Aug 14 09:20:01 2017
        Acct-Status-Type = Interim-Update
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003559916D76"
        Acct-Session-Time = 292
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:34:18 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:02"
        Alc-Acct-Triggered-Reason = regular
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x0001000000000000a5a0
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000001a8
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "433924b3304e2d9e"
        Timestamp = 1502727601
        Request-Authenticator = Verified

Mon Aug 14 09:20:11 2017
        Acct-Status-Type = Interim-Update
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Session-Time = 326
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:34:28 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Alc-Acct-Triggered-Reason = regular
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x0001000000000000b220
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000001c8
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502727611
        Request-Authenticator = Verified

Mon Aug 14 09:20:16 2017
        Acct-Status-Type = Interim-Update
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003259916D70"
        Acct-Session-Time = 313
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:34:33 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:01"
        Alc-Acct-Triggered-Reason = regular
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x0001000000000000b734
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000001d5
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "57659d99038f7c0d"
        Timestamp = 1502727616
        Request-Authenticator = Verified
```

```
Mon Aug 14 09:30:34 2017
        Acct-Status-Type = Stop
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000002E59916D5E"
        Acct-Session-Time = 949
        Acct-Terminate-Cause = User-Request
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:44:51 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:01"
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x00010000000000037974
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000008e5
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "fe5e55d7102d3f81"
        Timestamp = 1502728234
        Request-Authenticator = Verified

Mon Aug 14 09:30:34 2017
        Acct-Status-Type = Stop
        NAS-IP-Address = 192.0.2.2
        Framed-IP-Address = 10.10.1.1
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003259916D70"
        Acct-Session-Time = 931
        Acct-Terminate-Cause = User-Request
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:44:51 PDT"
        NAS-Port-Type = Ethernet
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-1"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:01"
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x000100000000000378ac
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000008e3
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "57659d99038f7c0d"
        Timestamp = 1502728234
        Request-Authenticator = Verified

Mon Aug 14 09:30:47 2017
        Acct-Status-Type = Stop
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003159916D66"
        Acct-Session-Time = 954
        Acct-Terminate-Cause = User-Request
        Acct-Multi-Session-Id = "D896FF0000002F59916D5E"
        Event-Timestamp = "Aug 14 2017 02:45:04 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.1"
        ADSL-Agent-Circuit-Id = "sub-1"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-1"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 1024-3023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 1024-3023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 1024-3023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:01:02"
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x00010000000000037974
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000008e5
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "cc1701cf565fcdf5"
        Timestamp = 1502728247
        Request-Authenticator = Verified

Mon Aug 14 09:30:47 2017
        Acct-Status-Type = Stop
        NAS-IP-Address = 192.0.2.2
        Service-Type = Framed-User
        Framed-Protocol = PPP
        Framed-IP-Address = 10.10.1.2
        NAS-Identifier = "BNG1"
        Acct-Session-Id = "D896FF0000003559916D76"
        Acct-Session-Time = 938
        Acct-Terminate-Cause = User-Request
        Acct-Multi-Session-Id = "D896FF0000003359916D70"
        Event-Timestamp = "Aug 14 2017 02:45:04 PDT"
        NAS-Port-Type = PPPoEoQinQ
        NAS-Port-Id = "1/1/5:1.2"
        ADSL-Agent-Circuit-Id = "sub-2"
        ADSL-Agent-Remote-Id = "host-2"
        Alc-Subsc-ID-Str = "sub-2"
        Alc-Subsc-Prof-Str = "sub-prof-1"
        Alc-SLA-Prof-Str = "sla-1"
        Alc-Nat-Port-Range = "172.16.3.3 3024-5023 router base l2-aware-base"
        Alc-Nat-Port-Range = "172.16.1.1 3024-5023 router 1 l2-aware-vrf1"
        Alc-Nat-Port-Range = "172.16.2.2 3024-5023 router 2 l2-aware-vrf2"
        Alc-Client-Hardware-Addr = "00:00:65:01:02:02"
        Alc-Acct-I-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Octets-64 = 0x000100000000000378ac
        Alc-Acct-I-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-I-Outprof-Pkts-64 = 0x000100000000000008e3
        Alc-Acct-O-Inprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Octets-64 = 0x00010000000000000000
        Alc-Acct-O-Inprof-Pkts-64 = 0x00010000000000000000
        Alc-Acct-O-Outprof-Pkts-64 = 0x00010000000000000000
        Acct-Unique-Session-Id = "433924b3304e2d9e"
        Timestamp = 1502728247
        Request-Authenticator = Verified

root@linux:/var/log/freeradius/radacct/192.168.114.2#
```

L2TP Network Server

This chapter provides information about L2TP network servers (LNS).

Topics in this chapter include:

Applicability

Initially, this chapter was initially written based on SR OS Release 11.0.R7, but the CLI in this edition is based on SR OS Release 16.0.R7.

Overview

The Layer 2 Tunneling Protocol (L2TP) is a tunneling protocol commonly used to transport PPP sessions from an initiator known as an L2TP Access Concentrator (LAC) to an L2TP Network Server (LNS). L2TP is typically used for wholesaling residential broadband services. In this scenario, the LAC resides in the wholesaler’s network and has a Layer-2 connection to an access concentrator such as a DSLAM. The LAC responds during the discovery phase (if PPPoE is used) and during PPP Link Control Protocol (LCP) negotiation. The LAC also performs an initial authentication of the subscriber. A successful authentication, typically from RADIUS, indicates to the LAC that PPP frames from this subscriber should be tunneled to an LNS at the indicated IP address. The LAC then tunnels the PPP frames from this subscriber over an L2TP tunnel to the LNS, where the PPP session is actually terminated. In other words, PPP **sessions** require L2TP **tunnels** for the sessions to get carried over.

L2TP uses two types of messages; control messages and data messages. Control messages are used in the establishment, maintenance, and tearing down of tunnels and sessions. In order to provide extensibility and maximize interoperability, the payloads of control messages are encoded using Attribute Value Pairs (AVPs), some of which are applicable to all control messages, and some of which are specific to particular control messages. The L2TP header contains sequence number fields that must be present in control messages to allow for a reliable L2TP control channel that guarantees delivery. Data messages are used to encapsulate PPP frames being carried over the tunnel. Data messages are not retransmitted if packet loss occurs.

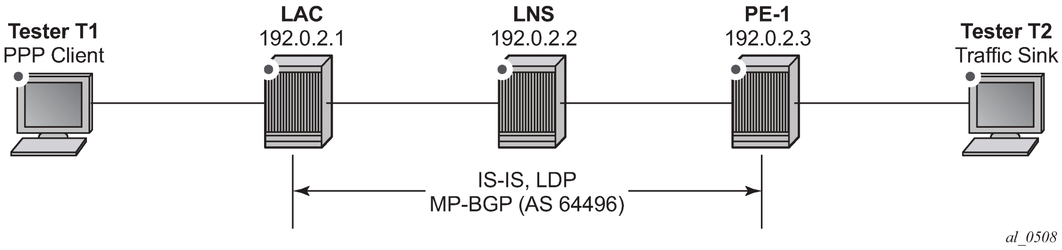
L2TP has a common fixed header format for both control and data messages, and a Type (T) bit in the header is used to indicate whether the packet is a control (1) or data (0) message. The L2TP packet is then carried in a transport protocol, and although the specifications allow L2TP to be directly encapsulated over Frame Relay, ATM, and UDP/IP, the latter is used almost exclusively.

The objective of this chapter is to provide a generic overview of how to configure the 7750 SR to support the LNS and LTS (L2TP Tunnel Switching) functions.

Example Topology

The example topology used through this chapter is shown in Figure 76: Example Topology. Both the LAC and the LNS participate in IS-IS and LDP, together with PE-1. All three devices form part of AS 64496 and peer using iBGP for the VPN-IPv4 address family. None of these protocols is mandatory for supporting LNS functionality; L2TP packets can ingress the system over any network interface as native IP or encapsulated as IP in MPLS, or through an IES/VPRN IP interface (SAP) as native IP. The MPLS data-plane within the example topology is chosen purely because of its simplicity and flexibility. Tester T1 simulates a DSLAM and one or more PPP clients, and is connected directly to the LAC. Although the LAC in this topology is a 7750 SR router, the configuration requirements of that device are beyond the scope of this example. Tester T2 provides a traffic source/sink capability and is connected directly to PE-1.

Figure 76: Example Topology



Hardware Configuration

To support LNS (and LTS) functionality, at least one MS-ISA card is required, which must be configured as MDA type isa-bb, and must be housed in the carrier IOM. The MS-ISA performs L2TP data-plane encapsulation and de-capsulation, whereas the subscriber processing (Enhanced Subscriber Management or ESM) for PPP sessions is implemented within the carrier IOM.

```
configure
    card 1
        card-type iom3-xp
        mda 1
            mda-type m4-10gb-xp-xfp
            no shutdown
        exit
        mda 2
            mda-type isa-bb
            no shutdown
        exit
        no shutdown
    exit
exit
```

The MS-ISA is then configured to become a member of an **lns-group**. Up to six MS-ISAs can be configured to belong to one or more lns-groups. When two or more MS-ISAs belong to the same lns-group, by default PPP sessions are load-balanced over those MS-ISAs on a per-session basis.

```
configure
    isa
        lns-group 1 create
            mda 1/2
            no shutdown
        exit
    exit
exit
```

Configuration

ESM Base Configuration

For completeness, the following outputs contain the base ESM configuration that is applied to subscribers instantiated at the LNS throughout this chapter. Deviations from these base parameters are mentioned explicitly.

The SLA-Profile and Sub-Profile configurations have a minimal set of parameters. The SLA-Profile uses the default ingress/egress QoS policy of 1, while the **no qos-marking-from-sap** command ensures that any subsequent marking is inherited from the egress QoS policy referenced in the SLA-profile, and not taken from the egress SAP. In order to do on-line accounting through RADIUS, the Sub-Profile also calls the relevant RADIUS accounting policy. Finally, the **sub-ident-policy** is configured with **use-direct-map-as-default** for the **sub-profile-map** and **sla-profile-map**, which means that the strings passed from RADIUS in the **Alc-Subs-Prof-Str** and **Alc-SLA-Prof-Str** Vendor Specific Attributes (VSAs) are interpreted verbatim so they are not used as string input to a mapping function.

```
configure
    subscriber-mgmt
        sla-profile "sla-profile-1" create
            egress
                no qos-marking-from-sap
            exit
        exit
        sub-profile "sub-profile-1" create
            radius-accounting
                policy "sm-acct-1"
            exit
        exit
        sub-ident-policy "all-subscribers" create
            sub-profile-map
                use-direct-map-as-default
            exit
            sla-profile-map
                use-direct-map-as-default
            exit
        exit
    exit
exit
```

Whilst it is entirely possible to authenticate subscribers locally using a local user database (LUDB), the more widely deployed approach is to use RADIUS, and this approach is therefore used throughout this chapter. The next output shows the authentication policy *sm-auth-1* and the RADIUS accounting policy *sm-acct-1*. Both policies reference the radius-server-policy *aaa-radius-1*, which provides the context to configure the source-address to use for RADIUS messages and an associated routing context. The RADIUS server policy then references a RADIUS server *radius-172.16.1.11*, which in turn allows for configuration of the server IP address, the secret key to be used for message exchanges, and any other optional port configuration. In this example also Change of Authorization (COA) is configured at RADIUS server level, through the **accept-coa** command.

The intention is not to provide a complete description of all of the RADIUS parameters as this would distract from the objective of this chapter.

```
configure
    router
        radius-server
            server "radius-172.16.1.11" address 172.16.1.11 secret vsecret1 create
                accept-coa
                pending-requests-limit 1024
            exit
        exit
    exit
    aaa
        radius-server-policy "aaa-radius-1" create
            servers
                router "Base"
                source-address 192.0.2.2
                server 1 name "radius-172.16.1.11"
            exit
        exit
    exit
    subscriber-mgmt
        authentication-policy "sm-auth-1" create
            pppoe-access-method pap-chap
            include-radius-attribute
                nas-port-id
                nas-identifier
                access-loop-options
                calling-station-id remote-id
            exit
            radius-server-policy "aaa-radius-1"
        exit
        radius-accounting-policy "sm-acct-1" create
            no queue-instance-accounting
            session-accounting interim-update host-update
            update-interval 120
            include-radius-attribute
                circuit-id
                framed-ip-addr
                nas-identifier
                nas-port-id
                nas-port-type
                sla-profile
                sub-profile
                subscriber-id
                std-acct-attributes
            exit
            session-id-format number
            radius-server-policy "aaa-radius-1"
        exit
    exit
exit
```

Basic LNS Configuration

To illustrate the building blocks that are required to implement LNS functionality, a VPRN is used on the LAC and the LNS supporting an L2TP tunnel and terminating PPP sessions at the LNS side of the L2TP tunnel. The required configuration for this VPRN at the LNS is shown in the following output. The unicast VPRN parameters such as **route-distinguisher** and **vrf-import**/**vrf-export** are not discussed here, only the parameters that are relevant to subscriber termination, which are equally applicable to VPRN and/or IES services.

The interface called *system* is a logical loopback interface and is used as the LNS endpoint address for L2TP signaling. The name of this interface is not important; this interface only must be a loopback interface. The LAC has a corresponding interface with IP address 192.168.0.1. The address of the interface *system* is also used in the **subscriber-interface** context as argument to the **unnumbered** command, meaning this IP address is used for the purpose of IPCP negotiation with incoming PPP sessions. Within the **subscriber-interface** context, the **group-interface** has a different definition than a conventional ESM group interface. A conventional group interface contains one or more SAPs belonging to the same port or LAG. However, in the context of LNS, there are no SAPs. The group interface also might terminate sessions within the same L2TP tunnel which are anchored on different MS-ISAs in a common lns-group. To accommodate this, the **group-interface** has the creation-time attribute **lns**. This attribute essentially means that the group interface can terminate subscribers from more than one port/LAG; where port/LAG is interpreted as different MS-ISAs.

The **group-interface** then provides a **sap-parameters** context that allows for configuration of **sub-sla-mgmt** parameters that would typically be found under a SAP. These parameters apply to all subscribers terminated on this group interface. In the example shown, only the **sub-ident-policy** is configured; meaning that other ESM parameters such as **sla-profile**, **sub-profile**, and **subscriber-id** must be obtained from a different source. In this chapter, they are obtained through RADIUS.

The static route black-holes prefix 10.48.127.0/24 ensures this prefix is added to the route-table. Subscribers are allocated /32 addresses from this range, which must be advertised upstream to PE-1 to ensure end-to-end IP connectivity. This is implemented through the **vrf-export policy** (not shown for conciseness).

Within the **l2tp** context, an hierarchy of groups and tunnels is defined. Groups reside directly under the **l2tp** context, and tunnels reside within the **group** context. Groups are intended to administratively organize tunnels that may share a common use or contain common parameters. The L2TP tunnel characteristics can be inherited from the group context, or overridden within the **tunnel** context. In the **group** context shown in the following output, the **lns-group 1** command refers to the **lns-group** previously configured at the ISA level. This is followed by the **local-address** command that defines the IP address to be used as a source address for L2TP signaling. The **ppp** context then defines the characteristics to be used when PPP sessions are established. In this case, the authentication mechanism is CHAP, and the previously configured RADIUS **authentication-policy** is used to authenticate the user. During the PPP session setup, the LAC negotiates LCP and authentication parameters with the subscriber. Two AVPs, the **Proxy LCP** AVP and the **Proxy Authentication** AVP allow this information to be forwarded by the LAC to the LNS. This information can be accepted by the LNS, allowing PPP to continue with negotiation of IPCP, or it can be rejected, in which case the LNS initiates a new round of NCP and PPP authentication. The **proxy-authentication** and **proxy-lcp** commands allow the information contained in these AVPs to be accepted.

Finally, the **tunnel** context provides the context for defining L2TP tunnel specific parameters. The **peer** command defines the far-end (LAC) IP address to which L2TP messages are addressed. The **password** is used to authenticate the far-end tunnel initiator, and is used in conjunction with the **challenge** parameter to implement a CHAP-like authentication mechanism. The default behavior is to never challenge the initiator (LAC); the **challenge always** command is the reverse of this behavior. The **remote-name** is used to provide an additional level of security. When the Start Control Connection Request (SCCRQ) is received from the LAC initiating the tunnel setup, it carries a mandatory **Host Name** AVP. The value of this AVP is compared with the name configured in the **remote-name**, and only tunnels with matching names are accepted. In a similar way, the **local-name** parameter is used to populate the Host Name AVP sent by the LNS in the SCCRP, and can be used as a similar security feature at the LAC.

When two or more MS-ISAs belong to the same lns-group, PPP sessions are load-balanced over those MS-ISAs on a per-session basis by default. Although it is not shown in the following configuration example, it is worth mentioning that within each L2TP group context, an option exists to load-balance the sessions on a per-L2TP tunnel basis using the **load-balance-method per-tunnel** command. This can be useful, for example, when multiple sessions are received from a single subscriber (for example, MLPPP member links) which must be handled within the same MS-ISA.

```
configure
    service
        vprn 1 customer 1 create
            vrf-import "vrf1-import"
            vrf-export "vrf1-export"
            route-distinguisher 64496:1
            auto-bind-tunnel
                resolution-filter
                    ldp
                exit
                resolution filter
            exit

            interface "system" create
                address 192.168.0.2/32
                loopback
            exit
            subscriber-interface "LNS-SUB-INT" create
                unnumbered 192.168.0.2
                group-interface "LNS-GROUP-INT" lns create
                    sap-parameters
                        sub-sla-mgmt
                            sub-ident-policy "all-subscribers"
                        exit
                    exit
                exit
            exit
            static-route-entry 10.48.127.0/24
                black-hole
                    no shutdown
                exit
            exit
            l2tp
                group "L2TP-GROUP-1" create
                    hello-interval 60
                    idle-timeout 600
                    lns-group 1
                    local-address 192.168.0.2
                    ppp
                        authentication chap
                        authentication-policy "sm-auth-1"
                        default-group-interface "LNS-GROUP-INT" service-id 1
                        keepalive 10 hold-up-multiplier 3
                        proxy-authentication
                        proxy-lcp
                    exit
                    tunnel "L2TP-TUNNEL-1" create
                        challenge always
                        local-name "LNS"
                        peer 192.168.0.1
                        remote-name "LAC"
                        password tunnelpwd
                        no shutdown
                    exit
                    no shutdown
                exit
                no shutdown
            exit
            no shutdown
        exit
    exit
exit
```

As previously described, RADIUS is used to authenticate the subscriber, which upon successful authentication returns the ESM parameters, Subscriber-ID (**Alc-Subsc-ID-Str**), SLA-Profile (**Alc-SLA-Prof-Str**), and Sub-Profile (**Alc-Subsc-Prof-Str**) as needed for instantiating the subscriber in SR OS. These parameters could be obtained locally on the LNS using the **def-sub-id**, **def-sla-profile** and **def-sub-profile** commands under the **group-interface>sap-parameters** context. This enables a mechanism to provide default parameters in the absence of another source. However, passing them from RADIUS has some benefits, such as:

- It is comparatively easy to provide different SLA- and Sub-Profiles to different users, which can be used to differentiate service levels.
- If RADIUS infrastructure is available and used to provide ESM parameters, it is relatively easy to extend that infrastructure to provide for mid-session changes of those parameters (such as **sla-profile** and **sub-profile**) using a Change of Authorization (CoA).

The following provides an example of a RADIUS users file entry for the test subscriber. In addition to the previously defined ESM parameters, the Alc-Serv-ID VSA is used to define the service number in which this subscriber must be terminated (in this case, VPRN 1 as previously configured), while the Alc-Interface VSA is used to define the relevant group interface within that service. If it is intended that all PPP sessions ingressing on a particular L2TP group are all to be terminated within a common service and group interface, it is not necessary for the RADIUS server to send the Alc-Serv-ID and Alc-Interface VSAs defining the service and group interface.Instead, a default service and group interface can be defined within the **ppp** context of the l2tp group using the parameter **default-group-interface** <*name*> **service-id** <*number*>. The remainder of the attributes in the output are well-known standard attributes.

```
subscriber1@isp.net     Cleartext-Password := "letmein"
                        Alc-Subsc-ID-Str = "subscriber1@isp.net",
                        Alc-Subsc-Prof-Str = "sub-profile-1",
                        Alc-SLA-Prof-Str = "sla-profile-1",
                        Alc-Serv-Id = "1",
                        Alc-Interface = "LNS-GROUP-INT",
                        Service-Type = Framed-User,
                        Framed-Protocol = PPP,
                        Framed-IP-Address = 10.48.127.27,
```

L2TP Tunnel Setup

Before the PPP session can be terminated at the LNS, an L2TP tunnel must be established between the LAC and LNS. This is achieved using a three-way control message exchange of Start Control Connection Request (SCCRQ), Start Control Connection Reply (SCCRP), and Start Control Connection Connected (SCCN). All of these messages are explicitly acknowledged by the peer using the sequence numbers (number sent, number received) in the L2TP header, thereby creating a reliable control channel. The acknowledgment can be piggybacked in a corresponding control message, or it can be an explicit acknowledgment using a control packet with only an L2TP header, known as a Zero Length Body (ZLB) message.

The SCCRQ is used to initiate the tunnel between LAC and LNS, and although it can be sent by either the LAC or LNS, it is typically sent by the LAC toward the LNS (as in this example). The SCCRQ contains a number of mandatory AVPs, denoted by the M-bit in the AVP header (set to 1), including Message Type, Protocol Version, Host Name, Framing Capabilities, and Assigned Tunnel ID. It can also contain a number of optional AVPs, such as Vendor Name, and Firmware Revision, which can be ignored by the recipient if they are unrecognized.

```
1 2019/05/23 16:27:13.662 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 0 session 0, ns 0 nr 0, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        StartControlConnectionRequest(1)
    AVP ProtocolVersion(0,2), flags: mandatory, reserved=0
        version=1, revision=0
    AVP HostName(0,7), flags: mandatory, reserved=0
        "LAC"
    AVP WindowSize(0,10), flags: mandatory, reserved=0
        64
    AVP FramingCapabilities(0,3), flags: mandatory, reserved=0
        sync=no, async=no
    AVP BearerCapabilities(0,4), flags: mandatory, reserved=0
        digital=yes, analogue=no
    AVP FirmwareRevision(0,6), flags:, reserved=0
        4096
    AVP VendorName(0,8), flags:, reserved=0
        "Nokia"
    AVP AssignedTunnelId(0,9), flags: mandatory, reserved=0
        10007"
```

The SCCRP is sent in response to the SCCRQ and is used to indicate that the parameters in the SCCRQ were acceptable and that the establishment of the L2TP tunnel can continue. The SCCRP contains the same mandatory AVPs and can contain the same optional AVPs as the SCCRQ, but an additional optional AVP is the Challenge AVP which is included as a result of the **challenge always** and **password** parameters configured within the **tunnel** context.

```
2 2019/05/23 16:27:13.662 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, egress)
"L2TP(v2, ctrl, egress): UDP 192.168.0.2:1701 -> 192.168.0.1:1701
tunnel 10007 session 0, ns 0 nr 1, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        StartControlConnectionReply(2)
    AVP ProtocolVersion(0,2), flags: mandatory, reserved=0
        version=1, revision=0
    AVP HostName(0,7), flags: mandatory, reserved=0
        "LNS"
    AVP WindowSize(0,10), flags: mandatory, reserved=0
        64
    AVP FramingCapabilities(0,3), flags: mandatory, reserved=0
        sync=no, async=no
    AVP BearerCapabilities(0,4), flags: mandatory, reserved=0
        digital=yes, analogue=no
    AVP FirmwareRevision(0,6), flags:, reserved=0
        4096
    AVP VendorName(0,8), flags:, reserved=0
        "Nokia"
    AVP AssignedTunnelId(0,9), flags: mandatory, reserved=0
        9265
    AVP Challenge(0,11), flags: mandatory, reserved=0
         0d 12 69 86 37 7b 35 7f d7 00 22 bd ca 25 ad ab
         59 d7 33 d5 a1 01 92 4f 22 a9 9a d8 b3 68 55 60
         98 96 8c f9 e4 0a 9c ce e3 b9 ed 48 d8 "
```

The response to the SCCRP, and the completion of the three-way message exchange is the SCCN. The only mandatory AVP for the SCCN is the Message Type, and since the SCCRP contained a Challenge AVP, the SCCN also contains an AVP Challenge Response. If this response is not satisfactory to the LNS, it generates a Stop Control Connection Notification (StopCCN) with a result code indicating that the requester is not authorized, and subsequently removes any associated tunnel state.

```
3 2019/05/23 16:27:13.664 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 9265 session 0, ns 1 nr 1, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        StartControlConnectionConnected(3)
    AVP ChallengeResponse(0,13), flags: mandatory, reserved=0
         e6 10 7b 11 36 7d ba fc e6 14 4d 30 3d f0 9a ba "
```

With a successful three-way exchange completed, the L2TP tunnel is established. A snapshot view of all L2TP tunnels within the relevant routing context can be displayed using the command in the following output. The Loc-Tu-ID and Rem-Tu-ID are the local and remote tunnel IDs passed in the Assigned Tunnel Id AVP in the SCCRP and SCCRQ respectively. The Conn ID, or connection Id, is a locally significant parameter used for identifying a particular tunnel, and is a 32-bit representation of the local tunnel Id (1442 * 65536 = 946470912). The connection ID is for example used in event log entries for this tunnel. If the state is shown as *established*, then one or more PPP sessions are running over the tunnel. The state can also be *establishedIdle* meaning that although the tunnel is up and established, there are no PPP sessions active within the tunnel.

```
*A:LNS# show router 1 l2tp tunnel
===============================================================================
Conn ID    Loc-Tu-ID Rem-Tu-ID State              Blacklist-state   Ses Active
  Group                                                             Ses Total
    Assignment
-------------------------------------------------------------------------------
607191040  9265      10007     establishedIdle    not-blacklisted   0
  L2TP-GROUP-1                                                      0
    L2TP-TUNNEL-1
-------------------------------------------------------------------------------
No. of tunnels: 1
===============================================================================
*A:LNS#
```

Once a tunnel is established, maintenance and health-checking of that tunnel is achieved using a keepalive mechanism that employs Hello control messages. The Hello message contains only one AVP, the **Message Type** AVP, which indicates it is a Hello message. The Hello messages operate asynchronously between the peers. There is no echo request and echo response function, but simply a Hello followed by an acknowledgment. The Hello is acknowledged in the same way as other control messages, using either piggybacking or ZLB acknowledgments. This asynchronous behavior allows for either end of the tunnel to be configured for different Hello intervals (they are not negotiated), or even for one end not send Hellos at all.

```
5 2019/05/23 16:28:11.374 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 9265 session 0, ns 2 nr 1, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        Hello(6)"

6 2019/05/23 16:28:11.374 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, egress)
"L2TP(v2, ctrl, egress): UDP 192.168.0.2:1701 -> 192.168.0.1:1701
tunnel 10007 session 0, ns 1 nr 3, flags:, reserved=0"
```

The Hello interval at the LNS is configurable under the l2tp, group, or tunnel contexts using the **hello-interval** parameter. The range is 60 to 3600 seconds, with the default being 60 seconds. The **hello-interval infinite** option suppresses sending of Hellos. If the system sends a Hello message and does not get an acknowledgment, it will retransmit the Hello message as many times as the value of the **max-retries-estab** parameter, each time with an exponential back-off. The **max-retries-estab** parameter can be configured in the l2tp, group, or tunnel contexts. The default value is 5, and if no acknowledgment is received before this value is exceeded, the tunnel is declared down and a StopCCN is sent toward the peer.

The retry interval starts with 1 second and doubles on each retry with a maximum-interval of 8 seconds. For example, using a max-retries-estab value of 7 results in a retry of [1, 2, 4, 8, 8, 8 seconds]

The StopCCN is a message that can be generated by either LAC or LNS and is used to inform its peer that the tunnel is being closed. This implicitly means that all PPP sessions carried within that tunnel are also being closed without any associated control messages for those sessions. The StopCCN must contain the **Message Type** and **Tunnel ID** AVPs, and additionally carries a **Result Code** AVP with result code and error code fields to indicate to the peer the reason for the tunnel closure.

```
9 2017/06/07 15:19:00.56 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 14442 session 0, ns 2 nr 3, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        StopControlConnectionNotification(4)
    AVP ResultCode(0,1), flags: mandatory, reserved=0
        result-code: "generalRequestToClearControlConnection"(1),
        error-code: "noGeneralError"(0)
        error-msg: "operator request"
    AVP AssignedTunnelId(0,9), flags: mandatory, reserved=0
        11458"

10 2017/06/07 15:19:00.56 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, egress)
"L2TP(v2, ctrl, egress): UDP 192.168.0.2:1701 -> 192.168.0.1:1701
tunnel 11458 session 0, ns 3 nr 3, flags:, reserved=0"
```

The tunnel **Connection Id** can be used as an additional argument to display the details of a particular tunnel when multiple tunnels are present. The following output is an example of this taken just after the L2TP tunnel has been closed by the LAC peer, and is intentionally taken at this time to illustrate the purpose of some of the fields shown in the output. The State is moved to *closedByPeer*, and the Stop CCN Result field and Error Message field respectively contain the result code and error code of the Result Code AVP received from the LAC in the StopCCN. Because the tunnel is now in a closedByPeer state, all state and information related to this tunnel is removed from the system after a period defined by the Destruct Timeout (shown in the output as Destruct TO). The intention of the Destruct Timeout is to retain information about the tunnel closure which might aid operational communities. The default value as shown is 60 seconds, but it can be configured using the **destruct-timeout** parameter in the l2tp, group, or tunnel contexts. The remainder of the fields in the output are the operational parameters of the tunnel and are self-explanatory.

```
*A:LNS# show router 1 l2tp tunnel detail

===============================================================================
L2TP Tunnel 607191040
===============================================================================

Connection ID: 607191040
Protocol     : v2
State        : closedByPeer
IP           : 192.168.0.2
UDP          : 1701
```

```
Peer IP      : 192.168.0.1
Peer UDP     : 1701
Tx dst-IP    : 192.168.0.1
Tx dst-UDP   : 1701
Rx src-IP    : 192.168.0.1
Rx src-UDP   : 1701
Name         : LNS
Remote Name  : LAC
Assignment ID: L2TP-TUNNEL-1
Group Name   : L2TP-GROUP-1
Acct. Policy : N/A
Error Message: operator request

                                        Remote Conn ID    : 655818752
Tunnel ID         : 9265                Remote Tunnel ID  : 10007
Preference        : 50                  Receive Window    : 64
Hello Interval (s): 300                 AVP Hiding        : never
Idle TO (s)       : 600                 Destruct TO (s)   : 60
Max Retr Estab    : 5                   Max Retr Not Estab: 5
Cfg'd Sess Limit  : unlimited           Oper Session Limit: 32767                      '
Transport Type    : udpIp               Challenge         : always
Time Started      : 05/23/2019 16:27:14 Time Idle         : N/A
Time Established  : 05/23/2019 16:27:14 Time Closed       : 05/23/2019 16:30:15
Stop CCN Result   : generalReq          General Error     : noError
Blacklist-state   : not-blacklisted
Set Dont Fragment : true

Failover
State             : not-recoverable
Recovery Conn ID  : N/A
Recovery state    : not-applicable
Recovered Conn ID : N/A
Recovery method   : mcs
Track SRRP        : (Not specified)
Ctrl msg behavior : handle
Recovery time (ms)
Requested         : N/A
Peer              : N/A
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
No. of tunnels: 1
===============================================================================
*A:LNS#
```

PPP Session Setup

Once the L2TP tunnel is created, the process of establishing a PPP session can start. Once again, a three-way control message exchange is used for establishing a session within an L2TP tunnel, consisting of the Incoming Call Request (ICRQ), Incoming Call Reply (ICRP), and Incoming Call Connected (ICCN). Given that they are control messages, they are all explicitly acknowledged using piggybacking or ZLB acknowledgments.

The ICRQ is sent from the LAC to the LNS to indicate that it has received an incoming call (PPP session) and that a session needs to be established between the two peers for this call. The ICRQ provides enough information about the call for the LNS to make a decision about whether it should answer the call or not. The ICRQ contains the Message Type and Assigned Session ID AVPs as well as a Call Serial Number AVP, which can be used on both the LAC and LNS as an easy reference to the call for troubleshooting purposes. The ICRQ can also carry optional AVPs including Calling Number and Access Line Information

AVPs (RFC 5515) such as Circuit ID, Remote ID, Actual Data Rate Upstream, and Actual Data Rate Downstream.

```
5 2019/05/23 16:38:49.730 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 12229 session 0, ns 2 nr 1, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        IncomingCallRequest(10)
    AVP AssignedSessionId(0,14), flags: mandatory, reserved=0
        15859
    AVP CallSerialNumber(0,15), flags: mandatory, reserved=0
        15550
    AVP CallingNumber(0,22), flags: mandatory, reserved=0
        "LAC 1/1/3:2"
    AVP AgentCircuitId(3561,1), flags:, reserved=0
        "circuit0"
    AVP AgentRemoteId(3561,2), flags:, reserved=0
        "remote0"
    AVP ActDataRateUp(3561,129), flags:, reserved=0
        2000000
    AVP ActDataRateDown(3561,130), flags:, reserved=0
        4000000"
```

The ICRP is sent by the LNS toward the LAC in response to the ICRQ to indicate that the parameters in the ICRQ were acceptable, and that the LAC should go ahead and proceed with the call. The ICRP contains only two AVPs; the Message Type and the Assigned Session ID. The Assigned Session ID values are local to each peer as opposed to a negotiated or agreed-upon value.

```
7 2019/05/23 16:38:49.731 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, egress)
"L2TP(v2, ctrl, egress): UDP 192.168.0.2:1701 -> 192.168.0.1:1701
tunnel 8826 session 15859, ns 1 nr 3, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        IncomingCallReply(11)
    AVP AssignedSessionId(0,14), flags: mandatory, reserved=0
        14338"
```

The final message in the three-way exchange used for establishing sessions within the tunnel is the ICCN. It is sent by the LAC to the LNS to indicate that the call has been answered, so the L2TP session is moved to the *established* state. It also provides additional information on parameters that were used to answer the call which may not have been available when the ICRQ was sent (although it is likely that in most cases they were available). At a minimum, the ICCN must contain the Message Type, Framing Type, and TX Connect Speed AVPs. The TX Connect Speed defines the speed in bits per second from the perspective of traffic flowing from the LAC toward the subscriber (i.e. the LAC downstream rate) and, for best accuracy, can be derived by the LAC from the PPP Broadband Forum Access Line Characteristic tags inserted by the access node (Appendix C TR-101). The TX Connect Speed can be useful for indirect setting of a Hierarchical QoS (H-QoS) aggregate rate. It is indirect because the LNS cannot infer and set an aggregate rate based directly on the TX Connect Speed AVP, but rather the TX Connect Speed is passed to the RADIUS server (using the **include-radius-attribute access-loop-option** parameter in the authentication policy), which in turn may pass the aggregate rate to the LNS in a QoS override VSA. This is described further in the QoS section.

A number of optional AVPs can also be present providing information from the LCP negotiation between the LAC and client. These include Initial Receive, Last Transmit and Last Receive LCP Config Requests, together with Proxy Authentication Type, Name, Challenge, and Response. These parameters allow the LNS to either force a renegotiation of LCP, or to continue with the PPP session and move to the IPCP phase. The final AVP present in the ICCN shown is the RX Connect Speed AVP, which is the opposite of the TX Connect Speed and defines the speed in bits per second from the perspective of traffic flowing from the subscriber toward the LAC.

```
9 2019/05/23 16:38:49.732 CEST MINOR: DEBUG #2001 vprn1 L2TP(v2, ctrl, ingress)
"L2TP(v2, ctrl, ingress): UDP 192.168.0.1:1701 -> 192.168.0.2:1701
tunnel 12229 session 14338, ns 3 nr 2, flags:, reserved=0
    AVP MessageType(0,0), flags: mandatory, reserved=0
        IncomingCallConnected(12)
    AVP FramingType(0,19), flags: mandatory, reserved=0
        sync=no, async=no
    AVP TxConnectSpeed(0,24), flags: mandatory, reserved=0
        4294967295
    AVP InitialRxLcpConfReq(0,26), flags:, reserved=0
         01 04 05 d4
         [1] MRU: 1492
    AVP LastTxLcpConfReq(0,27), flags:, reserved=0
         01 04 05 d4 03 05 c2 23 05 05 06 3c 32 9f a6
         [1] MRU: 1492
         [3] Authentication-Protocol: 0xc223 (CHAP), Algorithm = 5 (MD5)
         [5] Magic-Number: 0x3c329fa6
    AVP LastRxLcpConfReq(0,28), flags:, reserved=0
         01 04 05 d4
         [1] MRU: 1492
    AVP ProxyAuthenType(0,29), flags:, reserved=0
        chap(2)
    AVP ProxyAuthenName(0,30), flags:, reserved=0
         "subscriber1@isp.net"
    AVP ProxyAuthenChallenge(0,31), flags:, reserved=0
         d2 cb 6f 10 90 61 f5 bf 59 72 e0 d2 4b 8c c5 05
         02 5c c5 9e d3 c5 65 d9 f0 08 38 71 d0 a2 75 23
         6e 64 b3 7f 45 a9 3f 9e 9b a0 f0 e6 ac
    AVP ProxyAuthenId(0,32), flags:, reserved=0
        id=1, reserved=0
    AVP ProxyAuthenResponse(0,33), flags:, reserved=0
         d5 3e 86 2f 2b 50 0d 0b 01 15 5b 0f 6d ec aa fb
    AVP RxConnectSpeed(0,38), flags:, reserved=0
        4294967295"
```

On completion of the three-way control message exchange required for session setup, the LNS authenticates the user in the incoming call. In this example, RADIUS is used, which returns the standard and vendor-specific attributes previously defined in the users file. A successful authentication allows the LNS to move to the IPCP phase with the subscriber. In this example, RADIUS returns the IP address in the standard attribute Framed-IP-Address, but equally local pooling with a DHCP server could be used. For conciseness, the IPCP phase is not detailed within this example because the process is reasonably well-known and understood. However, on completion of the IPCP phase, the subscriber is instantiated and the L2TP session becomes active. The Tunnel-ID and Session-ID parameters are locally generated numbers that are passed in L2TP control messages. As previously described, the Connection Id is a locally significant parameter that is a 32-bit representation of the local tunnel Id (6482 * 65536 = 424804352). The ID field is again a locally significant parameter used to identify the L2TP session, and is again represented as a 32-bit number. It is derived from a sum of the Control Connection ID plus the Session ID (424804352 + 26255 = 424830607).

```
*A:LNS# show router 1 l2tp session

===============================================================================
L2TP Session Summary
===============================================================================
ID                   Control Conn ID     Tunnel-ID   Session-ID  State
-------------------------------------------------------------------------------
801454082            801439744           12229       14338       established
```

```
  subscriber1@isp.net
  interface: LNS-GROUP-INT
  service-id: 1
  10.48.127.27
-------------------------------------------------------------------------------
No. of sessions: 1
===============================================================================
*A:LNS#
```

The PPP session is also recorded in the subscriber-host table of VPRN 1 and a forwarding state of *Fwding* indicates that all attributes and resources associated with this subscriber are correctly installed and activated within the system. The subscriber username is shown, as is its MAC address and IP address. The IP address has an origin of IPCP. The fact that a MAC address is displayed here is somewhat misleading because this is a PPP over L2TP session, which does not have a MAC address present in any of its headers. When the MS-ISA removes the L2TP header, it converts the PPP packet to PPPoE for ease of subsequent processing. As a result of this, the MS-ISA generates a dummy MAC address, and this is the MAC address shown. The displayed SAP 1/2/lns-esm:1.259 is automatically generated by the system. Each operational MS-ISA that is part of the lns-group creates two internal objects, known as lns-net and lns-esm. These objects are the network-side (lns-net) and the subscriber-side (lns-esm) of each MS-ISA.

When the first L2TP session within this service is established, the system creates one lns-esm SAP where the first two digits indicate the MDA slot (1/2) where the MS-ISA is installed, and the last two numbers are the internal Q-in-Q tags used through the MS-ISA (1.259). This internal Q-in-Q tag value is of little relevance, but for informational purposes is derived from the group interface If index. If there is more than one MS-ISA active in the lns-group, a second session would be load-balanced onto this MS-ISA, and a second lns-esm SAP would be created, until a maximum of six SAPs is reached, which corresponds to the maximum number of supported active MS-ISA boards.

```
*A:LNS# show service id 1 subscriber-hosts
=============================================================
Subscriber Host table
=============================================================
Sap                    Subscriber
  IP Address
    MAC Address          PPPoE-SID Origin       Fwding State
-------------------------------------------------------------
[1/2/lns-esm:1.259]    subscriber1@isp.net
  10.48.127.27
    00:00:2f:c5:38:02    1         IPCP         Fwding
-------------------------------------------------------------
Number of subscriber hosts : 1
=============================================================
*A:LNS#
```

It is also possible to view the internal lns-net object, shown in the next output as interface name *_tmnx_lns-in-1/2* with port number *1/2/lns-net:1**. To further clarify (and reiterate), the lns-esm and lns-net are simply internal objects used to route L2TP traffic through the MS-ISA board. Upstream traffic (subscriber to LNS) ingresses through lns-net into the MS-ISA where the L2TP header is decapsulated before PPP packets are presented to the service group interface through lns-esm. Downstream traffic (LNS to subscriber) passes through lns-esm into the MS-ISA where the PPP packets are encapsulated in L2TP before egressing through lns-net and being routed toward the destination.

```
*A:LNS# show service id 1 interface
==============================================================================
Interface Table
==============================================================================
Interface-Name                   Adm       Opr(v4/v6)  Type    Port/SapId
   IP-Address                                                  PfxState
```

```
-------------------------------------------------------------------------------
system                           Up        Up/Down     VPRN    loopback
   192.168.0.2/32                                              n/a
LNS-SUB-INT                      Up        Up/Down     VPRN S* subscriber
   Unnumbered If[192.168.0.2]                                  n/a
LNS-GROUP-INT                    Up        Up/Down     VPRN G* bbg-5.lns-esm
_tmnx_lns-in-1/2                 Up        Up/Down     VPRN    1/2/lns-net:1*
   -                                                           -
-------------------------------------------------------------------------------
Interfaces : 4
===============================================================================
```

Wholesale/Retail

In the example configuration used so far, the L2TP tunnel/session is terminated in VPRN 1, and the subscriber is also terminated in the same VPRN 1. However, a common requirement is to build per-customer VRFs (VPRNs), particularly for business users. To meet this requirement, the so-called ‘Wholesale/Retail’ model is used, which provides a mechanism to terminate the subscriber in a different service context from the service which actually terminated the L2TP tunnel/session.

To achieve this, a second service is created which becomes the ‘Retail VRF’, or customer-specific VRF, and the previously defined VPRN 1 becomes the Wholesale VRF (which actually requires no further configuration). The necessary configuration for the Retail VRF is as follows and its parameters have been previously explained. Although they may seem obvious, there are a couple of points that are worth revisiting. The **vrf-import** and **vrf-export** parameters are used to reference policies to import/export VPN-IPv4/v6 prefixes with the customer-specific Route-Target Extended Communities. Given that a different routing context and unique Route-Targets are used for this Retail VRF, it is perfectly feasible to re-use the same IP address in VPRN 2 as was used in VPRN 1 for the unnumbered subscriber interface. The group interface has a different name from the group interface in VPRN 1, but this is simply for illustration purposes and both group interfaces can have the same name if a standard naming convention is required. More importantly, the group interface must have the creation-time attribute **lns** to allow subscriber termination without SAPs. The static route black-holes prefix 10.10.148.0/24, ensuring this prefix is added to the route-table. This IP address range is used to allocate addresses to subscribers, and is therefore advertised in VPN-IPv4.

```
configure
    service
        vprn 2 customer 1 create
            vrf-import "vrf2-import"
            vrf-export "vrf2-export"
            route-distinguisher 64496:2
            auto-bind-tunnel
                resolution-filter
                    ldp
                exit
                resolution filter
            exit
            interface "loopback" create
                address 192.168.0.2/32
                loopback
            exit
            subscriber-interface "VPRN2-SUB-INT" create
                unnumbered 192.168.0.2
                group-interface "VPRN2-GROUP-INT" lns create
                    sap-parameters
                        sub-sla-mgmt
                            sub-ident-policy "all-subscribers"
                        exit
                    exit
                exit
            exit
            static-route-entry 10.10.48.0/24
                black-hole
                    no shutdown
                exit
            exit
            no shutdown
        exit
    exit
exit
```

In addition to the Retail VRF configuration, the RADIUS entry for the subscriber returns **Alc-Serv-Id** VSA with a value of 2 to indicate the Retail VRF Service Id, while the **Alc-Interface** VSA refers to the group interface name within that Retail VRF.

```
subscriber2@isp.net     Cleartext-Password := "letmein"
                        Alc-Subsc-ID-Str = "subscriber2@isp.net",
                        Alc-Subsc-Prof-Str = "sub-profile-1",
                        Alc-SLA-Prof-Str = "sla-profile-1",
                        Alc-Serv-Id = "2",
                        Alc-Interface = "VPRN2-GROUP-INT",
                        Service-Type = Framed-User,
                        Framed-Protocol = PPP,
                        Framed-IP-Address = 10.10.148.22
```

In this Wholesale/Retail scenario, the high-level functions are as follows:

- The L2TP tunnel and session are terminated in the Wholesale VRF (in this example, VPRN 1).
- When the LNS receives the ICCN for the session, it authenticates the user (in this example via RADIUS).

RADIUS returns the Retail VRF service Id and group interface. If RADIUS returns IP address information this address is used for the purpose of IPCP negotiation with the subscriber within the Retail VRF (in this example, VPRN 2). If RADIUS does not return IP address information, it can be derived from either of the following:

- A DHCP client function within the group interface, which is used to obtain an IP address from a local or remote DHCP server.
- The local-address-assignment feature, which directly accesses a local DHCP server through an internal procedure call (the server pool name must be obtained through RADIUS, LUDB, or default-pool-name).

Once the subscriber is activated, the PPP session and subscriber-host can be seen in VPRN 2. The description field of the **show service id 2 ppp session** command is however somewhat misleading. It is automatically concatenated from the VPRN that terminated the L2TP tunnel, the tunnel Connection Id, the local tunnel Id, and the L2TP session Id. It should not be misinterpreted as meaning that the subscriber has been terminated in VPRN 1.

```
*A:LNS# show service id 2 ppp session

===============================================================================
PPP sessions for service 2
===============================================================================
User-Name
  Descr.
           Up Time       Type  Termination     IP/L2TP-Id/Interface-Id MC-Stdby
-------------------------------------------------------------------------------
```

```
subscriber2@isp.net
  vprn:1 connid:449518148 tid:6859 sid:6724
           0d 00:00:10   oL2tp local           10.10.148.22
-------------------------------------------------------------------------------
No. of PPP sessions: 1
===============================================================================
*A:LNS#
```

QoS

In the preceding examples, the subscriber PPP sessions terminated by the LNS have been instantiated using the default SAP-ingress/egress QoS policies (policy 1), with a single queue and no use of H-QoS. This section demonstrates the use of slightly more complex QoS policies that employ H-QoS, with the intention of providing an overview of those capabilities.

For subscriber termination in broadband networks, it is fairly commonplace to use one or more policers on ingress, and not apply an aggregate rate limit on ingress (upstream) traffic. Whilst this is possible in SR OS for general ESM subscriber termination, policers are not supported when the system is functioning as an LNS. It is therefore necessary to use one or more queues on ingress with the usual considerations with regard to the use of service-queuing or shared-queuing. Conversely, on egress (downstream) it is common to see more than one queue in use for different services, particularly for business services, with an aggregate rate applied to the subscriber through the use of H-QoS. For example, assume that there are three classes in use; Best-Effort (BE), Assured-Forwarding (AF), and Expedited Forwarding (EF). This section will look at two ways to achieve this. Firstly using a conventional H-QoS scheduler, and secondly using an egress Port-Scheduler.

The SAP-ingress QoS policy classifies traffic into three Forwarding Classes (FCs) and maps those FCs to a single queue. Ingress traffic is not rate-limited (default PIR in queue 1 is max), and queue 1 is not mapped to a parent H-QoS scheduler.

```
configure
    qos
        sap-ingress 10 create
            queue 1 create
            exit
            queue 11 multipoint create
            exit
            fc "af" create
                queue 1
            exit
            fc "be" create
                queue 1
            exit
            fc "ef" create
                queue 1
            exit
            dscp be fc "be"
            dscp ef fc "ef"
            dscp af31 fc "af"
        exit
    exit
exit
```

A scheduler policy is created having a single a tier 1 scheduler with a rate-limit of 8Mb/s.

```
configure
    qos
```

```
        scheduler-policy "Subscriber-Aggregate-Policy" create
            tier 1
                scheduler "Aggregrate-Rate" create
                    rate 8000
                exit
            exit
        exit
    exit
exit
```

The SAP-egress QoS policy performs egress classification and maps classified traffic to the relevant FC, which in turn is mapped to its own queue. All queues are mapped to the previously configured tier 1 scheduler *Aggregate-Rate* such that queue 3 (EF) is allocated bandwidth first, and queue 1 (BE) and 2 (AF) are allocated bandwidth next in a 1:4 ratio.

```
configure
    qos
        sap-egress 10 create
            queue 1 create
                parent "Aggregrate-Rate" level 2 weight 20
            exit
            queue 2 best-effort create
                parent "Aggregrate-Rate" level 2 weight 80
            exit
            queue 3 expedite create
                parent "Aggregrate-Rate" cir-level 3
                rate 1024 cir 1024
            exit
            fc af create
                queue 2
            exit
            fc be create
                queue 1
            exit
            fc ef create
                queue 3
            exit
            dscp be fc "be"
            dscp ef fc "ef"
            dscp af31 fc "af"
        exit
    exit
exit
```

To this point, the QoS configuration is no different from a typical SAP-level QoS application. To make it applicable to ESM, the previously configured SAP-ingress and SAP-egress QoS policies must be referenced in the ingress/egress contexts of the sla-profile, respectively. Equally, the H-QoS scheduler-policy must be referenced in the ingress/egress contexts of the sub-profile. In this example, H-QoS is only used on egress, and as a result the scheduler-policy is referenced only in the egress context.

```
configure
    subscriber-mgmt
        sla-profile "ESM-SLA-PROF" create
            ingress
                qos 10
                exit
            exit
            egress
                qos 10
                exit
                no qos-marking-from-sap
            exit
        exit
        sub-profile "ESM-SUB-PROF" create
            collect-stats
            radius-accounting
                policy "AAA-ACCT-POLICY"
            exit
            egress
                scheduler-policy "Subscriber-Aggregate-Policy"
                exit
            exit
        exit
    exit
exit
```

The queues assigned to the subscriber through the preceding SAP-ingress/egress QoS policies, together with accumulative statistics can be viewed using the **show service active-subscribers subscriber** <*name*> **detail** command (real time rates can be seen using the **monitor** command). The H-QoS scheduler hierarchy, with the SAP-egress queues mapped as child queues to a parent scheduler can be validated using the command **show qos scheduler-hierarchy subscriber** <*name*> **egress**. The **detail** argument as an extension of this command provides a significant amount of detail on real-time bandwidth allocated to each queue by the scheduler in the within-CIR and above-CIR passes. It also provides a useful snapshot on offered traffic loads in Kb/s on a per-queue basis.

```
*A:LNS# show qos scheduler-hierarchy subscriber "subscriber2@isp.net" egress

===============================================================================
Scheduler Hierarchy - Subscriber subscriber2@isp.net
===============================================================================
Egress Scheduler Policy : Subscriber-Aggregate-Policy
-------------------------------------------------------------------------------
Root (Egr)
| slot(1)
|--(S) : Aggregrate-Rate (Port 1/2/lns-esm Orphan)
|   |
|   |--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->3
|   |
|   |--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->2
|   |
|   |--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->1
|   |
|


===============================================================================
*A:LNS#
```

The advantage of using conventional H-QoS schedulers is that they can be applied universally on ingress and egress to provide a subscriber aggregate rate capability. The disadvantage of this approach is that the aggregate rate defined in the scheduler-policy (or overridden in the sub-profile) cannot be dynamically overridden from RADIUS using the QoS-override VSA (**Alc-Subscriber-QoS-Override**). If ingress H-QoS is not a requirement, but the ability to override the subscriber egress aggregate-rate is, then H-QoS should be implemented using an egress port-scheduler.

The egress port-scheduler is functionally the same as a conventional H-QoS scheduler in the manner with which it arbitrates bandwidth across its child queues. However, it has some notable differences:

- It is applied at the egress port level. Any queue that uses that egress port to which it is applied that is not explicitly mapped to a port-scheduler is considered an orphan queue. Orphan queues are not serviced by the port-scheduler until all of its child queues have been serviced.

- Unlike conventional H-QoS schedulers that include only Ethernet overhead, the port-scheduler includes Preamble and Inter-Frame Gap for each packet.
- It is supported only on Ethernet ports, and only on egress.
- The egress aggregate rate applied to the subscriber can be overridden from RADIUS.

The first bullet point above is significant from an LNS perspective. In general, after ESM handling, downstream traffic for subscribers egresses the system over a physical port. This is not the case for L2TP subscribers, for which the traffic is passed through to the MS-ISA for L2TP encapsulation before egressing the LNS (and in fact could egress the system on any number of physical ports). It is therefore not possible to apply the port-scheduler policy to the egress port in the conventional manner, and what is needed is a mechanism to apply the port-scheduler policy to the logical internal ports that interface to the MS-ISA. To achieve this, an intermediate object known as a **port-policy** is used, which, when configured, references the **port-scheduler** policy, and which subsequently is applied to the relevant **lns-group**.

Create the port-scheduler-policy.

```
configure
    qos
        port-scheduler-policy "egress-port-scheduler" create
        exit
    exit
exit
```

Create the port-policy and reference the previously configured port-scheduler policy.

```
configure
    port-policy "isa-port-policy" create
        egress-scheduler-policy "egress-port-scheduler"
    exit
exit
```

Attach the port-policy to the lns-group containing the MS-ISA.

```
configure
    isa
        lns-group 1 create
            shutdown
            port-policy "isa-port-policy"
            no shutdown
        exit
    exit
exit
```

Once the **port-scheduler** policy and **port-policy** are in place, the subscriber QoS can reference it. The QoS configuration previously used for conventional H-QoS schedulers differs in both the **sap-egress** policy and **sub-profile** when an egress **port-scheduler** is used. The queues within the **sap-egress** policy are each configured to be parented to the egress port-scheduler using the **port-parent** keyword (as opposed the **parent** keyword used for conventional H-QoS schedulers).

```
configure
    qos
        sap-egress 10 create
            queue 1 create
                port-parent level 2 weight 20
            exit
            queue 2 best-effort create
```

```
                port-parent level 2 weight 80
            exit
            queue 3 expedite create
                port-parent cir-level 3
                rate 1024 cir 1024
            exit
            fc af create
                queue 2
            exit
            fc be create
                queue 1
            exit
            fc ef create
                queue 3
            exit
            dscp be fc "be"
            dscp ef fc "ef"
            dscp af31 fc "af"
        exit
    exit
exit
```

The sub-profile contains no reference to scheduler policies, but instead contains a per-subscriber egress aggregate rate in Kb/s, defined through the **agg-rate-limit** parameter.

```
configure
    subscriber-mgmt
        sub-profile "sub-profile-1"
            egress
                no scheduler-policy
                agg-rate-limit 8000
            exit
        exit
    exit
exit
```

Once again, the queues assigned to the subscriber through the preceding SAP-ingress/egress QoS policies, together with accumulative statistics, can be viewed using the **show service active-subscribers subscriber** <*name*> **detail** command (real time rates can be seen using the **monitor** command). The scheduler SAP-egress queues mapped as child queues to a port-scheduler can be validated using the **show qos scheduler-hierarchy subscriber** <*name*> **egress** command. The **detail** argument provides a significant amount of detail on bandwidth allocated to each queue by the scheduler in the within-CIR and above-CIR passes. It also provides a useful snapshot on offered traffic loads in Kb/s on a per-queue basis. Alternatively, all of the child queues and orphans mapped to the port-scheduler can be displayed using the **show qos scheduler-hierarchy port** <*slot/mda/lns-esm*> command, again with the optional **detail** argument.

```
*A:LNS# show qos scheduler-hierarchy subscriber "subscriber2@isp.net" egress

===============================================================================
Scheduler Hierarchy - Subscriber subscriber2@isp.net
===============================================================================
Egress Scheduler Policy :
-------------------------------------------------------------------------------
Root (Egr)
| slot(1)
|--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->3  (Port
 1/2/lns-esm)
|
|--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->2  (Port
 1/2/lns-esm)
|
|--(Q) : Sub=subscriber2@isp.net:sla-profile-1 2->1/2/lns-esm:1.263->1  (Port
 1/2/lns-esm)
|


================================================================================
*A:LNS#
```

With the previously configured QoS policies and schedulers available, the aggregate rate limit in use for the subscriber can be viewed using the **show service active-subscribers subscriber** <*name*> **detail** command. There are three fields in this output that are of interest here. The **E. Agg Rate Limit** field shows the configured rate-limit in the sub-profile and is therefore relatively static. The **RADIUS Rate-Limit** field shows the aggregate rate received by RADIUS using the **Alc-Subscriber-QoS-Override** VSA, which overrides any rate limit statically configured in the sub-profile. Finally, the **Oper-Rate-Limit** shows the static or RADIUS-received rate-limit, minus any other H-QoS adjustments, such as Multicast H-QoS adjustment (snooping on IGMP joins) or ANCP line-rate adjustments.

```
*A:LNS# show service active-subscribers subscriber "subscriber2@isp.net" detail
             | match expression " E. Agg Rate Limit|RADIUS Rate-Limit|Oper-Rate-Limit"
E. Sched. Policy : N/A                              E. Agg Rate Limit: 8000
RADIUS Rate-Limit: N/A
Oper-Rate-Limit  : 8000
Hs-Oper-Rate-Limit   : Maximum
*A:LNS#
```

Overriding the **agg-rate-limit** defined in the sub-profile can be done as part of the RADIUS Access-Accept, or through a Change of Authorization (CoA), and as previously outlined uses the **Alc-Subscriber-QoS-Override** VSA. This override function can be used, for example, to reconcile the LNS aggregate rate with the subscriber downstream rate learned through the TxConnectSpeed AVP in the ICCN message from the LAC. This ensures that the LNS does not overwhelm any downstream access node, and ensures that the LNS is responsible for all QoS scheduling in the event of congestion. In the following example, an override of the aggregate rate to 10Mb/s is sent as a CoA.

```
63 2019/05/24 13:47:00.340 CEST MINOR: DEBUG #2001 Base RADIUS
"RADIUS: Receive
  Change of Authorization(43) id 178 len 66 from 172.16.1.11:38256 vrid 1
    SESSION ID [44] 22 020DFF000000225CE7D31B
    VSA [26] 16 Nokia(6527)
      SUBSCRIBER QOS OVERRIDE [126] 14 e:r:rate=10000

"

64 2019/05/24 13:47:00.341 CEST MINOR: DEBUG #2001 Base RADIUS
"RADIUS: Transmit
  Change of Authorization Ack(44) 172.16.1.11:38256 id 178 len 20 vrid 1

"
```

Re-issuing the **show service active-subscribers subscriber** <*name*> **detail** command after the CoA shows that the **RADIUS Rate-Limit** field and the **Oper-Rate-Limit** field both correctly show 10Mb/s.

```
*A:LNS# show service active-subscribers subscriber "subscriber2@isp.net" detail
         | match expression " E. Agg Rate Limit|RADIUS Rate-Limit|Oper-Rate-Limit"
E. Sched. Policy : N/A                              E. Agg Rate Limit: 8000
RADIUS Rate-Limit: 10000
Oper-Rate-Limit  : 10000
```

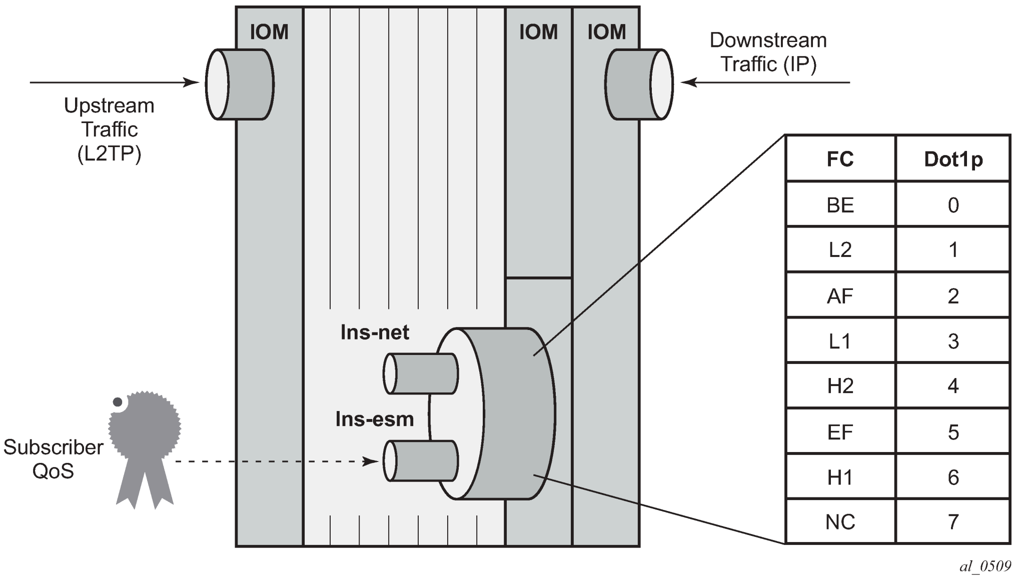
```
Hs-Oper-Rate-Limit   : Maximum
*A:LNS#
```

Propagating QoS Markings to L2TP/MPLS Headers

It is often desirable to mark the L2TP header (DCSP) or MPLS header (EXP) based on the class of service that is carried in the encapsulated subscriber IP payload. In general, when a packet is classified and mapped to an FC on ingress, that FC value is carried in the internal switch fabric header and is present when the packet is egressing the node. In the case of L2TP traffic however, the operation becomes a little more complex due to the fact that traffic transits the MS-ISA board with subscriber QoS implemented 'mid-chassis'.

In the upstream direction (from subscriber to LNS), traffic arrives encapsulated in L2TP at the ingress IOM, and is passed through the MS-ISA via the internal object lns-net. When traffic exits the MS-ISA through lns-esm as native IP, it is subject to subscriber ingress QoS implemented on the (MS-ISA) carrier IOM. Traffic is classified and mapped to an FC at this point, and that FC mapping is maintained in the switch fabric header. As a result, marking of traffic is effected by the network egress QoS policy.

Figure 77: Ingress/Egress QoS Processing



FC	Dot1p
BE	0
L2	1
AF	2
L1	3
H2	4
EF	5
H1	6
NC	7

In the downstream direction (from LNS to subscriber), traffic arrives at the ingress IOM as native IP and is diverted to the MS-ISA via the lns-esm internal object. At the lns-esm, the traffic is subject to subscriber egress QoS. When the traffic is passed through lns-esm to the MS-ISA for L2TP encapsulation, internal Q-in-Q VLAN tags are attached as previously described. As the lns-esm is effectively a SAP-egress, the internal switch fabric header containing the FC marking is removed at this point, and, as a result FC information is lost.

Therefore, in order to allow for FC-continuity through the MS-ISA, the system implements a queue-group at the ingress of lns-net that has a dot1p to FC mapping as shown in Figure 77: Ingress/Egress QoS Processing. Assuming a SAP-egress QoS policy that employs FCs BE, AF and EF, the QoS policy would include the additional configuration to implement the appropriate dot1p marking as shown in the following output. When traffic arrives at lns-net, it is classified and mapped into the appropriate FCs, and the associated FC mapping included in the switch fabric header. At network egress, the L2TP packet is then subject to marking as defined in the network egress QoS policy.

```
configure
    qos
        sap-egress 10 create
            queue 1 create
            exit
            queue 2 best-effort create
            exit
            queue 3 expedite create
            exit
            fc af create
                queue 2
                dot1p 2
            exit
            fc be create
                queue 1
                dot1p 0
            exit
            fc ef create
                queue 3
                dot1p 5
            exit
            dscp be fc "be"
            dscp ef fc "ef"
            dscp af31 fc "af"
        exit
    exit
exit
```

Framed-Route

The majority of residential services in broadband networks have a single registered 32-bit IPv4 address on the WAN side of the RG and a private (RFC 1918) network on the LAN side. Traffic from the LAN toward the BNG (and Internet) is thereafter subject to Network Address and Port Translation (NAPT). However, a common requirement for delivery of business services is the ability for the BNG to recognize one or more IP subnets on the LAN side of the RG that is not subject to NAT, and the subscriber prefix is a route to a network. This is achieved using the standard RADIUS **Framed-Route** attribute, or dynamic BGP peering. Both serve the function of allowing one or more subnets to be learned at the LNS with a next-hop IP address of the RG WAN.

To provide an example of the use of Framed-Route, the Retail VRF VPRN 2 is again used, and in fact requires no modification in order to support subscribers with Framed-Routes. In general ESM, where Framed-Route is used, there is a requirement to configure **anti-spoof type nh-mac**, but for LNS SAPs this is the default. The RADIUS users file is updated to also return a Framed-Route attribute for prefix 10.128.46.0/24 with a next-hop determined by the subscriber IP prefix. The prefix has a metric of 10, and has a tag of value 200, which may be used for example, for routing policy.

```
subscriber2@isp.net     Cleartext-Password := "letmein"
                        Alc-Subsc-ID-Str = "subscriber2@isp.net",
                        Alc-Subsc-Prof-Str = "sub-profile-1",
                        Alc-SLA-Prof-Str = "sla-profile-1",
                        Alc-Serv-Id = "2",
                        Alc-Interface = "VPRN2-GROUP-INT",
                        Service-Type = Framed-User,
                        Framed-Protocol = PPP,
                        Framed-IP-Address = 10.10.148.22,
                        Framed-Route = "10.128.46.0/24 0.0.0.0 10 tag 200",
```

In SR OS, a prefix learned through the Framed-Route attribute is known internally as a *Managed Route*. Once the subscriber is instantiated, the presence of the Managed Route can be verified as installed.

```
*A:LNS# show service id 2 ppp session detail | match "Managed Routes"
                                                               post-lines 5
Managed Routes
-------------------------------------------------------------------------------
IP Address                                  Status       Metric Tag        Pref
-------------------------------------------------------------------------------
10.128.46.0/24                              installed    10     200        0
-------------------------------------------------------------------------------
*A:LNS#
```

The Managed Route can also be seen present in the VPRN routing-table, learned through protocol *Managed*.

```
*A:LNS# show router 2 route-table protocol managed

===============================================================================
Route Table (Service: 2)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
10.128.46.0/24                                Remote  Managed   00h01m33s  0
       10.10.148.22                                                 10
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
*A:LNS#
```

L2TP Tunnel Switching (LTS)

In general, L2TP tunnels are established directly between LAC and LNS. However, if there are a large number of LAC devices (and therefore a large number of L2TP tunnels), it may be desirable to perform some aggregation of these tunnels before presenting them to the LNS. This is implemented by one or more LNS devices performing the function of an L2TP Tunnel Switch (LTS). The LTS terminates multiple L2TP tunnels from the LAC(s), and sources a single L2TP tunnel toward the target LNS, switching L2TP sessions from one tunnel to another tunnel accordingly.

SR OS supports LTS functionality, and from a configuration perspective, it requires no more than L2TP being enabled (no shutdown) in the required routing context if the relevant attributes are returned from RADIUS. As with LNS functions, at least one MS-ISA is required to support LTS functions. In this example, VPRN 1 is used with the previously defined configuration. To recap, this VPRN has a single L2TP group “L2TP-GROUP-1”, and within that group, a single tunnel defined “L2TP-TUNNEL-1” that terminates the tunnel from the LAC. To demonstrate LTS functionality, the LAC continues to function as a LAC, the LNS functions as an LTS, and PE-1 becomes the LNS.

The RADIUS users file for subscriber1@isp.net is modified to include a number of additional attributes and VSAs. The **Alc-Serv-Id** and **Alc-Interface** define the service ID and group interface where the subscriber is terminated, and this can be any IES or VPRN service. The **Alc-Tunnel-Serv-Id** VSA identifies the service where the L2TP tunnel is initiated. It can be different from the service where the subscriber is terminated, but id does not need to. If it is a different service, then the minimum requirement is that L2TP is placed in a no shutdown state. In this example VPRN 1 terminates the subscriber, and also initiates the L2TP tunnel. The other attributes are standard attributes defined in RFC 2868 for L2TP tunnel setup. The Tunnel-Assignment-Id attribute is used to maintain the concept of groups and tunnels, where Tunnel-Assignment-Id:0 is used to indicate the group name and Tunnel-Assignment-Id:1 is used to indicate the tunnel name. This provides sufficient information for the LTS to initiate an L2TP tunnel without any additional nodal configuration.

```
subscriber1@isp.net     Cleartext-Password := "letmein"
                        Alc-Subsc-ID-Str = "subscriber1@isp.net",
                        Alc-Subsc-Prof-Str = "sla-profile-1",
                        Alc-SLA-Prof-Str = "sub-profile-1",
                        Alc-Serv-Id = "1",
                        Alc-Interface = "LNS-GROUP-INT",
                        Alc-Tunnel-Serv-Id = 1,
                        Tunnel-Assignment-Id:0 = "RADIUS-returned-Tunnel-Group",
                        Tunnel-Type:1 += L2TP,
                        Tunnel-Medium-Type:1 += IP,
                        Tunnel-Server-Endpoint:1 += 192.168.0.3,
                        Tunnel-Password:1 += "letmein",
                        Tunnel-Assignment-Id:1 += "RADIUS-returned-Tunnel-Name",
                        Tunnel-Client-Auth-Id = "LTS",
```

The LAC forwards the PPP session into the LAC to LTS tunnel, and after the LTS receives the ICCN from the LAC, it proceeds in authenticating the subscriber. RADIUS returns the preceding attributes with sufficient information for the LTS to instantiate the subscriber and initiates an L2TP tunnel/session with PE-1, the target LNS. The LNS then authenticates the user once more, this time providing it with IP address information through IPCP negotiation. This interaction between PPP client and LNS is transparent to the LTS, which is responsible for switching PPP packets between L2TP sessions. However, the user is instantiated in the system as a full-fledged subscriber.

```
*A:LNS# show service id 1 ppp session

===============================================================================
PPP sessions for service 1
===============================================================================
User-Name
  Descr.
           Up Time       Type  Termination     IP/L2TP-Id/Interface-Id MC-Stdby
-------------------------------------------------------------------------------
subscriber1@isp.net
  vprn:1 connid:345328060 tid:5269 sid:18876
           0d 00:01:42   oL2tp lac             905456423
-------------------------------------------------------------------------------
No. of PPP sessions: 1
===============================================================================
*A:LNS#
```

Within VPRN 1, two L2TP tunnels are active. The entry with Connection Id 937033728 belongs to group *RADIUS-returned-Tunnel-Group* (obtained from the Tunnel-Assignment-Id:0 RADIUS attribute) and has tunnel name *RADIUS-returned-Tunnel-Name* (obtained from the Tunnel-Assignment-Id:1 RADIUS attribute). This is the tunnel from LTS to LNS, and it is in the *established* state and has one session active. The entry with Connection Id 560005120 is the statically defined tunnel from the LAC, belonging to the

CLI-configured group L2TP-GROUP-1 with tunnel name L2TP-TUNNEL-1. This tunnel is also in the *established* state, with one session active.

```
*A:LNS# show router 1 l2tp tunnel
===============================================================================
Conn ID    Loc-Tu-ID Rem-Tu-ID State              Blacklist-state   Ses Active
  Group                                                             Ses Total
    Assignment
-------------------------------------------------------------------------------
345309184  5269      12695     established        not-blacklisted   1
  L2TP-GROUP-1                                                      1
    L2TP-TUNNEL-1
905445376  13816     12071     established        not-blacklisted   1
  RADIUS-returned-Tunnel-Group                                      1
    RADIUS-returned-Tunnel-Name
-------------------------------------------------------------------------------
No. of tunnels: 2
===============================================================================
*A:LNS#
```

Equally, within VPRN 1, two L2TP sessions are active for subscriber subscriber1@isp.net. Session 937037337 is carried in Tunnel-ID 14298, which, as shown in the previous output, is the tunnel toward the LNS, while session 560030762 is carried in Tunnel-ID 8545, which is the tunnel toward the LAC.

```
*A:LNS# show router 1 l2tp session

===============================================================================
L2TP Session Summary
===============================================================================
ID                  Control Conn ID     Tunnel-ID   Session-ID  State
-------------------------------------------------------------------------------
345328060           345309184           5269        18876       established
  subscriber1@isp.net
  interface: LNS-GROUP-INT
  service-id: 1
  905456423
905456423           905445376           13816       11047       established
-------------------------------------------------------------------------------
No. of sessions: 2
===============================================================================
*A:LNS#
```

IPv6

The deployment of IPv6 into residential broadband networks dictates some design choices, or perhaps even some enforced IPv6 address allocation mechanisms:

- Bridged or Routed Residential Gateways (RGs).
- Numbered or unnumbered WAN.
- Stateful (DHCPv6) or stateless (Stateless Address Auto-Configuration, or SLAAC) address assignment.

The purpose of this example is not to show every possibility, but simply to demonstrate that enabling IPv6 is possible at the LNS, just as if this were a conventional BNG doing PPP Termination and Aggregation (PTA). This example uses a widely adopted approach of dual-stack Routed RG with DHCPv6 Prefix Delegation.

The configuration of VPRN 2 is modified to include some IPv6 parameters. In the **subscriber-interface** context, the **delegated-prefix-len** command is set to **variable** indicating that prefixes delegated to subscribers may be of varying length (the default delegated prefix length is /64). The **allow-unmatching-prefixes** command tells the subscriber interface to operate in an IPv6 unnumbered mode, allowing IPv6 addresses to be allocated to subscribers that do not fall within the range of any IPv6 subnet defined under the subscriber interface. Within the **group-interface** context, the **ipv6** context places router-advertisements into a no shutdown state and has the **managed-configuration** flag set indicating that stateful (DHCPv6) address configuration is to be used.

Also a **dhcp6 proxy-server** is enabled, providing an interworking function between the RADIUS server (where the Delegated Prefix is obtained from) and the DHCPv6 client. The proxy will take the RADIUS-provided prefix and responds to the clients Solicit message with an DHCPv6 Advertise message containing the delegated prefix (IA_PD). Because the DHCPv6 messages from the client need to be received over the subscriber PPP session, the proxy-server is configured to allow this using the **client-applications ppp** command. Finally, there is a static-route for black-holing the /48 IPv6 prefix. The client is allocated a /64 prefix from this range and this static-route is used to provide an aggregated upstream prefix advertisement.

```
configure
    service
        vprn 2
            subscriber-interface "VPRN2-SUB-INT" create
                ipv6
                    default-dns 2001:db8:2c41::56
                    delegated-prefix-len variable
                    allow-unmatching-prefixes
                exit
                group-interface "VPRN2-GROUP-INT" lns create
                    ipv6
                        router-advertisements
                            managed-configuration
                            no shutdown
                        exit
                        dhcp6
                            proxy-server
                                client-applications ppp
                                no shutdown
                            exit
                        exit
                    exit
                exit
            exit
            static-route-entry 2a00:8010:1b00::/48
                black-hole
                    no shutdown
                exit
            exit
            no shutdown
        exit
    exit
exit
```

The RADIUS users file entry for subscriber2@isp.net is also modified to return the IPv6 Delegated Prefix using the standard attribute **Delegated-IPv6-Prefix**.

```

subscriber2@isp.net     Cleartext-Password := "letmein"
                        Alc-Subsc-ID-Str = "subscriber2@isp.net",
                        Alc-Subsc-Prof-Str = "sub-profile-1",
                        Alc-SLA-Prof-Str = "sla-profile-1",
                        Alc-Serv-Id = "2",
                        Alc-Interface = "VPRN2-GROUP-INT",
                        Service-Type = Framed-User,
                        Framed-Protocol = PPP,
                        Framed-IP-Address = 10.10.148.22,
                        Delegated-IPv6-Prefix = 2001:db8:1b00:100::/64
```

After the PPP LCP phase and RADIUS authentication, the LNS is aware that the subscriber also has IPv6 enabled (in this case because it received the **Delegated-IPv6-Prefix** attribute). As a result, the LNS begins to negotiate both IPCP and IPv6CP with the client. For IPv6CP, only an Interface-ID is negotiated, for which the LNS uses an EUI-64 extended version of the chassis MAC address. Once IPv6CP negotiation is completed, the client can initiate a DHCPv6 Solicit for a delegated prefix (IA_PD option). After a successful Advertise/Request/Reply exchange the subscriber is instantiated as dual-stack IPv4/IPv6.

```
*A:LNS# show service active-subscribers subscriber "subscriber2@isp.net"

===============================================================================
Active Subscribers
===============================================================================
-------------------------------------------------------------------------------
Subscriber subscriber2@isp.net (sub-profile-1)
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
(1) SLA Profile Instance sap:[1/2/lns-esm:1.263] - sla:sla-profile-1
-------------------------------------------------------------------------------
IP Address
                MAC Address          Session        Origin       Svc        Fwd
-------------------------------------------------------------------------------
10.10.148.22
                00:00:14:95:29:18    PPP 1          IPCP         2          Y
2001:db8:1b00:100::/64
                00:00:14:95:29:18    PPP 1          DHCP6-PD     2          Y
-------------------------------------------------------------------------------

===============================================================================
*A:LNS#
```

Conclusion

SR OS offers a comprehensive feature set for LNS implementations. The MS-ISA provides the hardware-assist for L2TP encapsulation/de-capsulation while the carrier IOM implements conventional subscriber management functions.

Multi-Chassis IPSec Redundancy

This chapter provides information about multi-chassis IPSec redundancy configurations.

Topics in this chapter include:

- Applicability
- Overview
- Configuration
- Conclusion

Applicability

This initial version of this chapter was based on SR OS Release 10.0.R8, but the CLI in the current edition corresponds to SR OS Release 22.10.R2.

Overview

Multi-Chassis IPSec redundancy (MC-IPSec) is a stateful inter-chassis IPSec failover mechanism. IPSec tunnel states are synchronized between the primary and standby chassis. A tunnel group failure on the primary chassis or a primary chassis failure could trigger MC-IPSec failover to the standby chassis.

The following are some highlights of this feature:

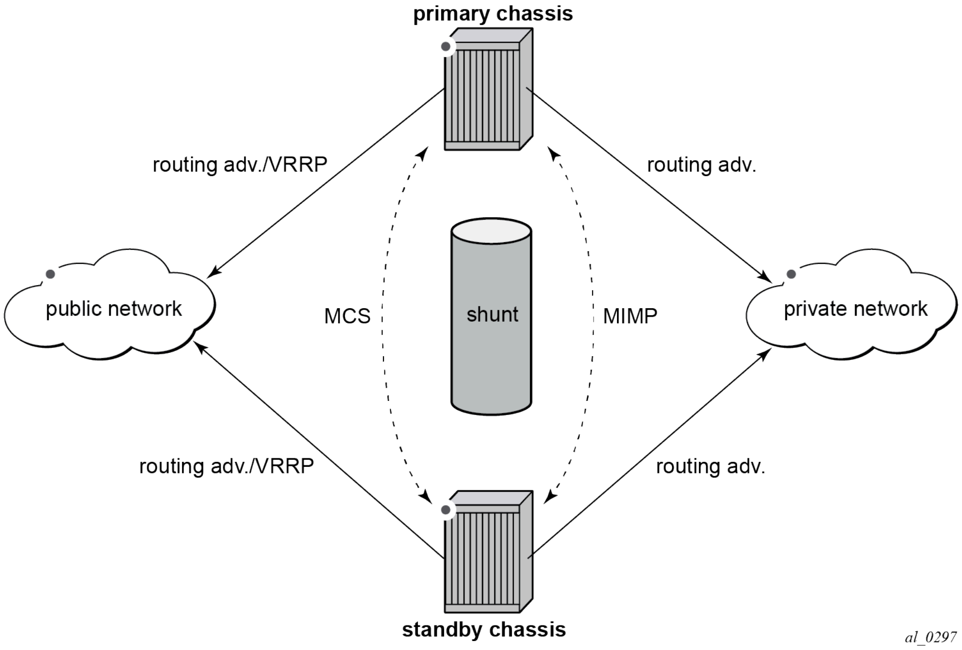
- Internet Key Exchange version 2 (IKEv2) only
- Multi-active tunnel group only
- The granularity of failover is tunnel group, which means a specific tunnel group could failover to the standby chassis independent of other tunnel groups on the primary chassis
- Both static and dynamic LAN-to-LAN tunnels are supported

This feature has the following building blocks:

- Primary chassis election: MC-IPSec mastership protocol (MIMP) runs between the chassis to elect a primary chassis with independent MIMP runs for each tunnel group
- Synchronization: multi-chassis synchronization (MCS) synchronizes the IPSec states between chassis
- Routing:
  - MC-IPSec-aware routing attracts traffic to the primary chassis
  - Shunting support
  - MC-IPSec-aware virtual router redundancy protocol (VRRP)

Figure 78: MC-IPSec architecture shows two redundant IPSec chassis in the middle: a primary chassis and a standby chassis.

Figure 78: MC-IPSec architecture



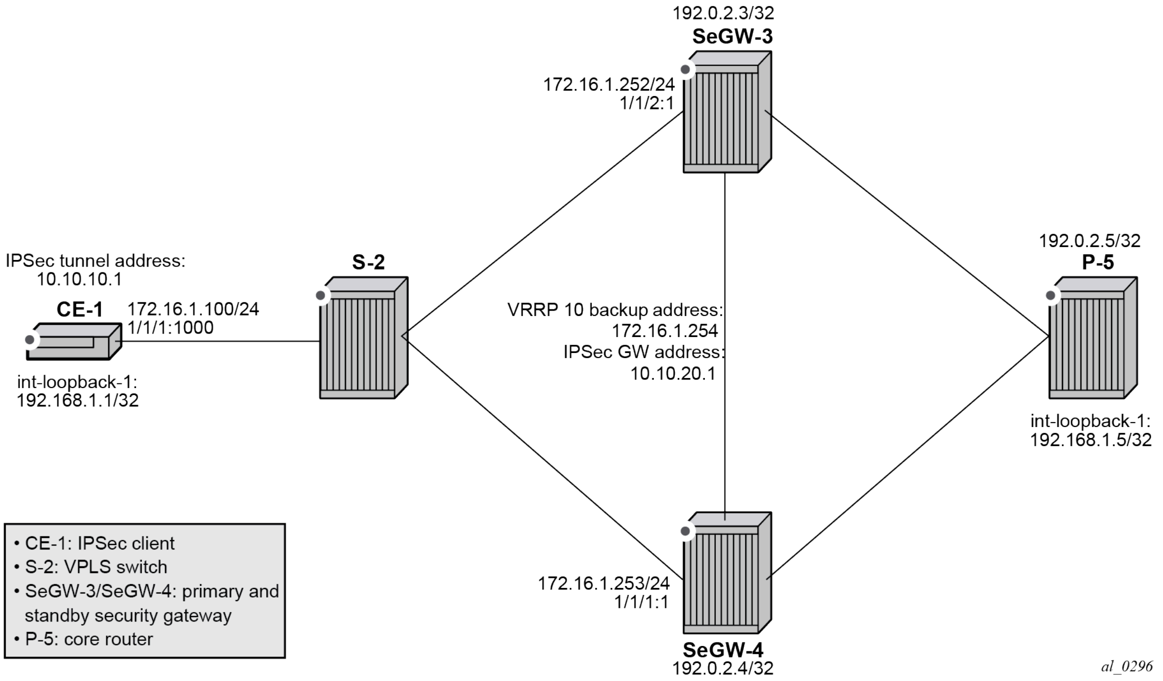
The fundamentals of MC-IPSec are:

- Only the primary chassis processes encapsulating security payload (ESP) and IKE traffic. If the standby chassis receives traffic, it shunts it to the primary chassis, if possible. The traffic is discarded if the standby chassis fails to shunt the traffic.
- The same local gateway address must be provisioned on both chassis.
- MC-IPSec does not synchronize configurations.
- MC-IPSec-aware routing attracts traffic to the primary chassis for both public and private services, which is achieved by exporting the corresponding IPSec routes to the routing protocol using a route policy and setting a different routing metric according to the MC-IPSec state.
- In case of a Layer 2 public network, MC-IPSec-aware VRRP can be used to trigger VRRP switchover upon MC-IPSec switchover.
- MCS synchronizes IPSec states between chassis so that existing IPSec tunnels do not need to be re-established upon switchover.
- MIMP elects mastership between two chassis, and it can also detect chassis failure and tunnel group failure; a central BFD session can be associated with MIMP to achieve fast chassis failure detection.

Configuration

The example topology is shown in Figure 79: Example topology.

Figure 79: Example topology



The example setup includes:

- an IPSec tunnel initiated by CE-1 and terminated on the primary chassis of the two SeGWs.
- a public IES service "IES-1" and a private VPRN service "VPRN-2" configured on CE-1, SeGW-3, and SeGW-4.
- VPRN 2 (also) configured on P-5.
- a static LAN-to-LAN tunnel with pre-shared key.
- a local VPLS service "VPLS-3" on S-2 to simulate a Layer 2 switch.
- VRRP 10 between SeGW-3 and SeGW-4 to provide a backup address 192.168.1.254, which is the default next hop for CE-1.
- VRRP policy 1 bound to VRRP 10 on the primary chassis SeGW-3 to change the in-use priority upon MC-IPSec switchover.
- OSPF as IGP running in the base routing instance between SeGW-3, SeGW-4, and P-5.
- MP-BGP running between SeGW-3, SeGW-4, and P-5 for the VPN-IPv4 address family.

A ping in VPRN 2 between loopback interface address 192.168.1.1 on CE-1 and 192.168.1.5 on P-5 is used to verify the connectivity over the IPSec tunnel.

The MC-IPSec configuration commands are shown below.

```
config>redundancy>multi-chassis>
     peer <ip-address> [create]
          sync
              ipsec
              tunnel-group <tunnel-group-id> sync-tag <tag-name> [create]
          mc-ipsec
              bfd-enable
```

```
               discovery-interval <interval-1> [boot <interval-2>]
               hold-on-neighbor-failure <multiplier>
               keep-alive-interval <interval>
               tunnel-group <tunnel-group-id> [create]
                   peer-group <tunnel-group-id>
                   priority <priority>
                   shutdown
```

```
config>router>policy-options>policy-statement>entry>from>
     state ipsec-master-with-peer|ipsec-non-master|ipsec-master-without-peer
     protocol ipsec
```

```
config>service>ies>if>
config>service>vprn>if>
     static-tunnel-redundant-next-hop <ip-address>
     dynamic-tunnel-redundant-next-hop <ip-address>
```

```
config>isa>tunnel-grp>
     ipsec-responder-only
```

```
config>vrrp>policy>priority-event>
     mc-ipsec-non-forwarding <tunnel-grp-id>
          hold-clear <seconds>
          hold-set <seconds>
          priority <priority-level> explicit
```

The parameters are the following:

- in the **configure redundancy multi-chassis** context:
  - **peer** <*ip-address*> **[create]** — This command creates or enters a multi-chassis peer. The peer address is by default the system address. This can be changed on the peer using the **configure redundancy multi-chassis peer source-address** command.
    - **sync** — This command enters the sync configuration context.
      - **ipsec** — This command enables MCS to synchronize IPSec states.
      - **tunnel-group** <*tunnel-group-id*> **sync-tag** <*tag-name*>**[create]** — This command enables MCS to synchronize the IPSec states of the specified tunnel group. The **sync-tag** parameter is used to match the tunnel group of the peer. The tunnel group states with the same **sync-tag** on both chassis will be synchronized.
    - **mc-ipsec** — This command enters the multi-chassis IPSec configuration context.
      - **bfd-enable** — This command enables tracking a central BFD session; if the BFD session goes down, then the system considers the peer as down and changes the MC-IPSec status of the configured tunnel group accordingly.

        The BFD session uses the source address of MCS as its source address and the MCS peer address as the destination address. Other BFD parameters are configured with the **bfd** command on the interface that the MCS source address resides on.

        The configuration of BFD is optional for MC-IPSec.
      - **discovery-interval** <*interval-1*> **[boot** <*interval-2*>**]** — This command specifies the time interval that the tunnel group stays in **discovery** state. Interval 1 is used as discovery interval when a new tunnel group is added to multi-chassis redundancy (**mp-ipsec**); interval 2 is used

as discovery interval after system boot-up. Interval 2 is optional, and when it is not specified, the value for interval 1 is used. Both intervals have a default value of 300 seconds.

  - **hold-on-neighbor-failure** *<2..25>* — This command specifies the number of keep-alive failures before considering the peer to be down. The default value is 3.
  - **keep-alive-interval** *<5..500>* — This command specifies the time interval of the mastership election protocol keep-alive packets in deciseconds. The default value is 10 deciseconds (1 s).
  - **tunnel-group** <*tunnel-group-id*> **[create]** — This command enables multi-chassis redundancy for the specified tunnel group, or enters an already configured tunnel group context. The configured tunnel groups can failover independently.

    — **peer-group** <*tunnel-group-id*> — This command specifies the corresponding tunnel group ID on the peer node. The peer tunnel group ID is not necessarily equal to local tunnel group ID.

    — **priority** <*priority*> — This command specifies the local priority of the tunnel group, this is used to elect a primary chassis, where the higher number prevails. If the priorities are the same, then the peer which has more active ISAs wins; if the priority and the number of active ISAs are same, then the peer with higher IP address wins. The range is from 0 to 255 and the default value is 100.

- in a **from** statement of a route policy entry:
  - **state ipsec-master-with-peer | ipsec-non-master | ipsec-master-without-peer** — These commands specify the MC-IPSec state in a **from** statement of a route policy entry:
    - **ipsec-master-with-peer**: The tunnel group is the primary chassis with a peer reachable.
    - **ipsec-master-without-peer**: The tunnel group is the primary chassis with peer unreachable.
    - **ipsec-non-master**: The tunnel group is not the primary chassis.
  - **protocol ipsec** — This command specifies IPSec as protocol in a **from** statement of a route policy entry. **protocol ipsec** refers to the /32 local gateway routes (of both static and dynamic tunnels) and reverse route of dynamic tunnel.
- on a public or private IPSec interface in an IES or VPRN service:
  - **static-tunnel-redundant-next-hop** <*ip-address*> and **dynamic-tunnel-redundant-next-hop** <*ip-address*> — These commands specify the redundant next hop address on a public or private IPSec interface (with public or private tunnel SAP) for a static and dynamic IPSec tunnel respectively. The specified next hop address is used by the standby chassis to shunt traffic to the primary chassis in case it receives any traffic. The next hop address is resolved in the routing table of the corresponding service.

    Note:

    - Shunting is supported over:
      - directly connected SAPs
      - spoke SDP terminated IP interfaces
    - Shunting over auto-bind tunnel is not supported.
    - Shunting does not work if the tunnel group is down.
- in the **isa tunnel-group <id>** context:

  - **ipsec-responder-only** — With this command enabled, the system only acts as IKE responder except for the automatic CHILD_SA rekey upon MC-IPSec switchover.

    This command is required for MC-IPSec support of static LAN-to-LAN tunnels.

- in the **vrrp policy <id> priority-event** context:
  - **mc-ipsec-non-forwarding** <*tunnel-grp-id*> — This command creates a VRRP policy priority event: *mc-ipsec-non-forwarding*, which is triggered whenever the specified tunnel group enters the non-forwarding state.
    - **hold-clear** <*seconds*> — This command configures the hold time before clearing the event. The range is from 0 to 86400 seconds and the default value is 0 s.
    - **hold-set** <*seconds*> — This command configures the hold time before setting the event. The range is from 0 to 86400 seconds and the default value is 0 s.
    - **priority** <*priority-level*> **explicit** — This command sets the VRRP in-use priority to the configured value upon the event. The range is from 0 to 254 and the default value is 0.

The initial configuration must include the following:

- The system time of SeGW-3 and SeGW-4 must be the same for the feature to work. Nokia recommends to use a time synchronization protocol such as NTP or SNTP.
- SeGW-3 and SeGW-4 must be IP reachable in the base routing instance because both MCS and MIMP run in the base routing instance.

Configuration of MC-IPSec

In this section, the following steps are described:

- configure CE-1
- configure S-2
- configure P-5
- configure IPSec tunnel on SeGW-3
- enable MC-IPSec for tunnel group on SeGW-3
- configure MC-IPSec-aware routing on SeGW-3
- configure MC-IPSec-aware VRRP on SeGW-3
- configure SeGW-4

Configure CE-1

On CE-1, the following is configured:

- a public IES service "IES-1" and a private VPRN service "VPRN-2".
- a static default route pointing to the VRRP backup address 172.16.1.254.
- a static IPSec tunnel "tunnel-1" with local address 10.10.10.1 and remote address 10.10.20.1.
- a loopback interface in VPRN 2 with address 192.168.1.1/32 to be used as source address for the ping command to verify the connectivity between CE-1 and P-5 over the IPSec tunnel.

The following base router configuration on CE-1 includes a static route with next hop 172.16.1.254, which is the VRRP backup address.

```
# on CE-1
configure
    router Base
        interface "int-CE-1-S-2"
            address 172.16.1.100/24
            port 1/1/1:1000
        exit
        interface "system"
            address 172.31.2.1/32
        exit
        autonomous-system 64496
        static-route-entry 0.0.0.0/0
            next-hop 172.16.1.254           # VRRP backup address
                no shutdown
            exit
        exit
```

IPSec is configured as follows:

```
configure
    ipsec
        ike-transform 1 create
        exit
        ike-policy 1 create
            ike-version 2
            dpd         # dead peer detection (on peer side; not on MC-IPSec chassis)
            ike-transform 1
        exit
        ipsec-transform 1 create
        exit
```

Tunnel group 1 is configured as follows:

```
configure
    isa
        tunnel-group 1 isa-scale-mode tunnel-limit-2k create
            primary 1/2
            no shutdown
        exit
```

The public IES service is configured as follows:

```
configure
    service
        ies 1 name "IES-1" customer 1 create
            interface "int-IPsec-Public-1" create
                address 10.10.10.254/24
                tos-marking-state untrusted
                sap tunnel-1.public:1 create
                exit
            exit
            no shutdown
        exit
```

The private VPRN service on CE-1 is configured as follows. Instead of configuring **delivery-service 1** for the IPSec tunnel, it is possible to configure **delivery-service-name "IES-1"**.

```
configure
    service
        vprn 2 name "VPRN-2" customer 1 create
            ipsec
                security-policy 1 create
                    entry 10 create
                        local-ip 192.168.1.1/32
                        remote-ip 192.168.1.5/32
                    exit
                exit
            exit
            interface int-loopback-1 create
                address 192.168.1.1/32
                loopback
            exit
            interface int-IPsec-private-1 tunnel create
                sap tunnel-1.private:1 create
                    ipsec-tunnel "tunnel-1" create
                        security-policy 1
                        local-gateway-address 10.10.10.1 peer 10.10.20.1 delivery-service 1
                        dynamic-keying
                            ike-policy 1
                            pre-shared-key "pass"
                            transform 1
                        exit
                        no shutdown
                    exit
                exit
            exit
            static-route-entry 192.168.1.5/32
                ipsec-tunnel "tunnel-1"
                    no shutdown
                exit
            exit
            no shutdown
```

Configure S-2

On S-2, a local VPLS service 3 simulates a Layer 2 switch between CE-1, SeGW-3, and SeGW-4:

```
# on S-2
configure
    service
        vpls 3 name "VPLS-3" customer 1 create
            sap 1/1/c1/1:1 create
                description "to SAP in IES 1 on SeGW-3"
            exit
            sap 1/1/c1/2:1000 create
                description "to router interface on CE-1"
            exit
            sap 1/1/c1/3:1 create
                description "to SAP in IES 1 on SeGW-4"
            exit
            no shutdown
        exit
```

Configure P-5

P-5 simulates the core network router, connecting to SeGW-3 and SeGW-4. The configuration on P-5 includes the following:

- a loopback interface with address 192.168.1.5/32 in VPRN 2 is the destination address of the ping traffic from CE-1.
- an MP-BGP session for the VPN-IPv4 address family between P-5, SeGW-3, and SeGW-4.
- GRE spoke SDPs to connect to SeGW-3 and SeGW-4.

On P-5, the following router interfaces are configured in the base router. OSPF is used as IGP.

```
# on P-5
configure
    router Base
        interface "int-P-5-SeGW-3"
            address 192.168.35.2/30
            port 1/1/c1/2:1000
        exit
        interface "int-P-5-SeGW-4"
            address 192.168.45.2/30
            port 1/1/c1/1:1000
        exit
        interface "system"
            address 192.0.2.5/32
        exit
        ospf 0
            area 0.0.0.0
                interface system
                exit
                interface "int-P-5-SeGW-3"
                exit
                interface "int-P-5-SeGW-4"
                exit
            exit
            no shutdown
        exit
```

On P-5, the following GRE SDPs are configured toward SeGW-3 and SeGW-4:

```
configure
    service
        sdp 53 create
            description "GRE SDP toward SeGW-3"
            signaling off
            far-end 192.0.2.3
            no shutdown
        exit
        sdp 54 create
            description "GRE SDP toward SeGW-4"
            signaling off
            far-end 192.0.2.4
            no shutdown
        exit
```

VPRN 2 is configured on P-5, as follows:

```
configure
    service
```

```
        vprn 2 name "VPRN-2" customer 1 create
            interface "int-loopback-1" create
                address 192.168.1.5/32
                loopback
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:2
                    vrf-target target:64496:2
                    no shutdown
                exit
            exit
            spoke-sdp 53:2 create
            exit
            spoke-sdp 54:2 create
            exit
            no shutdown
        exit
```

The BGP configuration on P-5 is as follows:

```
configure
    router Base
        autonomous-system 64496
        bgp
            group "MPBGP"
            family vpn-ipv4
            type internal
                neighbor 192.0.2.3
                exit
                neighbor 192.0.2.4
                exit
            exit
            no shutdown
        exit
```

Configure IPSec tunnel on SeGW-3

The configuration on SeGW-3 is described in four consecutive sections. In this first section, the following is configured:

- the tunnel group, which must be in multi-active mode before MC-IPSec can be enabled.
- an interface "int-Redundant-1", which is a spoke-SDP terminated interface used for shunting.
- GRE SDP 34 toward SeGW-4 and GRE SDP 35 toward P-5.
- IPSec tunnel "tunnel-1" is the tunnel to CE-1; both SeGW-3 and SeGW-4 use the same local gateway address: 10.10.20.1.

The following configures tunnel group 1 on SeGW-3:

```
# on SeGW-3
configure
    isa
        tunnel-group 1 isa-scale-mode tunnel-limit-2k create
            ipsec-responder-only
            multi-active
            mda 1/2
            no shutdown
        exit
```

On SeGW-3, the following router interfaces are configured in the base router. A static route is configured toward CE-1. OSPF is the IGP used between SeGW-3, SeGW-4, and P-5.

```
configure
    router Base
        interface "int-SeGW-3-P-5"
            address 192.168.35.1/30
            port 1/1/1:1000
            no shutdown
        exit
        interface "int-SeGW-3-SeGW-4"
            address 192.168.34.1/30
            port 1/1/3:1000
            no shutdown
        exit
        interface "system"
            address 192.0.2.3/32
            bfd 100 receive 100 multiplier 3
            no shutdown
        exit
        static-route-entry 10.10.10.0/24
            next-hop 172.16.1.100
                no shutdown
            exit
        exit
        ospf 0
            area 0.0.0.0
                interface "system"
                    no shutdown
                exit
                interface "int-SeGW-3-P-5"
                    no shutdown
                exit
                interface "int-SeGW-3-SeGW-4"
                    no shutdown
                exit
            exit
            no shutdown
        exit
```

The IPSec settings are as follows:

```
configure
    ipsec
        ike-transform 1 create
            isakmp-lifetime 172800
        exit
        ike-policy 1 create
            ike-version 2
            ipsec-lifetime 7200
            ike-transform 1
        exit
        ipsec-transform 1 create
        exit
    exit
```

The GRE SDPs are configured as follows:

```
configure
    service
```

```
        sdp 34 create
            description "GRE SDP toward SeGW-4"
            signaling off
            far-end 192.0.2.4
            keep-alive
                shutdown
            exit
            no shutdown
        exit
        sdp 35 create
            description "GRE SDP toward P-5"
            signaling off
            far-end 192.0.2.5
            keep-alive
                shutdown
            exit
            no shutdown
        exit
```

The public IES service is configured as follows. The VRRP configuration will be added in a later step.

```
configure
    service
        ies 1 name "IES-1" customer 1 create
            interface "int-SeGW-3-S-2" create
                address 172.16.1.252/24
                sap 1/1/2:1 create
                    description "SAP to switch S-2"
                exit
            exit
            interface "int-IPsec-Public-1" create
                address 10.10.20.254/24
                tos-marking-state untrusted
                sap tunnel-1.public:1 create
                exit
                static-tunnel-redundant-next-hop 192.168.34.2
            exit
            no shutdown
        exit
```

The private VPRN service is configured as follows:

```
configure
    service
        vprn 2 name "VPRN-2" customer 1 create
            ipsec
                security-policy 1 create
                    entry 10 create
                        local-ip 192.168.1.5/32
                        remote-ip 192.168.1.1/32
                    exit
                exit
            exit
            interface "int-IPsec-Private-1" tunnel create
                sap tunnel-1.private:1 create
                    ipsec-tunnel "tunnel-1" create
                        security-policy 1
                        local-gateway-address 10.10.20.1 peer 10.10.10.1 delivery-service 1
                        dynamic-keying
                            ike-policy 1
                            pre-shared-key "pass"
                            transform 1
                        exit
                        no shutdown
                    exit
                exit
                static-tunnel-redundant-next-hop 192.168.20.2
            exit
            interface "int-Redundant-1" create
                description "interface used for shunting"
                address 192.168.20.1/30
                spoke-sdp 34:20 create
                    ingress
                        vc-label 2049
                    exit
                    egress
                        vc-label 2048
                    exit
                    no shutdown
                exit
            exit
            static-route-entry 192.168.1.1/32
                ipsec-tunnel "tunnel-1"
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:2
                    vrf-target target:64496:2
                    no shutdown
                exit
            exit
            spoke-sdp 34:2 create
                description "SDP to SeGW-4"
            exit
            spoke-sdp 35:2 create
                description "SDP to P-5"
            exit
            no shutdown
        exit
```

Enable MC-IPSec for tunnel group 1 on SeGW-3

In this section, the following steps are described:

- Create a multi-chassis peer using the system address of SeGW-4.
- Enable MCS for IPSec and tunnel group 1.
- Enable MC-IPSec for the tunnel group with a configured priority 200.
- Bind a central BFD session to MC-IPSec from the system interface.

Create multi-chassis peer 192.0.2.4 and enable MCS and MC-IPSec for tunnel group 1:

```
# on SeGW-3
configure
    redundancy
        multi-chassis
            peer 192.0.2.4 create
                sync
                    ipsec
                    tunnel-group 1 sync-tag "tag-1" create
                    no shutdown
                exit
                mc-ipsec
                    bfd-enable
                    tunnel-group 1 create
                        peer-group 1
                        priority 200
                        no shutdown
                    exit
                exit
                no shutdown
            exit
        exit
```

BFD is enabled in the preceding configuration. On the system interface, the following BFD settings are configured:

```
configure
    router Base
        interface "system"
            address 192.0.2.3/32
            bfd 100 receive 100 multiplier 3
            no shutdown
        exit
```

Configure MC-IPSec-aware routing on SeGW-3

In this step, a route policy is defined and applied to VPRN 2.

Route policy "IPsec-to-MPBGP" exports static route 192.168.1.1/32 in VPRN 2 to P-5. This policy sets the local preference of the prefix 192.168.1.1/32 according to the MC-IPSec state:

- for the **ipsec-master-with-peer** state: local preference 200
- for the **ipsec-non-master** state: local preference 100
- for the **ipsec-master-without-peer** state: local preference 200

The state **ipsec-master-without-peer** can be used to attract traffic to the designated primary chassis in case of "dual master" (meaning two chassis lose the MIMP connection in the base routing instance). In this example, SeGW-3 has local preference 200 and SeGW-4 has local preference 100 for **ipsec-master-without-peer**.

The route policy is configured as follows:

```
# on SeGW-3:
configure
    router Base
        policy-options
            begin
            prefix-list "CE-1-Internal"
                prefix 192.168.1.1/32 exact
            exit
            community "vprn2"
                members "target:64496:2"
            exit
            policy-statement "IPsec-to-MPBGP"
                entry 10
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-master-with-peer
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 200
                    exit
                exit
                entry 20
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-non-master
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 100
                    exit
                exit
                entry 30
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-master-without-peer
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 200
                    exit
                exit
                default-action accept
                    community add "vprn2"
                exit
            exit
        commit
```

The BGP configuration on SeGW-3 is as follows:

```
configure
    router Base
        autonomous-system 64496
        bgp
            group "MPBGP"
            family vpn-ipv4
            type internal
                neighbor 192.0.2.4
                exit
                neighbor 192.0.2.5
                exit
            exit
            no shutdown
        exit
```

The route policy is applied as **vrf-export** in VPRN 2:

```
configure
    service
        vprn "VPRN-2"
            bgp-ipvpn
                mpls
                    vrf-export "IPsec-to-MPBGP"
                exit
            exit
        exit
```

Configure MC-IPSec-aware VRRP on SeGW-3

In this section, a VRRP policy is defined that uses the **mc-ipsec-non-forwarding** priority event to lower the in-use VRRP priority upon MC-IPSec switchover, which ensures VRRP and MC-IPSec have the same primary chassis. The VRRP instance needs to be in preempt mode.

This VRRP policy is only configured on the designated VRRP primary chassis SeGW-3, not on the standby chassis. The VRRP policy is applied to the interface "int-SeGW3-S-2" of IES 1.

VRRP policy 1 is configured as follows:

```
# on SeGW-3:
configure
    vrrp
        policy 1
            priority-event
                mc-ipsec-non-forwarding 1
                    priority 50 explicit
                exit
            exit
        exit
```

The VRRP policy is applied in VRRP instance 10 in the IES service:

```
configure
    service
        ies "IES-1"
            interface "int-SeGW-3-S-2"
                address 172.16.1.252/24
                vrrp 10
                    backup 172.16.1.254
                    priority 200
                    policy 1
                    ping-reply
                exit
---snip---
```

Configure SeGW-4

The configuration on the standby chassis SeGW-4 is similar, but with different priorities and without the VRRP policy.

The tunnel group is configured in multi-active mode:

```
# on SeGW-4
configure
    isa
        tunnel-group 1 create
            ipsec-responder-only
            multi-active
            mda 1/2
            no shutdown
        exit
```

The MCS and MC-IPSec configuration is as follows:

```
configure
```

```
    redundancy
        multi-chassis
            peer 192.0.2.3 create
                sync
                    ipsec
                    tunnel-group 1 sync-tag "tag-1" create
                    no shutdown
                exit
                mc-ipsec
                    bfd-enable
                    tunnel-group 1 create
                        peer-group 1
                        priority 150
                        no shutdown
                    exit
                exit
                no shutdown
            exit
        exit
```

The base router configuration on SeGW-4 includes the following router interfaces and a static route to CE-1. OSPF is used as IGP between SeGW-3, SeGW-4, and P-5.

```
configure
    router Base
        interface "int-SeGW-4-P-5"
            address 192.168.45.1/30
            port 1/1/2:1000
            no shutdown
        exit
        interface "int-SeGW-4-SeGW-3"
            address 192.168.34.2/30
            port 1/1/3:1000
            no shutdown
        exit
        interface "system"
            address 192.0.2.4/32
            bfd 100 receive 100 multiplier 3
            no shutdown
        exit
        static-route-entry 10.10.10.0/24
            next-hop 172.16.1.100
                no shutdown
            exit
        exit
        ospf 0
            area 0.0.0.0
                interface "system"
                    no shutdown
                exit
                interface "int-SeGW-4-SeGW-3"
                    no shutdown
                exit
                interface "int-SeGW-4-P-5"
                    no shutdown
                exit
            exit
            no shutdown
        exit
```

The IPSec configuration is as follows:

```
configure
    ipsec
        ike-transform 1 create
            isakmp-lifetime 172800
        exit
        ike-policy 1 create
            ike-version 2
            ipsec-lifetime 7200
            ike-transform 1
        exit
        ipsec-transform 1 create
        exit
```

The following route policy is configured on SeGW-4, The local preference is lower for the **ipsec-master-without-peer** state.

```
configure
    router Base
        policy-options
            begin
            prefix-list "CE-1-Internal"
                prefix 192.168.1.1/32 exact
            exit
            community "vprn2"
                members "target:64496:2"
            exit
            policy-statement "IPsec-to-MPBGP"
                entry 10
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-master-with-peer
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 200
                    exit
                exit
                entry 20
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-non-master
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 100
                    exit
                exit
                entry 30
                    from
                        prefix-list "CE-1-Internal"
                        state ipsec-master-without-peer
                    exit
                    action accept
                        community add "vprn2"
                        local-preference 100    # lower preference on standby SeGW
                    exit
                exit
                default-action accept
                    community add "vprn2"
                exit
            exit
            commit
```

The BGP configuration on SeGW-4 is as follows:

```
configure
    router Base
        autonomous-system 64496
        bgp
            group "MPBGP"
                family vpn-ipv4
                type internal
                neighbor 192.0.2.3
                exit
                neighbor 192.0.2.5
                exit
            exit
            no shutdown
        exit
```

The following GRE SDPs are configured:

```
configure
    service
        sdp 43 create
            description "GRE SDP toward SeGW-3"
            signaling off
            far-end 192.0.2.3
            keep-alive
                shutdown
            exit
            no shutdown
        exit
        sdp 45 create
            description "GRE SDP toward P-5"
            signaling off
            far-end 192.0.2.5
            keep-alive
                shutdown
            exit
            no shutdown
        exit
```

The public IES service is configured as follows:

```
configure
    service
        ies 1 name "IES-1" customer 1 create
            interface "int-SeGW-4-S-2" create
                address 172.16.1.253/24
                vrrp 10
                    backup 172.16.1.254
                    ping-reply
                exit
                sap 1/1/1:1 create
                    description "SAP toward switch S-2"
                exit
            exit
            interface "int-IPsec-Public-1" create
                address 10.10.20.254/24
                tos-marking-state untrusted
                sap tunnel-1.public:1 create
                exit
                static-tunnel-redundant-next-hop 192.168.34.1
            exit
            no shutdown
        exit
```

The private VPRN service is configured as follows:

```
configure
    service
        vprn 2 name "VPRN-2" customer 1 create
            ipsec
                security-policy 1 create
                    entry 10 create
                        local-ip 192.168.1.5/32
                        remote-ip 192.168.1.1/32
                    exit
                exit
            exit
            interface "int-IPsec-Private-1" tunnel create
                sap tunnel-1.private:1 create
                    ipsec-tunnel "tunnel-1" create
                        security-policy 1
                        local-gateway-address 10.10.20.1 peer 10.10.10.1 delivery-service 1
                        dynamic-keying
                            ike-policy 1
                            pre-shared-key "pass"
                            transform 1
                        exit
                        no shutdown
                    exit
                exit
                static-tunnel-redundant-next-hop 192.168.20.1
            exit
            interface "int-Redundant-1" create
                description "interface used for shunting"
                address 192.168.20.2/30
                spoke-sdp 43:20 create
                    ingress
                        vc-label 2048
                    exit
                    egress
                        vc-label 2049
                    exit
                    no shutdown
                exit
            exit
            static-route-entry 192.168.1.1/32
                ipsec-tunnel "tunnel-1"
                    no shutdown
                exit
            exit
            bgp-ipvpn
                mpls
                    route-distinguisher 64496:2
                    vrf-export "IPsec-to-MPBGP"
                    vrf-target target:64496:2
                    no shutdown
                exit
            exit
            spoke-sdp 43:2 create
                description "SDP to SeGW-3"
            exit
            spoke-sdp 45:2 create
```

```
                description "SDP to P-5"
            exit
            no shutdown
        exit
```

Verification

The following will be verified in this section:

- the MC-IPSec status and VRRP status on SeGW-3 and SeGW-4
- the status of the IPSec tunnel on CE-1
- the status of the IPSec tunnel on the SeGWs

Verify the MC-IPSec status on SeGW-3 and SeGW-4

The following is verified:

- SeGW-3 is the primary chassis and SeGW-4 is the standby for tunnel group 1 because SeGW-3 has the higher priority 200.
- SeGW-3 is the primary node for VRRP instance 10 and SeGW-4 is the backup.

SeGW-3 is the primary chassis in tunnel group 1 with priority 200:

```
*A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:37:51

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              200       Up             master
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
==============================================================================
```

SeGW-4 is the standby chassis in tunnel group 1 with priority 150:

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
```

```
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             standby
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

SeGW-3 is the primary node for VRRP instance 10:

```
*A:SeGW-3# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-3-S-2                   10    No  Up   Master       200       1
                                 IPv4      Up   1           200        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

SeGW-4 is backup for VRRP instance 10:

```
*A:SeGW-4# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Backup       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

Verify the IPSec tunnel on CE-1

The following is verified in this section:

- the connectivity between CE-1 and P-5
- the IPSec tunnel information

A ping command is launched from the loopback interface in VPRN 2 on CE-1 to the loopback interface in VPRN 2 on P-5:

```
*A:CE-1# ping router 2 192.168.1.5
PING 192.168.1.5 56 data bytes
64 bytes from 192.168.1.5: icmp_seq=1 ttl=63 time=3.97ms.
64 bytes from 192.168.1.5: icmp_seq=2 ttl=63 time=7.50ms.
64 bytes from 192.168.1.5: icmp_seq=3 ttl=63 time=2.84ms.
64 bytes from 192.168.1.5: icmp_seq=4 ttl=63 time=2.61ms.
64 bytes from 192.168.1.5: icmp_seq=5 ttl=63 time=2.67ms.

---- 192.168.1.5 PING Statistics ----
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min = 2.61ms, avg = 3.92ms, max = 7.50ms, stddev = 1.86ms
```

The following command shows the IPSec tunnel information.

```
*A:CE-1# show ipsec tunnel

===============================================================================
IPsec Tunnels
===============================================================================
TunnelName                       LocalAddress      SvcId        Admn   Keying
  SapId                            RemoteAddress     DlvrySvcId   Oper   Sec
                                                                         Plcy
-------------------------------------------------------------------------------
tunnel-1                         10.10.10.1        2            Up     Dynamic
  tunnel-1.private:1               10.10.20.1        1            Up     1
-------------------------------------------------------------------------------
IPsec Tunnels: 1
===============================================================================
```

Verify the IPSec tunnel on the SeGWs

In this section, the following is verified:

- the MCS database is in-sync, so the tunnel status is up on both chassis.
- P-5 receives two VPN-IPv4 routes for prefix 192.168.1.1/32: the route from SeGW-3 has local preference 200; the route from SeGW-4 has local preference 100.

On both SeGWs, the IPSec tunnel with local address 10.10.20.1 and remote address 10.10.10.1 is up:

```
*A:SeGW-3# show ipsec tunnel

===============================================================================
IPsec Tunnels
===============================================================================
TunnelName                       LocalAddress      SvcId        Admn   Keying
  SapId                            RemoteAddress     DlvrySvcId   Oper   Sec
                                                                         Plcy
-------------------------------------------------------------------------------
tunnel-1                         10.10.20.1        2            Up     Dynamic
  tunnel-1.private:1               10.10.10.1        1            Up     1
-------------------------------------------------------------------------------
IPsec Tunnels: 1
===============================================================================
*A:SeGW-4# show ipsec tunnel

===============================================================================
IPsec Tunnels
===============================================================================
TunnelName                       LocalAddress      SvcId        Admn   Keying
  SapId                            RemoteAddress     DlvrySvcId   Oper   Sec
                                                                         Plcy
-------------------------------------------------------------------------------
tunnel-1                         10.10.20.1        2            Up     Dynamic
  tunnel-1.private:1               10.10.10.1        1            Up     1
-------------------------------------------------------------------------------
IPsec Tunnels: 1
===============================================================================
```

MCS is in sync on both SeGWs:

```
*A:SeGW-3# show redundancy multi-chassis sync

===============================================================================
Multi-chassis Peer Table
===============================================================================
Peer
-------------------------------------------------------------------------------
Peer IP Address         : 192.0.2.4
Description             : (Not Specified)
Authentication          : Disabled
Source IP Address       : 192.0.2.3
Admin State             : Enabled
Warm standby            : No
Remote warm standby     : No
Sub-mgmt options        :
  DHCP lease threshold  : Inactive
    Local / Remote      : -- / --
-------------------------------------------------------------------------------
Sync-status
-------------------------------------------------------------------------------
Client Applications     : IPsec
Sync Admin State        : Up
Sync Oper State         : Up
Sync Oper Flags         :
DB Sync State           : inSync
Num Entries             : 2
Lcl Deleted Entries     : 0
Alarm Entries           : 0
OMCR Standby Entries    : 0
OMCR Alarm Entries      : 0
Rem Num Entries         : 2
Rem Lcl Deleted Entries : 0
Rem Alarm Entries       : 0
Rem OMCR Standby Entries: 0
Rem OMCR Alarm Entries  : 0
===============================================================================
===============================================================================
```

```
*A:SeGW-4# show redundancy multi-chassis sync

===============================================================================
Multi-chassis Peer Table
===============================================================================
Peer
-------------------------------------------------------------------------------
```

```
Peer IP Address         : 192.0.2.3
Description             : (Not Specified)
Authentication          : Disabled
Source IP Address       : 192.0.2.4
Admin State             : Enabled
Warm standby            : No
Remote warm standby     : No
Sub-mgmt options        :
  DHCP lease threshold  : Inactive
    Local / Remote      : -- / --
-------------------------------------------------------------------------------
Sync-status
-------------------------------------------------------------------------------
Client Applications     : IPsec
Sync Admin State        : Up
Sync Oper State         : Up
Sync Oper Flags         :
DB Sync State           : inSync
Num Entries             : 2
Lcl Deleted Entries     : 0
Alarm Entries           : 0
OMCR Standby Entries    : 0
OMCR Alarm Entries      : 0
Rem Num Entries         : 2
Rem Lcl Deleted Entries : 0
Rem Alarm Entries       : 0
Rem OMCR Standby Entries: 0
Rem OMCR Alarm Entries  : 0
===============================================================================
===============================================================================
```

The following command shows that P-5 received two VPN-IPv4 routes for prefix 192.168.1.1/32: one from SeGW-3 with local preference 200 and one from SeGW-4 with local preference 100:

```
*A:P-5# show router bgp routes vpn-ipv4
===============================================================================
 BGP Router ID:192.0.2.5        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
Flag  Network                                            LocalPref   MED
      Nexthop (Router)                                   Path-Id     IGP Cost
      As-Path                                                        Label
-------------------------------------------------------------------------------
u*>i  64496:2:192.168.1.1/32                             200         None
      192.0.2.3                                          None        10
      No As-Path                                                     524287
*i    64496:2:192.168.1.1/32                             100         None
      192.0.2.4                                          None        10
      No As-Path                                                     524287
u*>i  64496:2:192.168.20.0/30                            100         None
      192.0.2.3                                          None        10
      No As-Path                                                     524287
*>i   64496:2:192.168.20.0/30                            100         None
      192.0.2.4                                          None        10
      No As-Path                                                     524287
u*>i  64496:2:192.168.20.1/32                            100         0
```

```
      192.0.2.3                                          None        10
      No As-Path                                                     524287
u*>i  64496:2:192.168.20.2/32                            100         0
      192.0.2.4                                          None        10
      No As-Path                                                     524287
-------------------------------------------------------------------------------
Routes : 6
===============================================================================
```

MC-IPSec failover scenarios

Two MC-IPSec failover scenarios are described in this section:

- MC-IPSec failover when MS-ISA is disabled
- MC-IPSec failover when the primary chassis SeGW-3 reboots

Failover when MS-ISA is disabled

Initially, SeGW-3 is the primary chassis and SeGW-4 is the standby:

```
*A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:37:51

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              200       Up             master
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

```
*A:SeGW-3# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-3-S-2                   10    No  Up   Master       200       1
                                 IPv4      Up   1           200        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
```

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

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             standby
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

```
*A:SeGW-4# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Backup       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

The following command disables the MS-ISA on the primary chassis SeGW-3, which will trigger an MC-IPSec failover.

```
configure
    card 1
        mda 2
            shutdown
```

With MS-ISA disabled, the MC-IPSec state of tunnel group 1 on SeGW-3 becomes **notEligible**, which means that the tunnel group is down, see the *7450 ESS, 7750 SR, and VSR Multiservice ISA and ESA Guide* for details description of MIMP states.:

```
*A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
```

```
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:37:51

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              200       Up             notEligible
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

SeGW-3 is backup for VRRP instance 10 with in-use priority 50, as per the VRRP policy 1:

```
*A:SeGW-3# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-3-S-2                   10    No  Up   Backup       200       1
                                 IPv4      Up   1           50         No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

SeGW-4 is now the primary chassis in tunnel group 1. This is triggered by MC-IPSec failover, as per the **mc-ipsec-non-forwarding** event in VRRP policy 1.

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             master
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

SeGW-4 is primary for VRRP instance 10;

```
*A:SeGW-4# show router vrrp instance
```

```
===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Master       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

The situation is restored by enabling MS-ISA on SeGW-3:

```
configure
    card 1
        mda 2
            no shutdown
```

MC-IPSec failover when primary chassis reboots

The following **tools** command on SeGW-3 triggers an MC-IPSec switchover:

```
tools perform redundancy multi-chassis mc-ipsec force-switchover tunnel-group 1

A:SeGW-3# tools perform redundancy multi-chassis mc-ipsec force-switchover tunnel-group 1
WARNING! Forcing a mastership switchover may significantly impact traffic. Are you sure (y/n)?
 y
```

SeGW-3 is the primary chassis for tunnel group 1:

```
A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:37:51

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              200       Up             master
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

SeGW-3 is primary for VRRP instance 10:

```
A:SeGW-3# show router vrrp instance
```

```
===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-3-S-2                   10    No  Up   Master       200       1
                                 IPv4      Up   1           200        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

SeGW-4 is the standby chassis for tunnel group 1:

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             standby
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

The VRRP state on SeGW-4 is backup:

```
*A:SeGW-4# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Backup       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

The following command reboots the primary chassis SeGW-3:

```
admin reboot now
```

While SeGW-3 reboots, the IPSec state of SeGW-4 becomes **eligible**:

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             eligible
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

The VRRP state on SeGW-4 is primary (**master**):

```
*A:SeGW-4# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Master       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

When SeGW-3 comes up, the IPSec state of tunnel group 1 is **discovery**, which means that the system has not established the MIMP session with its peer yet.

```
A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/10/2023 12:17:46

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
```

```
1              1              200       Up             discovery
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

After a while, the preceding **show** command is repeated and the IPSec state for tunnel 1 on SeGW-3 is standby:

```
A:SeGW-3# show redundancy multi-chassis mc-ipsec peer 192.0.2.4

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.4
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:46:03

======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              200       Up             standby
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

The VRRP state on SeGW-3 is backup:

```
A:SeGW-3# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-3-S-2                   10    No  Up   Backup       200       1
                                 IPv4      Up   1           50         No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

SeGW-4 is the primary chassis in MC-IPSec tunnel group 1:

```
*A:SeGW-4# show redundancy multi-chassis mc-ipsec peer 192.0.2.3

===============================================================================
Multi-Chassis MC-IPsec
===============================================================================
Peer Name       : (Not Specified)
Peer Addr       : 192.0.2.3
Keep Alive Intvl: 1.0 secs            Hold on Nbr Fail     : 3
Discovery Intvl : 300 secs            Discovery Boot Intvl : 300 secs
BFD             : Enable
Last update     : 02/13/2023 09:38:21
```

```
======================================================================
Multi-Chassis IPsec Multi Active Tunnel-Group Table
======================================================================
ID             Peer Group     Priority  Admin State    Mastership
----------------------------------------------------------------------
1              1              150       Up             master
----------------------------------------------------------------------
Multi Active Tunnel Group Entries found: 1
======================================================================
===============================================================================
```

SeGW-4 is the primary node for VRRP instance 10:

```
*A:SeGW-4# show router vrrp instance

===============================================================================
VRRP Instances
===============================================================================
Interface Name                   VR Id Own Adm  State       Base Pri   Msg Int
                                 IP        Opr  Pol Id      InUse Pri  Inh Int
-------------------------------------------------------------------------------
int-SeGW-4-S-2                   10    No  Up   Master       100       1
                                 IPv4      Up   n/a         100        No
  Backup Addr: 172.16.1.254
-------------------------------------------------------------------------------
Instances : 1
===============================================================================
```

Configuration guidelines

The following is a list of guidelines for configuring MC-IPSec:

- To avoid high CPU load and issues in some complex cases, the following are suggestions for configuring the IKEv2 lifetime:
  - Both IKE_SA and CHILD_SA lifetime on MC-IPSec chassis (SeGW-3 and SeGW-4) should be around three times larger than on the IPSec peer (CE-1).
  - With the first rule, the lifetime of the side with smaller lifetime (IPSec peer CE-1) should not be too small (these being the default values):
    - IKE_SA: >= 86400 seconds
    - CHILD_SA: >= 3600 seconds
  - With the first rule, on the side with smaller lifetime (IPSec peer CE-1), the IKE_SA lifetime must be at least 3 times larger than CHILD_SA lifetime.
- The IKE protocol is the control plane of IPSec, so IKE packets must be treated as high QoS priority in the end-to-end path of the public service. On the public interface, a SAP ingress QoS policy must be configured to ensure that IKE packets get high QoS priority.
- Configure **ipsec-responder-only** under **tunnel-group** for static LAN-to-LAN tunnels.
- Enable dead peer detection (DPD) on the IPSec peer side (CE-1); disable DPD (default) on the MC-IPSec chassis side.
- The direct and redundant physical link between MC-IPSec chassis must be configured with sufficient bandwidth for MCS and shunting traffic, and proper QoS configuration to make sure the MIMP and MCS packets are treated as high priority traffic.

- The system time must be same on both MC-IPSec chassis.
- Make sure the protection status is **nominal** on both chassis before provoking a controlled switchover. The protection status can be displayed with the **show redundancy multi-chassis mc-ipsec peer <addr>** command.
- Wait at least five minutes between two consecutive switchovers if possible, to prevent a second switchover happening before the standby is ready to become the primary chassis.

Conclusion

MC-IPSec provides a stateful multi-chassis IPSec redundancy solution. This is very important in a carrier grade network, especially in applications such as mobile backhaul where high value mobile services run over IPSec tunnels.

NAT Stateless Dual-Homing

This chapter describes NAT stateless dual-homing.

Topics in this chapter include:

Applicability

The information and configuration in this chapter are based on SR OS Release 14.0.R4.

Overview

With the IPv4 address space almost consumed, many operators are deploying Network Address Translation (NAT) at centralized or semi-centralized points in their IP/MPLS networks. The NAT function is implemented using Carrier Grade NAT (CGN) nodes, which typically support tens of thousands of clients/ subscribers. Therefore, a failure of one of these nodes would be considered a significant event.

Many operators consider a stateful failover mechanism between CGN nodes to be too demanding with regard to control plane requirements and state synchronization of NAT bindings. A reasonable compromise appears to be a stateless failover mechanism, capable of providing failover between geo-redundant CGN devices, but with manageable control plane implications.

This chapter describes the NAT stateless dual-homing feature supported in SR OS. NAT stateless dual-homing is supported for Large-Scale NAT and NAT64, and this chapter describes how both can be supported, either independently or together.

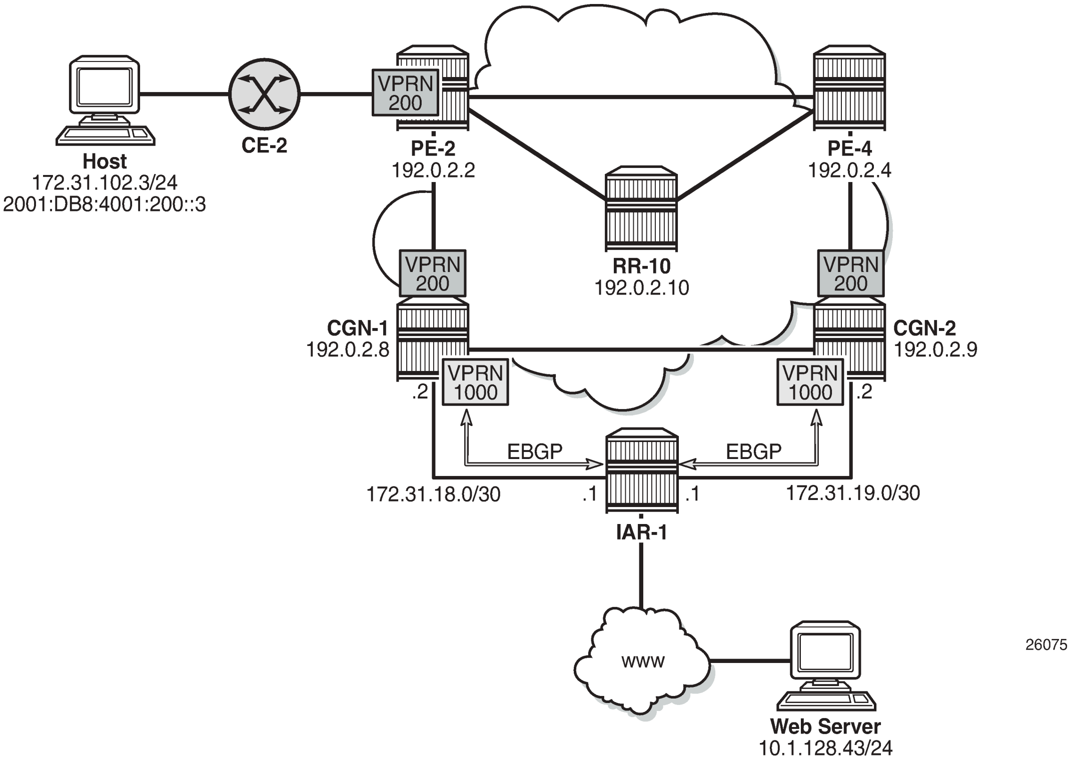
Example Topology

The topology shown in Figure 80: Example Topology is an example of the use of NAT stateless dual-homing. PE-2, PE-4, RR-10, CGN-1, and CGN-2 form part of Autonomous System 64496 and run IS-IS level 2 and LDP. PE-2, PE-4, CGN-1, and CGN-2 are clients of Route Reflector RR-10 and peer with the VPN-IPv4 and VPN-IPv6 address families.

IAR-1 acts as an Internet Service Provider (ISP) edge router belonging to AS 65535, and provides Internet access to AS 65596. CGN-1 and CGN-2 are configured with a VPRN (1000) that serves as a NAT outside routing VPN Routing and Forwarding (VRF) instance. In this VRF, both CGN-1 and CGN-2 peer with IAR-1 in EBGP for the IPv4 address family. CGN-1 and CGN-2 advertise the relevant NAT outside pools to IAR-1, and IAR-1 advertises a default route to both CGN-1 and CGN-2. IAR-1 has IP connectivity to a web server at 10.1.128.43/24.

CE-2 is connected to PE-2 and is part of VPRN 200. CE-2 has an IPv4 host (172.31.102.3) that serves to test NAT44 connectivity to the web server, and an IPv6 host (2001:DB8:4001:200::3) that serves to test NAT64 connectivity to the web server.

Figure 80: Example Topology



Configuration

To support NAT functionality, some form of Integrated Services Adapter (ISA) or Integrated Services Module (ISM) is required, as listed in the Applicability section. In this example, CGN-1 and CGN-2 both have a single MS-ISA card, which is configured as MDA type **isa-bb**, as follows:

```
configure
    card 1
        card-type iom3-xp-b
        mda 2
            mda-type isa-bb
            no shutdown
        exit
        no shutdown
    exit
```

The MS-ISA is then configured to become a member of one or more **nat-group**. Up to fourteen MS-ISAs can be configured to belong to up to four NAT groups. When more than one MS-ISA is configured in a NAT group, the MS-ISAs can work in active-standby mode or active-active mode.

In active-standby mode, one or more MS-ISAs act as standby, and in normal operation, are idle. If an active MS-ISA fails, one standby MS-ISA accepts the traffic from the failed card. In the active-standby scenario, the mapping between failed card and standby card can always be considered to be 1:1. In active-active mode, all of the MS-ISAs in the NAT group are active, and if one MS-ISA in the group fails, the load is distributed across the remaining active MS-ISAs in the group.

The default setting is active-standby. In both active-standby and active-active modes, the dynamically created NAT bindings are not synchronized between cards. Therefore, a failover will cause an interruption in traffic until the NAT bindings are re-initiated by the clients/subscribers behind the NAT.

The NAT group also requires that an **active-mda-limit** is configured, as follows, which allows the operator to specify how many MDAs (MS-ISAs) will be active in the group. Any operational MDAs above this configured limit will be considered spare MDAs. Finally, the **nat-group** must be placed in a **no shutdown** state.

```
configure
    isa
        nat-group 1 create
            active-mda-limit 1
            mda 1/2
            no shutdown
        exit
    exit
```

NAT Outside Context

The NAT **outside** function is responsible for creation of the NAT bindings, using outside IP addresses defined in outside pools (together with their associated ports), and for advertising the address ranges in those pools to upstream routers. The NAT stateless dual-homing redundancy mechanism is based on ownership of an outside pool, where each member of a redundant pair can assume either an active (master) or standby role for an outside pool. This active/standby role is determined by the presence of a **monitor** prefix. Both CGN nodes of a redundant pair implement the following:

1. Advertise a unique route into the routing instance that the NAT outside function resides in. This is known as the export route and may be advertised into the Global Routing Table (GRT) or a VPRN instance. For example, CGN-1 advertises (exports) prefix P1 while CGN-2 advertises prefix P2.
2. Check for the presence of a configured route in the routing instance that the NAT outside function resides in. This is known as the monitor route. Continuing the preceding example, CGN-1 monitors prefix P2 while CGN-2 monitors prefix P1.

Therefore, the export route of CGN-1 becomes the monitor route of CGN-2, and the export route of CGN-2 becomes the monitor route of CGN-1. The redundancy mechanism thereafter checks the (virtual) routing table of the NAT outside function for the presence of the monitor route. If it is not present, the redundancy state for the pool is set to active and the following occurs (subject to routing policy):

1. The redundancy export route is populated in the NAT outside routing instance and advertised externally in the relevant routing instance.
2. The outside pool address is populated in the NAT outside routing instance and advertised externally in the relevant routing instance.
3. Routes that need to become active in any associated NAT inside routing instances, to attract traffic to the active CGN node, are populated in the relevant routing tables. For example, NAT64 translator routes and/or NAT44 steering routes, both of which are described later in this chapter, are populated.

If the monitor route is present, the redundancy state for the pool is set to standby and the following occurs:

1. The redundancy export route is not populated in the NAT outside routing instance and is, therefore, not eligible to be advertised externally.
2. The outside pool address is not populated in the NAT outside routing instance and is, therefore, not eligible to be advertised externally.
3. Routes that need to become active in any associated NAT inside routing instances, to attract traffic to the active CGN node, are not populated in the relevant routing tables.

There are no configurable options for selection of active/standby CGN nodes. The status of a node is based on the presence of the monitor route. In the event of a collision during redundancy startup, hardcoded debounce timers ensure that only a single CGN node is selected as active.

The first of two following examples shows the configuration of VPRN 1000 (the NAT outside VRF) at CGN-2, with the second configuration showing the **vrf-export** and **vrf-import** policy statements. For advertising and importing routes from/to the VRF, there are two requirements: to advertise the export redundancy route and to import the monitor route. This is the purpose of the *"vrf1000-export"* and *"vrf1000-import"* policy statements. Other VPRN parameters are generic and, therefore, not discussed here.

The VPRN contains an interface, "to-IAR-1", which has an associated EBGP peering session to IAR-1. The export policy under the BGP **neighbor** context will be described later in this chapter. That policy contains sufficient logic to advertise the NAT outside pools.

The NAT pools are configured in the **nat>outside** context. The first configuration provides an example of a single pool, "4-to-4", which will be used for NAT44. As well as a name, the pool requires association with a NAT group, and definition of the **type** of NAT that will be configured; in this case, **large-scale**. The **mode** of the pool is set to **napt** to indicate N:1 NAT, and the **address-range** that will be used for outside addressing is 10.1.4.1 to 10.1.4.254.

The relevant and required parameters for stateless dual-homing are configured in the **redundancy** context. In this example, CGN-2 exports prefix 192.168.0.249/32 and monitors prefix 192.168.0.248/32. Conversely, CGN-1 exports prefix 192.168.0.248/32 and monitors prefix 192.168.0.249/32. The redundancy node and the pool must be placed into a **no shutdown** state.

```
*A:CGN-2#
configure
    service
        vprn 1000 customer 1 create
            vrf-import "vrf1000-import"
            vrf-export "vrf1000-export"
            autonomous-system 64496
            route-distinguisher 64496:1000
            auto-bind-tunnel
                resolution any
            exit
            interface "to-IAR-1" create
                address 172.31.19.2/30
                sap 1/1/2:200 create
                exit
            exit
            aggregate 10.1.4.0/24 summary-only
            bgp
                group "EBGP"
                    family ipv4
                    peer-as 65535
                    split-horizon
                    neighbor 172.31.19.1
                        authentication-key <password>
                        export "vrf1000-ebgp-export"
                    exit
                exit
                no shutdown
            exit
            nat
                outside
                    pool "4-to-4" nat-group 1 type large-scale create
                        redundancy
                            export 192.168.0.249/32
                            monitor 192.168.0.248/32
                            no shutdown
                        exit
                        mode napt
                        address-range 10.1.4.1 10.1.4.254 create
                        exit
                        no shutdown
                    exit
                exit
            exit
            service-name "NAT-Outside"
            no shutdown
        exit
    exit
```

```
*A:CGN-2#
configure
   router
        policy-options
            begin
            prefix-list "vrf1000-nat-export"
                prefix 192.168.0.249/32 exact
            exit
            community "vrf1000-export" members "target:64496:1000"
            community "vrf1000-import" members "target:64496:1000"
            policy-statement "vrf1000-export"
                entry 10
                    from
                        protocol nat
                        prefix-list "vrf1000-nat-export"
                    exit
                    to
                        protocol bgp-vpn
                    exit
                    action accept
                        community add "vrf1000-export"
                    exit
                exit
            exit
            policy-statement "vrf1000-import"
                entry 10
                    from
                        community "vrf1000-import"
                    exit
                    action accept
                    exit
                exit
                default-action drop
                exit
            exit
```

After the redundancy node and pool are placed into a **no shutdown** state, the master and standby can be elected, based on the previously described criteria. In this example, CGN-2 becomes the active CGN node for the pool "4-to-4". This is shown using the following command, where the Active field shows true. Conversely, the same output at CGN-1 shows the Active field as false.

```
*A:CGN-2# show router 1000 nat pool "4-to-4"

===============================================================================
NAT Pool 4-to-4
===============================================================================
Description                           : (Not Specified)
ISA NAT Group                         : 1
Pool type                             : largeScale
Applications                          : (None)
Admin state                           : inService
Mode                                  : napt
Port forwarding dyn blocks reserved   : 0
Port forwarding range                 : 1 - 1023
Port reservation                      : 128 blocks
Block usage High Watermark (%)        : 90
Block usage Low Watermark (%)         : 20
Subscriber limit per IP address       : 65535
Active                                : true
Deterministic port reservation        : (Not Specified)
Last Mgmt Change                      : 10/06/2016 11:29:41
===============================================================================

===============================================================================
NAT address ranges of pool 4-to-4
===============================================================================
Range                                                          Drain Num-blk
-------------------------------------------------------------------------------
10.1.4.1 - 10.1.4.254                                                0
-------------------------------------------------------------------------------
No. of ranges: 1
===============================================================================

===============================================================================
NAT members of pool 4-to-4 ISA NAT group 1
===============================================================================
Member                                                        Block-Usage-% Hi
-------------------------------------------------------------------------------
1                                                             < 1          N
-------------------------------------------------------------------------------
No. of members: 1
===============================================================================

===============================================================================
Dual-Homing
===============================================================================
Type                                  : Leader
Export route                          : 192.168.0.249/32
Monitor route                         : 192.168.0.248/32
Admin state                           : inService
Dual-Homing State                     : Active
===============================================================================

===============================================================================
Dual-Homing fate-share-group
===============================================================================
Router         Pool                                                   Type
-------------------------------------------------------------------------------
vprn1000       4-to-4                                                 Leader
-------------------------------------------------------------------------------
No. of pools: 1
```

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

Although the entire 10.1.4.0/24 block is allocated for NAT outside addressing, the address range shown in the preceding output does not include the network address (10.1.4.0/24) or broadcast address (10.1.4.255/24). Therefore, the address range does not include the entire /24 prefix and has to be fragmented into a number of longer prefixes, known through protocol NAT, to cover the 10.1.4.1-10.1.4.255 range. This is shown in the route table of VPRN 1000, following, and is due to the whole subnet not being defined in the address-range configuration. (If the address range was 10.1.4.0-10.1.4.255, there would be a single entry of 10.1.4.0/24 in the route table of VPRN 1000.)

```
*A:CGN-2# show router 1000 route-table 10.1.4.0/24 longer

===============================================================================
Route Table (Service: 1000)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
10.1.4.0/24                                   Blackh* Aggr      00h01m34s  130
       Black Hole                                                   0
10.1.4.1/32                                   Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.2/31                                   Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.4/30                                   Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.8/29                                   Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.16/28                                  Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.32/27                                  Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.64/26                                  Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.128/26                                 Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.192/27                                 Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.224/28                                 Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.240/29                                 Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.248/30                                 Remote  NAT       00h01m34s  0
       NAT outside to mda 1/2                                       0
10.1.4.252/31                                 Remote  NAT       00h01m36s  0
       NAT outside to mda 1/2                                       0
10.1.4.254/32                                 Remote  NAT       00h01m36s  0
       NAT outside to mda 1/2                                       0
-------------------------------------------------------------------------------
No. of Routes: 15
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
* indicates that the corresponding row element may have been truncated.
```

CGN-2 advertising all of these longer prefixes to the edge router of the ISP is not wanted. The ISP may even enforce a minimum /24 prefix length. Therefore, the preceding configuration of VPRN 1000 shows an **aggregate** command for the 10.1.4.0/24 prefix with the argument **summary-only**. When at least one of the more-specific prefixes in the 10.1.4.0/24 range is populated in the route table of VPRN 1000, the aggregate becomes active and can be used by the route policy for exporting to IAR-1, while suppressing the more-specific routes.

The following output shows that outside pool prefixes are not populated in the NAT outside routing context at the standby CGN node (CGN-1). Even with the aggregate command configured at both CGN nodes, the aggregate prefix will only become active at the active CGN node. Therefore, only the active CGN node will advertise that aggregate prefix upstream.

```
*A:CGN-1# show router 1000 route-table 10.1.4.0/24 longer

===============================================================================
Route Table (Service: 1000)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                    Metric
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
No. of Routes: 0
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested
===============================================================================
```

So far, only one pool has been defined that could be used for stateless dual-homing for both NAT44 and NAT64. However, to allow for independent address management of each of these functions, a separate outside pool is created for each. The following configuration shows two pools in the **nat>outside** context of VPRN 1000: the pool "4-to-4", which is for NAT44 purposes, and the pool "6-to-4", which is for NAT64 purposes.

The address range defined in the "6-to-4" pool is 10.1.6.1 to 10.1.6.254. As with the range 10.1.4.1 to 10.1.4.254 in the "4-to-4" pool, an aggregate command is configured to advertise the 10.1.6.0/24 prefix, while suppressing the more-specific prefixes.

With the redundancy node in the "6-to-4" pool, there are instances where traffic from a NAT inside routing context may be mapped to multiple outside pools, which in a stateless dual-homed environment may cause the NAT function to fail. For example, assume a NAT outside context has two pools, P1 and P2, where pool P1 is active and pool P2 is standby. An active pool can trigger the advertisement of inside and outside prefixes, and traffic will be attracted to this CGN node. When traffic arrives, it may be mapped to pool P1 on the active CGN node, due to the NAT mapping criteria. However, traffic may also be mapped to pool P2, due to the mapping criteria; this traffic will fail because the pool P2 is active on the redundant CGN node.

To ensure that this traffic failure does not happen, a group of pools accessed by the same inside routing context must all be active on the same CGN node simultaneously. To achieve this, SR OS uses a Pool Fate Sharing Group (PFSG). The PFSG ensures that all co-located pools accessed by the same inside routing context are either active or standby; not a combination of both. This is achieved by having a *leader* pool and *follower* pools.

If the leader pool is active, all follower pools are active. If the leader pool is standby, all follower pools are standby. This is enabled in the **redundancy** context of the "6-to-4" pool using the **follow** command. The **follow** command configures the pool as a follower and allows the user to access the routing context and outside pool of the leader pool. In the following example, pool "4-to-4" is a leader pool and pool "6-to-4" is a follower pool, which always assumes the same state as that of the leader.

```
*A:CGN-2#
configure
    service
        vprn 1000 customer 1 create
```

```
            aggregate 10.1.4.0/24 summary-only
            aggregate 10.1.6.0/24 summary-only
            nat
                outside
                    pool "4-to-4" nat-group 1 type large-scale create
                        redundancy
                            export 192.168.0.249/32
                            monitor 192.168.0.248/32
                            no shutdown
                        exit
                        mode napt
                        address-range 10.1.4.1 10.1.4.254 create
                        exit
                        no shutdown
                    exit
                    pool "6-to-4" nat-group 1 type large-scale create
                        redundancy
                            follow router 1000 pool "4-to-4"
                        exit
                        mode napt
                        address-range 10.1.6.1 10.1.6.254 create
                        exit
                        no shutdown
                    exit
                exit
            exit
```

The following output shows a PFSG with leaders and followers in the operational state of pool “6-to-4” at CGN-2. The pool state is active, but the Dual-Homing Type field indicates that this pool is a follower. Therefore, the state is derived from the state of the leader pool, which is pool “4-to-4” in router 1000. The output also contains a list of all the pools that are part of the same PFSG.

```
*A:CGN-2# show router 1000 nat pool "6-to-4"

===============================================================================
NAT Pool 6-to-4
===============================================================================
Description                           : (Not Specified)
ISA NAT Group                         : 1
Pool type                             : largeScale
Applications                          : (None)
Admin state                           : inService
Mode                                  : napt
Port forwarding dyn blocks reserved   : 0
Port forwarding range                 : 1 - 1023
Port reservation                      : 128 blocks
Block usage High Watermark (%)        : 90
Block usage Low Watermark (%)         : 20
Subscriber limit per IP address       : 65535
Active                                : true
Deterministic port reservation        : (Not Specified)
Last Mgmt Change                      : 10/06/2016 13:09:07
===============================================================================

===============================================================================
NAT address ranges of pool 6-to-4
===============================================================================
Range                                                          Drain Num-blk
-------------------------------------------------------------------------------
10.1.6.1 - 10.1.6.254                                                0
-------------------------------------------------------------------------------
No. of ranges: 1
===============================================================================
===============================================================================
NAT members of pool 6-to-4 ISA NAT group 1
===============================================================================
Member                                                        Block-Usage-% Hi
-------------------------------------------------------------------------------
1                                                             < 1           N
-------------------------------------------------------------------------------
No. of members: 1
===============================================================================

===============================================================================
Dual-Homing
===============================================================================
Type                                  : Follower
Follow-pool                           : "4-to-4" router 1000
Dual-Homing State                     : Active
===============================================================================

===============================================================================
Dual-Homing fate-share-group
===============================================================================
Router         Pool                                                   Type
-------------------------------------------------------------------------------
vprn1000       4-to-4                                                 Leader
vprn1000       6-to-4                                                 Follower
-------------------------------------------------------------------------------
No. of pools: 2
===============================================================================
```

NAT Policies

NAT policies allow for definition of NAT attributes such as:

- filtering behavior (endpoint-independent or address-and-port-dependent)
- NAT mapping timeouts
- per-user session/flow limits
- configuration of Application Level Gateway (ALG) protocols
- high/low resource watermarks.

These attributes are generic NAT configuration parameters that are beyond the scope of this chapter.

A NAT policy also references the routing context and name of the outside pool used for the creation of NAT bindings associated with the policy. Therefore, if multiple outside pools are needed, multiple NAT policies must also be used. The following shows the configuration of the required NAT policies at CGN-2.

Because two outside pools exist in VPRN 1000 (the pool "4-to-4" for NAT44 and the pool "6-to-4" for NAT64), two policies are created using the **nat-policy** parameter. The **nat-policy** "NAT44" uses the **pool** keyword to access the "4-to-4" pool in **router** 1000, while the **nat-policy** "NAT64" uses the same **pool** keyword to access the "6-to-4" pool in **router** 1000. In this example, the same outside routing context is used for both NAT policies, but the outside routing contexts can also be different for each policy.

```
configure
    service
        nat
            nat-policy "NAT44" create
                pool "4-to-4" router 1000
            exit
            nat-policy "NAT64" create
                pool "6-to-4" router 1000
            exit
        exit
```

NAT Inside Context

The NAT inside routing context is the interface toward the customer or end user. There can be multiple NAT inside routing contexts mapped to a single NAT outside context (the relationship can be 1:1 or N:1). This is possible even if overlapping addresses are used in two or more NAT inside routing contexts because the NAT flow mapping tuple consists of the parameters {routing-instance, inside-IP, inside-port} mapped to {routing-instance, outside-IP, outside-port}.

The NAT inside routing context is responsible for two main functions:

1. Diverting some or all traffic toward the NAT function (ISA board).
2. Attracting traffic that should be subject to NAT toward it. This should be conditional because only the master CGNAT node should attract traffic toward itself.

For diverting traffic toward the NAT function, there are two approaches:

1. The first approach is to use IP filters with **action nat** to divert matched traffic into the ISA. Traffic subject to NAT can have a different inside and outside routing context, or the same routing context can be used for both inside and outside.
2. The second approach is a routing-based approach using a **destination-prefix** in the **nat>inside** context. Any traffic with a destination address matching the defined destination-prefix is forwarded to the ISA for NAT. When the destination-prefix approach is used, different routing contexts must be used for inside and outside.

For NAT44, both the IP filter and destination-prefix approaches are permitted. For NAT64, the diversion to NAT is only supported using IPv6 filters. The example setup in this chapter consists of both NAT44 and NAT64. Therefore, for the purpose of standardization across the NAT44 and NAT64 functions, the IP filter-based approach is used for both.

In Figure 80: Example Topology, CE-2 is connected to PE-2 and is part of VPRN 200. To provide Internet access with stateless NAT dual-homing to VPRN 200, it is extended to both CGN-1 and CGN-2 as a NAT inside VRF. The following shows the configuration of VPRN 200 at CGN-2 with a similar configuration also applied at CGN-1.

Because one of the main functions of the NAT inside routing context is to attract traffic toward the (active) CGN node, this is configured in the **nat>inside** context. The **redundancy** parameter provides a context for the configuration of NAT44 redundancy when the diversion to NAT is implemented with IP filters. In this **redundancy** context, the **peer** command is used to configure the address of the redundant peer (in this case CGN-1). If upstream traffic that is subject to NAT inadvertently arrives at the CGN node that is standby for the outside pool used for the NAT mapping, this parameter provides a forwarding address for that traffic. However, if destination-prefix based redirect to NAT is used instead of IP filters, only a **nat-policy** and **destination-prefix** need to be configured in the NAT inside routing context.

The **steering-route** command is optional. It allows for configuration of a (non-default) prefix/length that is only active in the routing table of the NAT inside routing context of the active CGN node. When this steering route is active in the routing table, it can either be advertised directly using the route-policy framework, or it can be used as an indirect next-hop to advertise some other prefix. This latter approach is used in the following configuration example, where the **steering-route** of 192.168.203.1/32 is used as an indirect next-hop for the **static-route-entry** of 0.0.0.0/0. This creates the following dependencies:

- If the CGN node is active, the steering route of 192.168.203.1/32 becomes active in the routing table of VPRN 200.
- When 192.168.203.1/32 is active, it becomes a valid indirect next-hop for 0.0.0.0/0, so this route also becomes active in the routing table.
- When the default route is active, it can be exported to the rest of the VPN using the route-policy framework.

In the first of three following configurations, the **vrf-export** command accesses a policy with the name *vrf200-export*. The second configuration shows the contents of that policy, where entry 10 accesses a prefix-list (vrf200-lsn44-default) containing the default route and advertises it into BGP-VPN, with the relevant Route-Target (target:64496:200) and Origin (origin:64496:200) Extended Communities attached.

The Origin Extended Community is used by the redundant CGN peer to drop the default route, as shown in entry 10 of the corresponding *vrf200-import* policy in the same configuration. The reason for dropping the default route at the standby CGN node is that the *vrf200-export* policy only requires that a default route is present in the routing table in order to source/advertise a default route itself. If the standby CGN node imports the default route from the active CGN node into the routing table, the standby will also attempt to advertise a default route, which is not wanted.

The **nat64** command provides the context to configure the NAT64 redundancy parameters. In the case of NAT64, the CGN node becomes a translator between the IPv6 and IPv4 address families and needs to advertise the NAT64 translator address that will be used by IPv6 clients to embed IPv4 addresses. In the first configuration, the address 2001:DB8:122:344::/96 is used as the NAT64 translator address. As with NAT44, the advertisement of this address is conditional and must only be advertised by the active CGN node. This is ensured because the prefix is only present in the routing table of the active CGN, not the standby CGN.

Entry 20 of the *vrf200-export* policy shown in the second configuration provides the relevant policy rules to ensure that this IPv6 prefix is advertised into BGP-VPN with the relevant Route-Target value when the prefix is present in the routing table.

The last parameter in the **nat>inside** context is the **nat-policy**, which is known as the default NAT policy and must exist in the **nat>inside** context. When multiple NAT policies are used in a single NAT inside routing context, the default NAT policy is used for any traffic that is not matched (using the destination-prefix for NAT44 or IPv4/IPv6 filters for NAT44/NAT64) and associated with an explicit NAT policy. The default NAT policy can reference a separate NAT policy, or it can reference a NAT policy that is already in use.

As previously described, the intention is to use IP filters to implement the diversion to NAT. The relevant IPv4 and IPv6 filter IDs (ID number 200 in both cases) are shown in the third configuration. The IPv4 filter has no match criteria in this example; it has **action nat** using **nat-policy** "NAT44", which accesses the outside pool "4-to-4" in VPRN 1000. The IPv6 filter also has no match criteria and has **action nat**, but distinguishes between DSLite (DSLite is not supported for NAT stateless dual-homing) and NAT64, using the **nat-type** argument. The **nat-policy** that should be used is the policy "NAT64", which accesses the outside pool "6-to-4" in VPRN 1000. These IP filters need to match traffic that ingresses the redundant CGN nodes from the MPLS side of VPRN 200 and are, therefore, applied in the **network>ingress** context in the first configuration.

The remainder of the VPRN parameters are generic and are not explained here.

```
*A:CGN-2#
configure
    service
        vprn 200 customer 1 create
            vrf-import "vrf200-import"
            vrf-export "vrf200-export"
            route-distinguisher 64496:200
            auto-bind-tunnel
                resolution any
                exit
            exit
            static-route-entry 0.0.0.0/0
                indirect 192.168.203.1
                    no shutdown
                exit
            exit
            nat
                inside
                    nat-policy "NAT44"
                    nat64
                        prefix 2001:db8:122:344::/96
                        no shutdown
                    exit
                    redundancy
                        peer 192.0.2.8
                        steering-route 192.168.203.1/32
                    exit
                exit
            exit
            network
                ingress
                    filter ip 200
                    filter ipv6 200
                exit
            exit
            no shutdown
        exit
```

```
*A:CGN-2#
configure
    router
        policy-options
            begin
            prefix-list "vrf200-lsn44-default"
                prefix 0.0.0.0/0 exact
            exit
            prefix-list "vrf200-nat64-translator"
                prefix 2001:db8:122:344::/96 exact
            exit
            community "vrf200-soo" members "origin:64496:200"
            community "vrf200-export" members "target:64496:200"
            community "vrf200-import" members "target:64496:200"
            policy-statement "vrf200-export"
                entry 10
                    from
                        prefix-list "vrf200-lsn44-default"
                    exit
                    to
                        protocol bgp-vpn
                    exit
                    action accept
                        community add "vrf200-soo" "vrf200-export"
                    exit
                exit
                entry 20
                    from
                        prefix-list "vrf200-nat64-translator"
                    exit
                    to
```

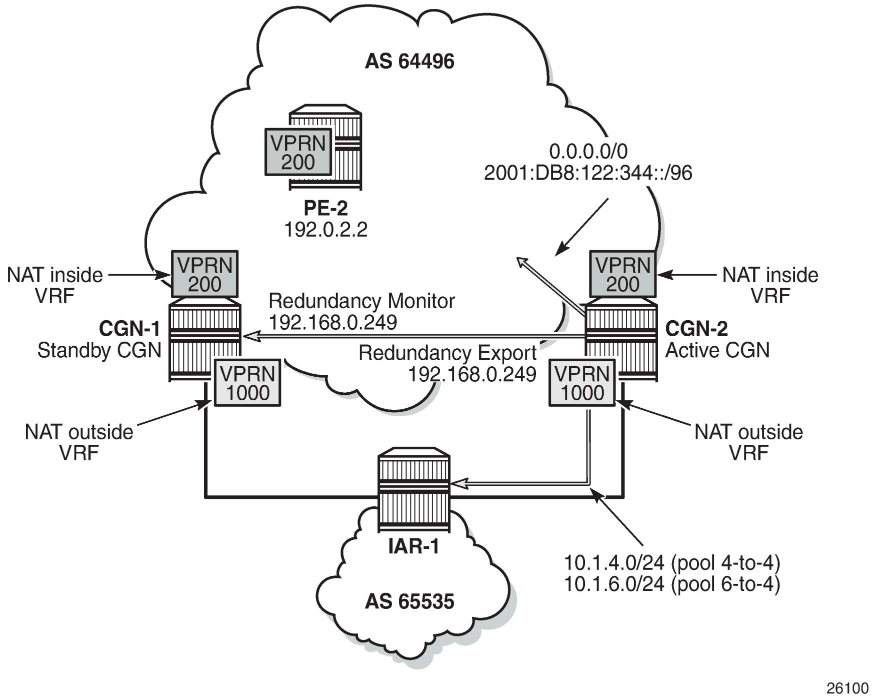
```
                        protocol bgp-vpn
                    exit
                    action accept
                        community add "vrf200-export"
                    exit
                exit
            exit
            policy-statement "vrf200-import"
                entry 10
                    from
                        community "vrf200-soo"
                    exit
                    action reject
                exit
                entry 20
                    from
                        community "vrf200-import"
                    exit
                    action accept
                    exit
                exit
            exit
```

```
*A:CGN-2#
configure
    filter
        ip-filter 200 create
            entry 10 create
                action
                    nat nat-policy "NAT44"
                exit
            exit
        exit
        ipv6-filter 200 create
            entry 10 create
                action
                    nat nat-type nat64 nat-policy "NAT64"
                exit
            exit
        exit
    exit
```

Verification of the Active CGN Node

After the configuration of the inside and NAT outside routing contexts, with the associated NAT policies, the state of the stateless redundant CGN nodes can be verified; see Figure 81: Redundancy Status.

Figure 81: Redundancy Status



The following two outputs show that the pool "4-to-4" is a leader at both CGN-1 and CGN-2 (for which pool "6-to-4" is a follower), and that CGN-2 is the active CGN node.

```
*A:CGN-2# show router 1000 nat pool "4-to-4" | match "Dual-Homing State" pre-lines 6
Dual-Homing
===============================================================================
Type                                    : Leader
Export route                            : 192.168.0.249/32
Monitor route                           : 192.168.0.248/32
Admin state                             : inService
Dual-Homing State                       : Active
```

```
*A:CGN-1# show router 1000 nat pool "4-to-4" | match "Dual-Homing State" pre-lines 6
Dual-Homing
===============================================================================
Type                                    : Leader
Export route                            : 192.168.0.248/32
Monitor route                           : 192.168.0.249/32
Admin state                             : inService
Dual-Homing State                       : Standby
```

Although the redundancy export route is populated in the NAT outside routing table and advertised externally by the active CGN node (if permitted by route-policy), it is not populated in the NAT outside routing table of the standby CGN, which is not advertised externally. The following two outputs show that CGN-2 advertises its export route (192.168.0.249/32) into IPv4 BGP-VPN, but CGN-1 does not advertise its own export route because the monitor route (192.168.0.249/32) is present.

```
*A:CGN-2# show router bgp routes vpn-ipv4 rd 64496:1000 hunt
---snip---
```

```
-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 192.168.0.249/32
Nexthop        : 192.0.2.9
Route Dist.    : 64496:1000             VPN Label      : 262128
Path Id        : None
To             : 192.0.2.10
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : 0
AIGP Metric    : None
Connector      : None
Community      : target:64496:1000
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 192.0.2.10
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : N/A
Orig Validation: N/A
Source Class   : 0                      Dest Class     : 0

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

```
*A:CGN-1# show router 1000 route-table 192.168.0.249/32

===============================================================================
Route Table (Service: 1000)
===============================================================================
Dest Prefix[Flags]                            Type    Proto     Age        Pref
      Next Hop[Interface Name]                                   Metric
-------------------------------------------------------------------------------
192.168.0.249/32                              Remote  BGP VPN   06d03h21m  170
       192.0.2.9 (tunneled)                                        0
-------------------------------------------------------------------------------
No. of Routes: 1
Flags: n = Number of times nexthop is repeated
       B = BGP backup route available
       L = LFA nexthop available
       S = Sticky ECMP requested


*A:CGN-1# show router bgp routes vpn-ipv4 rd 64496:1000 hunt
---snip---
-------------------------------------------------------------------------------

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

As well as the export/monitor routes, the outside pools are populated in the NAT outside routing context and advertised by the active CGN node. The outside pools are summarized as 10.1.4.0/24 (pool "4-to-4") and 10.1.6.0/24 (pool "6-to-4") using the **aggregate** command in VPRN 1000. Because the (more-explicit) NAT outside pool addresses are only populated in the route table of the active CGN node, the aggregate will also only be populated in the routing table of the active CGN node. Therefore, the following policy is applied to the EBGP peering session with IAR-1 at both CGN-1 and CGN-2. The output following the policy example shows that CGN-2 advertises both of the NAT outside pools to IAR-1, while CGN-1 advertises no outside pool prefixes to IAR-1.

```
configure
    router
        policy-options
            begin
            policy-statement "vrf1000-ebgp-export"
                entry 10
                    from
                        protocol aggregate
                    exit
                    to
                        protocol bgp
                    exit
                    action accept
                        origin igp
                    exit
                exit
                default-action drop
                exit
            exit
            commit
```

```
*A:CGN-2# show router 1000 bgp neighbor 172.31.19.1 advertised-routes
===============================================================================
 BGP Router ID:192.0.2.9        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP IPv4 Routes
===============================================================================
Flag  Network                                            LocalPref   MED
      Nexthop (Router)                                   Path-Id     Label
      As-Path
-------------------------------------------------------------------------------
i     10.1.4.0/24                                        n/a         None
      172.31.19.2                                         None        -
      64496
i     10.1.6.0/24                                        n/a         None
      172.31.19.2                                         None        -
      64496
-------------------------------------------------------------------------------
Routes : 2
===============================================================================

*A:CGN-1# show router 1000 bgp neighbor 172.31.18.1 advertised-routes
===============================================================================
 BGP Router ID:192.0.2.8        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP IPv4 Routes
===============================================================================
Flag  Network                                            LocalPref   MED
```

```
      Nexthop (Router)                                   Path-Id     Label
      As-Path
-------------------------------------------------------------------------------
No Matching Entries Found
===============================================================================
```

In the NAT inside routing instance, any routes that are used to attract traffic are populated in the relevant route tables of the active CGN node and must be advertised externally. For the NAT44 function, the steering route 192.168.203.1/32 populates the route table of the active CGN node, and this route is used as an indirect next-hop for a static-route-entry to a default route. For the NAT64 function, the NAT64 translator address 2001:DB8:122:344::/96 is used to attract IPv6 traffic with IPv4-embedded addresses.

The first of the two following outputs shows that CGN-2 is advertising the NAT44 default route as a VPN-IPv4 prefix and the NAT64 translator address as a VPN-IPv6 prefix. Both routes are advertised with the relevant Route-Target for VPRN 200 (target:64496:200) and, therefore, will be imported by PE-2 and subsequently advertised to CE-2. The second output shows the same commands entered at CGN-1 and verifies that because CGN-1 is the standby CGN node, it is not advertising either VPN-IPv4/VPN-IPv6 prefix.

```
*A:CGN-2# show router bgp routes vpn-ipv4 rd 64496:200 hunt
===============================================================================
 BGP Router ID:192.0.2.9        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
---snip---
-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 0.0.0.0/0
Nexthop        : 192.0.2.9
Route Dist.    : 64496:200              VPN Label      : 262141
Path Id        : None
To             : 192.0.2.10
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : None
AIGP Metric    : None
Connector      : None
Community      : target:64496:200
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 192.0.2.10
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : N/A
Orig Validation: N/A
Source Class   : 0                      Dest Class     : 0

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


*A:CGN-2# show router bgp routes vpn-ipv6 rd 64496:200 hunt
===============================================================================
 BGP Router ID:192.0.2.9        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv6 Routes
===============================================================================
---snip---
-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
Network        : 2001:db8:122:344::/96
Nexthop        : ::ffff:192.0.2.9
Route Dist.    : 64496:200              VPN Label      : 262141
Path Id        : None
To             : 192.0.2.10
Res. Nexthop   : n/a
Local Pref.    : 100                    Interface Name : NotAvailable
Aggregator AS  : None                   Aggregator     : None
Atomic Aggr.   : Not Atomic             MED            : 0
AIGP Metric    : None
Connector      : None
Community      : target:64496:200
Cluster        : No Cluster Members
Originator Id  : None                   Peer Router Id : 192.0.2.10
Origin         : IGP
AS-Path        : No As-Path
Route Tag      : 0
Neighbor-AS    : N/A
Orig Validation: N/A
Source Class   : 0                      Dest Class     : 0

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

```
*A:CGN-1# show router bgp routes vpn-ipv4 rd 64496:200 hunt
===============================================================================
 BGP Router ID:192.0.2.8        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv4 Routes
===============================================================================
---snip---
-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------


*A:CGN-1# show router bgp routes vpn-ipv6 rd 64496:200 hunt
===============================================================================
 BGP Router ID:192.0.2.8        AS:64496       Local AS:64496
===============================================================================
 Legend -
 Status codes  : u - used, s - suppressed, h - history, d - decayed, * - valid
                 l - leaked, x - stale, > - best, b - backup, p - purge
 Origin codes  : i - IGP, e - EGP, ? - incomplete

===============================================================================
BGP VPN-IPv6 Routes
===============================================================================
---snip---
-------------------------------------------------------------------------------
RIB Out Entries
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
```

Verification of Data Path

The host connected to CE-2 and the web server accessible from IAR-1 are used to verify the end-to-end data path for both NAT44 and NAT64.

NAT44

The host connected to CE-2 initiates an IPv4 UDP session toward the web server with a source address of 172.31.102.3 and a destination address of 10.1.128.43. The source port used is 1357, and the destination port is 80.

A two-way data transfer is verified as successful. The following output shows the details of the NAT44 binding at CGN-2, the active CGN node. The inside IP address and port are as described, while the allocated outside IP address is 10.1.4.254 using outside port 1047.

```
*A:CGN-2# tools dump nat sessions inside-ip 172.31.102.3

===============================================================================
Matched 1 session on Slot #1 MDA #2
===============================================================================
Owner               : LSN-Host@172.31.102.3
Router              : 200
Policy              : NAT44
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 172.31.102.3
Inside Port         : 1357
Outside IP Addr     : 10.1.4.254
Outside Port        : 1047
Foreign IP Addr     : 10.1.128.43
Foreign Port        : 80
Dest IP Addr        : 10.1.128.43
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
===============================================================================
```

NAT64

The host connected to CE-2 also initiates an IPv6 UDP session toward the web server with a source address of 2001:DB8:4001:200::3 and a destination address of 2001:DB8:122:344::A01:802B. The destination address represents the NAT64 translator address (2001:DB8:122:344::/96) and the embedded IPv4 address (10.1.128.43) translated into colon-hexidecimal format (A01:802B). The source port is 2468 and the destination port is 80.

A two-way data transfer is verified as successful. The following output shows the details of the NAT64 binding at CGN-2. Again, the inside IPv6 address and port are as described, while the allocated outside IP address is 10.1.6.254 using outside port 1032.

```
*A:CGN-2# tools dump nat sessions inside-ip 2001:db8:4001:200::3

===============================================================================
Matched 1 session on Slot #1 MDA #2
===============================================================================
Owner               : NAT64-Sub@2001:db8:4001:200::3
Router              : 200
Policy              : NAT64
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 2001:db8:4001:200::3
                                        Inside Port         : 2468
Outside IP Addr     : 10.1.6.254
Outside Port        : 1032
Foreign IP Addr     : 10.1.128.43
Foreign Port        : 80
Dest IP Addr        : 10.1.128.43
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
===============================================================================
```

Failover

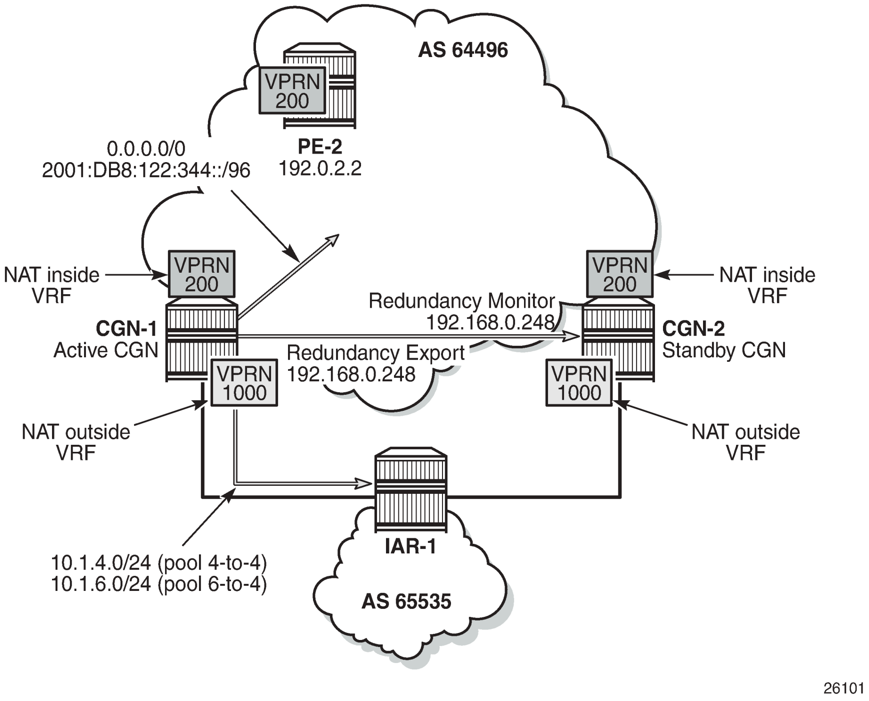
Before simulating a failover test, an IPv4 UDP session is established between the host connected to CE-2 and the web server, to ensure data-path continuity during the failure.

CGN-2 is the active CGN node; to simulate a failure of the MS-ISA board, it is placed into a shutdown state.

```
*A:CGN-2# configure card 1 mda 2 shutdown

1 2016/10/13 15:10:21.39 UTC WARNING: SNMP #2004 Base 1/2/nat-in-ip
"Interface 1/2/nat-in-ip is not operational"

2 2016/10/13 15:10:21.39 UTC MINOR: NAT #2024 Base NAT
"The state of NAT group 1 changed to out-of-service."
```

Figure 82: Post-Failover Redundancy State shows the example topology in the post-failover redundancy state.

Figure 82: Post-Failover Redundancy State



Because the MS-ISA in slot 1/2 is the only MDA in **nat-group 1**, it is sufficient to force the NAT group down. After **nat-group 1** is declared down at CGN-2, the following actions take place:

1. CGN-2 transitions the active state to false for the leader outside pool and any follower pools because its admin state changed to down, as follows:

```
22 2016/10/13 15:10:21.40 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to false for pool "4-to-4" - state changed
 to "Down"."

23 2016/10/13 15:10:21.40 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to false for pool "6-to-4" - state changed
 to "Down"."
```

2. CGN-2 withdraws the redundancy export route (192.168.0.249/32) from the NAT outside routing context (VPRN 1000). This means that the monitor route is no longer present in the routing table of CGN-1. Therefore, CGN-1 transitions to an active state and advertises its own export route. In this example, where BGP is used to advertise monitor routes, the Minimum Route Advertisement Interval is configured for 1 second, to reduce re-convergence times. However, when the monitor route is withdrawn at the standby CGN node, the system will wait for 10 seconds to ensure that this is not a route flap before declaring itself active. This is a non-configurable timer.

```
6 2016/10/13 15:10:31.53 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to true for pool "4-to-4" - state changed
to "Active"."

8 2016/10/13 15:10:31.53 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to true for pool "6-to-4" - state changed
to "Active"
```

3. Because CGN-2 is now standby, the outside pool addresses are no longer present in the routing table of the outside routing context (VPRN 1000). Therefore, they are withdrawn by CGN-2 in the EBGP peering session to IAR-1. Conversely, because CGN-1 is now active, it advertises the outside pools to IAR-1.
4. The NAT44 redundancy steering route (192.168.203.1/32) is no longer active in the NAT inside (VPRN 200) routing table at CGN-2. Therefore, the default route no longer has a valid next-hop, so the route is withdrawn. Conversely, the steering route is now present in the routing table of VPRN 200 at CGN-1. Therefore, the default route becomes active, and is advertised into VPRN 200 as a VPN-IPv4 prefix.
5. The NAT64 translator address is no longer active in the NAT inside (VPRN 200) IPv6 routing table at CGN-2, so the address is withdrawn. Conversely, the NAT64 translator address is now present in the IPv6 routing table of VPRN 200 at CGN-1 and is advertised into VPRN 200 as a VPN-IPv6 prefix.

The operational state of the outside pools can be verified at CGN-1, as follows:

```
*A:CGN-1# show router 1000 nat pool "4-to-4" | match "Dual-Homing" post-lines 12
Dual-Homing
===============================================================================
Type                                  : Leader
Export route                          : 192.168.0.248/32
Monitor route                         : 192.168.0.249/32
Admin state                           : inService
Dual-Homing State                     : Active
===============================================================================

===============================================================================
Dual-Homing fate-share-group
===============================================================================
Router         Pool                                                   Type
-------------------------------------------------------------------------------
vprn1000       4-to-4                                                 Leader
vprn1000       6-to-4                                                 Follower
-------------------------------------------------------------------------------
No. of pools: 2
===============================================================================
```

Finally, the integrity of the IPv4 UDP session between the host connected to CE-2 and the web server is verified, and the associated NAT binding is shown at CGN-1, as follows:

```
*A:CGN-1# tools dump nat sessions inside-ip 172.31.102.3

===============================================================================
Matched 1 session on Slot #1 MDA #2
===============================================================================
Owner               : LSN-Host@172.31.102.3
Router              : 200
Policy              : NAT44
FlowType            : UDP               Timeout (sec)       : 300
Inside IP Addr      : 172.31.102.3
Inside Port         : 1357
Outside IP Addr     : 10.1.4.254
Outside Port        : 1049
Foreign IP Addr     : 10.1.128.43
Foreign Port        : 80
Dest IP Addr        : 10.1.128.43
Nat Group           : 1
Nat Group Member    : 1
-------------------------------------------------------------------------------
===============================================================================
```

When the failure is resolved at CGN-2 and the MS-ISA comes back online, the failover mechanism is non-revertive. This is because CGN-2 already has the CGN-1 export route present in the routing table of the NAT outside routing context (VPRN 1000) as its monitor route. The following output at CGN-2 shows the MS-ISA and NAT group 1 transitioning to in-service, followed by the active state of the outside pools changing from down to standby.

```
*A:CGN-2# configure card 1 mda 2 no shutdown


9 2016/10/13 15:55:10.39 UTC MINOR: NAT #2024 Base NAT
"The state of NAT group 1 changed to in-service."

10 2016/10/13 15:55:10.39 UTC MINOR: NAT #2025 Base NAT
"The NAT group 1 is not degraded."

17 2016/10/13 15:55:10.40 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to false for pool "4-to-4" - state changed
 to "Standby"."

18 2016/10/13 15:55:10.39 UTC WARNING: NAT #2017 vprn1000 NAT redundancy
"The Large Scale NAT activity changed to false for pool "6-to-4" - state changed
 to "Standby"."

19 2016/10/13 15:55:10.39 UTC MINOR: NAT #2020 Base NAT
"The NAT MDA 1/2 is now active in group 1."
```

Conclusion

NAT stateless dual-homing provides a compromise between a lack of redundancy and the protocol and state synchronization requirements for stateful NAT redundancy. This is particularly true when CGN nodes provide a gateway to the Internet where Service Level Agreements (SLAs) are often difficult to guarantee.

This chapter provides an example of how NAT stateless dual-homing is configured and describes how SR OS provides the redundancy mechanism for NAT44 and NAT64. The example in this chapter does not represent the only way that NAT stateless dual-homing can be delivered. It uses VPRNs in both the inside and outside routing contexts, but the GRT is also an option for either. It uses IP filtering for NAT diversion for both NAT44 and NAT64, but a routing-based approach using destination-prefix is also option for NAT44. It uses BGP in the NAT outside routing context and BGP-VPN in the NAT inside routing context to advertise redundancy routes externally, but any routing protocol that can be accessed through the route-policy framework is applicable.

Customer document and product support

Customer documentation

Customer documentation welcome page

Technical support

Product support portal

Documentation feedback

Customer documentation feedback