Application Assurance

In This Section

This section provides an overview of Alcatel-Lucent's implementation of the Application Assurance service model.

Topics include:

Application Assurance (AA) Overview

Network operators are transforming broadband network infrastructures to accommodate unified architecture for IPTV, fixed and mobile voice services, business services, and High Speed Internet (HSI), all with a consistent, integrated awareness and policy capability for the applications using these services.

As bandwidth demand grows and application usage shifts, the network must provide consistent application performance that satisfies the end customer requirements for deterministic, managed quality of experience (QoE), according to the business objectives for each service and application. Application Assurance (AA) is the enabling network technology for this evolution in the service router operating system.

Application Assurance, coupled with subscriber and/or VPN access policy control points enables any broadband network to provide application-based services. For service providers, this unlocks:

- The opportunity for new revenue sources.
- Content control varieties of service.
- Control over network costs incurred by various uses of HSI.
- Complementary security aspects to the existing network security.
- Improved quality of service (QoS) sophistication and granularity of the network.
- The ability to understand and apply policy control on the transactions traversing the network.

Application Assurance: Inline Policy Enforcement

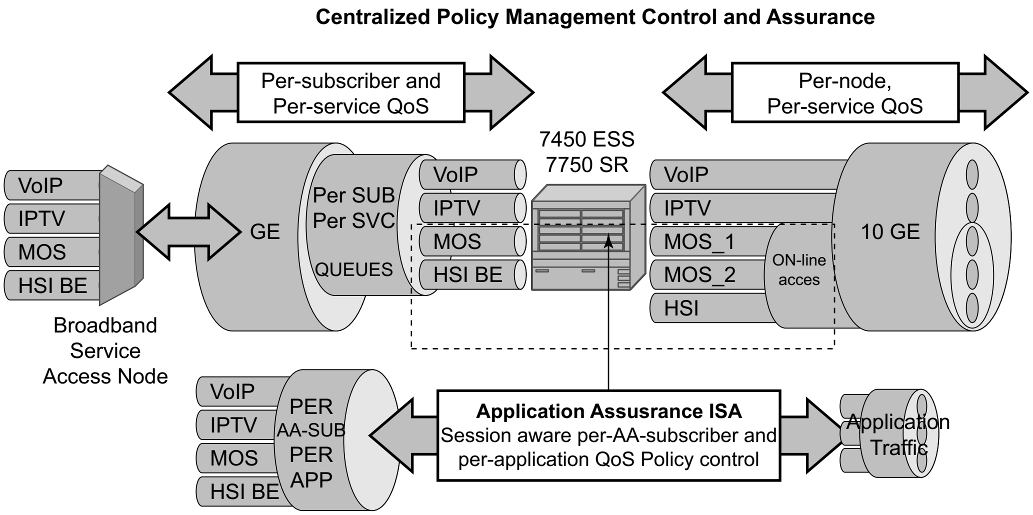

Centralized Policy Management Control and Assurance
Per-subscriber and Per-service QoS
Per-node, Per-service QoS
7450 ESS
7750 SR
VoIP
IPTV
MOS
HSI BE
Broadband Service Access Node
GE
Per SUB
Per SVC
QUEUES
VoIP
IPTV
MOS
HSI BE
VoIP
IPTV
MOS_1
MOS_2
HSI
ON-line acces
10 GE
VoIP
IPTV
MOS
HSI BE
PER
AA-SUB
PER
APP
Application Assusrance ISA
Session aware per-AA-subscriber and per-application QoS Policy control
Application Traffic

Figure 4: AA ISA Inline Identification, Classification and Control

The integrated solution approach for Application Assurance recognizes that a per-AA-subscriber and per-service capable QoS infrastructure is a pre-condition for delivering application-aware QoS capabilities. Enabling per-application QoS in the context of individual subscriber’s VPN access points maximizes the ability to monetize the application service, because a direct correlation can be made between customers paying for the service and the performance improvements obtained from it. By using an integrated solution there is no additional cost related to router port consumption, interconnect overhead or resilience to implement in-line application-aware policy enforcement.

AA Integration in Subscriber Edge Gateways

Multiple deployment models are supported for integrating application assurance in the various subscriber edge and VPN PE network topologies. In all cases, application assurance can be added by in-service upgrade to the installed base of equipment rather than needing deploy and integrate a whole new set of equipment and vendors into the network for Layer 4-7 awareness.

Integrating Layer 4-7 application policy with the 7750 SR or 7450 ESS subscriber edge policy context is the primary solution to address both residential broadband edge or Layer 2/Layer 3 application aware business VPN. Placement of Layer 4-7 analysis at the distributed subscriber edge policy point simplifies AA deployments in the following ways:

- For residential markets, CO-based deployment allows deployment-driven scaling of resources to the amount of bandwidth needed and the amount of subscribers requiring application-aware functionality.
- For AA business VPNs, a network deployment allows large scale application functionality at a VPN provider edge access point, vastly reducing complexity, cost, and time to market required to offer application-aware VPN services.
- Traffic asymmetry is avoided. Any subscriber traffic usually passes through one CO subscriber edge element so there is no need for flow paths to be recombined for stateful analysis.
- PE integration provides a single point of policy enforcement.
- SeGW integration provides firewall protection for NMS, MME and SGW.

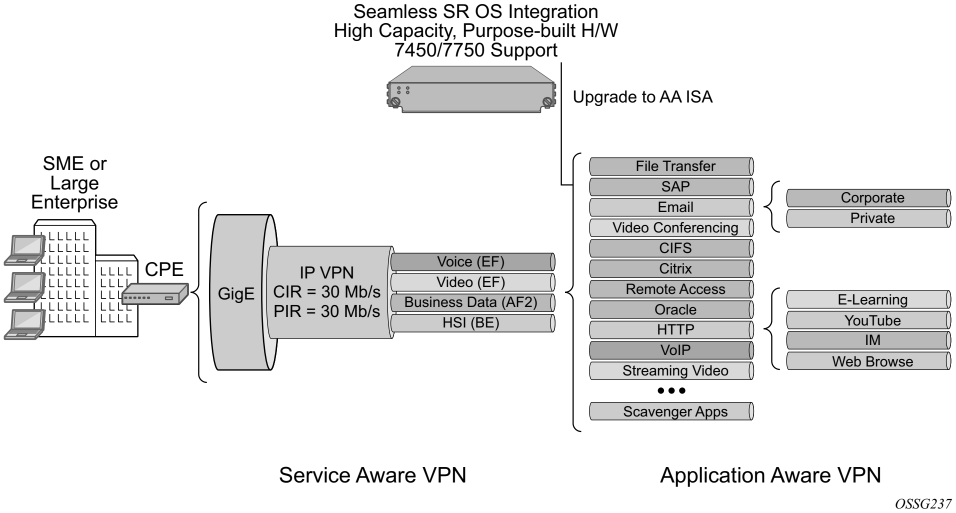


Figure 5: AA Deployment Topologies

There are residential topologies where it is not possible or practical to distribute ISAs into the same network elements that run ESM, including for legacy edge BRASs that still need Application Assurance policy (reporting and control) for the same Internet services, and which needs to be aligned and consistent with the ESM AA policy. This is supported using transit AA subscribers, typically in the first routed element behind the legacy edge.

Application Assurance enables per AA-subscriber (a residential subscriber, or a Layer 2/Layer 3 SAP or spoke SDP), per application policy for all or a subset of AA subscriber's applications. This provides the ability to:

- Implement Layer 4-7 identification of applications using a multitude of techniques from a simple port-based/IP address based identification to behavioral techniques used to identify, for example, encrypted or evasive applications.
- Once identified, to apply QoS policy on either an aggregate or a per-AA-subscriber, per-application basis.
- Provide reports on the identification made, the traffic volume and performance of the applications, and policies implemented.

An integrated AA module allows the SR/ESS product families to provide application-aware functions that previously required standalone devices (either in residential or business environment) at a fraction of cost and operational complexity that additional devices in a network required.

A key benefit if integrating AA in the existing IP/MPLS network infrastructure (as opposed to an in-line appliance) is the ability to select traffic for treatment on a granular, reliable basis. Only traffic that requires AA treatment is simply and transparently diverted to the ISA. Other traffic from within the same service or interface will follow the normal forwarding path across the fabric. In the case of ISA failure, ISA redundancy is supported and in the case no backup ISAs are available the AA traffic reverts to the normal fabric matrix forwarding, also known as “fail to fabric”.

Table 3: Traffic Diversion to the ISA

Deployment Case	System Divert ID	AA-Sub Type	App-Profile on:
Residential Edge	ESM Sub-ID	ESM	ESM sub (All IPs, not per-host)
Wireless LAN GW	DSM-ID	DSM	DSM
Business Edge	L2/L3 SAP	SAP	SAP (Aggregate)
Residential Transit	Parent L3 SAP/Spoke-SDP	Transit AA	Transit Sub
Spoke Attached Edge	Spoke SDP	Spoke SDP	Spoke SDP (Aggregate)
SeGW	Parent SAP/Spoke-SDP or L2/L3 SAP	Transit AA SAP	Transit AA SAP

Fixed Residential Broadband Services

Fixed residential HSI services as a single edge Broadband Network Gateway (BNG) or as part of the Triple Play Service Delivery Architecture (TPSDA) are a primary focus of Application Assurance performance, subscriber and traffic scale.

To the service provider, application-based service management offers:

- Application aware usage metering packages (quotas, 0-rating etc.)
- New revenue opportunities to increase ARPU (average revenue per user) (for gaming, peer-to-peer, Internet VoIP and streaming video, etc.).
- Fairness: Aligns usage of HSI network resources with revenue on a per-subscriber basis.
- Operational visibility into the application usage, trends, and pressure points in the network.

To the C/ASP, service offerings can be differentiated by improving the customer’s on-line access experience. The subscriber can benefit from this by gaining a better application experience, while paying only for the value (applications) that they need and want.

Dual-Stack Lite – DS-Lite

Dual Stack Lite is an IPv6 transition technique that allows tunneling of IPv4 traffic across an IPv6-only network. Dual-stack IPv6 transition strategies allow service providers to offer IPv4 and IPv6 services and save on OPEX by allowing the use of a single IPv6 access network instead of running concurrent IPv6 and IPv4 access networks. Dual-Stack Lite has two components: the client in the customer network (the Basic Bridging BroadBand element (B4)) and an Address Family Transition Router (AFTR) deployed in the service provider network.

Dual-Stack Lite leverages a network address and port translation (NAPT) function in the service provider AFTR element to translate traffic tunneled from the private addresses in the home network into public addresses maintained by the service provider. On the 7750 SR, this is facilitated through the Carrier Grade NAT function.

When a customer's device sends an IPv4 packet to an external destination, DS-Lite encapsulates the IPv4 packet in an IPv6 packet for transport into the provider network. These IPv4-in-IPv6 tunnels are called softwires. Tunneling IPv4 over IPv6 is simpler than translation and eliminates performance and redundancy concerns.

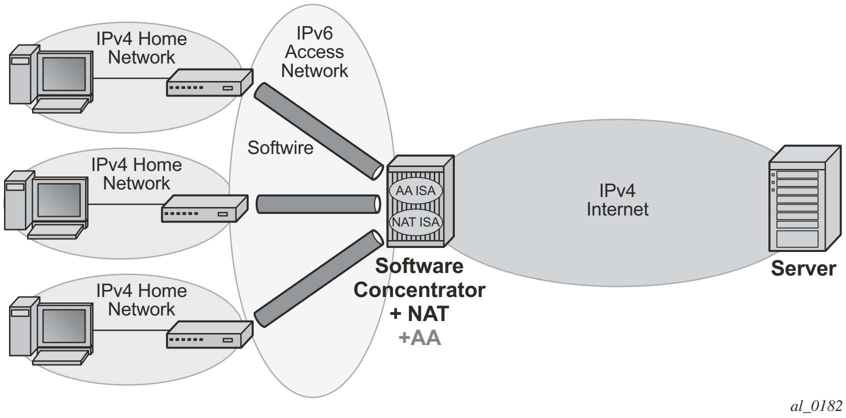


Figure 6: DS-Lite Deployment

The IPv6 source address of the tunnel represents a unique subscriber. Only one tunnel per customer (although more is possible), but the IPv6 addresses cannot overlap between different customers. When encapsulated traffic reaches the softwire concentrator, the device treats the source-IP of the tunnel to represent a unique subscriber. The softwire concentrator performs IPv4 network address and port translation on the embedded packet by re-using Large Scale NAT and L2-Aware NAT concepts.

Advanced services are offered through Application Assurance multi service ISA to the DS-Lite connected customers. Subscribers' traffic (ESMs or transit-ip) are diverted to AA ISA for L3-L7 identification / classifications, reporting and control based on the IPv4 packets (transported within the IPv6 DS-Lite tunnel). This AA classification, reporting and control of subscribers' traffic take effect before any NAT44 functions. In other words, AA sites on the subscriber side of NAT44.

The absence of a control protocol for the IP-in-IP tunnels simplifies the operational/management model, since any received IPv6 packet to the AA ISA can be identified as a DS-Lite tunneled packet if:

- protocol 4 in the IPv6 header, and
- the embedded IP packet is IPv4 (inside).

Fragmented IPv4 are supported only if tunneled through non-fragmented IPv6 packets.

Fragmentation at the IPv6 layer is not supported by AA ISA (when used to tunnel fragmented or non-fragmented IPv4 packets). These packets are cut-through with sub-default policy applied with a possibility of re-ordering.

If DSCP AQP action is applied to DS-Lite packet, both IPv4 and IPv6 headers are modified. AQP mirroring action is applied at the IPv6 layer. All collected statistics include the tunnel over-head bytes (also known as IPv6 header size).

6to4 /6RD

6RD/6to4 tunneling mechanism allows IPv6 sites to communicate over an IPv4 network without the need to configure explicit tunnels, as well as and for them to communicate with native IPv6 domains via relay routers. Effectively, 6RD/6to4 treats the wide area IPv4 network as a unicast point-to-point link layer. Both ends of the 6RD/6to4 tunnel are dual-stack routers. Because 6RD/6to4 does not build explicit tunnels, it scales better and is easier to manage after setup

6to4 encapsulates an IPv6 packet in the payload portion of an IPv4 packet with protocol type 41. The IPv4 destination address for the encapsulating IPv4 packet header is derived from the IPv6 destination address of the inner packet (which is in the format of 6to4 address) by extracting the 32 bits immediately following the IPv6 destination address's 2002:: prefix. The IPv4 source address in the encapsulating packet header is the IPv4 address of the outgoing interface (not system IP address).

6RD is very similar to 6to4. The only difference is that the fixed 2002 used in 6to4 prefix is replaced by a configurable prefix.

An important deployment of 6RD/6to4 deployment is in access network as shown in Figure 7.

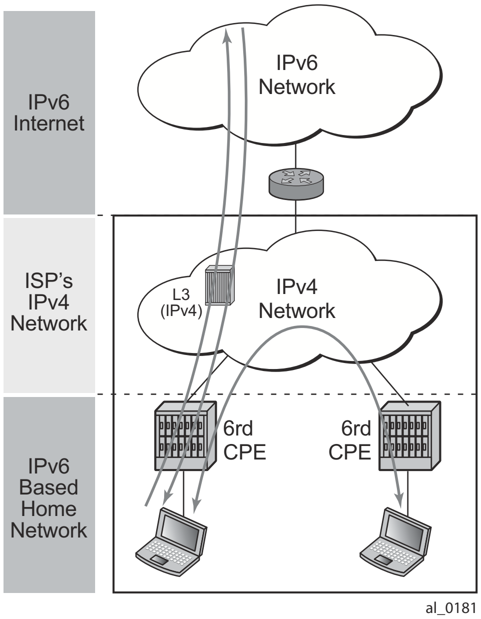


Figure 7: 6to4 in Access Network Deployment

To provide IPv6 services to subscribers, 6RD is deployed in these access networks to overcome the limitations of IPv4 only access network gear (for example, DSLAMs) with no dual stack support.

From an AA ISA point of view, deployment of 6RD in the access network is similar to that of the general deployment case between IPv6 islands with the added simplification that each 6RD tunnel carries traffic of a single subscriber.

When AA ISA sees an IPv4 packet with protocol type 41 and a payload that includes IPv6 header, it detects that this is a 6rd/6to4 tunneled packet.

AA ISA detects, classifies, reports, and applies policies to 6rd/6to4 packet for ESM, SAP, spoke-SDP, and transit-IP (ip-policy) AA subscriber types.

Fragmented IPv6 are supported only if tunneled through non-fragmented IPv4 packets.

Fragmentation at the IPv4 layer is not supported by AA ISA (when used to tunnel fragmented or non-fragmented IPv6 packets). These packets are cut-through with sub-default policy applied with a possibility of re-ordering.

If the packet has IPv4 options then AA ISA will not look into the IPv6 header. The packet will be classified as IPv4 “unknown TCP/UDP”. Furthermore, TCP/UDP checksum error detection is only applied for IPIPE and routed services.

If the DSCP AQP action is applied to 6RD6to4 packets, both IPv4 and IPv6 headers are modified. AQP mirroring action is applied at the IPv4 layer. All collected statistics include the tunnel overhead bytes, aka. IPv4 header size.

Wireless LAN Gateway Broadband Services

Application Assurance enables a variety of use cases important for Wireless LAN Gateway deployments in residential, public WiFi or VPN wireless LAN services. These include:

- HTTP redirect for subscriber authentication with HTTP whitelist — Redirects all non-authenticated subscriber HTTP traffic to an authentication portal and blocks the rest of Internet access, but allows user access to specific whitelisted websites, download Apps and software needed to authenticate.
- HTTP redirect by policy — URL or application blocking or usage threshold notification. Redirects some or all subscriber HTTP traffic to an portal landing site based on static or dynamic policy. This can be done while not interrupting selected HTTP based services such as streaming video.
- Inline HTTP browser notification — Provides messaging in the form of web banners, overlays, or http-redirection. This can always be enabled, One-time per sub at authentication (greeting message "Welcome to our WiFi Service"), one time per COA, or periodically.
- ICAP for large scale URL filtering — ICAP client in AA interacts with offline ICAP URL filtering services, for parental control or large blacklists. Reduces cost as only URLs for specific flows are sent to server, rather than full inline traffic.
- Analytics — Provides operator insight into the following: Application and App-group volume usage by time of day/day of week, top subs, devices, etc.
- Traffic control for fair use policy — Prevents some users of the hotspot from consuming a disproportionate amount of resources by limiting to volume of such use across all subscribers as a traffic management tool, or on a per-subscriber basis.
- Stateful Firewall — Prevents unsolicited sessions from attacking devices.

Application-Aware Business VPN Services

AA for business services can be deployed at the Layer 2 or Layer 3 network provider edge (PE) policy enforcement point for the service or at Layer 2 aggregation policy enforcement point complimentary to the existing Layer 3 IP VPN PE. In a business environment, an AA-subscriber represents a VPN access point. A typical business service can have a much larger average bandwidth rate then the residential service and is likely to have a smaller AA-subscriber count than a residential deployment.

Up to seven active ISAs can be deployed per PE, each incrementally processing up to 10Gbps. The in-network scalability is a key capability that allows a carrier to be able to grow the service bandwidth without AA throughput affecting the network architecture (more edge nodes, application-aware devices).

Application-aware Layer 2 and Layer 3 VPNs implemented using AA ISA equipped 7750 SR/ 7450 ESS together with rich network management (5620 SAM, 5750 RAM, end customer application service portals) give operators a highly scalable, flexible, and cost effective integrated solution for application-based services to end customers. These services may include:

- Rich application reporting with VPN, access site visibility.
- Right-sizing access pipes into a VPN service to improve/ensure application performance.
- Application-level QoS (policing, session admission, remarking, etc.) to ensure application-level performance, end-customer QoE objectives are met.
- Value-added services such as application verification, new application detection, application mirroring.
- Performance reporting for real time (RTP) and non-real time (TCP) based applications.
- Dual Stack IPv4 – IPv6 support.
- Control unauthorized or recreational applications by site, by time of day.

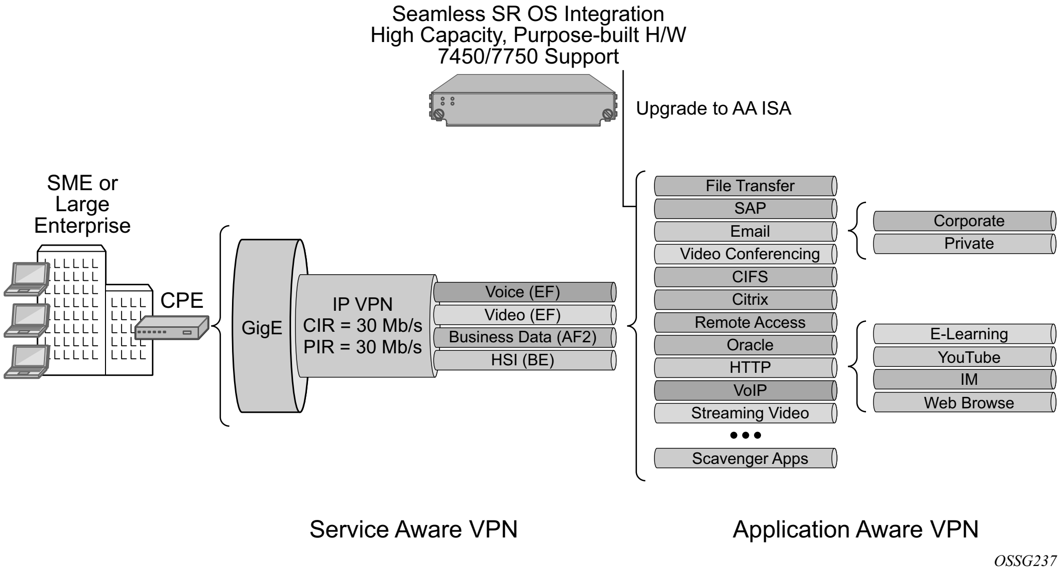


Figure 8: AA BVS Services Integrated into the Provider Edge

SeGW Firewall Service

Application Assurance deployed within a 7750-SR Security gateway in ultra-broadband access networks (3G/4G/Femto) provides the operator with back-end core network security protection. AA Firewall protection includes:

1. S1-MME (SCTP) traffic
2. S1- U (GTP-U) traffic
3. OAM traffic

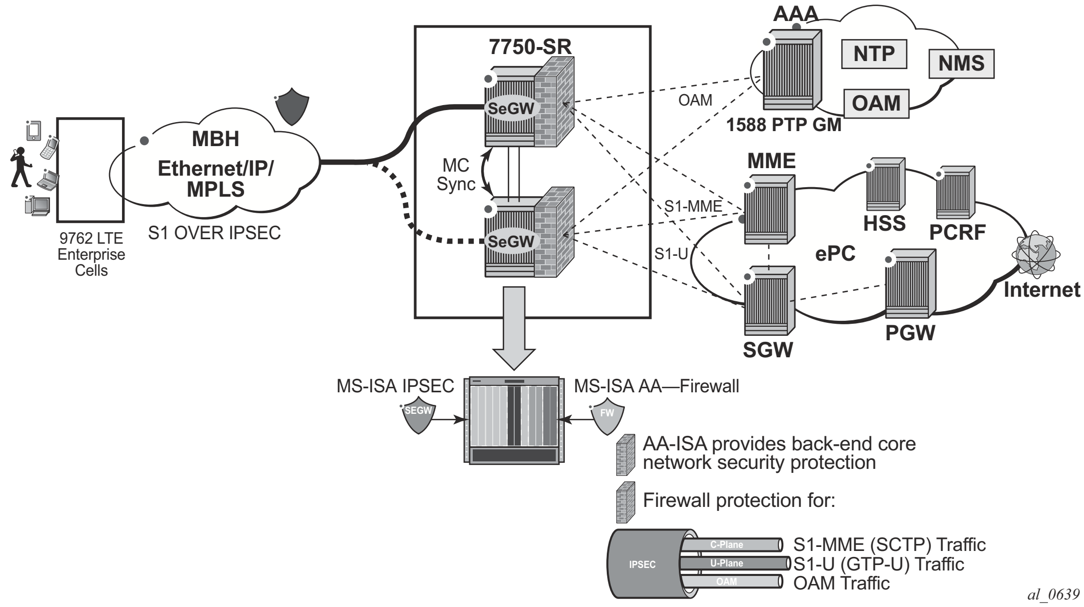


Figure 9: SeGW Firewall Deployment

OAM Traffic Protection

The aim of AA Firewall protection is to protect and prevent 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/Femto Access Points (FAP).

Ports Scan attacks: Using AA FW stateful session filters, operators can allow traffic only on certain IP address(s) and port number(s).

1. **Ports Scan Attacks**: Using AA FW stateful session filters, operators can allow traffic only on certain IP address(s) and port number(s).
   - → For example, operator can configure AA to only allow traffic that is initiated by NMS towards the FAPs. Hence, a compromised FAP cannot initiate an attack on NMS infrastructure.
   - → Operator can limit the type of traffic allowed based on L3 — L7 classification.. Operator can allow only HTTP with a certain URL/domain, DNS, PTP, FTP (independent of the port number used) and block all other traffic.
2. Flood Attacks: The operator can limit the type of traffic allowed based on Layer 3 — Layer 7 classification. The operator can allow only HTTP with a certain URL/domain, DNS, PTP, FTP. Note that the AA ISA provides configurable flow policers that can act on FW permitted sessions. These policers, once configured prevent all sort of flooding attacks, such as ICMP PING flooding, UDP flooding, SYN Flood Attack, etc., of the port number used) and block all other traffic.
   - → These policers provide protection at multiple levels; per system per application/application groups and per FAP (or per NMS) per applications/applications groups.
   - → There are three types of AA ISA policers:
     - – Flow setup rate policers to limit the number of new flows.
     - – Flow count policers to limit the total number of active flows.
     - – Bandwidth policers to limit the total OAM bandwidth allowed by a given FAP towards NMS.
3. Malformed Packets Attacks: In order to protect Hosts and network resources, AA FW performs validation on IP packets, at the IP layer and TCP/UDP layer, to ensure that the packets are valid. Invalid packets are discarded (a configurable option). This provides protection against well known attacks such as LAND attack. See SeGW Firewall Service on page 41 for a complete description. AA allows the operator to optionally drop fragmented or out-of-order fragmented IP packets.

In addition, for OAM traffic, all AA functionalities including Layer 7 analytics and control as well as Application Layer Gateway (ALG) are supported.

S1-MME Traffic Protection

The aim of AA Firewall (FW) in this deployment is to protect the MME(s) infrastructure against an attack from a compromised eNB/FAP.

Network flooding attacks, malformed packets and port scans are examples of DoS attacks that can be carried out using a compromised eNB/ Femto Access Points (FAP).

AA FW provides inspection of SCTP – protocol used to communicate to MME. Such inspection includes checking for SCTP protocol ID, source /destination ports, PPID, SCTP chunk checking and malformed SCTP packet (such as checksum validation).

For S1-MME traffic, the operator can configure various AA actions:

- Drop packets with invalid checksum, src/dest IP and/or port numbers (malformed packet protection)
- PPID Filtering according to configured rules
- Rate limit the amount of S1-MME traffic (flooding protection) in terms of Bandwidth (bits/sec).
- Limit the number of concurrent SCTP flows (flooding protection)
- Limit the SCTP flow setup rate (flows/sec) to protect against DoS flooding.
- Drop fragmented packets or drop out-of-order fragmented packets.

The actions above can be applied per eNB/FAP IP address and /or per MME (to control aggregate traffic per MME).

S1-U GTP Traffic Protection

The aim of AA Firewall (FW) in this deployment is to protect the SGW/SGSN infrastructure against an attack from a compromised eNB/FAP. AA FW supports:

- Protection against malformed GTP packets attack:
  - → Packet sanity checks: includes GTP mandatory, optional and extension header checks. As well as checks for invalid reserved IE and missing IEs.
- Protection against un-supported GTP messages
  - → Filter messages based on message type(s) and/or message length.
- Protection against flooding attack:
  - → GTP Traffic bandwidth shaping: limits the GTP-U bandwidth that a FAP can send to the core (SGW)
  - → GTP tunnel limiting: limit the number of concurrent GTP tunnels and/or setup rate of these tunnels from a FAP to the core network.

  - → In order to prevent the shared resources of bandwidth and the GSN’s processor from being consumed by an attacker, GTP rate limiting is recommended.

- Protection against IP Fragmentation based attacks:
  - → Drop Rules for IP fragmentation of GTP messages

Application Assurance System Architecture

AA ISA Resource Configuration

AA ISAs are flexible embedded, packet processing resource cards that require configuration such that services may be associated with the resources. This includes assigning ISAs to groups, optionally defining group partitions, and setting the redundancy model. Load balancing is affected by how ISAs are grouped.

AA ISA Groups

An AA ISA group allows operators to group multiple AA ISAs into a single logical group for consistent management of AA resources and policies across multiple AA ISA cards configured for that group.

AA ISA Groups

An AA ISA group allows operators to group multiple AA ISAs into one of several logical groups for consistent management of AA resources and policies across multiple AA ISA cards configured for that group. The following operations can be performed at the group level:

- Define one or multiple AA ISA groups to allow AA resource partitioning/reservation for different types of AA service.
- Define the AA subscriber scale mode for the group. Residential, VPN and distributed subscriber management modes are supported.
- Assign physical AA ISAs to a group.
- Select forwarding classes to be diverted for inspection by the AA subscribers belonging to the group and select the AA policy to be applied to the group.
- Configure redundancy and bypass mode features to protect against equipment failure.
- Configure QoS on IOMs which host AA ISAs for traffic toward AA ISAs and from AA ISAs.
- Configure ISA capacity planning using low and high thresholds.
- Enable partitions of a group.
- Configure the ISA traffic overload behavior for the group to either backpressure to the host IOM (resulting in possible network QoS-based discards) or to cut-though packets through the ISA without full AA processing. Cut-through is typically enabled for AA VPN groups but not for residential groups.

Residential services is an example where all AA services might be configured as part of a single group encompassing all AA ISAs, for operator-defined AA service. This provides management of common applications and reporting for all subscribers and services, with common or per customer AQP (using ASOs characteristics to divide AA group's AQP into per app-profile QoS policies).

Multiple groups can be further used to create separate services based on different sets of common applications, different traffic divert needs (such as for capacity planning) or different redundancy models. Cases where multiple groups might be used can include:

- For mix of residential and business customers.
- Among different business VPN verticals.
- For business services with a common template base but for different levels of redundancy, different FC divert, or scaling over what is supported per single group.
- System level status statistics have AA ISA group/partition scope of visibility.

AA Group Partitions

VPN-specific AA services are enabled using operator defined partitions of an AA Group into AA policy partitions, typically with one partition for each VPN-specific AA service. The partition allows VPN specific custom protocols/application/application group definition, VPN specific policy definition and VPN specific reporting (some VPNs with volume-only reports, while others with volume and performance reports). Each partition's policy can be again divided into multiple application QoS policies using ASOs.

The use of ISA groups and partitions also improves scaling of policies, as needed with VPN-specific AA policies.

If partitions are not defined, all of the AA group acts as a single partition. When partitions are configured, application identification, policy and statistics configuration applies only to the given partition and not any other partitions configured under the same AA group.

The definition of application profiles (and related ASO characteristics/values) are within the context of a given partition (however, application profiles names must have node-wide uniqueness)

The definition of applications, application groups and AQP are also specific to a given partition. This allows:

- The definition of unique applications and app-groups per partition.
- The definition of AQP policy per partition.
- The definition of common applications and app-groups per partition with per partition processing and accounting.

Partitions also enable accounting/reporting customization for every AA subscriber associated with a partition, for example:

- The ability to define different types of reporting/accounting policies for different partitions in a single AA group, such as uniquely define which application, protocols, app groups are being reported on for every AA subscriber that uses a given partition.
- AA group level protocol statistics with partition visibility (for example, protocol counts reported for each partition of the group separately).

The system provides independent editing and committing of each partition config (separate begin/commit/abort commands).

Policer templates allow group-wide policing, and can be referenced by partition policies.

Bypass Modes

If no active AA ISA is available (for example, due to an operational failure, misconfiguration) the default behavior is to forward traffic as if no AA was configured, the system does not send traffic to the AA ISA (equivalent to fail to closed). Alarms are raised to flag this state externally. There is an optional “fail to open” feature where AA ISA service traffic is dropped if no active AA ISA is present (such as no AA ISA is present and operationally up).

Redundancy

AA ISA group redundancy is supported, to protect against card failure and to minimize service interruption during maintenance or protocol signature upgrades.

No AA ISA Group Redundancy

AA can be configured with no ISA redundancy within the AA group. All AA ISAs are configured as primary with no backup (up to the limit of active AA ISAs per node). There is no fault state indicating that a spare AA ISA is missing. If a primary is configured but not active, there will be a "**no aa-isa**" fault.

Failure to Fabric

In the event that no ISA redundancy is deployed or insufficient ISAs are available for needed sparing, the system implements "failure to fabric". When the ISA status shows the ISA is not available and there is no redundant ISA available, the ingress IOMs simply do not divert the packets that would have been sent to that ISA, but instead these proceed to the next hop directly across the fabric. When the ISA becomes available, the divert eligible packets resume divert through the ISA. This behavior is completely internal to the system, without affecting the forwarding or routing configuration and behavior of the node or the network.

N+1 AA ISA Card Warm Redundancy

The system supports N+1 AA ISA equipment warm redundancy (N primary and 1 backup). If a backup is configured and there is no ISA available (a primary and backup failed), there will be a "**no aa-isa**" fault. The backup AA ISA is pre-configured with isa-aa.tim and the group policies. Datapath traffic is only sent to active AA ISAs, so the backup has no flow state. If a backup ISA is unavailable, there will be a "backup missing" fault.

An AA subscriber is created and assigned to a primary AA ISA when an application profile is assigned to a subscriber, SAP, or spoke SDP. By default, AA subscribers are balanced across all configured primary AA ISAs.

Upon failure of a primary AA ISA, all of its AA subscribers and their traffic are operationally moved to the newly active backup AA ISA but the current flow states are lost (warm redundancy). The new AA ISA will identify any session-based active flows at a time of switchover as an **existing** protocol, while the other flows will be re-identified. The **existing** protocol-based application filters can be defined to ensure service hot redundancy for a subset of applications. Once the backup AA ISA has taken control, it will wait for operator control to revert activity to the failed primary AA ISA module.

The user can disable a primary AA ISA for maintenance by triggering a controlled AA ISA activity switch to do the AA ISA software field upgrade (a shutdown of an active AA ISA is recommended to trigger an activity switch).

The activity switch experiences the following AA service impact:

- All flow states for the primary ISA are lost, but existing flows can be handled with special AQP rules for the existing flows by the newly active backup AA ISA until sessions end.
- All statistics gathered on the active AA ISA since the last interval information that was sent to the CPM will be lost.

ISA Load Balancing

Capacity-cost based load balancing allows a cost to be assigned to diverted AA subscribers (by the app-profile). Load Balancing uses the total allocated costs on a per-ISA basis to assign the subscriber to the lowest sum cost ISA resource. Each ISA supports a threshold as the summed cost value that notifies the operator if or when capacity planning has been exceeded.

The load balancing decision is made based on the AA capacity cost of an AA subscriber. The capacity cost is configured against the app-profile. When assigning a new diverted aa-sub to an ISA, the ISA with the lowest summed cost (that also has sufficient resources) is chosen. Examples of different load-balancing approaches that may be implemented using this flexible model include:

- aa-sub count balancing — Configure the capacity cost for each app-profile to the same number (for example, 1).
- aa-sub stats resource balancing — Configure the capacity cost to the number of stats collected for AA subscribers using the app-profile. This might be used if different partitions have significantly different stats requirements.
- Bandwidth balancing — Configure the capacity cost to the total bandwidth in both directions (in kbps) expected for those AA subscribers. This might be used if different AA subscribers have highly varying bandwidth needs.

Load balancing operates across ISAs with in an AA group, and will not balance across groups. The system will ensure that app-profiles assigned to AA subscribers (ESM subscribers, SAPs and spoke SDPs) that are within a single VPLS/Epipe/IES/VPRN service are all part of the same AA group (partitions within an AA group are not checked/ relevant).

Users can replace the app-profile assigned to an AA subscriber with another app-profile (from the same group/partition) that has a different capacity cost.

Regardless of the preferred choice of ISA, the system takes into account.

- Per previous releases, resource counts have per-ISA limits. If exceeded on the ISA of choice, that ISA cannot be used and the next best is chosen.
- Divert IOM service queuing resources may limit load-balancing. If queuing resources are exhausted, the system attempts to assign the aa-sub to the ISA where the first AA subscriber within that service (VPLS/Epipe) or service type (IES/VPRN) was allocated.

For prefix transit AA subscriber deployments using the remote-site command, traffic for the remote transit subs are processed a second time. The ISA used by the parent AA-sub will be used by all transits within the parent. In remote-site cases there may be a need to increase capacity cost of parent since the transits stay on same ISA as the parent.

Prefix transit AA-subs are all diverted to the same Group and partition as the parent SAP.

Asymmetry Removal

Asymmetry removal is a means of eliminating traffic asymmetry between a set of multi-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. Asymmetry means that the two directions of traffic for a given flow (to-sub and from-sub) take different paths through the network. Asymmetry removal ensures that all packets for each flow, and all flows for each AA subscriber are diverted to a given AA ISA.

Traffic asymmetry is created when there are multi-homed links for a given service, and the links are simultaneously carrying traffic. In this scenario packets for flows will use any reachable paths, thus creating dynamic and changing asymmetry. Single node or multi-chassis asymmetry removal is used for any case where traffic for an AA subscriber may be forwarded over diverse paths on active AA divert links in a multi-homed topology. This includes support for SAP/spoke AA subscribers as well as business and residential transit AA subscribers within the diverted service.

Asymmetry removal must be implemented in the first routed hop on the network side of the subscriber management point, such that there will be a deterministic and fixed SAP / spoke-SDP association between the downstream subscriber management the parent divert service.

Asymmetry removal allows support for the SAP or spoke SDPs to the downstream element to be multi-homed on active links to redundant PE AA nodes as shown in Figure 10.

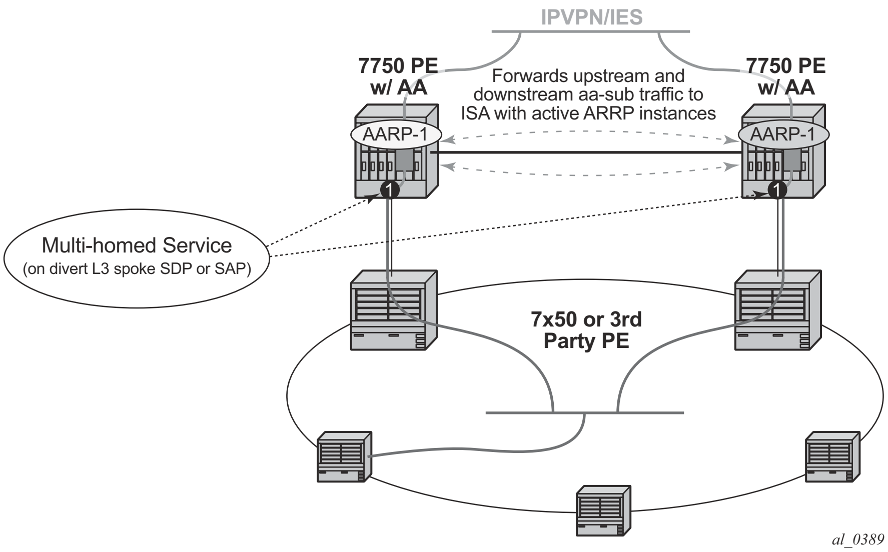

IPVPN/IES
7750 PE w/ AA
Forwards upstream and downstream aa-sub traffic to ISA with active ARRP instances
7750 PE w/ AA
AARP-1
AARP-1
1
1
Multi-homed Service
(on divert L3 spoke SDP or SAP)
7x50 or 3rd Party PE
al_0389

Figure 10: Transit Sub SAP/Spoke SDP Multi-Homing with Asymmetry

AA for transit-ip subscribers is commonly deployed behind the point of the subscriber policy edge after aggregation. This includes cases where AA needed is behind:

- Any 7x50 node running ESM but where there is not desire, need or space to deploy distributed AA ISAs.
- Legacy BRAS that do not support integrated application policy.

Asymmetry removal also allows a VPN site (Figure 11) to be connected with multi-homed, dual-active links while offering AAN services with the ISA.

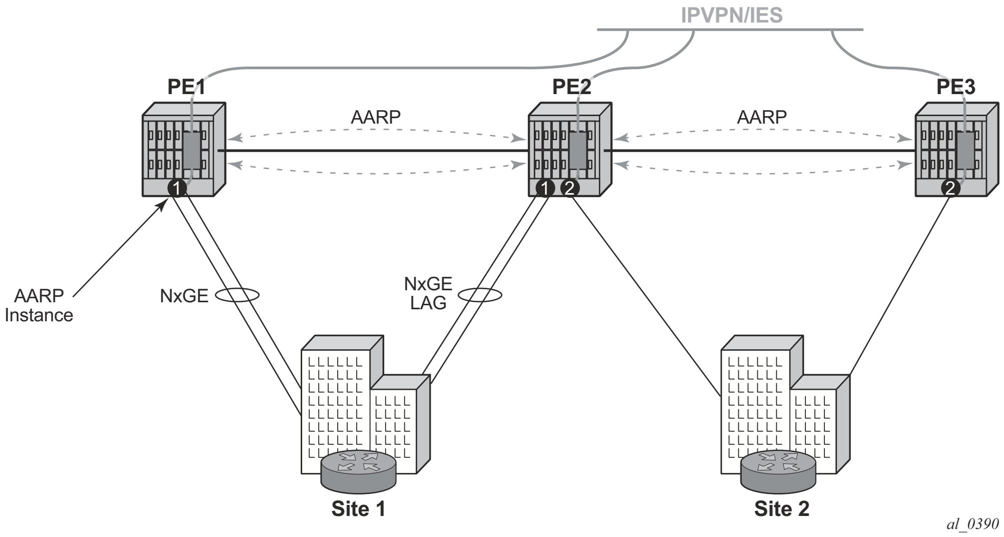


Figure 11: VPN Site Multi-Homing with Asymmetry

Asymmetry removal is supported for Layer 3 AA divert services:

- IES SAP and spoke SDP
- VPRN SAP and spoke SDP

When asymmetry exists between multi-chassis redundant systems, Ipipe spoke SDPs are used to interconnect these services between peer nodes over an Inter-Chassis Link (ICL).

Asymmetry removal supports multiple endpoints of a service with a common AARP instance, with a primary endpoint assigned the app-profile (and transit policy for transit subs). The remaining endpoints are defined as secondary endpoints of the AARP instance. All SAP or spoke endpoints within a given AARP instance are load balanced to the same ISA in that node. Multi-endpoint AARP instances allow single-node asymmetry removal and multi-chassis asymmetry removal with multiple active links per node.

Asymmetry Removal Overview

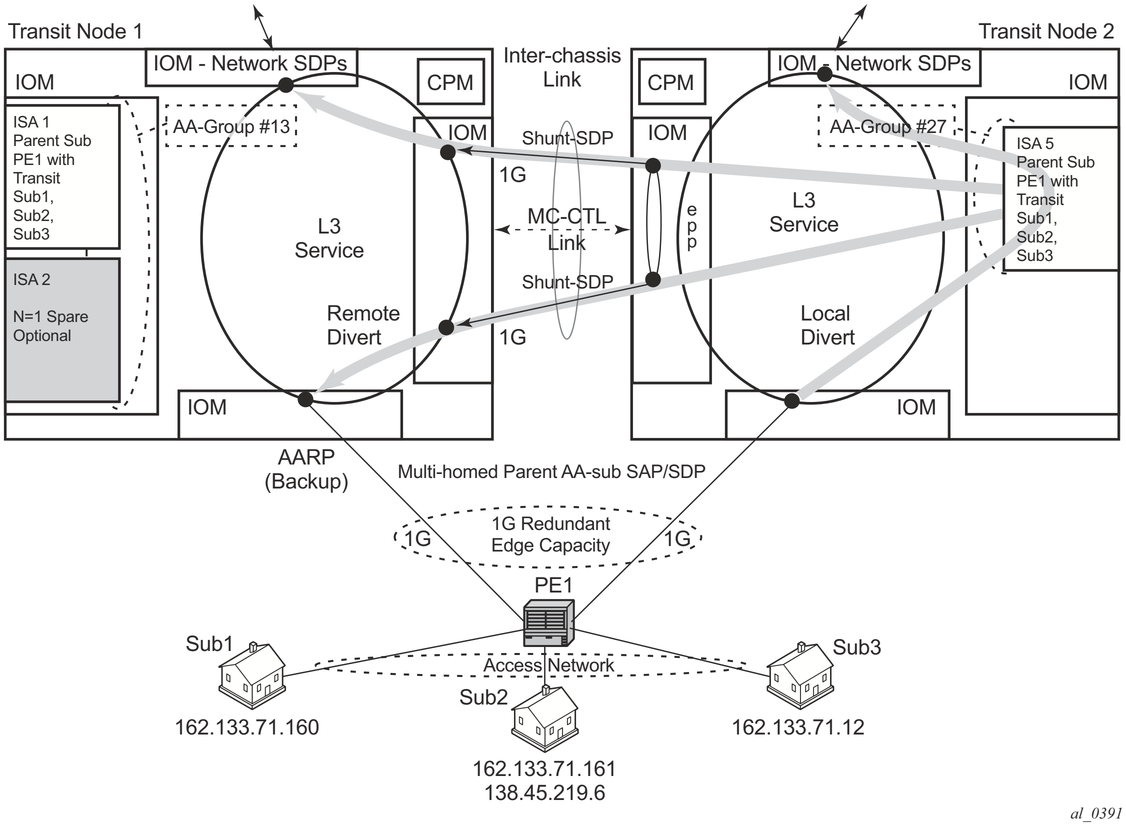


Figure 12: Multi-Chassis Asymmetry Removal Functional Overview

For a Multi-homed parent AA-sub, the parent SAP/SDP that is Active/Inactive per chassis is agreed by the inter-chassis AA Redundancy Protocol (AARP). For single node multi-homed endpoints, the AARP state is determined within a single node, as explained later in the AARP operational states section.

- Divert AA-subs are cost-based load balanced across ISAs in each chassis/AA group (node-local decision).
- Divert AA-sub multi-homed pairing is supported by AA Redundancy Protocol (AARP) over inter-chassis link.
  - → The same AARP ID is assigned to the divert SAP in both nodes.
  - → AARP state in one node is master when all AARP conditions are met.
  - → Packets arriving on node with the master AARP ID divert locally to ISA.

  - → From sub packets on a node with backup AARP ID remote diverted over the subscriber side shunt, appearing to the ISA as if it was a local packet from the AA-sub and returned to the network side interface spoke SDP shunt after ISA processing.o To-sub packets on node with backup AARP ID remote divert over the network side shunt, appearing to the ISA as if it was a local network side divert packet for the AA-sub, then returned to the subscriber side interface spoke-sdp shunt after ISA processing.
  - → All packets are returned to the original node for system egress (sent back over the inter-chassis shunts).
- If ISA N+1 sparing is available in a node, ISA sparing activates before AARP activity switch.
- Supports asymmetry for business SAPs and spoke SDPs, with or without transit AA subs.
- The AARP master-selection-mode is in minimize-switches mode by default, which is non-revertive and does not factor endpoint status. This can be configured per AARP instance using the master-selection-mode. The priority-rebalance configuration will rebalance based on priority once the master failure condition is repaired. The inter-chassis-efficiency mode does priority based rebalance and includes the EP status for cases where an AARP activity switch is preferred to sustained ICL traffic load (when peer nodes are geographically remote).

Failure Modes

Failure modes include the following:

- AARP infrastructure failure including shunts: For AARP to remove asymmetry, the AARP link must be synchronized between peers and all components of the Shunts (iPipe shunts and interface shunts) must be up and operational. If any of those components has failed, each AARP Id operates as standalone and diverts locally. Asymmetry is not removed.
- Failure of one of the interfaces to the dual homed site: routing will move all traffic to the remaining link/node, if this is the master AARP peer node no action is required. For any traffic the backup node, inter-chassis shunting will be used. There is no change to the AARP master/backup state. Traffic will still be processed by the same ISA as before the failure.
- Network reachability fails to master AARP node: AARP node loses reachability on the network side. This does not trigger an AARP activity switch, the shunt is used to move traffic from the backup node to the master node for the duration of the reachability issue. Routing should take care of traffic reconvergence. However, if the peer AARP is also not reachable, both nodes go on standalone mode and there is no asymmetry removal.
- Master AA ISA failure: AARP activity will flip for all the master AARP instances linked to this local ISA if there is no local spare available. Any traffic arriving on the node with the failed ISA will use the shunt to reach the master ISA.

AARP Peered Node/Instance Configuration

The multi-homed diverted AA-sub in each peer node must be configured with the following parameters set in each node of the peer pair as follows:

- Service ID — Node specific
- Interface — Node specific
- SAP or spoke/SDP ID — Node specific
- AA-group ID — Node specific
- App-profile name — Content must be same in both peers to not affect behavior, recommend using same name and content
- Transit policy ID — Same in both (only applies if transits are used)
- AARP ID — Same in both
- shunt-sdp *sdp-id:vc-id* — Node specific but must properly cross-connect the local AA-sub service with the peer Ipipe/service shunt interface in order to operate properly for asymmetry removal for remote divert traffic. Peer AARPs will stay in standalone mode until cross-connect is configured properly.
- Master-selection mode — same in both.
- Other ISA-AA group configuration — Same in both, including fail-to, divert FC, etc.
- IOM traffic classification into a FC — Same in both (can affect AA divert since this is conditioned by the FC). This includes sub side, network side and shunt IOMs.

AARP operation has the following required dependencies:

- For multi-chassis, shunt links are configured and operationally Up.
- For multi-chassis, peer communications established.
- Dual-homed SAP/spoke configured.
- app-profile configured against SAP/spoke with divert (making the sub an aa-sub). This endpoint is called the primary endpoint if more than one endpoint is configured for an AARP instance.
- All endpoints within an AARP instance must be of the same type (SAP or spoke).
- All endpoints with an AARP instance must be within the same service.

Multi-Chassis Control Link (MC-CTL)

A multi-chassis control link is automatically established between peer AARP instances to exchange configuration and status information. Information exchanged includes configured service, protecting sap/spoke, redundant-interface name, shunt-sdp, app-profile, priority and operational states.

AARP requires configuration of the peer IPv4 system address in order to establish a session between the two node’s system IPv4 addresses.

Multi-Chassis Datapath Shunts

When traffic needs to be remotely diverted it flows over shunts that are provisioned as *sdp-id:vc-id* between the dual-homed aa-sub local service an a remote vc-switching Ipipe.

Subscriber to Network Direction

The traffic is either handled locally (diverted to a local ISA when the AARP state is Master) or at the peer 7750 SR (redirect over the shunt Ipipe when the local AARP state is Backup or Remote). When traffic arrives on the subscriber side spoke SDP of the shunt-Ipipe, the system uses the AARP ID of the Ipipe to associate with an app-profile, hence triggering Ipipe divert. It is diverted to the same ISA used to service the dual-homed SAP/spoke SDP. The ISA then treats this traffic the same as if it was received locally on the dual-homed SAP/spoke SDP context. After ISA processing, the traffic returns on the network side of the Ipipe to the peer. When the traffic returns to the original 7750 SR, the shunt Ipipe terminates into the routed service and it makes a new routing decision.

Network to Subscriber Direction

The traffic is either handled locally (diverted to a local ISA when the AARP state is Master) or at the peer 7750 SR (remote divert over the shunt Ipipe when the local AARP state is Backup or Remote). When traffic arrives on the shunt Ipipe from the peer with an AARP ID and associated app-profile, it is diverted through AA on the way to the subscriber-side spoke SDP. After AA processing, the traffic returns on the subscriber side of the Ipipe to the peer. When the traffic returns to the original 7750 SR, the shunt Ipipe terminates into the routed service and it makes a new routing decision to go out the dual-homed SAP/spoke SDP.

AARP Operational States

In single node operation, there are 2 operational states, Master or Standalone. A single node AARP instance is Master when all these conditions are met, otherwise AARP is in the standalone state with is no asymmetry removal occurring:

- Dual-homed (primary) and dual-homed-secondary endpoints are configured
- Divert Capability is Up
- App Profile is diverting
- AA-Sub is configured

With multi-chassis operation there are 4 operational states for an AARP instance: master, backup, remote and standalone.

- Master — In multi-chassis operation, an AARP instance can only become operationally Master when the inter-chassis link datapath is operational and the control path is or was up, the received peer node status indicating the peer's AARP instance and similar dependencies is or was up, and the AARP priority is higher than the peer. When the priority is equal then higher system interface IP address is used as a tiebreak.

  The Master state will be immediately switched to Remote for AARP related failures that result in the instance being not ready. ICL datapath shunt SDP failures will cause the peer AARP go standalone. A shunt/Ipipe SDP failure is determined by the failure detection protocol used (BFD on routes, keep-alive on SDPs, LDP/RSVP, etc.).

  When a SAP/spoke SDP with an AARP instance is shutdown nothing changes for AARP, as packets can still use the AARP interface. When the SAP/spoke SDP is deleted, AARP will be disassociated the from the spoke SDP/SAP before deleting. The AARP instance will still exist after deleting the sap/spoke but without an association to an aa-sub, the AARP state will to go standalone.
- Backup — Backup is the AARP state when all required conditions of the AARP instance are met except the master/backup priority evaluation.
- Remote — When an AARP instance is operating with remote divert set for the protecting SAP/spoke aa-sub. The peer AARP instance is the Master, there is no Backup as the local system is not ready. This state is entered as a result of a failure in a local resource on the AARP instance, which triggers the divert traffic to the remote peer, such as a ISA failure without ISA backup). AA-sub traffic is diverted over shunts to the peer.
- Standalone — AARP is not operational between the multi-chassis pair, with AA operating with local AA divert to the ISAs within that node. There is no Master or Backup. This is the starting initial state for the AARP instance, or as a result of a failure in a dependant ICL resource (MC-CTL communication link or shunt down).

An AARP instance activity switch is when one node moves from Master to remote or backup mode, with the peer node becoming Master. This can occur on a per-instance basis using the re-evaluate tool, or for all instances on an ISA that fails. On an AARP activity switch, AA divert changes from local to remote (or vice versa) such that any given packet will not been seen by both nodes, resulting in no missed packet counts or double counts against the aa-sub.

AARP activity is non-revertive, in order to maximize the ID accuracy of flows. When an AARP instance toggles activity, packets are diverted to the newly active divert ISA and are processed as new flows, which for mid-session flows will often result in “unknown” traffic ID until those flows terminate. When the condition that triggered the AARP activity switch is resolved and the instance remains in backup state, in order to not cause an additional application ID impacting event. This is consistent with AA N+1 ISA activity switches.

Because AA ISA availability is a criteria for AARP switches, any ISA failure or shutdown will move all AARP instance activity to ISAs in the master peer nodes, such as during software upgrades of ISAs. Depending on the nature of the failure or sequence of an upgrade procedure, all AA traffic may be processed by ISAs in one of the peers with no traffic being processed by ISAs on the other node.

If it is desired to rebalance the ISA load between the peer nodes, there is a **tools perform application-assurance aarp** *aarp-id* **force-evaluate** command will re-run AARP activity evaluation on a per-ISA basis to determine Master/Backup AARP based on configured priority.

Table 4 shows the interaction and dependencies between AARP states between a local node and its peer:

Table 4: Interaction and Dependencies Between AARP States

Local AARP Operation State	Peer AARP Operational State	Description
Master	Backup	• Inter-Chassis Link (ICL) Communication established between AARP peers. • AARP dependent resources are up (to-sub/from-sub shunt, aarp control link, dual-homed SAP/spoke SDP). • AARP instances have negotiated initial state assignment using configured priority/system IP address. • AA service is available for the dual-homed SAP/spoke subscriber. • All to-sub/from-sub traffic specific to the dual-homed SAP/spoke SDP will be serviced on the local node. • Peer node is available to takeover in the event of a AA service failure on the local node. • Asymmetry is removed for the dual-homed SAP/spoke subscribers, serviced by AA on the local (master) node.
Master	Remote	Same as Master/Backup except: • AA service is available on the local node. AA service is unavailable on the peer node.
Standalone	Standalone	Initial state of the AARP instances upon creation or a result of a failure in any of the AARP dependent resources. • All to-sub/from-sub traffic for the dual-homed sap/spoke will be serviced on each node independently. • aarp instance operational state is outOfService on both sides. • Asymmetry is not removed for the dual-homed SAP/spoke subscribers (traffic ID is not optimal).

ISA Overload Detection

Capacity cost resource counting does not have a hard per-ISA limit, since the cost values are decoupled from actual ISA resources. However, the value of the total summed cost per-ISA can be reported, and a threshold value can be set which will raise an event when exceeded.

ISA capacity overload detection and events are supported within the system resource monitoring / logging capabilities if the traffic and resource load crosses the following high and low load thresholds on a per-ISA basis:

- ISA capacity cost
- Flow table consumption (number of allocated flows)
- Flow setup rate
- Traffic volume
- Host IOM egress weighted average shared buffer pool use (within the egress QoS configuration for each group). These thresholds are also used for overload cut-through processing

While an app-profile is assigned to AA subscribers, the capacity-cost for that app-profile can be modified. The system makes updates in terms of the load balancing summary, but this does not trigger a re-balance.

In the absence of user configuration, the App-profile default capacity cost is 1. The range for capacity cost is 1 — 65535 (for example, for bandwidth based balancing the value 100 could represent 100kbps). Note that 0 is an invalid value.

If the re-balancing of AA subscribers is required (for instance after the addition of new ISAs), there is a **tools** command to rebalance AA subscribers between ISAs within a group. Rebalance affects which AA subs divert to which ISAs based on capacity cost. Transit subs cannot be rebalanced independent of the parent (they move with the parent divert), and DSM subs cannot be load balanced as all subs on an ISA-AA are from the associated ISA-BB pair. The system attempts to move aa-subs from the most full ISA to the least full ISA based on the load balancing mode. If the load becomes balanced or an aa-sub move fails due to ISA resources or divert IOM service queuing resources, the load balancing terminates.

Alternatively, load balancing can be manually accomplished by the AA subscriber being removed and re-added. This will trigger a load balancing decision based on capacity-cost. For ESM, SAP, and spoke-sdp subscriber types, this can be accomplished by removing and re-applying the AA subscriber's app-profile. In the case of ESM AA subscribers, shutting down and re-enabling either sub-sla-mgmt or the host(s) will have the same effect. Dynamic ESM AA subscribers will re-balance naturally over time as subscribers come and go from the network.

For transit AA subscriber deployments, the parent divert SAPs are load-balanced based on AA capacity cost from the app-profile configured against the SAP/SDP. The parent capacity cost should be configured to represent the maximum expected cost when all transit subs are present.

All traffic not matching a configured transit subscribers is dealt with as a member of the parent SAP and according to its app-profile.

AA Packet Processing

There are four key elements of Application Assurance packet processing (Figure 13):

1. Divert: Selection of traffic to be diverted to the AA ISA.
2. Identification of the traffic on a per flow (session) basis.
3. Reporting of the traffic volume and performance.
4. Policy treatment of the identified traffic.

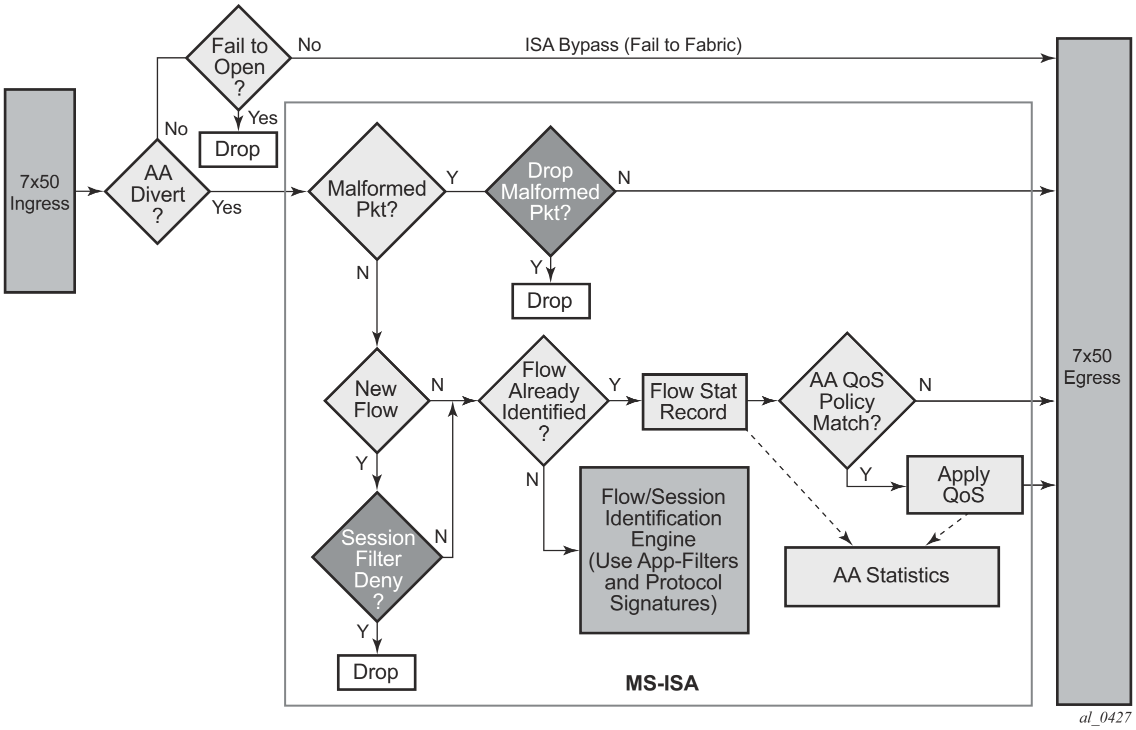


Figure 13: Application Assurance High Level Functional Components

Divert of Traffic and Subscribers

Any traffic can be diverted for application-aware processing. Application Assurance is enabled through the assignment of an application profile as part of either an enhanced subscriber management or static configuration. This process enables the AA functionality for all traffic of interest to and from a given subscriber/SAP/spoke SDP. Which traffic is deemed of interest, is configured through an AA ISA group-specific configuration of forwarding classes (FCs) to be diverted to AA and enabled on a per subscriber/SAP/spoke SDP using application profiles.

Figure 14 shows the general mechanism for filtering traffic of interest and diverting this traffic to the appropriate AA ISA module residing on an IOM (referred as the host IOM). This traffic management divert method applies to both bridged and routed configurations.

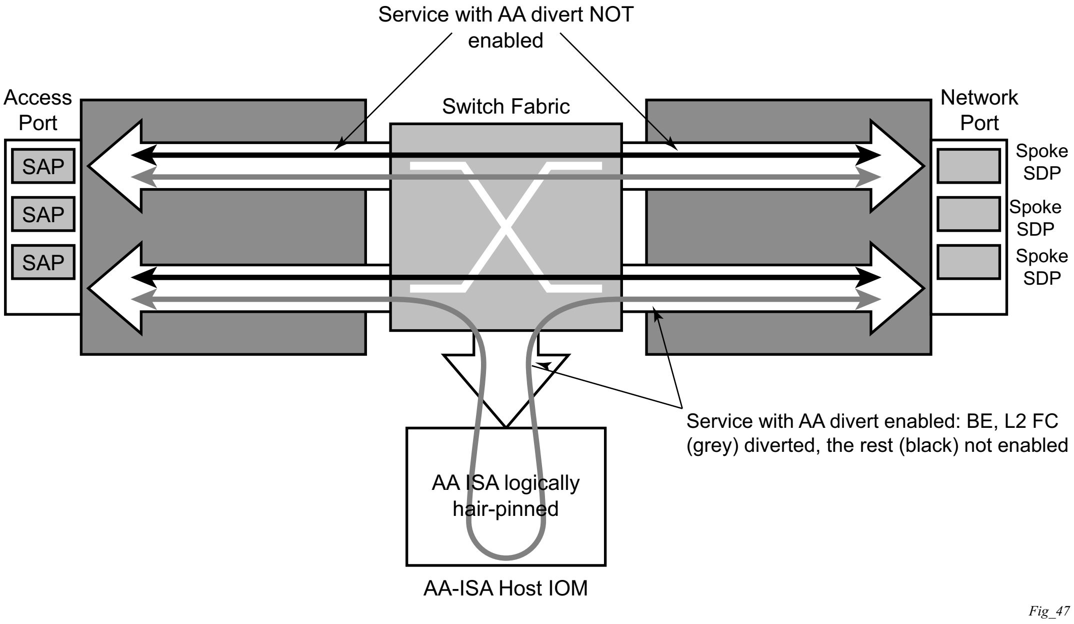

Service with AA divert NOT enabled
Access Port
SAP
SAP
SAP
Switch Fabric
Network Port
Spoke SDP
Spoke SDP
Spoke SDP
Service with AA divert enabled: BE, L2 FC (grey) diverted, the rest (black) not enabled
AA ISA logically hair-pinned
AA-ISA Host IOM
Fig_47

Figure 14: Application Assurance Ingress Datapath

For a SAP, subscribers with application profiles enabling AA, the traffic is diverted to the active AA ISA using ingress QoS policy filters, identifying forwarding and sub-forwarding classes that could be diverted to the Application Assurance. Only single point (SAP, ESM, or DSM subscriber, spoke SDP) configuration is required to achieve divert for both traffic originated by and destined to a given AA subscriber. Diversion (divert) to the AA ISA is conditional based on the AA ISA status (enabled, failed, bypassed, etc.).

Unless the AA subscriber's application profile is configured as "divert" using Application Profiles and the FC is selected to be diverted as well, the normal ingress forwarding occurs. Traffic that is filtered for divert to AA ISAs is placed in the appropriate location for that system's AA ISA destination.

Users can leverage the extensive QoS capabilities of the router when deciding what IP traffic is diverted to the Application Assurance system for inspection. Through AA ISA group-wide configuration, at least one or more QoS forwarding classes with the "divert" option can be identified. The forwarding classes can be used for any AA subscriber traffic the service provider wants to inspect with Application Assurance.

Services and AA Subscribers

The 7750 SR/7450 ESS AA ISA provides the Layer 3-7 packet processing used by the Application Assurance feature set. Application Assurance is applied to IPv4 and IPv6 traffic on a per AA subscriber basis, where an AA subscriber is one of:

- ESM subscriber
- Distributed sub management (DSM) subscriber
- SAP/spoke
- Transit

Non-IPv4 and IPv6 traffic is not diverted to AA and forwarded as if AA was not configured where an AA subscriber may be contained in the following services:

- IES
- VPLS
- VLL — Epipe and Ipipe
- VPRN

Application Assurance is supported with:

- Bridged CO
- Routed CO
- Multi-homed COs
- Layer 2/Layer 3 VPN service access points and spoke SDPs

The AA ISA feature set uses existing 7750 SR/7450 ESS QoS capabilities and further enhances them to provide application-aware traffic reporting and management on per individual AA subscriber, AA subscriber-type or group. A few examples of per-application capabilities within the above AA subscriber contexts include:

- Per AA subscriber, application traffic monitoring and reporting.
- Per application bandwidth shaping/policing/prioritization.
- Throttling of flow establishment rate.
- Limiting the number of active flows per application (such as BitTorrent, video or teleconference sessions, etc.).
- Application-level classification to provide higher or lower (including drop) level traffic management in the system (for example, IOM QoS) and network.

The following restrictions are noted — Application Assurance is not supported for tunneled transit traffic (GRE or L2TP tunnels using PPP or DHCP based policy) destined for a remote BRAS.

Spoke SDPs

AA on spoke SDP services allows AA divert of the spoke SDP, logically representing a remote service point, typically used where the remote node does not support AA. A given SAP/spoke can be assigned and app-profile, and when this app-profile is enabled for **divert** all packets to and from that SAP/spoke will be diverted to an AA ISA (for forwarding classes that are configured as divert eligible).

Table 5 shows spoke SDP divert capabilities.

Table 5: Spoke SDP Divert

	Connected to Service				
Access Node Service (spoke SDP type)	**Epipe**	**VPLS**	**IES**	**VPRN**	**Ipipe**
Epipe (Ethernet spoke)	Y	Y	Y	Y	Y
Ipipe (IP spoke)	N/A	N/A	Y	Y	Y
VPLS (Ethernet spoke)	N/A	Y	Y	Y	N

The following restriction is noted.

- Spoke SDP divert is only supported on services to/from IOM3-XP or newer IOMs or IMMs.

Transit AA Subs

A transit AA sub is an ISA local AA sub contained within a parent AA sub. There are two types of transit AA subs:

- Transit IP AA-subs: defined by Transit IP Policy as one or more /32 IP addresses per sub
- Transit Prefix AA-subs: defined by Transit Prefix Policy as one or more prefix IP addresses, used in business VPNs

A transit AA-sub incorporates the following attributes:

- Name
- IP address (one or more hosts)
- App-profile (note that the divert/no divert and capacity cost setting of the app-profile does not affect transit AA-subs since divert occurs only against the parent SAP).

When a SAP or spoke-SDP diverted to AA is configured with transit subs, that SAP or Spoke-SDP is referred to as the parent AA subscriber. Transit AA subs are supported on the following parent Layer 3 SAPs or spoke SDPs that support AA divert:

Table 6: Transit AA Subs Support

Transit Subscriber Type	Epipe	VPLS	IES	VPRN	Ipipe
Transit IP	N/A	N/A	Y	Y	N/A
Transit Prefix	Y	Y	Y	Y	Y

The transit AA-subs within a given parent AA sub can be displayed using the **show aa group transit policy** or **transit-prefix policy** command.

For transit IP subscribers all packets are accounted for once in the ISA records. Therefore, transit IP AA sub counts do not count against the parent SAP in reporting. For transit prefix AA subscriber deployments using the remote-site command, traffic for the remote transit subs are processed and counted for both the local parent and the remote transit subscriber.

Transit AA-Sub App-Profile

The app-profile assigned to the aa-sub-id affects both stats and control of the policy. App-profiles are assigned to the transit AA-subs either explicitly when the transit-aa-sub is created, or by default (when not specified) according to a default app-profile configured in a transit-ip-policy or transit-prefix-policy. This allows transit AA subs to be treated with a different default app-profile than the app-profile (default or specified) set against the parent aa sub. The number of aa-sub stats used per ISA is proportional to the number of AA subscribers including transit subscribers subs are added.

ASO policy override is supported for static transit subs.

Transit IP Policy and Transit Prefix Policy

A transit policy is associated with the parent (divert) SAP/SDP to define how transit AA subs are created within that parent. The transit policy must be defined in the config>app-assure>group context before it can be assigned to a parent. Transit IP subs can be created by the following methods:

- Static — CLI/SNMP configuration of a transit aa-sub is done within the transit-ip-policy
- Dynamic - DHCP authentication
- Dynamic - RADIUS accounting to PCRF

Transit prefix subs are created by static CLI/SNMP configuration of a transit aa-sub within the transit-prefix-policy. The transit prefix policy follows IP filter conventions for first match and ordering of entries. While for residential /32 transits if there is an IP address conflict between any static prefix transit subs, the latter config will be blocked, for business transit subs multiple overlapping address entries are allowed to enable longest match within subnets. IP addresses for a VPN site as an AA-sub are configured with the transit prefix policy. There are two options:

- aa-sub-ip is used when the site is on the same side of the system as the parent SAP
- aa-network-ip is used when the site is on the same opposite of the system as the parent SAP

A given transit prefix subscriber may only have either aa-sub-ip entries or aa-network-ip entries but not both.

The IP addresses defined in the transit-ip-policy for a transit sub are full /32 IP addresses. The IP addresses defined in the transit-prefix-policy for a transit sub are any length from /0 to /32.

Multiple IP addresses (from any prefix/pool) can be assigned to a single transit AA sub. IP addresses must be unique within a transit policy, but can be re-used in separate policies (since they have parent specific context).

The transit policy contains the default app-profile for the transit sub if a transit policy is created but app-profile is not specified. An app-profile can be later explicitly assigned to the transit sub after the sub is created (using RADIUS COA, DHCP or static).

For dynamic transit ip subs, a sub-ident-policy (also used by ESM to associate sub ID policies to a SAP) can now also be associated with the AA-sub parent by defining the sub-ident policy in the transit IP policy. This determines how sub identifying strings are derived from DHCP option 82 fields. The policy also contains app-profile-map which maps the strings to the defined app-profiles. Transit subs do not use the sla-profile or sub-profile aspects of the sub-ident-map.

In the case of multi-homed transit subs, the transit-ip-policy must be the same on both nodes of the multi-homed parent link to ensure consistency of sub context and policy.

There are no configurable limit hosts per sub per sub (this is similar to lease-populate which limits the number of dynamic hosts per SAP), or, limit the number of transit subs per transit ip policy (parent). This is a function for the PE doing subscriber management.

If transit sub resource limits are exceeded (hosts per sub, or subs per ISA) the transit sub creation is blocked (for both static and dynamic models).

There is a per-ISA group/partition show list of AA-subs in a transit-ip-policy which includes a parent field for transit subs (static versus dynamic identified).

Persistent AA statistics is supported dynamic transit AA subs, ensuring that accounting usage information is not lost when the sub disconnects prior to reporting interval end.

static-remote-aa-sub Command

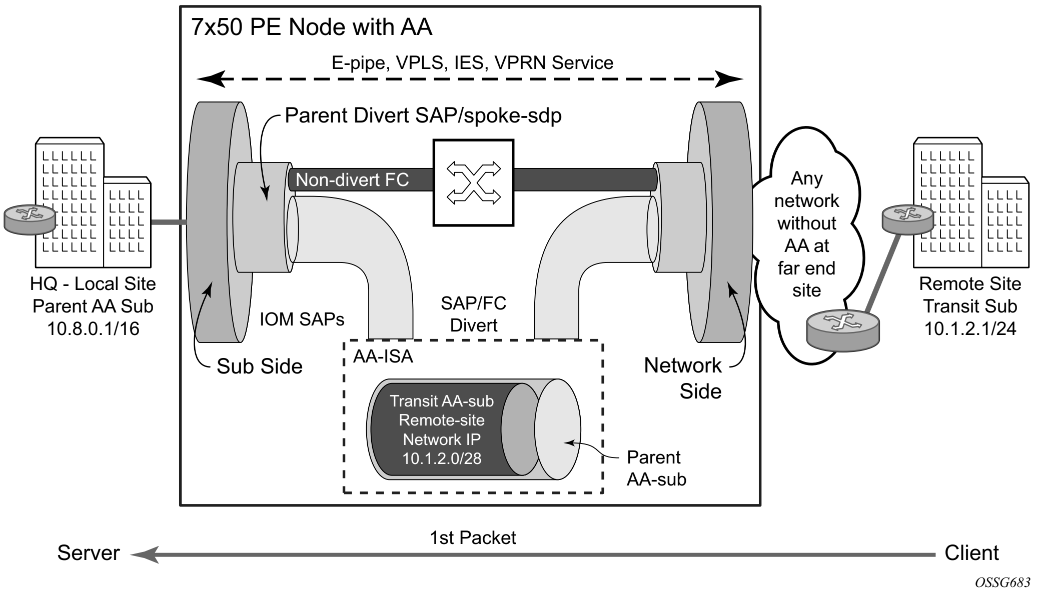


Figure 15: static-remote-aa-sub Usage Topology

This command enables unique ISA treatment of transit prefix subscribers configured on the opposite (remote) side of the system from the parent SAP/spoke SDP. Provisioning a transit sub as remote-aa-sub within a transit prefix policy enables the ISA to treat any network IP-based transit subs in the following ways:

- Treat packets for the parent aa-sub independent of whether transits are also configured (stats and policers for parent work as usual).
- Subsequently treat the same packet as a transit-sub packet when matching to a configured transit sub (stats, policers).
- Allows natural direction of the packet for both the parent aa-sub and the transit-aa-sub, as shown in Figure 15, where a packet from a remote client to a local server will be seen as to-sub for the parent, and from-sub for the transit sub that is logically at the far end site.
- Correct directionality of packet ID for all aa subs allows proper operation of app-filter flow-setup-direction

Static Transit AA-Sub Provisionings

Static (through CLI/SNMP) provisioning of transit AA-subs is supported. A profile policy override to set policy characteristics by ASO (as opposed to within an app-profile) is supported only for statically configured transit AA subs.

If there is an IP address conflict between a static and dynamic transit sub, the static takes precedence (per ESM). If the static is configured first, the dynamic transit sub will be rejected. If the dynamic is created first, a warning is provided before removing the dynamic transit sub and notifying the sub-manager by COA failure.

DHCP Transit IP AA-Subs at DHCP Relay Node

DHCP-based transit sub creation provides a sub ID and lease time for IP addresses correlated to the ESM/subscriber context in the PE.

The 7750 DHCP relay agent creates dynamic DHCP AA-subs when the DHCP ACK is received from the DHCP server, including the sub name, IP address and app-profile from DHCP Option 67 (if present) when the DHCP ACK messages passes through AA node to the downstream subscriber-edge node. If there is no app-profile assigned when the transit aa-sub is created, a default transit aa-sub app-profile is used (configured in the transit-ip-policy assigned against the divert parent aa-sub).

This is compatible with the ESM 7x50 edge as well as third-party BRAS and CMTS.

Dynamic AA-sub stats records are persistent across modem reset/session releases. The end of accounting records are created when transit subs are released.

Multiple IPs per transit AA sub are determined by seeing a common the DHCP Option 82 cct ID.

RADIUS Transit AA-Subs

Transit subs can be dynamically provisioned by RADIUS accounting start messages forwarded by the RADIUS AAA server to a RADIUS sub-manager function at the OSS layer (5780 DSC). This RADIUS sub manager manages dynamic transit AA subs on the appropriate ISA and transit-ip-policy based on the RADIUS accounting information. The interface for the sub manager to configure transit AA subs is RADIUS COA messages, which are acknowledged with a COA success message to the sub manager.

If a dynamic transit sub cannot be created as requested by a COA due to resource constraints or conflicts, the node replies to the sub manager with a COA fail message so that retries will not continue. This message should contain information as to the cause of the rejection. Multiple IPs per sub are allowed when common sub-ID names are seen, but with differing IP hosts.

When a RADIUS update/COA message is seen, it could contain a modified IP address or app-profile for an existing transit sub which is accepted without affecting transit AA subscriber statistics. These transit AA-subs are removed by the sub manager when a RADIUS accounting stop message is received.

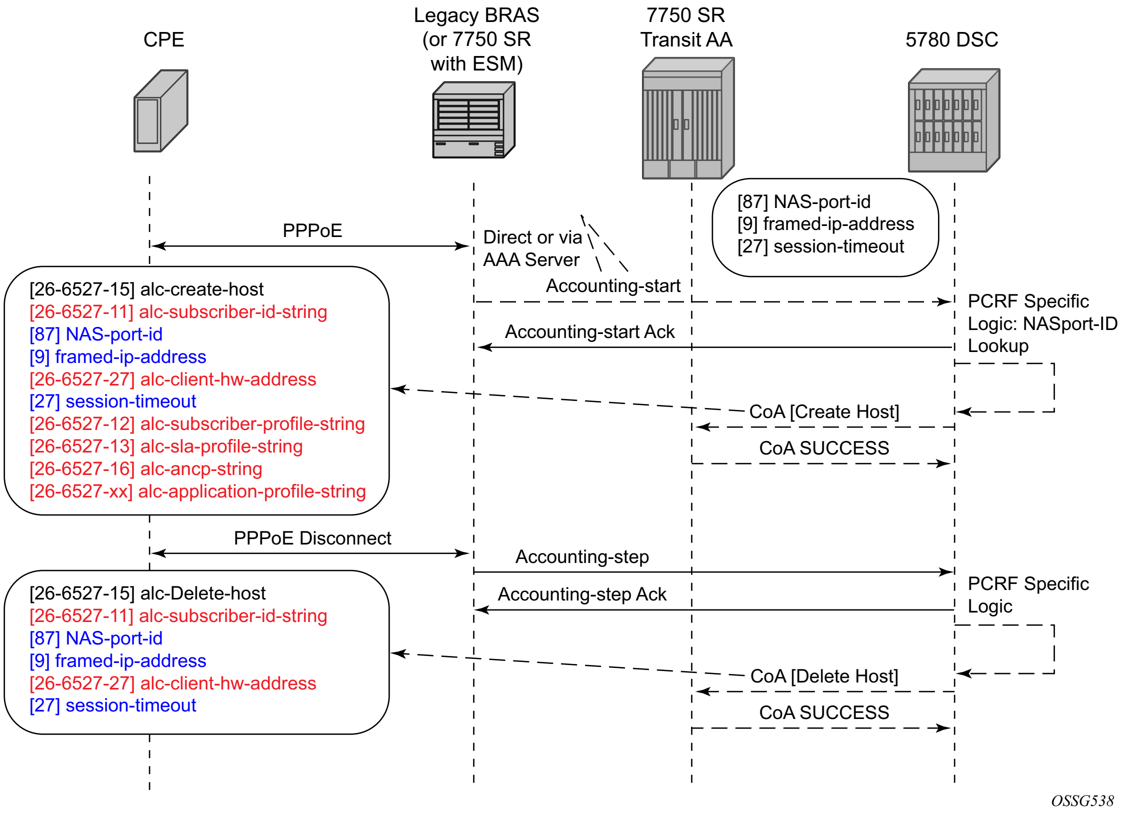

CPE
Legacy BRAS (or 7750 SR with ESM)
7750 SR Transit AA
5780 DSC
[87] NAS-port-id
[9] framed-ip-address
[27] session-timeout
PPPoE
Direct or via AAA Server
Accounting-start
PCRF Specific Logic: NASport-ID Lookup
Accounting-start Ack
[26-6527-15] alc-create-host
[26-6527-11] alc-subscriber-id-string
[87] NAS-port-id
[9] framed-ip-address
[26-6527-27] alc-client-hw-address
[27] session-timeout
[26-6527-12] alc-subscriber-profile-string
[26-6527-13] alc-sla-profile-string
[26-6527-16] alc-ancp-string
[26-6527-xx] alc-application-profile-string
CoA [Create Host]
CoA SUCCESS
PPPoE Disconnect
Accounting-step
PCRF Specific Logic
Accounting-step Ack
[26-6527-15] alc-Delete-host
[26-6527-11] alc-subscriber-id-string
[87] NAS-port-id
[9] framed-ip-address
[26-6527-27] alc-client-hw-address
[27] session-timeout
CoA [Delete Host]
CoA SUCCESS
OSSG538

Figure 16: RADIUS COA Example

The attributes in RADIUS COA that identify the downstream transit AA-subs are:

- Downstream BRAS/ CMTS: NAS-port-ID
- IP address: framed-ip-address
- Subscriber ID: per RADIUS accounting sub-id-string

Seen-IP RADIUS Notification

Seen-IP transit subscriber notification provides RADIUS Accounting Start notification of the IP addresses and location of active subscribers within a parent AA service.This allows a PCRF to dynamically manage RADIUS AA subscriber policy (create, modify, delete) without requiring static network topology mapping of a subscriber edge gateway to the parent transit service.

When detect-seen-IP is enabled within a transit policy, the ISA will detect IP flows on a AA parent subscriber that do not map to an existing transit AA-subscriber. It will then use a simple RADIUS Accounting Start notification from the transit AA node to the PCRF to initiate subscriber creation, providing information on the location of the transit subscriber traffic. This provides notice for subscriber authentication changes such as new subscriber session, or new host IP addresses added to an existing aa-sub, while being independent of the network topology for how the BNG is homed into the transit AA nodes.

The RADIUS Accounting Start message is sent to the RADIUS Server referenced by the specified seen-ip-radius-acct-policy. This RADIUS message contains the following information about the flow:

- subscriber-side IP address
- Parent SAP/Spoke-SDP ID (NAS Port ID)
- IP address of node making the request
- Peer SAP/Spoke-SDP ID (NAS Port ID) - if configured
- Peer IP address of SR making the request - if configured
- AARP ID - if configured

Transit AA-Sub Persistence

Transit AA subs can be persistent within a single node, since, in some cases, there is not a dual-node BNG subscriber redundancy configuration. This allows a single node that has dynamically created transit subs to retain the subscriber state, context, and stats across a node or ISA reboot.

If dynamic transit AA subs are released, renewed or otherwise changed during an outage or reboot of a transit AA node, the sub manager will notify the transit node of these changes.

Prefix transit subs are not affected by persistence since they can only be statically configured.

Policers for Transit AA-Subs

AA-sub per-subscriber policers can provide per SAP policing for the parent SAP, with transit AA-subs each supporting distinct per-sub policers within the parent (packets are only processed once against one aa-sub – the parent or the transit sub). Packets matching transit AA subs and policers will not be included in a parent policer.

There is no policer hierarchy unless system wide policers are referred to by both the parent aa-sub and transit aa-sub. When the remote-site configuration is not used, system policers can be used to police all traffic for a site containing transits, subject to constraints on system policer scale.

When the remote-aa-sub config is used, the parent owns all packets for stats and policing, so any transit sub configuration within the parent does not affect the stats or policers. AA policers are supported on a transit subscriber basis, across all (multiple) IP prefixes per sub.

ISA Host IOM for Transit Subs

The AA divert IOM is not impacted by transit AA subs in the divert parent. The ISA host IOM egress datapath functions to convert the parent SAP into transit AA-subs that are then handled by the ISA consistent with all other AA-sub features. The ISA itself treats all AA-subs equally regardless of whether the AA sub is from ESM, from DSM, from an SAP, or from a transit subscriber in a parent SAP/spoke.

Prefix transit subs can only be created on IOM3-xp as host IOM, or with MS-ISM as host for ISA2. Asymmetry removal requires IOM3-xp or MS-ISM as host and IOM3-xp or newer (IMM) as divert IOM.

AA Subscriber Application Service Definition

Application Profile

Application profiles enable application assurance service for a given ESM or DSM subscriber, Service Access Point or spoke SDP (AA subscriber). Each application profile is unique in the system and defines the AA service that the AA subscriber will receive. An ESM subscriber can be assigned to an application profile which affects every host of the particular subscriber. For SAP or spoke SDP AA subscribers, an application profile can be assigned which affects all traffic originated/destined over that SAP or spoke SDP. By default, ESM and DSM subscribers, SAPs or spoke SDPs are not assigned an application profile.

The following are main properties of application profiles:

- One or more application profiles can be configured in the system.
- Application profiles specify whether or not AA subscriber's traffic is to be diverted to Application Assurance.
- Application profiles are defined by an operator can reference the configured application service options (ASO) characteristics (see Application Service Options (ASOs) on page 78.
- Application profiles must only be assigned once AA resources (AA ISA cards) are configured.
- App-profiles can be assigned a capacity cost used for subscriber load balancing among ISAs within the AA group. (See ISA Load Balancing on page 51.)

ESM and DSM policy includes an application profile string. The string points to an application profile pre-provisioned within the router and is derived by:

- Parsing the DHCP Option 82 sub-option 1 circuit ID payload, vendor specific sub-option 9, or customer-defined option different from option 82, during authentication and the DHCPDISCOVER, as well as re-authentication and the subscriber’s DHCPREQUEST.
- RADIUS using a new VSA. [26-6527-xx] alc-application-profile-string
- DIAMETER using “AA-profile-name” AVP under ADC rule.
- Inherited by defaults in the **sap>sub-sla-mgmt** context, to allow default application profile assignment if no application profile was provided.

- Static configuration.

Mid-session (PPP/DHCP) changes to the application profile string allows:

- Modification of the application profile a subscriber is mapped to and pushes the change into the network as opposed to waiting for the subscriber to re-authenticate to the network.
- Change to the subscribers application profile inline, without a need for the subscriber to re-authenticate to RADIUS or perform any DHCP message exchange (renew or discover) to modify their IP information.

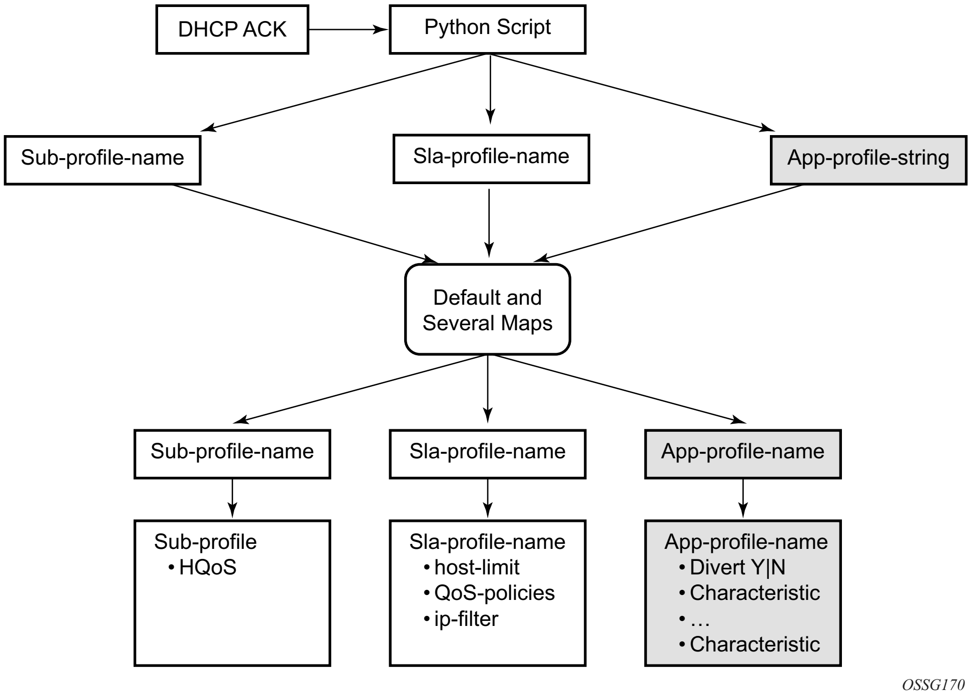


Figure 17: Determining the Subscriber Profile, SLA Profile and Application Profile of a Host

Application Profile Map

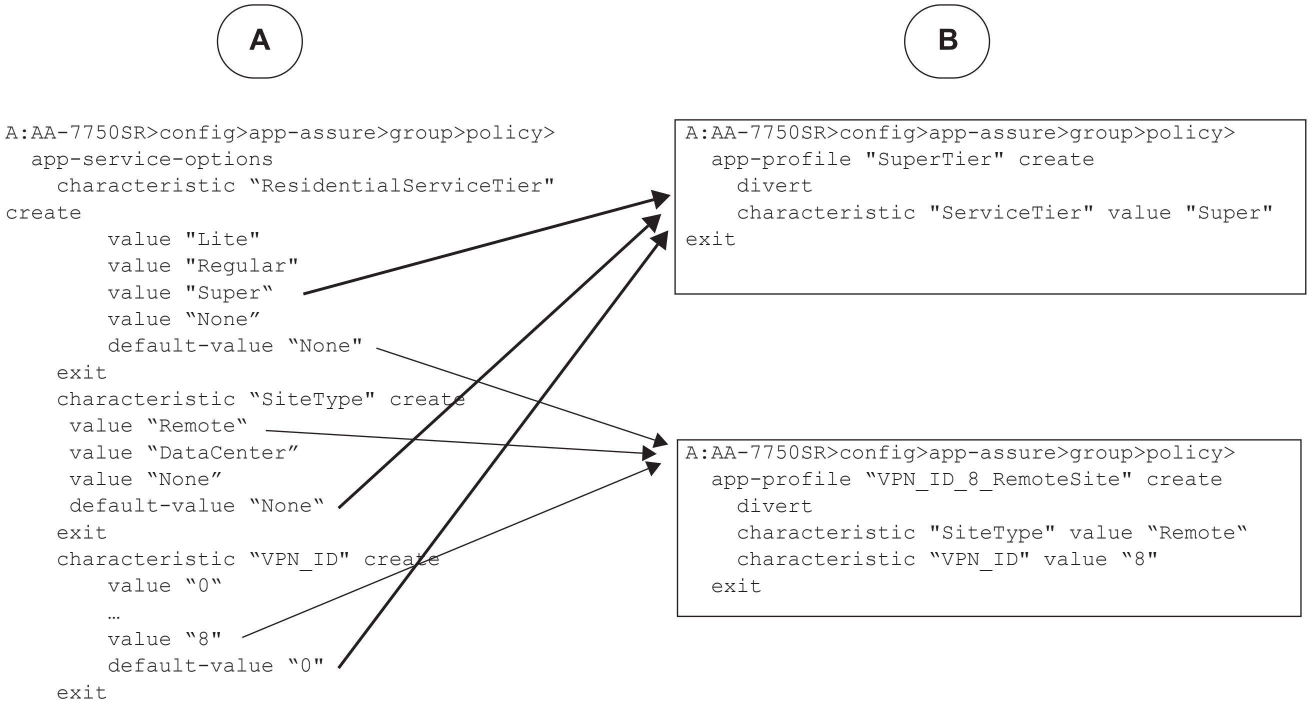
Application Assurance adds new map (app-profile-map) application profile command to associate an *app-profile-string* from dynamic subscriber management to a specific application profile using its app-profile-name that has been pre-provisioned. The application profile map is configured in the **config>subscr-mgmt>sub-ident-pol** context.

The pre-defined subscriber identification policy has to be assigned to a SAP, which determines the sub-id, sub, sla, and app-profiles.

Application Service Options (ASOs)

ASOs are used to define service provider and/or customer visible network control (policy) that is common between sets of AA subscribers (for example, upstream/downstream bandwidth for a tier of AA service). ASO definition decouples every AA subscriber from needing subscriber-specific entries in the AQP for standard network services.

As an example, an operator can define an ASO called ServiceTier to define various HSI services (Super, Lite, etc.) (Figure 18-A). The operator can then reference these defined ASOs when creating the App Profiles that are assigned to AA-subscribers (Figure 18-B).



```
A:AA-7750SR>config>app-assure>group>policy>
  app-service-options
    characteristic “ResidentialServiceTier"
create
        value "Lite"
        value "Regular"
        value "Super“
        value “None”
        default-value “None"
    exit
    characteristic “SiteType" create
     value “Remote“
     value “DataCenter”
     value “None”
     default-value “None“
    exit
    characteristic “VPN_ID" create
        value “0“
        …
        value “8"
        default-value “0"
    exit
```

```
A:AA-7750SR>config>app-assure>group>policy>
  app-profile "SuperTier" create
    divert
    characteristic "ServiceTier" value "Super"
exit
```

```
A:AA-7750SR>config>app-assure>group>policy>
  app-profile “VPN_ID_8_RemoteSite" create
    divert
    characteristic "SiteType" value “Remote“
    characteristic “VPN_ID" value “8"
  exit
```

Figure 18: Configuration Example

Then, the defined ASOs are used in the AQP definition to determine the desired treatment / policy (Figure 19).

```
app-qos-policy
    entry 50 create
        description "Limit downstream b/w for Super sub-
scribers"
        match
            traffic-direction network-to-subscriber
            characteristic "ServiceTier" eq "Super"
        exit
        action
            bandwidth-policer "SuperDown"
        exit
        no shutdown
    exit
    entry 110 create
        match
            application-group eq “Tunneling"
            characteristic "SiteType" eq “Remote"
        exit
        action
            remark fc af
        exit
        no shutdown
    exit
```

Figure 19: AQP Definition Example

Alternatively, if ASOs were not used in the previous example, then the operator would have to define a unique AQP entry for every subscriber. Each of these AQPs will have its “match” criteria setup to point to the subscriber ID, while the action for all of these unique AQPs will be the same for the same service (for Tier 1 service, the policer bandwidth will be the same for all Tier 1 AA subscribers) (Figure 20).

```
7750SR>config>aa>group>policy>aqp>
  entry 100 create
    match
      aa-sub eq “ sub_1”
    exit
    action
      bandwidth-policer “superDown”
    exit
    no shutdown
  exit

  entry 101 create
    match
      aa-sub eq “ sub_2”
    exit
    action
      bandwidth-policer “superDown”
    exit
    no shutdown
  exit

  entry 102 create
    match
      aa-sub eq “ sub_3”
    exit
    action
      bandwidth-policer “superDown”
    exit
    no shutdown
```

Figure 20: Single ASO Example

The example in Figure 20, shows how the use of just a single ASO can save the user from having to provision an AQP entry every time a subscriber is created.

Other example uses of ASO entries include:

- Entry per application group that is to be managed, such as VoIP, P2P, HTTP.
- Several entries where specific applications within an application group can individually be managed as service parameters, for example, HTTP content from a specific content provider, or streaming video from network television or games.
- HSI tiers (for example, Gold, Silver, and Bronze for specifying bandwidth levels).
- VPN customer ID.

Application characteristics are defined as specific to the services offered within the operator□ network. The operator defines ASO characteristics and assigns to each ASO one or more values to define service offering to the customers.

The following are the main elements of an ASO:

- A unique name is applied to each characteristic.
- The name is unique to the group-partition-policy, but the expectation is that characteristics will be consistent network wide.
- Operator-defined values (variables) are defined for each characteristic and are unique to each characteristic. A default value must be specified from the set of the values configured.

The following lists how ASO characteristics are used:

- Application service options are used as input to application profiles.
- AQP rule sets also use the ASO characteristics to influence how specific traffic is inspected and policies applied.
- Multiple ASO characteristic values are allowed in a single rule.

Syntax checking is performed when defining application profiles and AQPs that include application characteristics. This ensures:

- The characteristic is correctly identified.
- In an app-profile and app-qos-policy when specifying a characteristic, the value must be specified. The “default-value” applies if a characteristic is not specified within an app-profile.

ASO Overrides

This feature enables individual attributes/values to be set against an aa-sub complementary to using app-profiles. The aa-sub types supported that can have ASO overrides by CLI/SNMP are provisioned business AA SAPs and spoke SDPs, and statically-provisioned transit AA subs. Dynamic AA subscribers (ESM, DSM, and transit subs) can have ASO overrides applied by RADIUS override VSAs.

Application profile assignment is still used to obtain the following information:

- The application-assurance group (and partition) to which the AA-sub is being assigned to
- Whether or not the traffic should be diverted
- Capacity-cost (for load balancing to a multi-isa group)

The information configured in the app-profile is also used, but the following can be overridden:

- ASO characteristics and values (these are from the policy defined in the group and partition)

The overrides are specific to a single aa-sub. An ASO override does affect any other aa-sub or the app-profile config itself.

Typically the ASO characteristics in the app-profile would not be specified, thus leaving all characteristics at their default values. This is not mandatory though and the app-profile could specify any ASO characteristic and non-default value.

The AA app-qos-policy has entries that can refer to ASO characteristics (attributes) and values in their match criteria. In the absence of any individual attribute/value override, an aa-sub will continue to work as before - using the ASO characteristics/values defined inside the app-profile assigned to them. With overrides, the aa-sub attributes used in app-qos-policy lookups are the combination of the following:

- The characteristics/values from the app-profile,
- Any specific characteristics and values overridden for that aa-sub.

Show command output display the combined set of attributes that apply to the aa-sub.

The **override** commands can only be used if there is already an app-profile assigned to the aa-sub, otherwise, the overrides are rejected.

The app-profile attribute override is assigned to a specific aa-sub (SAP, spoke SDP) within the AA Group:partition with where the subscriber exists. While subscriber names are unique, the Group:partition policy context where apps, app-profiles and ASO characteristics are defined is relevant to the override context. Override for ESM subscribers can be triggered via DIAMETER or RADIUS.

AA-Sub Scale Mode

An AA VPN policy is generally administered using a per-site (aa-subscriber) policy attribute assignment (ASO override), as opposed to a service profile based model commonly used for residential services.   Due to this, the number of attributes and values of ASOs that can be needed in an AA VPN service will be much larger than ASO scale needed for residential uses.

On the other hand, the number of AA subscribers needed per node and per ISA is much smaller for VPN services, and the size of each in bandwidth is generally much larger than residential.

In conjunction with App-profile ASO override, a new capability is added to place an AA-group into a mode optimized for VPN scale requirements:

```
config>aa>aa-group>aa-sub-scale {residential|vpn} (residential is default)
```

Application Identification

This section discusses the following topics:

- Application Assurance Identification Components on page 86
- Protocol Signatures on page 88
- Custom Protocols on page 89
- Protocol Shutdown on page 90
- Supported Protocol Signatures on page 90
- Application Groups on page 90
- Applications on page 91
- Application Filters on page 92
- HTTP on page 93
- Charging Groups on page 91
- AA IP Prefix Lists on page 94

Application identification means there is sufficient flow information to provide the network operator with a view to the underlying nature and value of the content. Application ID does not include:

- Anti-virus signatures per IPS/UTM.
- Content inspection (e-mail, text, picture, or video images). The payload data content of flows is typically not examined as part of the application identification.

Application Assurance can identify and measure non-encrypted IP traffic flows using any available information from Layer 2-Layer 7, and encrypted IP traffic flows using heuristic techniques.

Application Assurance attempts to positively identify the protocols and applications for flows based on a pattern signature observation of the setup and initial packets in a flow. The system correlates control and data flows belonging to the same application. In parallel, statistical and behavioral techniques are also used to identify the application. Until identified, the flow will not have a known application and will be treated according to the default policies (AQP policies defined using all or any ASO characteristics, subscriber Id and traffic direction as match criteria) for traffic for that AA subscriber, app-profile and direction (packets will be forwarded unless an action is configured otherwise). If the identification beyond OSI Layer 2is not successful, the flow will be flagged as an unknown protocol type, (for example unknown_tcp or unknown_udp). The unknown traffic is handled as part of all application statistics and policy, including generation of stats on the volume of unknown traffic.

Application Assurance allows operators to optionally define port-based applications for “trusted” TCP or UDP ports. Operators must explicitly identify a TCP/UDP port(s) in an application filter used for “trusted” port application definition and specify whether a protocol signature-based application identification is to be performed on a flow or not. Two options are available:

- If no protocol signature processing is required (expected to be used only when (A) AQP policy must be performed from the first packet seen, (B) the protocol signature processing requires more than 1 packet to positively identify a protocol/application, and (C) no other application traffic runs over a given TCP/UDP port), the first packet seen by AA ISA for a given flow on that TCP/UDP port will allow application identification. The traffic for a given flow will be identified as “trusted tcp/trusted_udp” protocols.
- If protocol signature verification of an application is required (expected to be used only when (a) AQP policy must be performed from the first packet seen, (b) the protocol signature processing requires more than 1 packet to positively identify a protocol/application, but (c) other application traffic may run over a given TCP/UDP port, for example TCP port 80), the first packet seen will identify the application but protocol signature-based analysis continues. Once the identification completes, the application is re-evaluated against the remaining application filters allowing detection and policy control of unexpected applications on a “trusted” port.

At Application Assurance system startup or after an AA ISA activity switch, all open flows are marked with the “existing” protocol signature and have a policy applied according to an application based on the “existing” protocol until they end or the identification of an in-progress flow is possible. Statistics are generated.

From the first packet of a flow, a default per AA subscriber AQP policy is applied to every packet. Once an application is identified, subsequent packets for a flow will have AA subscriber and application-specific AQP applied. The AA-generated statistics for the flow with AA subscriber and application context are collected based on the final determination of the flow's application. A subset of the applications may be monitored on an ongoing basis to further refine the identification of applications carried with the traffic flow and to identify applications using an external application wrapper to evade detection.

Application Assurance Identification Components

Figure 21 shows the relationship between the Application Assurance system components used to identify applications and configure Application Assurance related capabilities. Each ID-related component is defined as follows:

- Protocol signatures
- Application filters
- Applications
- Application groups
- Charging groups

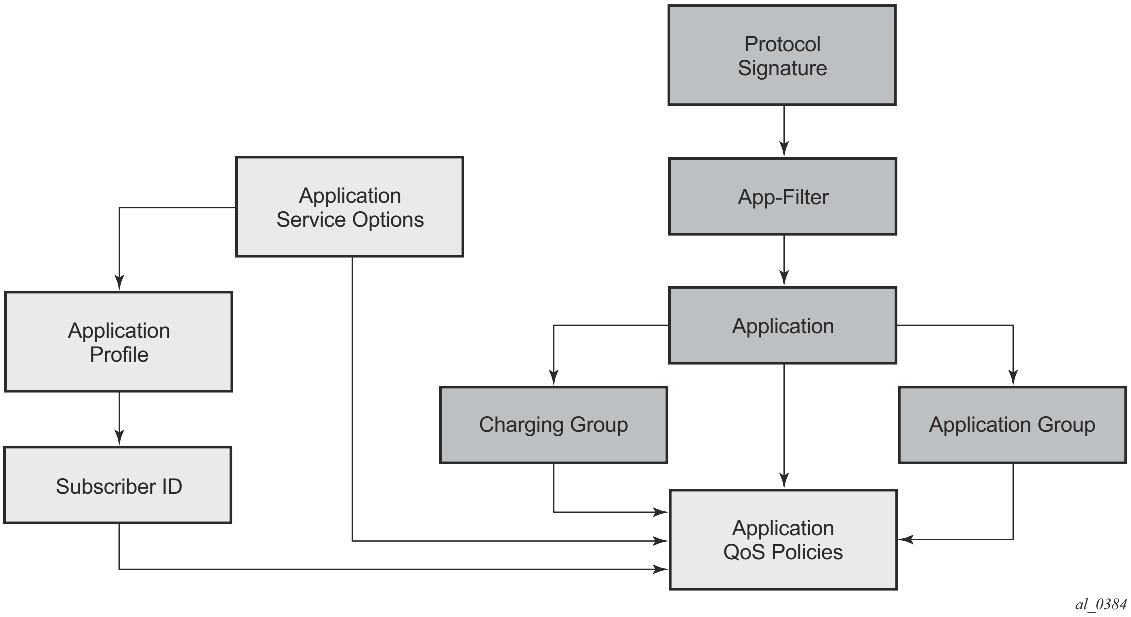


Figure 21: Policy Structure

Table 7 provides an overview of how those various components used in Application Assurance to recognize types of flows/sessions.

Table 7: AA Flows and Sessions

Term	Definition	Examples
Protocol Signature	Alcatel-Lucent's proprietary component of AA flow identification provided as part of AA S/W load to identify protocols used by clients. Where a protocol is defined as an agreed upon format for transmitting data between two devices.	Tftp, iMap, msn-msgr, RTP, emule, http_video, bittorrent, SIP **Note**: Alcatel-Lucent's protocol signatures do not rely on IP port numbers to identify a TCP/UDP port based protocols / applications in order to avoid eliminate false-positives but allow operators to define application filters if a port-based identification is deemed adequate (see an example below).
Application Filter	Operator configurable, optional component of AA flow identification that uses any combination of protocol signatures, server IP address and port, flow set-up direction, configurable expressions (for example an HTTP string match) to identify user's traffic.	http_video + IP address of partner's video server or http_video + an HTTP string to identify partner's video content TCP or UDP + TCP/UDP port number to identify a TCP or UDP based protocol or application.
Application	Operator configurable, optional component of AA flow identification that allows defining any specific forms of traffic to and from end user clients by combining application filter entries	Google Talk, POP3, YouTube, iTunes, Shoutcast
Application Charging Group	Operator configurable, optional component of AA flow identification that allows grouping of similar end use applications using operator defined names and groups.	IM, Mail, Multimedia, P2P, Tunneling, Web, Other
Clients	End user programs that generate user traffic for applications and protocols, and that are used in a process of AA flow identification verification.	The list of clients is constantly evolving as new clients or versions are introduced in the marketplace. The following example illustrates clients that may be used to generate Application traffic matching BitTorrent application defined using BitTorrent and DHT protocol signatures: Limewire, BitTorrent, Azureus, Ktorrent, Transmission, Utorrent

Protocol Signatures

The set of signatures used to identify protocols is generated by Alcatel-Lucent and included with the Application Assurance software load. The signature set includes:

- The protocols that can be identified with this load, using a combination of pattern and behavioral techniques. The protocols are used in generating statistics by protocol, and are used as input in combination with other information to identify applications.
- Pattern signatures are the set of pattern-match signatures used in analysis.
- Behavior signatures are the set of diagnostic techniques used in analysis.

Dynamic upgrades of the signatures in the system are implemented by invoking an **admin application-assurance upgrade** command and then performing AA ISA activity switches.

The protocol signatures are included in aa-isa.tim software load which is not tightly coupled with software releases allowing for protocol signature updates without upgrading and impacting of routing/forwarding engines as part of an ISSU upgrade that updates only the AA ISA software. Refer to upgrade procedures described in the 7750 SR and/or 7450 ESS Release Notes for detailed information.

Since protocol signatures are intended to be the most basic block of Application Identification, other AA components like Application Filters are provided to further customize Protocol Signatures allowing operators to customize their applications and to reduce a need for a new Protocol Signature load when a new Application may need to be identified. This architecture gives operators more flexibility in responding to ever changing needs in application identifications.

Signature upgrade without a router upgrade is allowed within a major router release independently of system ISSU limits. An AA ISA signature upgrade is supported before the first ISSU router release (for example, operators can upgrade signatures for pre-ISSU minor releases).

In addition, any router release from ISSU introduction release can run any newer aa-isa.tim image within the same major release by performing an aa-isa.tim single step upgrade. For example, release 8.4 may be upgraded in a single step to run release 8.14 of isa-aa.tim.

Each protocol, except internal protocols used for special-case processing statistic gathering (like “cut-though”, for example), can be referenced in the definition of one or multiple applications (through the App-Filter definition). Assignment of a supported protocol to an app-filter or application is not mandatory. Protocols not assigned to an application are automatically mapped by the system to the default “Unknown” application.

Custom Protocols

Custom protocols are supported using configurable strings (up to 16 hex octets) for pattern-matched application identification in the payload of TCP or UDP based applications (mutually exclusive to other string matches in an app-filter).

The match is specified for the “client-to-server”, “server-to-client”, or “any” direction for TCP based applications, and in the “any” direction for UDP based applications.

There is a configurable description and custom protocol id for a protocol, with configurable shutdown. When disabled, traffic is identified as if the protocol was not configured.

Custom protocols and ALU-provided protocols are functionally equivalent. Custom protocols are used in application definition without limitations (all app-filter entries except strings are supported). Collection of custom protocol statistics on a partition/ISA group/special study sub level is supported.

Protocol Shutdown

The protocol **shutdown** feature provides the ability for signature upgrades without automatically affecting policy behavior, especially if some or even all new signatures are not required for a service. All new signatures are disabled on upgrade by default to ensure no policy/service impact because of the signature update.

All protocols introduced at the R1 stage of a given release are designated as "Parent" signatures for a given release and cannot be disabled.

Within a major release, all protocols introduced post-R1 of a major release as part of any isa-aa.tim ISSU upgrade are by default **shutdown**. They must be enabled on a per-protocol basis (system-wide) to take effect.

When shutdown, post R1-introduced protocols do not change AA behavior (app-id, policy, statistics are as before the protocol introduction), for example, traffic maps to the parent protocol on which the new signature is based. In cases where there is more than one parent protocol, all traffic is mapped to a single, most-likely, parent protocol. For example if 80% of a new protocol has traffic mapping to unknown_tcp, and 20% mapping to another protocol(s), unknown_tcp would be used as parent.

Enabling/disabling of a new protocol takes affect for new flows only. The current status (enabled/shutdown) of a signature and the parent protocol is visible to an operator as part of retrieving protocol information through CLI/SNMP.

Supported Protocol Signatures

Protocol signatures are release independent and can be upgraded independently from the router's software and without impacting router's operations as part of an ISSU upgrade. A separate document outlines signatures supported for each signature software load (isa-aa.tim). New signature loads are distributed as part of the SR/ESS maintenance cycle. Traffic identified by new signatures will be mapped to an "Unknown" application until the AA policy configuration changes to make use of the newly introduced protocol signatures.

Application Groups

Application groups are defined as a container for multiple applications. The only application group created by default is **Unknown**. Any applications not assigned to a group are automatically assigned to the default **Unknown** group. Application groups are expected to be defined when a common policy on a set of applications is expected, yet per each application visibility in accounting is required. The application group name is a key match criteria within application QoS policy rules.

Charging Groups

Charging Groups allow usage accounting by application and/or app groups in a manner that does not affect app to app-group mapping. For example, AA app groups statistics for “Streaming Video” includes all streaming apps, independent of whether any specific application is 0-rated for charging. AA charging groups are used for charging related statistics.

As with app-groups, charging groups are defined under an AA policy context for an AA group or partition. Once defined, individual apps and app-groups can be associated with the desired charging group. The charging group name is a key match criteria within application QoS policy rules.

A default charging group can be specified for the AA policy to associate a charging group to any applications or app-groups that are not explicitly assigned to a charging group.

Charging groups are also assigned an export-id number for accounting export purposes.

If no export-id is assigned, that charging group cannot be added to the aa-sub stats RADIUS export-type. Once a charging group index is referenced, it cannot be deleted without removing the reference.

Applications

The application context defines and assigns a description to the application names supported by the application filter entries, and assigns applications to application groups.

- Application name is a key match criteria within application QoS policy rules, which are applied to a subscribers IP traffic.
- Each application can be associated with one of the application groups provided by Application Assurance.

The Application Assurance system provides no pre-defined applications other than **Unknown**. Applications must be explicitly configured. Any protocols not assigned to an application are automatically assigned to the default **Unknown** application. Alcatel-Lucent provides sets of known-good application/app-group configurations upon request. Contact the technical support staff for further information.

The applications are used by Application Assurance to identify the type of IP traffic within the subscriber traffic.

The network operator can:

- Define unique applications.

- Associate applications with an application group. The application group must already be configured.

Application Filters

Application filters (app-filter) are provided as an indirection between protocols and applications to allow the addition of variable parameters (port number, IP addresses, etc.) into an application definition. An application filter is a numbered rule entry that defines the use of protocol signatures and other criteria to define an application. Multiple rules can be used to define what constitutes an application but each rule will map to only one application definition.

The system concept of application filters is analogous to IP filters. Match of a flow to multiple rules is possible and is resolved by picking the rule with the lowest entry number that matches. A flow will only ever be assigned to one application.

The following criteria can be assigned to an application filter rule entry:

- Unique entry ID number
- Application name
- Flow setup direction
- Server IP address (or server IP filter list)
- Server port
- Protocol signature
- IP protocol number
- String matches against Layer 5+ protocol header fields (for example, a string expression against HTTP header fields)

The application must be pre-configured prior to using it in an app-filter. Once defined, the new application names can be referenced.

HTTP

HTTP Protocol

The Hyper-Text Transfer Protocol (HTTP) has become the most significant protocol used on the Internet and has expanded its role beyond web browsing with a large number of applications using HTTP for a variety of functions on both desktop and mobile devices.

Application Assurance provides the tools required by residential, mobile and business VPN service providers to accurately classify any web-based applications regardless of where the content is stored and how it is delivered. This is done by using either the default protocol signatures delivered with the AA ISA software or by defining string based signatures from the HTTP header information fields included in the HTTP request messages to further refine the detection.

HTTP Session Persistency

HTTP can use both non persistent connections and persistent connections. Non-persistent connection uses one TCP connection per HTTP request while persistent connection can reuse the same TCP connection for multiple HTTP request to the same server.

Nowadays most applications are using HTTP/1.1 and persistent connection but HTTP/1.0 and non-persistent connections remains on older software and mobile devices.

HTTP flows are classified in a particular application using the first HTTP request of the flow only by default. Optionally, the MS-ISA offers the flexibility to classify each HTTP request within the same flow independently using **http-match-all-request** feature.

HTTP Proxy Support

Application Assurance also supports traffic classification of HTTP between a subscriber and a web proxy. This feature is enabled by default, the ISA monitors and detects HTTP proxy flows automatically, each request within the same persistent connection to the proxy server is classified independently.

AA IP Prefix Lists

AA ISA allows the match section of session filters, AQPs entries and application filters to include matching against a configured IP filter list(s). Each IP filter list (aka IP pools) can have up to 64 IP address entries with a configurable mask for each entry.

Statistics and Accounting

Application Assurance statistics provide the operator with information to understand application usage within a network node.

Application Assurance XML record accounting aggregates the flow information into per application group, per application, per protocol reports on volume usage during the last accounting interval. This information is then sent to a statistics collector element for network wide correlation and aggregation into customized graphical usage reports. Application Assurance uses and benefits from the rich 7750 SR/7450 ESS accounting infrastructure and the functionality it provides to control accounting policy details.

The following types of accounting volume records are generated and can be collected:

- Per ISA group and partition record for each configured application group
- Per ISA group and partition record for each configured application
- Per ISA group and partition record for each configured protocol
- Per each AA subscriber record with operator-configurable field content using custom AA records for operator-selected subset of protocols, applications and application groups
- Per AA subscriber per each configured application record (special study mode)
- Per AA subscriber per each supported protocol record (special study mode)
- Per ISA AA-performance record, containing information about the traffic and resources of each ISA
- Per AA partition stats record for counts of traffic by Layer 3 protocol used to transport L4 protocols. This includes TCP, UDP and NonTcpUdp carried by IPv4, IPv6, DS_Lite, 6to4/6RD and Teredo protocols

Application Assurance supports RADIUS accounting export of per AA subscriber charging group statistics.

Each AA group:partition can be configured for AA-subscribers stats export by referencing both an accounting policy (for XML statistics) and/or a RADIUS accounting policy. In order to determine how to export various counters for subscriber AA statistics, an export-using keyword is used when enabling aa-sub level stats export to specify the export method to be used for each, whether accounting-policy or radius-accounting-policy and/or diameter-based usage monitoring.

Per AA flow statistics are provided as described in the cflowd section.

Refer to the 7750 SR/7450 ESS OS System Management Guide for information on general accounting functionality.

Per-AA-Subscriber Special Study

The system can be configured to generate statistical records for each application and protocol that the system identifies for specific AA subscribers. These capabilities are disabled by default but can be enabled for a subset of AA subscribers to allow detailed monitoring of those AA subscriber's traffic.

Per-aa-sub per-application and per-aa-sub per-protocol records are enabled by assigning individual AA subscribers to "special study" service lists. The system and ISA group limit the number of AA subscribers in this mode to constrain the volume of stats generated. When an AA subscriber is in a special study mode, one record for every application and/or one record for every protocol that are configured in the system are generated for that subscriber. For example, if 500 applications are configured and 200 protocols are identified, 700 records per AA subscriber will be generated, if the AA subscriber is listed in both the per-aa-sub-application and per-aa-sub protocol lists.

System Aspects

Application Assurance uses the existing redundant accounting and logging capability of the 7750 SR/7450 ESS for sending application and subscriber usage information, in-band or out-of-band. Application Assurance statistics are stored using compressed XML format with other system and subscriber statistics in compact flash modules on the redundant SF/CPMs. A large volume of statistics can be expected under scaled scenarios when per-AA-subscriber statistics/accounting is enabled.

AA accounting and statistics can be deployed as part of other system functionality as long as the system's function is compatible with AA accounting or as long as the system-level statistics can become application-aware due to, for example, AA ISA-based classification. An example of this feature interaction includes volume and time-based accounting where AA-based classification into IOM queues with volume and time accounting enabled can, for instance, provide different quota/credit management for off-net and on-net traffic or white/grey applications.

Application Assurance XML Volume Statistics and Accounting

Application Assurance is configured to collect and report on the following statistics when at least one AA ISA is active. The default Application Assurance statistics interval is 15 minutes.

Statistics to be exported from the node are aggregated into accounting records, which must be enabled in order to be sent. By default, no records are sent until enabled. Each record template type is enabled individually to control volume of statistics to the desired level of interest. Only non-zero records are written to the accounting files for all AA subscriber based statistics to reduce the volume of data.

The operator can further select a subset of the fields to be included in per-AA-subscriber records and whether to send records if no traffic was present for a given protocol or application, for example, sending only changed records.

Each record generated contains the record fields as described in Table 8. The header row represents the record type.

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
Application Group	Name	X						data name
Application	Name		X			X		data name
Protocol	Name			X			X	data name
Aggregation Type ID	ID (can be protocol, application, charging group or application group record)				X			agg-type-name
# Active Subscribers	# of subscribers who had a flow of this category during this interval	X	X	X				nsub
# allowed flows from-sub	# of new flows that were identified and allowed	X	X	X	X	X	X	sfa
# allowed flows to-sub	As above in opposite direction	X	X	X	X	X	X	nfa

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
# denied flows from-sub	the # of new flows that were identified and denied	X	X	X	X	X	X	sfd
# denied flows to-sub	As above in opposite direction	X	X	X	X	X	X	nfd
# Active flows from-sub	# of flows that were either: closed, opened & closed, opened, or continued during this interval	X	X	X	X	X	X	saf
# active flows to-sub	As above, in opposite direction	X	X	X	X	X	X	naf
Total packets from-sub		X	X	X	X	X	X	spa
Total packets to-sub		X	X	X	X	X	X	npa
Total bytes from-sub		X	X	X	X	X	X	sba
Total bytes to-sub		X	X	X	X	X	X	nba
Total discard packets from-sub		X	X	X	X	X	X	spd
Total short flows	Number of flows with duration <= 30 seconds that completed up to the end of this interval	X	X	X	X	X	X	sdf
Total medium flows	Number of flows with duration <= 180 seconds that completed up to the end of this interval	X	X	X	X	X	X	mdf
Total long flows	Number of flows with duration > 180 seconds that completed up to the end of this interval	X	X	X	X	X	X	ldf
Total discard packets to-sub		X	X	X	X	X	X	npd
Total discard bytes from-sub		X	X	X	X	X	X	sbd
Total discard bytes to-sub		X	X	X	X	X	X	nbd

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
Total flows completed	# of to- and from-subscriber flows that have been completed up to the reported interval.	X	X	X	X	X	X	tfc
Total flow duration	Duration, in seconds, of all flows that have been completed up to the reported interval.	X	X	X	X	X	X	tfd
From AA Sub: Maximum throughput byte count	Maximum of all total byte counts recorded for throughput intervals within this accounting interval for traffic originated by AA subscriber for a given application/app-group. AA ISA discarded traffic is not included.				X			sbm
From AA Sub: Packet count corresponding to the max. throughput byte count interval.	Packet count for the throughput interval with the maximum byte count value for traffic originated by AA subscriber for the application/ app-group. AA ISA discarded traffic is not included.				X			spm
To AA Sub: Max throughput time slot index	UTC time that corresponds to the end of the 5-minute throughput interval where the max throughput byte count was detected.				X			smt
From AA Sub: Forwarding-class	Observed forwarding-class bits.	X	X	X	X	X	X	sfc
To AA Sub: Forwarding-class	Observed forwarding-class bits.	X	X	X	X	X	X	nfc

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
To AA Sub: Maximum throughput byte count	Maximum of all total byte counts recorded for throughput intervals within this accounting interval for traffic originated from Network towards AA subscriber for a given application/app-group. AA ISA discarded traffic is not included.				X			nbm
To AA Sub: Packet count corresponding to the max. Throughput byte count interval.	Packet count for the throughput interval with the maximum byte count value for traffic originated from network towards AA subscriber for a given application / app-group. AA ISA discarded traffic is not included.				X			npm
From AA Sub: Max throughput time slot index	UTC time that corresponds to the end of the 5-minute throughput interval where the max throughput byte count was detected.				X			nmt
From AA Sub: Forwarding-class	Observed forwarding-class bits.	X	X	X	X	X	X	X
From AA Sub: Maximum throughput byte count	Maximum of all total byte counts recorded for throughput intervals within this accounting interval for all traffic originated by AA subscriber. AA ISA discarded traffic is not included.				X			sbm

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
From AA Sub: Packet count corresponding to the max. Throughput byte count interval.	Packet count for the throughput interval with the maximum byte count value for traffic originated by AA subscriber. AA ISA discarded traffic is not included.				X			spm
From AA Sub: Max throughput time slot index	UTC time that corresponds to the end of the 5-minute throughput interval where the max throughput byte count was detected.				X			smt
To AA Sub: Maximum throughput byte count	Maximum of all total byte counts recorded for throughput intervals within this accounting interval for traffic originated from network towards AA subscriber. AA ISA discarded traffic is not included.				X			nbm
To AA Sub: Packet count corresponding to the max. Throughput Byte Count interval.	Packet count for the throughput interval with the maximum byte count value for traffic originated from network towards AA subscriber. AA ISA discarded traffic is not included.				X			npm
To AA Sub: Max throughput time slot index	UTC time that corresponds to the end of the 5-minute throughput interval where the max throughput byte count was detected.				X			nmt

Table 8: Application Assurance Statistics Fields Generated per Record (Accounting File)

Record Fields	Description	Group/Partition App Group	Group/Partition Application	Group/Partition Protocol	AA-Sub Custom	AA-Sub Special Study per App	AA-Sub Special Study Protocol	XML Name
Forwarding Class		X						fc
App-Profile	AA-Sub App-Profile name				X			app-profile
App-Service-Options	List of the app-service-options characteristics and values per AA-Sub				X			app-service-option

The records are generated per ISA group and partition, with an ISA group identified by the group ID (XML field name “aaGroup”), partition identified by the partition ID (XML field name “aaPart name” and per AA subscriber (if applicable) with the AA subscriber identified by the ESM, DSM, or transit subscriber name, SAP ID (XML field name “subscriber name”, “sap name” or “spoke SDP ID” respectively).

The date, time, and system ID for the records will be visible as part of the existing accounting log capability, thus does not need to be contained inside the Application Assurance records themselves.

The Forwarding Class is included in AA XML records as generally a VPN interconnection SLA is a combination of Bandwidth connection at the site level and Forwarding Class to transport the traffic over the MPLS network, by mapping the end-customer DSCP or 802.1P traffic value into a given FC.

AA accounting stats of the application/application-group volume usage per forwarding class shows the exact volume of each application at the per FC level and better ties the AA reports to the VPN services and SLA.

This can also identify key applications using a non optimal FC over a given VPN/Site and allow the option for AA to remark these into a higher traffic class, with reporting per FC to show resulting use.

AA Partition Traffic Type Statistics

AA-ISA provides, at the AA partition level, traffic volume visibility of the Layer 3 protocols used to transport the different Layer 4 protocols. These include a traffic volume break down of TCP, UDP and Non-TCP-UDP carried by IPv4 and IPv6 protocols.

Traffic-type statistics are broken down by "family" and "protocol":

- Family: IPv4, IPv6
- Protocol: TCP, UDP, Other

Therefore, AA-ISA traffic type record provides a collection of 15 sets of traffic volume (bytes).

Statistics figures as follows:

- IPv4 — TCP, UDP, Other
- IPv6 — TCP, UDP, Other

These statistics are always counted. There is no configuration required to enable/disable tracking. However, the operator has the option to enable/disable export of these statistics via XML

Table 9 lists the statistic record fields per AA partition.

Table 9: AA-Partition traffic type statistics

Record name	Type	Description
sba	cumulative	octets admitted (from-sub)
spa	cumulative	packets admitted (from-sub)
sbd	cumulative	octets denied (from-sub)
spd	cumulative	packets denied (from-sub)
nba	cumulative	octets admitted (to-sub)
npa	cumulative	packets admitted (to-sub)
nbd	cumulative	octets denied (to-sub)
npd	cumulative	packets denied (to-sub)
sfa	cumulative	flows admitted (from-sub)
sfd	cumulative	flows denied (from-sub)
saf	intervalized	active flows (from-sub)
nfa	cumulative	admitted flows (to-sub)
nfd	cumulative	flows denied (to-sub)
naf	intervalized	active flows (to-sub)
tfc	cumulative	total terminated flows
tfd	cumulative	total terminated flow duration
sdf	cumulative	short duration flows

Table 9: AA-Partition traffic type statistics (Continued)

Record name	Type	Description
mdf	cumulative	medium duration flows
ldf	cumulative	long duration flows
sfc	cumulative	forwarding-class bitmap (from-sub)
nfc	cumulative	forwarding-class bitmap (to-sub)
tet	cumulative	num of subscribers tethered
nte	cumulative	num of subscribers not tethered

Configurable AA-Subscriber Statistics Collection

Existing average volume statistics collected over an accounting interval are extended to provide the maximum volume (bytes/packets) recorded for a throughput measurement period (5 minutes) within an accounting interval. These additional statistics improve accuracy for the access-pipe right-sizing service.

Maximum throughput statistics can be enabled for the selected applications and/or application groups enabled for custom per AA statistics. In addition, the operator can enable (disabled by default) per AA-subscriber "Max-throughput" statistics for total (/aggregate) subscriber traffic, independent of defined applications/application-groups.

Maximum throughput statistics records are allocated from the 2048K records available for use for per subscriber records.

Maximum throughput statistics are not provided for the protocols enabled for custom per AA statistics.

AA-Performance Record for ISA Load

The AA-performance statistics record provides visibility of ISA loading related statistics to allow operational monitoring and planning of ISA overload:

1. Provides end of reporting interval snapshot of current values of the parameters listed in below into a per AA ISA Planning record. “Current” is the value of a counter at the end of the reporting interval, for rate based values this is the ~10sec short term current rate used in CLI statistics.
2. Provides time-based averages during record interval of the above values: Average(I)
3. Provides peak values of the above values in the reporting interval: Peak(I)

The 5670 RAM provides further analysis and thresholding triggers based on these ISA statistics, suitable for long-range planning trends such as average number of subs or peak numbers of flows.

The node per-ISA planning record values are cleared on accounting read (per all accounting records). Not reading the records means that the average and peak values are the values for the last reporting interval. The time last read is indicated in the record.

The AA performance planning record contains the following fields:

Table 10: AA Performance Planning Record Fields

Parameter	Current	Average(I)	Peak(I)
ISA ID			
active flows	# flows	# flows	# flows
flow setup rate	flows/sec	flows/sec	flows/sec
traffic rate	bits/sec	bits/sec	bits/sec
Packet rate	packets/sec	packets/sec	packets/sec
active subs	# subs	# subs	# subs
downloaded subs	# subs	# subs	# subs
flow resources in use (active flows + wildcard flows)	# flows		
ISA AA sub stats resource allocation	# stats records		
ISA capacity cost	sum of cost of active AA subs		
ISA Transit Subs	# subs		
diverted traffic	(packets, octets)		
entered ISA	(packets, octets)		

Table 10: AA Performance Planning Record Fields (Continued)

Parameter	Current	Average(I)	Peak(I)
policy discards in ISA	(packets, octets)		
congestion discards in ISA	(packets, octets)		
error discards in ISA	(packets, octets)		
policy bypass errors	(packets, octets)		
returned traffic	(packets, octets)		
Volume cflowd			
Records reported	# records		
Reports dropped	# records		
Packets sent	# packets		
Comprehensive cflowd			
Records reported	# records		
Reports dropped	# records		
Packets sent	# packets		
TCP performance cflowd			
Flows not allocated	#flows		
Records reported	# records		
Reports dropped	# records		
Packets sent	# packets		
RTP performance cflowd			
Flows not allocated	#flows		
Records reported	# records		
Reports dropped	# records		
Packets sent	# packets		
Num of synchronization sources that had to be aborted	#SSRC aborted		

Table 10: AA Performance Planning Record Fields (Continued)

Parameter	Current	Average(I)	Peak(I)
Records sent (Note that the data name=collector address and port inserted in XML record.	#records to be collected (referenced by XML name)		
AA-Subs Created			
AA-Subs Deleted			
AA-Subs Modified			
seen-ip - requests sent	#requests		
seen-ip - requests dropped	#requests		
subscribers created	#subs		
subscribers deleted	#subs		
subscribers modified	#subs		
transit-prefix v4 address count	#addresses		
transit-prefix v6 address count	#addresses		
transit-prefix v6 remote address count	#addresses		

Table 11: Per ISA AA Performance Record Fields

Record Name	Type	Description	MIB object
tmo	cumulative	octets to MDA	tmnxBsxGrpStatusOctsToMda
tmp	cumulative	packets to MDA	tmnxBsxGrpStatusPktsToMda
fmo	cumulative	octets from MDA	tmnxBsxGrpStatusOctsFromMda
fmp	cumulative	packets from MDA	tmnxBsxGrpStatusPktsFromMda
dco	cumulative	octets discarded due to congestion in MDA	tmnxBsxGrpStatusOctsDisCongMda
dcp	cumulative	packets discarded due to congestion in MDA	tmnxBsxGrpStatusPktsDisCongMda
dpo	cumulative	octets discarded due to policy in MDA	tmnxBsxGrpStatusOctsDiscPolicy
dpp	cumulative	packets discarded due to policy in MDA	tmnxBsxGrpStatusPktsDiscPolicy
deo	cumulative	octets discarded due to error	tmnxBsxGrpStatusOctsDiscErrors
dep	cumulative	packets discarded due to error	tmnxBsxGrpStatusPktsDiscErrors
pbo	cumulative	octets policy bypass	tmnxBsxGrpStatusOctsPolicyByps
pbp	cumulative	packets policy bypass	tmnxBsxGrpStatusPktsPolicyByps
nfl	cumulative	number of flows	tmnxBsxGrpStatusFlows
caf	intervalized	current active flows	tmnxBsxGrpStatusFlowsCurrent
aaf	intervalized	average active flows	tmnxBsxGrpStatusFlowsAverage
paf	intervalized	peak active flows	tmnxBsxGrpStatusFlowsPeak
cfr	intervalized	current flow setup rate	tmnxBsxGrpStatusFlowSetupRate
afr	intervalized	average flow setup rate	tmnxBsxGrpStatusFlowSetupRateAvg
pfr	intervalized	peak flow setup rate	tmnxBsxGrpStatusFlowSetupRatePk
ctr	intervalized	current traffic rate	tmnxBsxGrpStatusTrafficRate
atr	intervalized	average traffic rate	tmnxBsxGrpStatusTrafficRateAvg
ptr	intervalized	peak traffic rate	tmnxBsxGrpStatusTrafficRatePeak
cpr	intervalized	current packet rate	tmnxBsxCflowdStatusPktRateCurr
apr	intervalized	average packet rate	tmnxBsxGrpStatusPacketRateAvg

Table 11: Per ISA AA Performance Record Fields (Continued)

Record Name	Type	Description	MIB object
ppr	intervalized	peak packet rate	tmnxBsxGrpStatusPacketRatePeak
cas	intervalized	current active subscribers (with flows)	tmnxBsxGrpStatusSubsCurrent
aas	intervalized	average active subscribers (with flows)	tmnxBsxGrpStatusSubsAverage
pas	intervalized	peak active subscribers (with flows)	tmnxBsxGrpStatusSubsPeak
cds	intervalized	current diverted subscribers	tmnxBsxGrpStatusSubsDiverted
ads	intervalized	average diverted subscribers	tmnxBsxGrpStatusSubsDivertedAvg
pds	intervalized	peak diverted subscribers	tmnxBsxGrpStatusSubsDivertedPk
rfi	intervalized	flows in use	tmnxBsxGrpStatusFlowResInUse
rcc	cumulative	ISA capacity cost	tmnxBsxGrpMdaCapacityCost
rss	cumulative	subscriber statistics count	tmnxBsxGrpMdaStatsResourceCount
rti	cumulative	transit-ip address count	tmnxBsxGrpMdaTransitIpAddrs
rtp4	cumulative	transit-prefix v4 address count	tmnxBsxGrpMdaTransPrefV4Entr
rtp6	cumulative	transit-prefix v6 address count	tmnxBsxGrpMdaTransPrefV6Entr
rtp6r	cumulative	transit-prefix v6 remote address count	tmnxBsxGrpMdaTransPrefV6RemEntr
srs	cumulative	seen-ip - requests sent	tmnxBsxGrpStatusHCSeenIpReqSent
srd	cumulative	seen-ip - requests dropped	tmnxBsxGrpStatusHCSeenIpReqDrop
tsc	cumulative	total subscribers created	tmnxBsxGrpStatusHCSubsCreated
tsd	cumulative	total subscribers deleted	tmnxBsxGrpStatusHCSubsDeleted
tsm	cumulative	total subscribers modified	tmnxBsxGrpStatusHCSubsModified
vrr	cumulative	volume cflowd - records reported	tmnxBsxCflowdStatusRecReported
vrd	cumulative	volume cflowd - records dropped	tmnxBsxCflowdStatusRecDropped
vps	cumulative	volume cflowd - packets sent	tmnxBsxCflowdStatusPktsSent
crr	cumulative	comprehensive cflowd - records reported	tmnxBsxCflowdStatusRecReported
crd	cumulative	comprehensive cflowd - records dropped	tmnxBsxCflowdStatusRecDropped
cps	cumulative	comprehensive cflowd - packets sent	tmnxBsxCflowdStatusPktsSent

Table 11: Per ISA AA Performance Record Fields (Continued)

Record Name	Type	Description	MIB object
trr	cumulative	tcp-performance cflowd - records reported	tmnxBsxCflowdStatusRecReported
trd	cumulative	tcp-performance cflowd - records dropped	tmnxBsxCflowdStatusRecDropped
tps	cumulative	tcp-performance cflowd - packets sent	tmnxBsxCflowdStatusPktsSent
tfn	cumulative	tcp-performance cflowd - flows but no cflowd resources available	tmnxBsxCflowdStatusFlowsNoRes
rrr	cumulative	rtp-performance cflowd - records reported	tmnxBsxCflowdStatusRecReported
rrd	cumulative	rtp-performance cflowd - records dropped	tmnxBsxCflowdStatusRecDropped
rps	cumulative	rtp-performance cflowd - packets sent	tmnxBsxCflowdStatusPktsSent
rfn	cumulative	rtp-performance cflowd - flows but no cflowd resources available	tmnxBsxCflowdStatusFlowsNoRes
rsr	cumulative	rtp-performance cflowd - num of synchronization sources that had to be aborted	tmnxBsxCflowdStatusHCUSupSSRCSt
res	cumulative	srflow collector - records sent	tmnxBsxCflowdCollStatRecSent
hrs	cumulative	url-filter http-requests sent	tmnxBsxUrlFltrStatsHttpRequests
hre	cumulative	url-filter - http-request errors	tmnxBsxUrlFltrStatsHttpReqErrors
hri	cumulative	url-filter - http-requests dropped	n/a
hrp	cumulative	url-filter - http-requests permitted	tmnxBsxUrlFltrStatsHttpRespAllow
hrrt	cumulative	url-filter - http-requests redirected	mnxBsxUrlFltrStatsHttpRespRedir
hrb	cumulative	url-filter - http-requests blocked	tmnxBsxUrlFltrStatsHttpRespBlock
hda	cumulative	url-filter - http default actions	tmnxBsxUrlFltrStatsHttpRespDef
irs	cumulative	icap - icap requests	tmnxBsxIcapServerStatsRequests
ire	cumulative	icap - icap request errors	tmnxBsxIcapServerStatsReqErrors
irp	cumulative	icap - icap permits	tmnxBsxIcapServerStatsRespAllow
irr	cumulative	icap - icap redirects	tmnxBsxIcapServerStatsRespRedir
ird	cumulative	icap - icap drops	tmnxBsxIcapServerStatsRespBlock
ilr	cumulative	icap - icap late responses	tmnxBsxUrlFltrStatsIcapLateResp

Table 11: Per ISA AA Performance Record Fields (Continued)

Record Name	Type	Description	MIB object
irt	cumulative	icap - icap average rtt	tmnxBsxIcapServerStatsRoundTrip
itc	cumulative	icap - icap tcp connections	tmnxBsxIcapServerStatsConnEst
ifs	cumulative	url-filter - subscriber count	n/a
rtp4r	cumulative	transit-prefix v4 remote address count	tmnxBsxGrpMdaTransPrefV4RemEntr
lrp	cumulative	url-list permits	tmnxBsxUrlFltrStatsHttpRespAllow
lrr	cumulative	url-list redirects	tmnxBsxUrlFltrStatsHttpRespRedir
lrd	cumulative	url-list drops	tmnxBsxUrlFltrStatsHttpRespBlock

AA Partition Traffic Type Statistics

AA ISA provides, at the AA partition level, traffic volume visibility of the Layer 3 protocols used to transport the different Layer 4 protocols. These include a traffic volume break down of TCP, UDP and Non-TCP-UDP carried by IPv4, IPv6, DS_Lite, 6to4/6RD and Teredo protocols.

Traffic-type statistics are broken down by family and protocol:

- Family: IPv4, IPv6, DS-Lite, 6RD/6to4, Teredo
- Protocol: TCP, UDP, Other

Therefore, AA ISA traffic type record provides a collection of 15 sets of traffic volume (Bytes) statistics figures as follows:

- IPv4 — TCP, UDP, Other
- IPv6 — TCP, UDP, Other
- DS-Lite — TCP, UDP, Other (IPv4 tunneled inside IPv6)
- 6to4/6RD — TCP, UDP, Other (IPv6 tunneled inside IPv4)
- Teredo — TCP, UDP, Other (IPv6 tunneled inside IPv4 and UDP)

These statistics are always counted. There is no configuration required to enable/disable tracking. However, the operator has the option to enable/disable export of these statistics via XML.

Table 12: Per AA Partition Stats Record Fields

Record Name	Type	Description	MIB object (if applicable)
sba	cumulative	octets admitted (from-sub)	tmnxBsxTrafStatOctsAdmFmSb
spa	cumulative	packets admitted (from-sub)	tmnxBsxTrafStatPktsAdmFmSb
sbd	cumulative	octets denied (from-sub)	tmnxBsxTrafStatOctsDnyFmSb
spd	cumulative	packets denied (from-sub)	tmnxBsxTrafStatPktsDnyFmSb
nba	cumulative	octets admitted (to-sub)	tmnxBsxTrafStatOctsAdmToSb
npa	cumulative	packets admitted (to-sub)	tmnxBsxTrafStatPktsAdmToSb
nbd	cumulative	octets denied (to-sub)	tmnxBsxTrafStatOctsDnyToSb
npd	cumulative	packets denied (to-sub)	tmnxBsxTrafStatPktsDnyToSb
sfa	cumulative	flows admitted (from-sub)	tmnxBsxTrafStatFlwsAdmFmSb
sfd	cumulative	flows denied (from-sub)	tmnxBsxTrafStatFlwsDnyFmSb

Table 12: Per AA Partition Stats Record Fields (Continued)

Record Name	Type	Description	MIB object (if applicable)
saf	intervalized	active flows (from-sub)	tmnxBsxTrafStatActFlwsFmSb
nfa	intervalized	active flows (to-sub)	tmnxBsxTrafStatActFlwsToSb
nfd	cumulative	flows denied (to-sub)	tmnxBsxTrafStatFlwsDnyToSb
naf	intervalized	active flows (from-sub)	tmnxBsxTrafStatActFlwsFmSb
tfc	cumulative	total terminated flows	tmnxBsxTrafStatTermFlws
tfd	cumulative	total terminated flow duration	tmnxBsxTrafStatTermFlwDur
sdf	cumulative	short duration flows	tmnxBsxTrafStatShrtDurFlws
mdf	cumulative	medium duration flows	tmnxBsxTrafStatMedDurFlws
ldf	cumulative	long duration flows	tmnxBsxTrafStatLngDurFlws
sfc	cumulative	forwarding-class bitmap (from-sub)	n/a
nfc	cumulative	forwarding-class bitmap (to-sub)	n/a

RADIUS Accounting AA Records

AA RADIUS accounting provides per aa-subscriber level charging group statistics as well as application-group (AG) support into RADIUS accounting infrastructure and application group support. The primary use of this is to allow RADIUS accounting to be enhanced with AA information useful for usage-based billing plans, providing flexibility to charge and rate application content using IP subnets, HTTP URLs, SIP URIs and other AA identified applications.

The system can export AA accounting statistics using accounting policy records exported with RADIUS accounting. AA RADIUS accounting provides total volume broken out by charging groups which are mapped by application. AA RADIUS accounting is aa-sub-ID based, where the AA-sub context IPv4 and IPv6 host addresses for the sub are not reflected in RADIUS accounting.

AA RADIUS accounting is implemented using ALU Vendor Specific Attributes (VSAs). This provides all charging group counters for a given subscriber to be exported with a common accounting session ID. The following statistics are included in each record. Accounting values are for forwarded packets (after AA policy):

- input octets (from-sub)
- input packets (from-sub)
- output octets (to-sub)
- output packets (to-sub)

RADIUS accounting is supported for all AA-sub types. RADIUS accounting is used to export of AA CG and AG (App-group) values at according to the RADIUS accounting policy interval. Charging groups statistics are exported in RADIUS accounting independent of app-groups (either or both can be enabled).

AA GX Based Usage Monitoring

Using 3GPP (third generation Partnership Project) diameter (Gx) functionality, AA ISA upon receiving requests from Policy and Charging Rules Function (PCRF), can monitor application usage at the subscriber's level and report back to PCRF whenever the usage exceeds the threshold(s) set by the PCRF.

Usage-monitoring can be used by operators to report to PCRF when:

a.) AA ISA detects the start of a subscriber application (by setting usage threshold to be very low)

b.) 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.

AA ISA Gx-based usage monitoring is restricted to AA ESM subscribers' type.

The AA ISA Gx usage monitoring feature builds on 3GPP Release 11 defined Application Detection and Control (ADC) Gx attributes. In addition, AA ISA is compliant with 3GPP Release 12, whereby the ADC rule functionality is integrated in the PCC rules.

AA ISA reports accumulated usage when:

- A usage threshold is reached
- The PCRF explicitly disables usage monitoring
- The PCRF requests for a report
- When the ADC or PCC rule associated with the monitoring instance is removed or deactivated
- When a session is terminated

An AA defined application, application group and/or charging group is automatically allowed to be referenced by a an ADC rule for the purpose of usage monitoring only if:

a.) It is already selected for either XML or Radius per subscriber accounting or

b.) It is explicitly enabled by the operator for per sub statistics collection and

c.) Usage monitoring is enabled for the given AA group:partition.

Figure 22 illustrates the different messaging /call flows involved in application level usage monitoring. For details about the supported AVPs used in these messages, see section Supported AVPs.

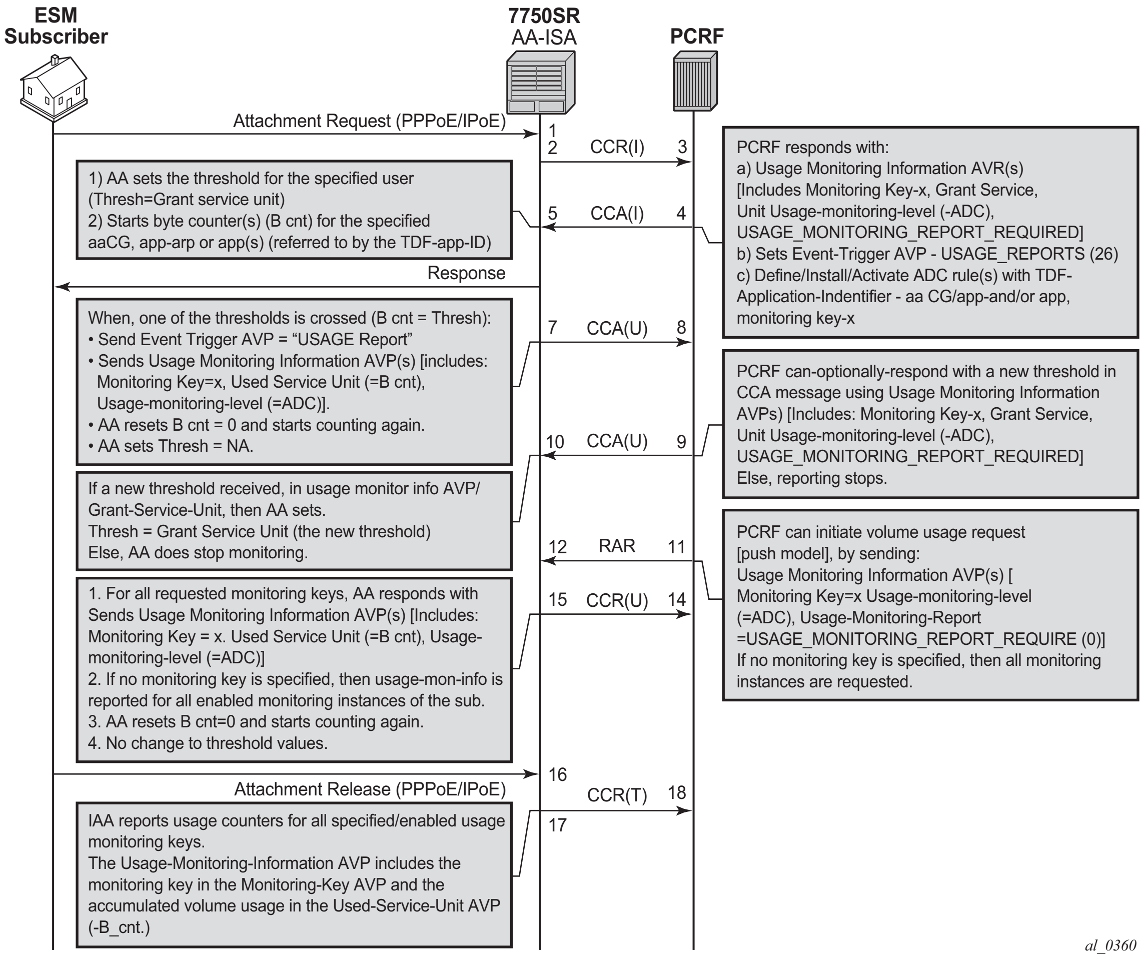

ESM Subscriber
7750SR AA-ISA
PCRF
Attachment Request (PPPoE/IPoE)
1
2
CCR(I)
3
1) AA sets the threshold for the specified user (Thresh=Grant service unit)
2) Starts byte counter(s) (B cnt) for the specified aaCG, app-arp or app(s) (referred to by the TDF-app-ID)
5
CCA(I)
4
PCRF responds with:
a) Usage Monitoring Information AVR(s) [Includes Monitoring Key-x, Grant Service, Unit Usage-monitoring-level (-ADC), USAGE_MONITORING_REPORT_REQUIRED]
b) Sets Event-Trigger AVP - USAGE_REPORTS (26)
c) Define/Install/Activate ADC rule(s) with TDF-Application-Indentifier - aa CG/app-and/or app, monitoring key-x
Response
When, one of the thresholds is crossed (B cnt = Thresh):
• Send Event Trigger AVP = "USAGE Report"
• Sends Usage Monitoring Information AVP(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 = NA.
7
CCA(U)
8
PCRF can-optionally-respond with a new threshold in CCA message using Usage Monitoring Information AVPs) [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 received, in usage monitor info AVP/ Grant-Service-Unit, then AA sets.
Thresh = Grant Service Unit (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 Information AVP(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)
CCR(T)
18
17
IAA 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_0360

Figure 22: Usage Monitoring

AA ISA (the PCEF) supports **Usage-Thresholds** AVPs that refer to the thresholds (in byte) at which point an event needs to be sent back to the PCRF (Figure 22).

No time based thresholds are supported.

AA supports **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 22 (T=7), AA sends a 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 PCRF). The counter(s) however does not stop counting newly arriving traffic. AA counters only include “admitted” packets. Any packets that got discarded by AA due to –say- policing actions- are not counted for usage-monitoring purposes.

The TDF-Application-Identifier AVP–within the ADC or PCC rule- refers to AA Charging group, AA application group or to an AA application.

TDF-Application-Identifiers (such as charging-groups) have to be manually entered at the PCRF to match AA charging groups configured on the 7750 SR.

If the TDF-Application-Identifiers refers to a name that is used for both a charging group and an application (or application group), AA monitors the charging group. In other words, AA charging group has higher precedence, than AA application group.

Supported AVPs

ADC Rule AVP

The ADC Rule install appears in the CCA and RAR messages from PCRF towards AA ISA.

- For installing a new ADC rule or modifying an ADC rule already installed, ADC-Rule-Definition AVP shall be used.
- For activating a specific predefined ADC rule, ADC-Rule-Name AVP shall be used as a reference for that ADC rule..

```
ADC-Rule-Definition ::= < AVP Header: 1094 >
                            { ADC-Rule-Name }
                            [ TDF-Application-Identifier ]
                                 ; AA charging group /application group / application name
                            [ Flow-Status ]*
                            [ QoS-Information ]*
                            [ Monitoring-Key ]
                            [ Redirect-Information ] ::= < AVP Header: 1085 >*
                                 [ Redirect-Support ] ; *
                                 [ Redirect-Address-Type ];*
                                 [ Redirect-Server-Address ];*
                            [ Mute-Notification ]*

                            *[ AVP ]
```

Note— The AVPs marked by an asterisk in the above example are not supported by AA ISA.

The TDF-application-Identifier field specifies a predefined AA charging group, application group or application name for which usage monitoring functionality is required (for a given subscriber).

The Monitoring-Key AVP (AVP code 1066), refers to a predefined (by PCRF) USAGE Monitoring AVP.

The value of the monitoring key is random. However, it should be noted that a monitoring key instance can only be used in a single ADC rule (for example, single app/app-grp/chg-grp). While the standards allow for a monitoring instance to be referenced by one or more ADC rules, AA ISA implementation restricts this to one ADC rule. Hence, if a monitoring key is referenced in one ADC rule, it cannot be referenced by another.

PCC Rule AVP

The PCC rule install appears in the CCA and RAR messages from PCRF towards AA-ISA.

- For installing a new PCC rule or modifying an PCC rule already installed, the ADC-Rule-Definition AVP shall be used.
- For activating a specific predefined ADC rule, ADC-Rule-Name AVP shall be used as a reference for that ADC rule.

```
Charging-Rule-Definition ::= < AVP Header: 1003 >
           { Charging-Rule-Name }
           [ TDF-Application-Identifier ]
           [ Monitoring-Key]
             ……….. 
          *[ AVP ]
```

Charging-Rule-Name — The name of the charging rule that contains a rule related to usage monitoring of a TDF_application_id has to start with:"AA-UM:" e.g. AA-UM: Peer to peer traffic for APN x"

TDF-application-Identifier — This field specifies a predefined AA charging group, application group or application name for which usage monitoring functionality is required (for a given subscriber).

The Monitoring-Key AVP (AVP code 1066) refers to a predefined (by PCRF) USAGE Monitoring AVP.

The value of the monitoring key is random. However, it should be noted that a monitoring key instance can only be used in a single PCC rule (e.g. single app/app-grp/chg-grp). i.e. while the standards allow for a monitoring instance to be referenced by one or more PCC rules, AA ISA implementation restricts this to one PCC rule. Hence, if a monitoring key is referenced in one PCC rule, it cannot be referenced by another.

Usage-Monitoring-Information AVP

The Usage-Monitoring-Information AVP (AVP code 1067) is of type Grouped, and it contains the usage monitoring control information.

The Monitoring-Key AVP identifies the usage monitoring control instance.

```
Usage-Monitoring-Information::= < AVP Header: 1067 >
                             [ Monitoring-Key ]
                             [ Granted-Service-Unit ]
                             [ Used-Service-Unit ]
                             [ Usage-Monitoring-Level ]
                             [ Usage-Monitoring-Report ]
                             [ Usage-Monitoring-Support ]
                            *[ AVP ]
```

Monitoring-Key-AVP

The Monitoring-Key AVP (AVP code 1066) is of type OctetString and is used for usage monitoring control purposes as an identifier to a usage monitoring control instance.

Granted-Service-Unit AVP

The Granted-Service-Unit AVP shall be used by the PCRF to provide the threshold level to the PCEF.

The CC-Total-Octets AVP shall be used for providing threshold level for the total volume, or the CC-Input-Octets and/or CC-Output-Octets AVPs shall be used for providing threshold level for the uplink volume and/or the downlink volume.

```
    Granted-Service-Unit ::= < AVP Header: 431 >
                             [ Tariff-Time-Change ]*
                             [ CC-Time ]*
                             [ CC-Money ]*
                             [ CC-Total-Octets ]
                             [ CC-Input-Octets ]
                             [ CC-Output-Octets ]
                             [ CC-Service-Specific-Units ]*
                            *[ AVP ]*
```

Note— The AVPs marked by an asterisk in the above example are not supported by AA ISA.

Used-Service-Unit AVP

This AVP is used by AA_ISA (the PCEF) to provide the measured usage to the PCRF. Reporting is done, as requested by the PCRF, in CC-Total-Octets, CC-Input-Octets an/or CC-Output-Octets AVPs of Used-Service-Unit AVP.

The Used-Service-Unit AVP contains the amount of used units measured from the point when the service became active or, if interim interrogations are used during the session, from the point when the previous measurement ended.

```
Used-Service-Unit ::= < AVP Header: 446 >
                            [ Tariff-Change-Usage ]*
                            [ CC-Time ]*
                            [ CC-Money ]*
                            [ CC-Total-Octets ]
                            [ CC-Input-Octets ]
                            [ CC-Output-Octets ]
                            [ CC-Service-Specific-Units ]*
                           *[ AVP ]*
```

Note— The AVPs marked by an asterisk in the above example are not supported by AA ISA

CC-Total-Octets AVP — The CC-Total-Octets AVP (AVP Code 421) is of type Unsigned64 and contains the total number of requested, granted, or used octets regardless of the direction (sent or received).

CC-Input-Octets AVP — The CC-Input-Octets AVP (AVP Code 412) is of type Unsigned64 and contains the number of requested, granted, or used octets that can be/have been received from the end user.

CC-Output-Octets AVP —

The CC-Output-Octets AVP (AVP Code 414) is of type Unsigned64 and contains the number of requested, granted, or used octets that can be/have been sent to the end user.

Usage-Monitoring-Level AVP

The Usage-Monitoring-Level AVP (AVP code 1068) is of type Enumerated and is used by the PCRF to indicate the level to which the usage monitoring instance applies.

If Usage-Monitoring-Level AVP is not provided, its absence shall indicate the value PCC_RULE_LEVEL (1).

The following values are defined (by the standard):

SESSION_LEVEL (0)—Not applicable for AA-ISA

PCC_RULE_LEVEL (1) —This value, if provided within an RAR or CCA command by the PCRF, indicates that the usage monitoring instance applies to one or more PCC rules. This is used in 3GPP Release 12 by the AA Usage Monitoring feature.

ADC_RULE_LEVEL (2) — This value, if provided within an RAR or CCA command by the PCRF, indicates that the usage monitoring instance applies to one or more ADC rules. This is used in 3GPP Release 11 by the AA Usage Monitoring feature.

Usage-Monitoring-Report AVP

The Usage-Monitoring AVP (AVP code 1069) is of type Enumerated and is used by the PCRF to indicate that accumulated usage is to be reported by AA ISA (the PCEF) regardless of whether a usage threshold is reached for certain usage monitoring key (within a Usage-Monitoring-Information AVP) .

The following values are defined:

> USAGE_MONITORING_REPORT_REQUIRED (0)
>
> This value, if provided within an RAR or CCA command by the PCRF indicates that accumulated usage shall be reported by the PCEF.

Note— If no monitoring keys are set, AA ISA reports all enabled monitoring instances for the subscriber.

Usage-Monitoring-Support AVP

The Usage-Monitoring-Support AVP (AVP code 1070) is of type Enumerated and is used by the PCRF to indicate whether usage monitoring shall be disabled for certain Monitoring Key.

The following values are defined:

> USAGE_MONITORING_DISABLED (0)
>
> This value indicates that usage monitoring is disabled for a monitoring key.

Event-Trigger AVP (All Access Types)

The Event-Trigger AVP (AVP code 1006) is of type Enumerated. When sent from the PCRF to the PCEF (AA ISA) the Event-Trigger AVP indicates an event that can cause a re-request of ADC rules. When sent from the PCEF to the PCRF the Event-Trigger AVP indicates that the corresponding event has occurred at the gateway.

> USAGE_REPORT (26)
>
> This value is used in a CCA and RAR commands by the PCRF when requesting usage monitoring at the PCEF (AA ISA). The PCRF also provides in the CCA or RAR command the Usage-Monitoring-Information AVP(s) including the Monitoring-Key AVP and the Granted-Service-Unit AVP.

When used in a CCR command, this value indicates that AA ISA (the PCEF) generated the request to report the accumulated usage for one or more monitoring keys. AA_ISA provides the accumulated usage volume using the Usage-Monitoring-Information AVP(s) including the Monitoring-Key AVP and the Used-Service-Unit AVP.

Note— The usage_report event must be set by the PCRF, otherwise AA ISA will not report usage-monitoring when a threshold is crossed.

Usage-Monitoring Disabled

Once enabled, the PCRF may explicitly disable usage monitoring as a result of receiving a CCR from AA ISA which is not related to reporting usage, but related to other external triggers (such as subscriber profile update), or a PCRF internal trigger.

Note that when the PCRF disables usage monitoring, AA ISA reports the accumulated usage which has occurred while usage monitoring was enabled since the last report.

To disable usage monitoring for a monitoring key, the PCRF sends the Usage-Monitoring-Information AVP including only the applicable monitoring key within the Monitoring-Key AVP and the Usage-Monitoring-Support AVP set to USAGE_MONITORING_DISABLED.

When the PCRF disables usage monitoring in a RAR or CCA command, AA ISA sends a new CCR command with CC-Request Type AVP set to the value UPDATE_REQUEST and the Event-Trigger AVP set to USAGE_REPORT to report accumulated usage for the disabled usage monitoring key(s).

Termination Session

At AA ISA subscriber’s session termination, AA ISA sends the accumulated usage information for all monitoring keys for which usage monitoring is enabled in the CCR command with the CC-Request-Type AVP set to the value TERMINATION_REQUEST.

Cflowd AA Records

AA ISA allows cflowd records to be exported to an external cflowd collector. The cflowd collector parameters (such as IP address and port number) are configured per application assurance group. All cflowd records collected for both volume and per-flow performance (TCP and/or audio video) are exported to the configured collector(s). AA ISA supports cflowd Version 10/ IPFIX.

A cflowd record is only exported to the collector once the flow is closed/terminated.

TCP Application Performance

AA ISA allows an operator to collect per flow TCP performance statistics to be exported through cflowd v10/IPFIX.

The operator can decide to collect TCP performance for sampled flows within a TCP enabled group-partition-application/application-group. The flow sampling rate is configurable on per ISA-group level. For example a flow sample rate of 10 means that every 10th TCP flow is selected for TCP performance statistics collection. Anytime a flow is sampled (selected for TCP performance statistics collection) its mate flow in reverse direction is also selected. This allows collectors to correlate the results from the two flows and provide additional statistics (such as round-trip delay). Per-flow cflowd TCP performance records are exported to the configured collector(s) upon flow closure. The system can gather per flow TCP performance statistics for up to 307,200 concurrent flows.

Two configurable TCP flow sampling rates are available per AA ISA group. Applications and/or Application groups selected for TCP performance monitoring can use of one these two sampling rates. For example, important applications are assigned high sampling rates, while other TCP applications are subjected to TCP performance monitoring using a lower flow sampling rate.

Per-flow TCP performance can be enabled (or disabled), using one of two configurable sampling rates, per application/app-group per partition per AA ISA-group.

Volume Statistics

AA ISA allows an operator to collect per flow volume statistics to be exported for any group partition. The packet sampling rate is configurable per AA- ISA-group level. For example, a packet sample rate of 10 means that one of every 10 packets is selected for volume statistics collection. If a flow has at least a single packet sampled for cflowd volume statistics, its per-flow cflowd volume record is exported to the configured collector upon flow closure.

Comprehensive Statistics

AA ISA allows an operator to collect per flow comprehensive statistics to be exported through cflowd v10/IPFIX.

This record type facilitates two deployments scenarios:

1. HTTP host and device info — Using the new performance cflowd, operators can collect statistics regarding the host names (used, for example, in HTTP GET methods) and device types being used in different flows within the network. These per flow statistics are exported via IPFIX v10 cflowd formatted records to a cflowd collector (such as RAM DCP) to enable intelligent reporting on devices and host fields.
2. Scaling of cflowd — In some situation, operators are mainly interested in augmenting the 5 Tuple IP flow information with AA classification of the flows in terms of application/ application group. While AA volume cflowd provides such a function, however it is enabled at AA-partition level, covering all traffic within a partition, which then prohibits the use of high sampling rates. Using AA comprehensive flow- sampled cflowd mechanism, operators can target (or exclude), within an AA partition, certain applications (/application groups) for sampling. Hence providing finer control at the application/ application group level, rather than at the partition level (case of volume cflowd).

The operator can decide to collect comprehensive statistics for sampled flows within an enabled group-partition-application/application-group. Parameters such as flow's applications/application groups, host fields (applicable to HTTP traffic only), subscriber's device type (when available), along with other general statistics such as flow's bytes/packets counts are collected in a comprehensive cflowd record.

The flow sampling rate is configurable on per ISA group level. For example, a flow sample rate of 10 means that every 10th flow is selected for comprehensive statistics collection. Any time a flow is sampled (selected for comprehensive statistics collection) its mate flow in reverse direction is also selected. The two flows are exported in a single cflowd record.

Per-flow comprehensive can be enabled (or disabled), using one of two configurable sampling rates, per application/app-group per partition per AA ISA-group.

Applications and/or Application groups selected for comprehensive statistics gathering can use one these two sampling rates. For example, important applications are assigned high sampling rates, while other applications are subjected to a lower flow sampling rate.

Audio/Video (A/V) Application Performance

AA ISA integrates a third party audio/video performance measurement software stack to perform VoIP and video conferencing MOS-related measurements for RTP based A/V applications.

A passive monitoring technology estimates transmission quality of voice and video over packet technologies by considering the effects of packet loss, jitter and delay in addition to the impairments caused by encoding/decoding technology. A rich set of diagnostic data is provided that can be used to help network managers identify a variety of problems that could impact the quality of voice and video streams and/or service level agreements (SLAs).

This feature provides:

- Call quality analysis using optimized ITU-T G.107, such as listening and conversational quality MOS and R-factor scores – MOS-LQ, MOS-CQ R-LQ and R-CQ.
- Measurements of perceptual effects of burst packet loss and recency using ETSI TS 101 29-5 Annex E Extensions
- Reporting of RTCP XR (RFC 3611, *RTP Control Protocol Extended Reports (RTCP XR)*) VoIP metrics payloads.

Once a flow terminates, AA ISA formats the flow MOS parameters into a cflowd record and forwards the record to a configured IPFIX /10 cflowd collector (such as 5670 RAM). The collector then summarizes these records using route of interest information (source/destinations). In addition, RAM provides the user with statistics (min/max/avg values) for the different performance parameters that are summarized.

Two configurable RTP flow sampling rates are available per AA ISA group. Applications and/or Application groups selected for RTP performance monitoring can use one of these two sampling rates. For example, important applications (such as Cisco's Telepresence video conferencing or operator's VoIP service) are assigned high sampling rates, while other RTP applications are subjected to RTP performance monitoring using a lower flow sampling rate.

Like TCP performance, per flow audio/video performance can be enabled (or disabled), using one of two configurable sampling rates, per application/app-group per partition per AA ISA-group.

The operator can decide to collect RTP A/V performance for sampled RTP flows within an RTP A/ V enabled group-partition-application/application-group. The two available flow sampling rates is are configurable on per ISA group level. For example a flow sample rate of 10 means that every 10th RTP flow is selected for RTP performance statistics collection. Anytime a flow is sampled (selected for RTP A/V performance statistics collection) its mate flow reverse direction is also selected. When RTP dynamic payload types (RTP "PT") are used, only flows that use SIP to signal RTP codec can be selected for RTP performance measurement. Flows that use static RTP payload types can be selected for performance measurement regardless of the signaling channel used to setup the call. The system can gather per flow RTP A/V performance statistics for up to 6000 voice calls.

Application QoS Policy (AQP)

An AQP is an ordered set of entries defining application-aware policy (actions) for IP flows diverted to a given AA ISA group. The IP flow match criteria are based on application identification (application or application group name) but are expected to use additional match criteria such as ASO characteristic value, IP header information or AA subscriber ID, for example.

When application service option characteristic values are used in application profiles, the characteristics values can be further used to subdivide an AQP into policy subsets applicable only to a subset of AA subscribers with a given value of an ASO characteristic in their profile. This allows to, for example, subdivide AQP into policies applicable to a specific service option (MOS iVideo Service), specific subscriber class (Broadband service tier, VPN, Customer X), or a combination of both.

A system without AQP defined will have statistics generated but will not impact the traffic that is flowing through the system. However, it is recommended that an AQP policy is configured with at least default bandwidth and flow policing entries to ensure a fair access to AA ISA bandwidth/ flow resources for all AA subscribers serviced by a given AA ISA.

AQP rules consist of match and action criteria:

- Match: Refers to application identification determined by application and application group configuration using protocol signatures and user-configurable application filters that allow customers to create a wide range of identifiable applications. To further enhance system-wide per subscriber/service management user configurable application groups are provided.

  An AQP consists of a numbered and ordered set of entries each defining match criteria including AND, NOT and wildcard conditions followed by a set of actions.

  ```
  AQP Entry <#> = <Match Criteria> AND <Match Criteria> <action>
  <action>
  ```

  OR match conditions are supported in AQP through defining multiple entries. Multiple match criteria of a single AQP entry form an implicit AND function. An AQP can be defined for both recognized and unrecognized traffic. IP traffic flows that are in the process of being identified have a default policy applied (AQP entries that do not include application identification or IP header information). Flows that do not match any signatures are identified as unknown-tcp or unknown-udp and can have specific policies applied (as with any other protocol).
- Actions: Defines AA actions to be applied to traffic, a set of actions to apply to the flows like bandwidth policing, packet discards, QoS remarking and flow count or/and rate limiting.

AQP Match Criteria

Match criteria consists of any combination of the following parameters:

- The source/destination IP address and port, or IP-prefix list
- Application name
- Application group name
- Charging group name
- One or more application service option characteristic and value pairs
- Direction of traffic (subscriber to network, network to subscriber, or both, or spoke SDP)
- DSCP name
- AA subscriber (ESM, DSM, or transit subscriber, SAP or spoke SDP)
- ip-protocol-num field, which when used in AQP matches allows more precise control of match criteria, e.g. to specify port or IP address matches specifically for either TCP or for UDP.

AQP entries with match criteria that exclusively use any combination of ASO characteristic and values, direction of traffic, and AA subscriber define default policies. All other AQP entries define application aware policies. Both default and application aware policies. Until a flow's application is identified only default policies can be applied.

AQP Actions

An AQP action consists of the following action types. Multiple actions are supported for each rule entry (unlike ip-filters):

- Dual or single-bucket bandwidth rate limit policer
- Drop (discard)
- Error drop
- Flow count limit policer
- Flow setup rate limit policer
- Fragment drop
- HTTP enrichment
- HTTP error redirect
- HTTP notification
- HTTP redirect
- Source mirror for an existing mirror service

- Remark QoS (one or a combination of discard priority, forwarding class name, DSCP). When applied, ingress marked FC and discard priority is overwritten by AA ISA and the new values are used during egress processing (for example, egress queueing or egress policy DSCP remarking). For MPLS class-based forwarding, ingress-marked FC is still used to select an egress tunnel.
- None (monitor and report only)
- Session filter
- URL-Filter (ICAP Category Based URL Filtering)
- GTP filter
- SCTP filter

Any flow diverted to an ISA group is evaluated against all entries of an AQP defined for that group at flow creation (default policy entries), application identification completion (all entries), and an AA policy change (all flows against all entries as a background task). Any given flow can match multiple entries, in which case multiple actions will be selected based on the AQP entry's order (lowest number entry, highest priority) up to a limit of:

- 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;
  - → up to 12 flow policers.

AQP entries the IP flow matched, that would cause the above per-IP-flow limits to be exceeded are ignored (no actions from that rule are selected).

Examples of some policy entries may be:

- Limit the subscriber to 20 concurrent Peer To Peer (P2P) flows max.
- Rate limit upstream total P2P application group to 400kbps.
- Remark the voice application group to EF.

Application Assurance Policers

The rate limit (policer) policy actions provide the flow control mechanisms that enable rate limiting by application and/or AA subscriber(s).

There are four types of policers:

- Flow rate policer monitors a flow setup rate.
- Flow count limits control the number of concurrent active flows
- Single-rate bandwidth policers monitor bandwidth using a single rate and burst size parameters.
- Dual-rate bandwidth rate policers monitor bandwidth using CIR/PIR and CBS/MBS. These can only be used at the per-subscriber granularity.

  Once a policer is referred to by an AQP action for one traffic direction, the same policer cannot be referred to in the other direction. This also implies that AQP rules with policer actions must specify a traffic direction other than the "both" direction.

Table 13 illustrates a policer's hardware rate steps for AA ISA:

Table 13: Policer's Hardware Rate Steps for AA ISA

Hardware Rate Steps	Rate Range (Rate Step x 0 to Rate Step x 127 and max)
0.5Gb/sec	0 to 64Gb/sec
100Mb/sec	0 to 12.7Gb/sec
50Mb/sec	0 to 6.4Gb/sec
10Mb/sec	0 to 1.3Gb/sec
5Mb/sec	0 to 635Mb/sec
1Mb/sec	0 to 127Mb/sec
500Kb/sec	0 to 64Mb/sec
100Kb/sec	0 to 12.7Mb/sec
50Kb/sec	0 to 6.4Mb/sec
10Kb/sec	0 to 1.2Mb/sec
8Kb/sec	0 to 1Mb/sec
1Kb/sec	0 to 127Kb/sec

Policers are unidirectional and are named with these attributes:

- Policer name
- Policer type: single or dual bucket bandwidth, flow rate limit, flow count limit.
- Granularity: select per-subscriber or system-wide
- Parameters for flow setup rate (flows per second rate)
- Parameters for flow count (maximum number of flows)
- Rate parameters for single-rate bandwidth policer (PIR)
- Parameters for two-rate bandwidth policer (CIR, PIR)
- PIR and CIR adaptation rules (min, max, closest)
- Burst size (CBS and MBS)
- Conformant action: allow (mark as in-profile)
- Non-conformant action: discard, or mark with options being in profile and out of profile

Policers allow temporary over subscription of rates to enable new sessions to be added to traffic that may already be running at peak rate. Existing flows are impacted with discards to allow TCP backoff of existing flows, while preventing full capacity from blocking new flows.

Policers can be based on an AQP rule configuration to allow per-app-group, per-AA-sub total, per AA profile policy per application, and per system per app-group enforcement.

Policers are applied with two levels of hierarchy (granularity):

- Per individual AA subscriber
  - → Per-AA-sub per app group/application or protocol rate
  - → Per-AA-sub per application rate limit for a small selection of applications
  - → Per-AA-sub PIR/CIR. This allows the AA ISAAA ISA to emulate IOM ingress policers in from-sub direction.
- Per system (AA ISA or a group of AA subscribers)
  - → Total protocol/application rate
  - → Total app group rate

Flows may be subject to multiple policers in each direction (from-subscriber-to-network or from -network-to-subscriber).

In Figure 23, Application Assurance policers are applied after ingress SAP policers. Configuration of the SAP ingress policers can be set to disable ingress policing or to set PIR/CIR values such that AA ISA ingress PIR/CIR will be invoked first. This enables application aware discard decisions, ingress policing at SAP ingress is application blind. However, this is a design/implementation guideline that is not enforced by the node.

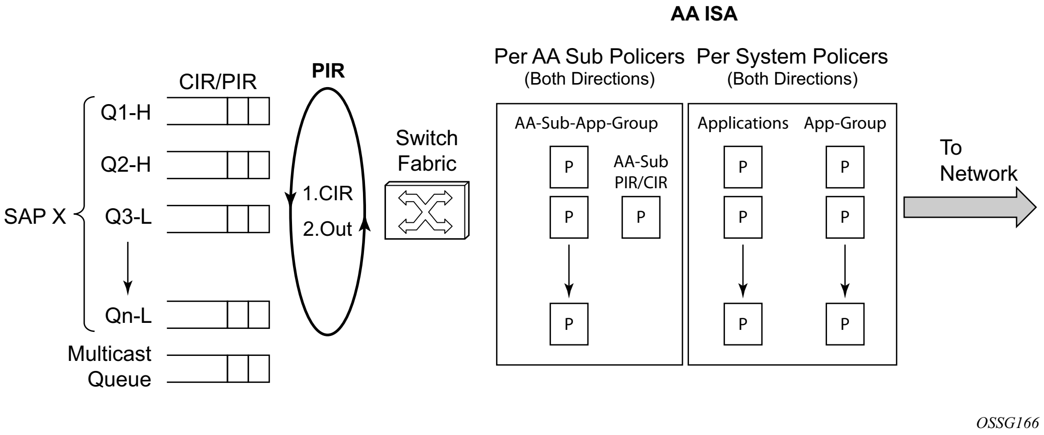


Figure 23: From-AA-Sub Application-Aware Bandwidth Policing

In the to-aa-sub direction (Figure 24), traffic hits the AA ISA policers before the SAP egress queuing and scheduling. This allows application aware flow, AA subscriber and node traffic policies to be implemented before the Internet traffic is mixed with the other services at node egress. Note that AA ISA policers may remark out-of-profile traffic which allows preferential discard at an IOM egress congestion point only upon congestion.

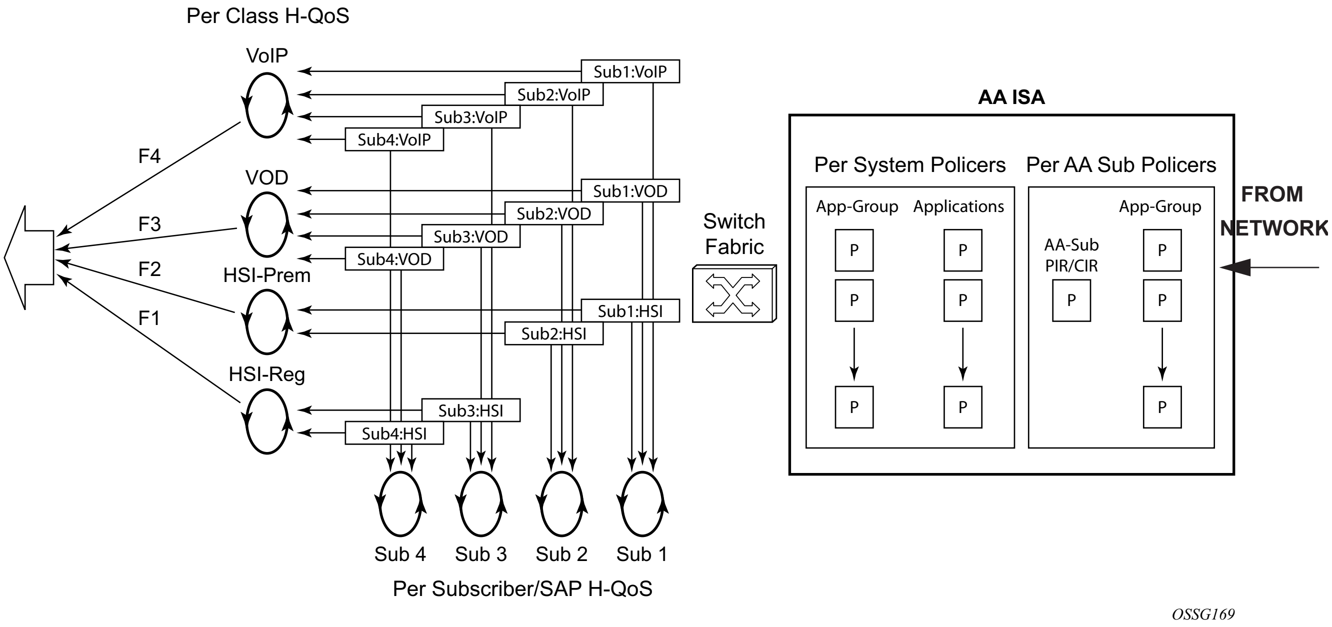


Figure 24: To-AA-Sub Application-Aware Bandwidth Policing

Time of Day Policing Adjustments

Time-of-day changes to Application Assurance policing rates are supported through the use of time-of-day override in the policers, up to eight overrides. Up to eight overrides can be configured per policers each using either a daily or weekly time-range. The adjusted policing limits are applied immediately to any pre-existing or new flows.

Application Assurance HTTP Redirect

AA HTTP Policy Redirect

With AA ISA HTTP policy based redirect feature, when HTTP flows are blocked, the user is directed to a web portal that displays relevant messages to indicate why the HTTP traffic is blocked, such as.: time of day policy to block youtub.com, top-up request, etc.

Without HTTP policy redirect, when HTTP flows are blocked, the subscriber application retries and before it times-out. The subscriber in most cases is unaware of the cause of this timeout. Frustration builds up and leads to increase calls to IT / operators help desk. That in turn, results in an increase in operator's OPEX. Hence, with HTTP policy redirect, the operator realizes savings related to decrease in network loading associated with retries as well as IT HELP desk OPEX savings. Above all, the operator retains happier and less frustrated customers / clients.

AA ISA provides full customer control to configure an AQP action that redirects traffic that matches the AQP match criteria. Hence, the HTTP redirect service can applied at any level (application, application group, specific subscribers, specific source IP addresses) or any other AQP match criteria.

To illustrate, say the operator configures www.poker.com as a "gambling" app-group.

The operator configures AA_ISA to drop and redirect all HTTP traffic classified under "gambling" app-group to www.redirect.isp.com. When a client/subscriber initiates an HTTP GET for www.poker.com. Traffic to poker.com is dropped at the AA ISA. AA ISA issues a redirect to the client. [in the redirect, information about the user are encoded in the PATH message, such as www.poker.com, sub-ID, sub-type, reason for redirect (=AQP drop action) AA application name]. Client, unaware of the drop, responds to the redirect.

Redirect landing page explains to the user why the page at www.poker.com is not accessible. See Figure 25.

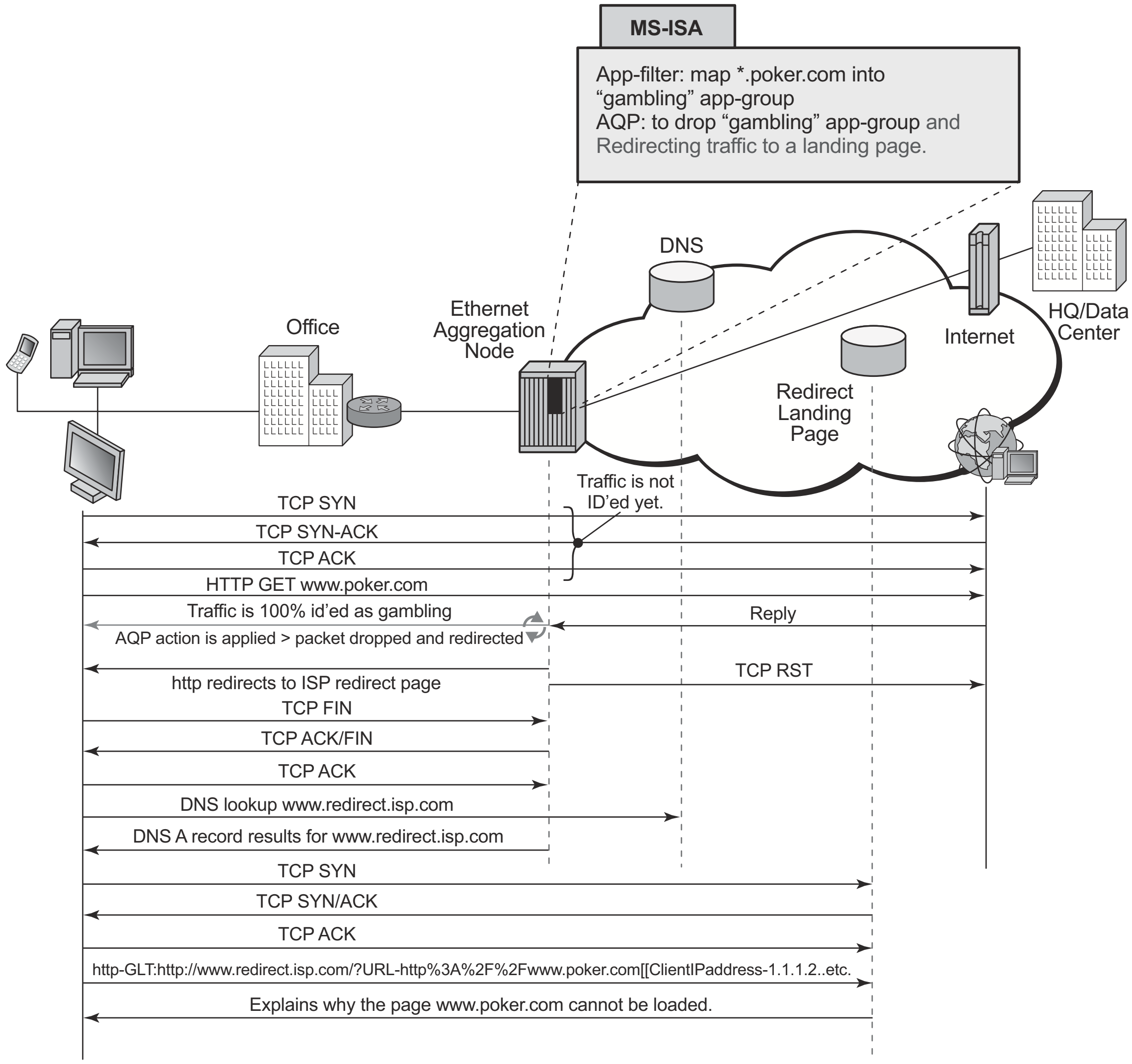

MS-ISA
App-filter: map *.poker.com into
“gambling” app-group
AQP: to drop “gambling” app-group and
Redirecting traffic to a landing page.
DNS
Office
Ethernet
Aggregation
Node
Internet
HQ/Data
Center
Redirect
Landing
Page
Traffic is not
ID’ed yet.
TCP SYN
TCP SYN-ACK
TCP ACK
HTTP GET www.poker.com
Traffic is 100% id’ed as gambling
AQP action is applied > packet dropped and redirected
Reply
TCP RST
http redirects to ISP redirect page
TCP FIN
TCP ACK/FIN
TCP ACK
DNS lookup www.redirect.isp.com
DNS A record results for www.redirect.isp.com
TCP SYN
TCP SYN/ACK
TCP ACK
http-GLT:http://www.redirect.isp.com/?URL-http%3A%2F%2Fwww.poker.com[[ClientIPaddress-1.1.1.2..etc.
Explains why the page www.poker.com cannot be loaded.

Figure 25: HTTP Redirect Due To URL Block

AA ISA allows the operator to configure HTTP redirect policies. An HTTP redirect policy contains, most importantly, the HTTP host to be used for the redirect. Within the AQP actions, such polices can be linked (like policers). Redirect will take place only if the AQP configured matching criteria is met and the HTTP flow is dropped (due to other AQP actions, such as “drop”, flow-count/rate policers). Obviously, redirect only applies to HTTP traffic. Non-HTTP flows (even if the conditions above are met) are not redirected (no redirect for RTSP traffic).

The HTTP redirect policy includes an option for TCP-client-reset. This is used to improve the end-user experience when TCP traffic that cannot be HTTP redirected is blocked. Resetting the client TCP session avoids the client waiting for tcp session timeout. The ISA will initiate a TCP reset towards the client if the AA policy results in an http-redirect with packet drop but the HTTP redirect cannot be delivered.   Scenarios for this include blocked HTTPs (TLS) sessions, blocking of non-HTTP traffic, and blocking of existing flows after a policy re-evaluate of an existing subscriber.   A mid-session policy change to redirect & block traffic for a sub will cause a TCP reset of existing non-http tcp sessions when the next packet for those sessions arrives. For example, when the packet is dropped.

AA HTTP 404 Redirect

HTTP status code-based redirect feature provides error resolution and search technology that enhances the Internet experience for end customers while generating new revenue stream for the ISP.

Alcatel-Lucent’s AA ISA HTTP status code-based redirect feature, along with its partners Barefruit or Xercole, replaces unhelpful DNS and HTTP error messages with relevant alternatives, giving the user a search solution rather than a no direction. Customers benefit from an improved surfing experience as they are served relevant results that can help them find what they were looking for. The ISP, on the other hand, receives a share of the search revenue.

Every time an end-user clicks on a broken link (Page Not Found), an error page displays. Frequently, a search provider produces results, through a browser plug-in, for that user. This generates revenue for the search provider if the user clicks on a paid link.

With AA ISA HTTP status code-based redirect feature, the user sees high-quality, relevant search results. In addition, instead of the search provider receiving all of the revenue, the ISP is paid every time a user clicks on a sponsored link.

AA ISA provides full customer control to configure an AQP action that redirects traffic that matches the AQP match criteria (Figure 26). Hence, the HTTP redirect service can applied at any level (application, application group, specific subscribers, specific source IP addresses) or any other AQP match criteria.

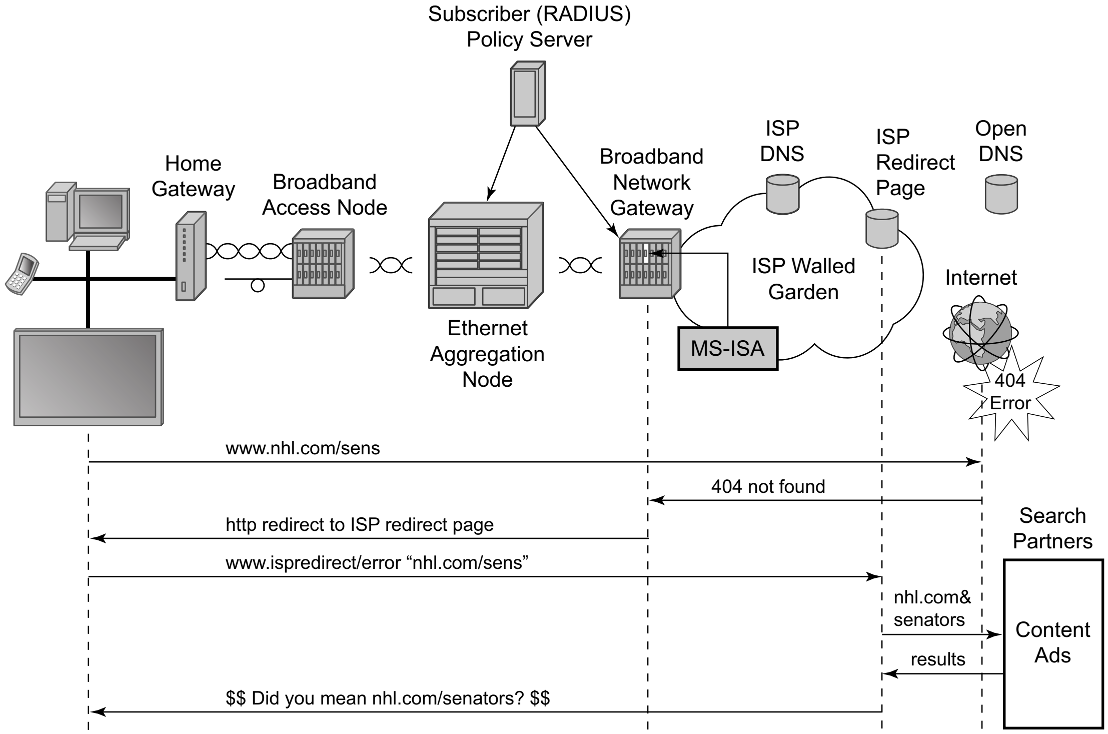


Figure 26: AQP Actions

HTTP headers are intercepted by AA ISA on the return path from the requested web site. If the HTTP status code is a non custom 404, then the response is replaced with JavaScript that redirects the client to the Contextual Analysis Servers (Barefruit server). This redirect contains details of the original URI that gave rise to the 404 error.

The operator can configure AA ISA HTTP 404 redirect to use either Barefruit or Xerocole partner contextual analysis servers. A redirect policy can be defined once at the AA group level (similar to policers), and then referenced as many times as needed in AQP actions. The system allows a maximum of one HTTP 404 redirect policy per AA group.

ICAP - Large Scale Category based URL Filtering

Large scale URL filtering is a common content filtering requirement from broadband, mobile, and business vpn operators that allows them to solve various use cases such as:

- Category based URL filtering: typically offered as an opt-in service by broadband or mobile operators to protect the subscribers from accessing selected category of URLs, such as, gambling, drugs, pornography, racism etc.
- Managed URL filtering service for Business VPN to prevent employee from accessing specific content.

Application Assurance provides both a cost efficient and best of breed content filtering solution to solve these use cases by enabling offline dedicated web filtering servers though the Internet Content Adaptation Protocol (ICAP). Using application assurance the operator does not need to deploy costly inline filtering appliances or a limited client software solution requiring maintenance and updates for a growing number of computing devices and operating systems (for example, laptop, smartphone, smartTV, tablets).

A high level packet flow diagram of the solution is shown in Figure 27. The AA ISA is the ICAP client and performs inline L7 packet processing functions while the ICAP application server is used for URL filtering offline, thus the application server does not need to be inserted in the data flow:

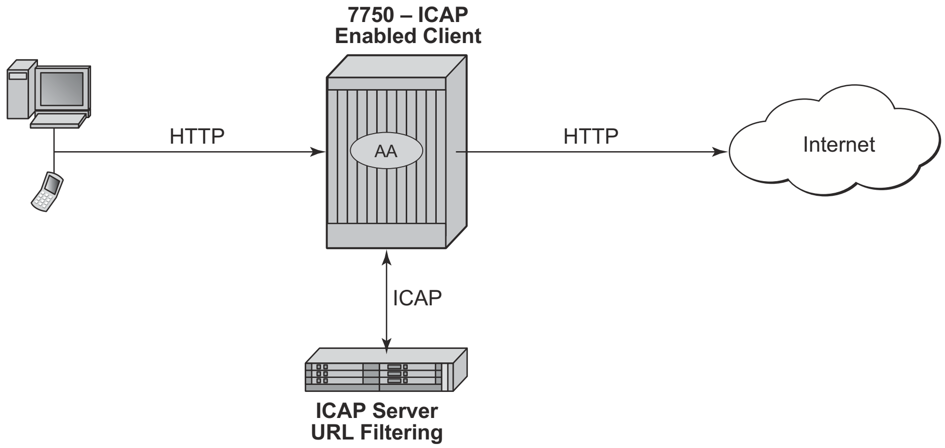


Figure 27: ICAP High Level Flow Diagram

The 7750 uses the Application Assurance capabilities to extract the URL from the subscriber's HTTP/HTTPs request and send an ICAP rating request to the ICAP server along with the subscriber-id information. The ICAP server can then return an accept or redirect response based on various criteria such as subscriber profile, URL categories, whitelist, blacklist, time of the day.

The ICAP response received by the 7750 ICAP client is used to either accept, redirect, or block the flow.

Note:

- Each HTTP request within a TCP flows are sent to the ICAP server for rating.
- HTTPs (SSL/TLS) ICAP Url-Filtering is limited to the domain name information.

Local URL-List Filtering

Service providers may need to apply network wide URL filtering policies to prevent subscribers from accessing illegal content in the following context:

- Court order URL takedown
- Child pornography related content
- Government mandated URL takedown list

Operators can use AA to comply with these regulatory requirements typically driven by government or court order to control the access to specific URLs hosting illegal content. In the context of child protection the operator may be required or incited to provide this filtering.

Local url-list filtering is applied network-wide to all subscribers. This solution provides a cost-efficient method by storing the list of URLs to be filtered on the system compact flash. Therefore, using the AA-ISA ICAP functionality along with an external server is not necessary.

The ISA-AA url-filter local url-list filtering policy provides URL control capability using a list of URLs contained in a file stored on one of the system's compact flash cards. The 7x50 uses the Application Assurance capabilities to extract the URL from the subscriber's HTTP request and compares it to the list of URLs contained in this local file. If a match is found the subscriber flow 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 3DES encrypted file formats to protect the content of the list.

URL-List Update

The system supports a flexible mechanism to upgrade a local url-list automatically using either CRON or the 5620 SAM to comply with the regulatory requirements in terms of list upgrade frequency.

HTTP/HTTPS

Each HTTP request within a TCP flow is filtered by the AA ISA. For HTTPS traffic, the system extracts the domain name information contained in the SSL/TLS server name.

HTTP Header Enrichment

AA ISA supports modifications of the HTTP header for traffic going to specific user configured sites (URLs/IPs); in order to add Network based information, such as subscriber ID to the HTTP header. These sites use this information to authenticate the user and/or present the user with user-specific information and services.

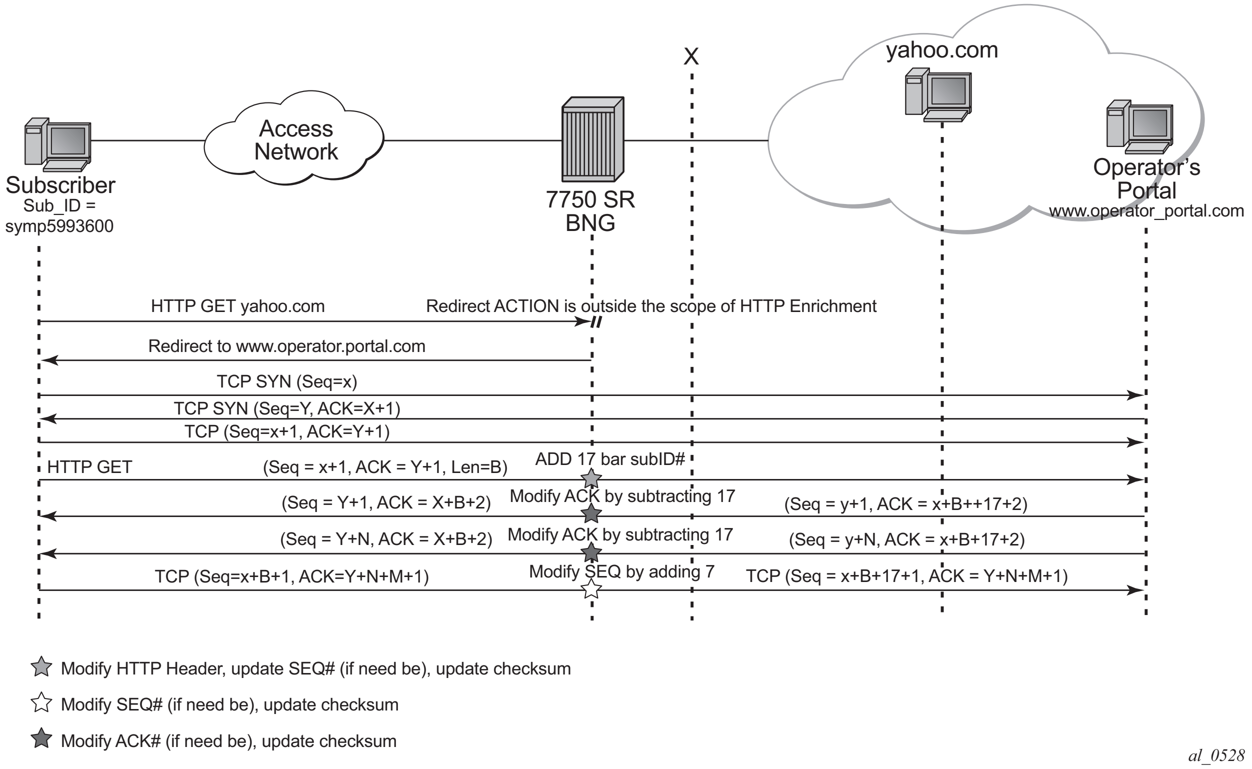

X
yahoo.com
Access Network
Subscriber
Sub_ID = symp5993600
7750 SR BNG
Operator's Portal
www.operator_portal.com
HTTP GET yahoo.com
Redirect ACTION is outside the scope of HTTP Enrichment
Redirect to www.operator.portal.com
TCP SYN (Seq=x)
TCP SYN (Seq=Y, ACK=X+1)
TCP (Seq=x+1, ACK=Y+1)
HTTP GET
(Seq = x+1, ACK = Y+1, Len=B)
ADD 17 bar subID#
(Seq = Y+1, ACK = X+B+2)
Modify ACK by subtracting 17
(Seq = y+1, ACK = x+B++17+2)
(Seq = Y+N, ACK = X+B+2)
Modify ACK by subtracting 17
(Seq = y+N, ACK = x+B+17+2)
TCP (Seq=x+B+1, ACK=Y+N+M+1)
Modify SEQ by adding 7
TCP (Seq = x+B+17+1, ACK = Y+N+M+1)
Modify HTTP Header, update SEQ# (if need be), update checksum
Modify SEQ# (if need be), update checksum
Modify ACK# (if need be), update checksum
al_0528

Figure 28: HTTP Enrichment

In Figure 28, the operator configures the AA ISA to insert the subscriber ID into the HTTP header for all the HTTP traffic destined to the operator portal (designated by server IP and/or HTTP host name). Traffic going to other destinations, such as yahoo.com, does not get enriched. To support this, AA introduces a new AQP action called **HTTP_enrich** that allows the operator to enrich traffic that satisfies the AQP-matching conditions.

The operator can configure multiple HTTP enrichment policies that get applied to traffic going to different destinations. For example, HTTP traffic destined to “xyz.com”, gets User’s IP inserted into the header, while traffic going to “billing.xyz.com” gets enriched with subscriber ID and user’s ip address.

AA ISA supports insertion of one or more of the following parameters/fields into the HTTP header: User’s IP@, subscriber ID and configurable static string fields. The text preceding the inserted field is fully configurable. For example, sub-ID = 1243534666 or x-sub-ID = 1243534666.

AA supports enrichment of all HTTP methods, such as GET, POST, etc. AA enriches HTTP traffic without having to terminate the TCP session (for example, does not act as a proxy). In this way, AA enrichment function does not intervene with other TCP acceleration functions/appliances that could be deployed by the operator.

For configured enriched fields, operators can optionally configure AA ISA to perform MD5 hashing and/or anti-spoofing. Anti-spoofing, once enabled, ensures that only the fields enriched by AA are valid. Anti-spoofing is applicable only to subscriber-id field.

AA statistics reflect post header enrichment packet sizes.

AA HTTP enrichment functionality has the following caveats:

- To handle the case of TCP retransmission, AA ISA implements an enrichment window of size =5. If a retransmission of a packet occurs outside the last 5 enriched packets, no enrichment takes place.
- No enrichments of corrupted packets, AA ISA-cut-through and/or out-of-order fragmentation
- Out of sequence packets are not enriched. For example, if AA –ISA receives out-of-sequence HTTP requests: REQ2,REQ1,REQ3; only REQ2 and REQ3 can be enriched
- No enrichment takes place if by enriching, the resulting packet size will exceed the configured MTU size. AA ISA does not perform fragmentation.
- AA ISA does not support header enrichment for WAP1.x, RTSP or SIP headers.
- AA TCP performance measurements cannot co-exist with HTTP enrichment. Enriched flows are ineligible for TCP performance sampling. If a flow is selected for TCP performance measurements and is later enriched, then TCP performance measurements cease to continue.

- Enrichment can be applied as an action to any AQP entry, subject to:
  - → The matching conditions for the AQPs cannot include a specific HTTP protocol (such as, protocol eq HTTP_video). In other words, applications which require a specific HTTP protocol type (video/flash) are not considered for enrichment.
  - → Within the same AQP entry, the enrichment action cannot co-exist with any other AQP action (such as mark/police, etc.).
  - → AQP hit counter is not updated based on executing an HTTP enrichment action of an AQP.

HTTP In Browser Notification

The AA ISA HTTP notification policy-based feature enables the operator to send in browser notification messages to their subscribers. The notification format can either be an overlay, a web banner, or a splash page, which makes HTTP notification less disruptive than standard HTTP redirection for the subscriber; both the original content and the notification message can be displayed at the same time while browsing.

There is a wide range of notification use cases in Broadband and Wifi networks to use this functionality such as fair use policy threshold warning, marketing and monetization messages, late bill payment notice, copyright infringement notice and operational outages.

The solution is based on two primary components, the AA ISA responsible for specific packet manipulation and a messaging server. The messaging server controls the message format and its content while the AA ISA modifies selected HTTP flows so that the subscriber transparently downloads a script located on the messaging server. This script is then executed by the web browser to display the notification message. The AA ISA only select specific HTTP request flows meeting the criteria of a web browser session compatible with in browser notification messages.

A high level view of the typical network elements involved in HTTP in browser notifications are describe in Figure 29:

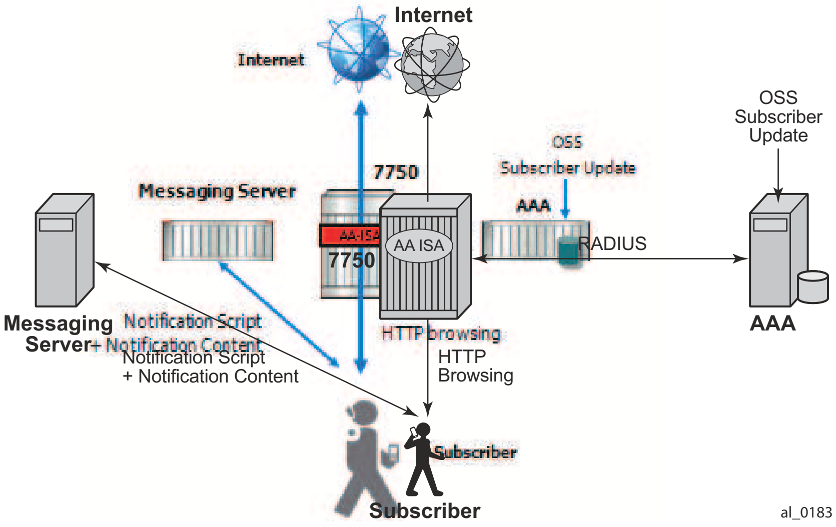


Figure 29: HTTP in Browser Notification - High Level

The AA ISA provides full subscriber control to configure an AQP action enabling HTTP notification policy based on specific app-profiles attributes (ASO characteristics), application, or application group. The operator can dynamically modify the subscriber policy from the policy manager to enable/disable HTTP notification during the lifetime of the subscriber.

Notification Interval

The notification can be configured to be displayed either once during the lifetime of the subscriber or at configured minimum interval (in minutes). When an interval in minutes is selected, the subscriber will continue to receive notifications messages while browsing.

Success Verification

The system identifies successful and failed notifications. In the event the notification is not successful, the system will automatically retry notifying the subscriber at the next flow that meets the criteria of a web browser session.

HTTP Notification Example

To illustrate how HTTP notification works, the steps below describe a typical usage quota notification example with a subscriber reaching its monthly quota:

- AAA identifies that a particular subscriber is now over its quota.
- A Radius CoA message is sent from the AAA to the 7750 to modify the subscriber app-profile in order to enable HTTP notification.
- The AA ISA modifies the subscriber profile and enable HTTP notification for this subscriber.
- The notification message is displayed in the subscriber web browser while browsing (in the form of an overlay or web banner). The content of the notification includes a link to the operator web portal to acknowledge the reception of the overage notification.
- Until the subscriber clicks on the acknowledgment link, the AA ISA will continue to execute the same policy so that notification messages are displayed in the subscriber web browser at the configured interval.
- Once the subscriber has clicked on the link provided in the notification message, the provider OSS system updates the AAA which then sends a new CoA message to the 7750 in order to modify the subscriber app-profile.
- The AA ISA modifies the subscriber app-profile and disables HTTP notifications for this subscriber.

HTTP Notification Customization through Radius VSA

The operator can customize the notification message per subscriber through the use a new radius VSA returned either at the subscriber creation time or within a CoA. This new VSA is a string appended automatically at the end of the script-url request made by the subscriber towards the messaging server, and it is not interpreted by the AA ISA. When received by the messaging server, it can be used to return specific content to the subscriber.

As an example, the HTTP Notification can be customized using the RADIUS VSA to display location based information, and the messaging server can use this data to display content based on the desired location:

- Alc-AA-Sub-Http-Url-Param RADIUS VSA: location=SohoStation
- 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=<aa-subscriber-id>&VSA=&location=SohoStation

AA Mirroring to Offline Processing

Some deployments require specialized offline processing not provided by the 7750 SR/7450 ESS AA. An example of such processing is Lawful Intercept (LI) traffic content processing or using an offline appliance. To enable such capabilities in a highly-scalable fashion that minimizes traffic seen by the offline device, the 7750 SR/7450 ESS AA allows operators to use an AQP action to mirror traffic conditioned by both application and AA subscriber context, so detailed content processing can be performed only for AA subscribers and applications of choice. The content processing equipment generally needs to see the entire traffic stream for a given application, therefore, the entire application's traffic is mirrored including packets that have not yet been identified. Optionally, only traffic positively identified can be mirrored as well.

Although similar functionality could be achieved by mirroring service or a SAP, the total bandwidth and added complexity that an offline appliance would need to handle extra bandwidth makes such a solution more costly and harder to scale.

Since the application mirroring is an additional function independent from all other AA functions provided on the ESS/SR, the in-line deployed AA ISA modules not only reduce the amount of traffic the offline device must see, but also allows in-line policy enforcement actions with application awareness once the offline devices triggers such a policy change. For example, AA subscriber traffic for an application or applications being mirrored can be quarantined while the remaining traffic remains unaffected.

Figure 30 depicts an example of application mirroring to a specialized offline appliance for further processing.

1. AA subscriber traffic contains applications requiring specialized offline appliance processing that requires Layer 2 — Layer 7 application identification.
2. AA ISA with AQP configured:

   Match:

   → Application for offline processing for selected subscribers (downstream only, upstream only, or both).

   Action:

   → Mirror source for application's IP packets into a mirror service configured on a router.
3. Specialized appliance sees only the required traffic and performs the desired offline processing.

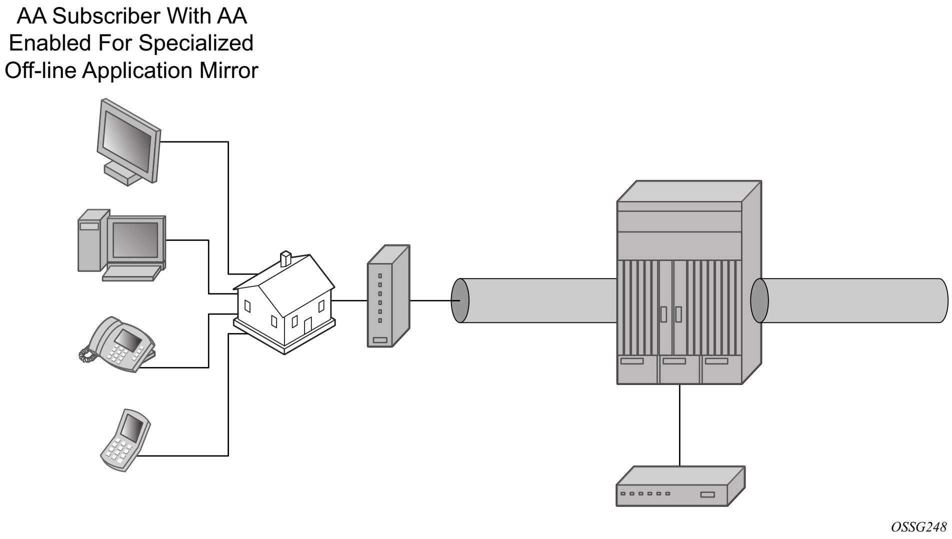


Figure 30: AA Mirroring for Offline Specialized Appliance Processing

Application Assurance Firewall

The Application Assurance firewall (FW) feature extends AA ISA application level analysis to provide an in-line integrated stateful service that protects subscribers from malicious security attacks. Using the AA stateful packet filtering feature combined with AA Layer 7 classifications and control empowers operators with advanced, next generation firewall functionalities that integrated are within. AA stateful firewall and application firewall run on the AA ISA. In a stateful inspection, the AA FW does not only inspect 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 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:

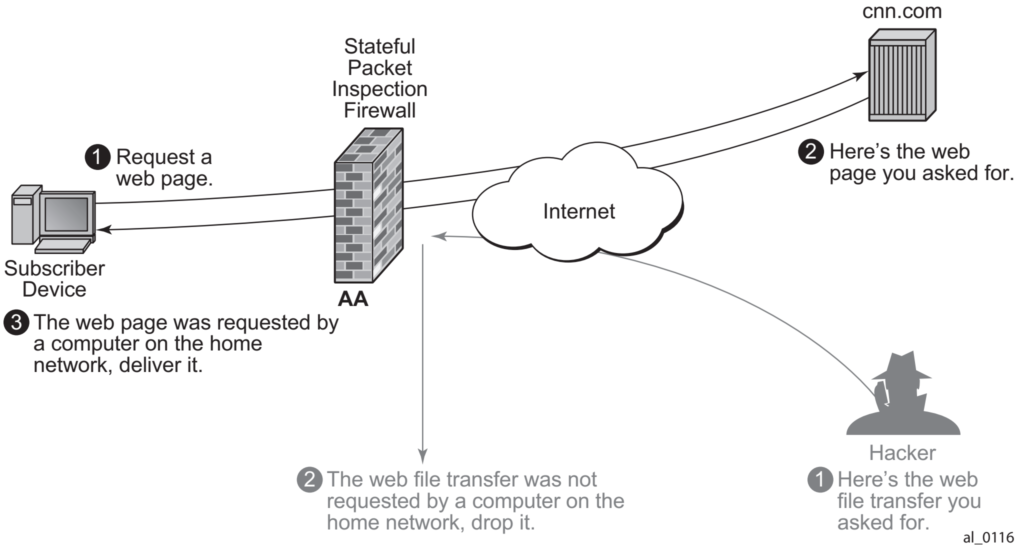
- BNG (ESM)
- WLAN Gateway (ESM or DSM)
- Transit-subscriber
- Business AA.

AA FW enabled solution provides:

- Stateful /Stateless Packet Filtering and Inspection with Application-Level Gateway (ALG) Support
- Security Gateway — SeGW Firewall Protection S1-MME (/SCTP), S1-U (GTP) and OAM traffic protection.

Stateful /Stateless Packet Filtering and Inspection with Application-Level Gateway (ALG) Support

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, if 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.



cnn.com
Stateful Packet Inspection Firewall
1 Request a web page.
2 Here's the web page you asked for.
Internet
Subscriber Device
AA
3 The web page was requested by a computer on the home network, deliver it.
Hacker
2 The web file transfer was not requested by a computer on the home network, drop it.
1 Here's the web file transfer you asked for.
al_0116

Figure 31: Stateful Firewall

Stateless packet filtering does not take note of session initiator and hence, it discards or allows packets independent of the any previous packets. Stateless packet filtering can be performed in the system using IOM ACLs.

AA FW inspection of packets at Layer 7 offers Application Layer Gateway functionality for the following applications:

- rtsp
- sip
- h323 (IPv4 only)
- googletalkvoice
- ftp
- tftp
- pptp
- citrix
- sybase
- msexchange
- skinny
- ares
- bittorrent

- dns
- irc
- mailru
- qvod
- R commands
- sc2
- socks
- vudu
- winmx
- xunlei

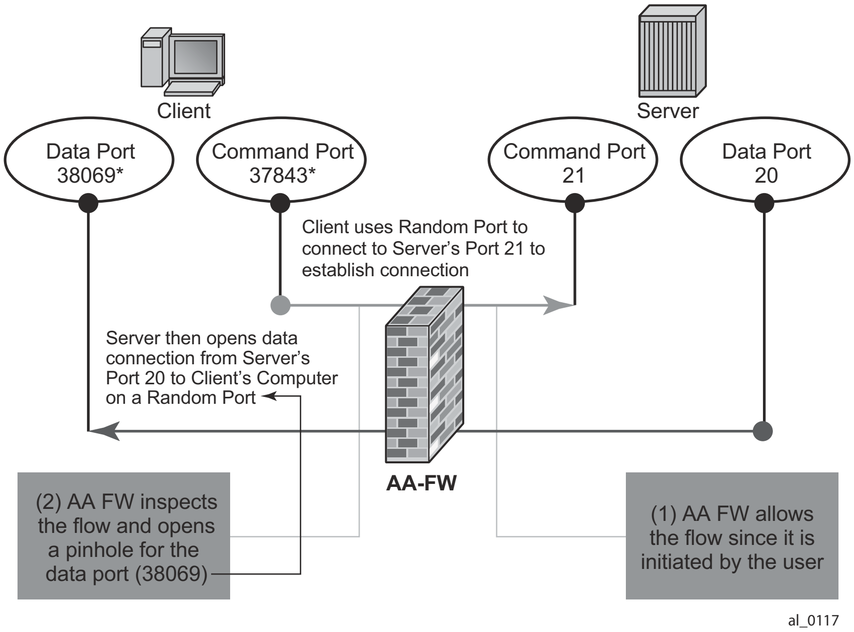

Client
Server
Data Port 38069*
Command Port 37843*
Command Port 21
Data Port 20
Client uses Random Port to connect to Server's Port 21 to establish connection
Server then opens data connection from Server's Port 20 to Client's Computer on a Random Port
AA-FW
(2) AA FW inspects the flow and opens a pinhole for the data port (38069)
(1) AA FW allows the flow since it is initiated by the user
al_0117

Figure 32: Application Layer Gateway Support

These applications make use of control channels/flows that spun other flows. AA FW inspects the payload of these control flows so that it can open a pinhole for the associated required flows.

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 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 all sort of flooding attacks (for example, ICMP PING flooding, UDP flooding, SYN Flood Attack). These policers provide protection at multiple levels; per system per application/ application groups and per subscriber per applications/applications groups. AA ISA flow policers has two flavors; 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.

To protect hosts and network resources, AA_FW validates/checks the following parameters, if any fails, it declares the packet to be invalid:

- IP layer Validation:
  → IP version is not 4 nor 6

  → Checksum error (IPv4)

  → Header length check

  → Packet length check

  → TTL/Hop limit (not equal to zero) check

  → IPv4 source address checks:
    - class D/E (>=224.0.0.0)
    - BCAST 255.255.255.255 (multicast source address)
      - 127.x.x.x (invalid source address)
    - invalid subnet (subnet, 0) [unless /31 point-to-point interface]
    - invalid subnet multicast (subnet, -1) unless /31 point-to-point interface

  → IPv4 destination address checks:

  BCAST 255.255.255.255, 0.x.x.x,127.x.x.x

  → IPv6_source address check

  multicast source address (FFxx:xxxx:……:xxxx)

  → IPv6_destination address check

  invalid destination address (=::)

- TCP/UDP validation:
  - → header checksum
  - → Source or destination ports (not equal to zero) check

    (only dest port is checked for UDP)
  - → Invalid TCP flags:
    - – TCP FIN Only: only the FIN flag set.
    - – TCP No Flags: no flags are set.
    - – TCP FIN RST: both FIN and RST are set.
    - – TCP SYN URG: both SYN and URG are set.
    - – TCP SYN RST: both SYN and RST are set.
    - – TCP SYN FIN: both SYN and FIN are set

The above complements ESM enhanced security features, such as IP (or mac) anti-spoofing protection (for example, protecting against “LAND attack”) and network protocols DoS protections. The combination provides a world class carrier grade FW function.

Policy Partitioned AA FW

AA FW can provide up to 128 virtual/partitioned FWs, each with its own FW policies. This is achieved through the use of AA-Partitions. Different VPNs can have different FW policies/rules.

Configuring AA FW

AA ISA AQPs are enhanced with few new AQP actions that provide session filtering functionality. As is the case of AQPs, these have partition level scope. This allows different FW polices to be implemented by utilizing AA partitions concepts within the same AA ISA group. Hence, multiple virtual AA FW instances can be realized. There is no 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** and **action** per entry. Actions are **deny** or **permit**. A **deny** action results in packets discarded without creating a session /flow context. **match** conditions include IP 5 tuples info. An overall default action is also configurable, in case of a no match to any session filter entry.

Note: AQPs with session filter actions, need to have, as a matching condition, traffic direction, ASOs and/or subscriber name. It cannot have any references to applications and/or application groups.

AA FW offers, in addition to session-filter actions, a variety of AQP actions to that are aimed to allow or deny: errorred/malformed packets, fragmented packets and/or first packet out-of-order fragments and overload traffic.

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.

A session-filter hit-count counter is maintained by AA ISA and can be viewed via CLI. There is no current support for export of session-filter entry hit counters via XML to SAM.

AA FW Logging

AA ISA can be configured, per AQP or per session filter, to log events related to how the packets are processed (allowed/denied). AA ISA FW logs contain the following information:

- Grou: partition
- Timestamp
- 5-tuple
- Direction
- Subscriber info (if available)
- Log source/type (session-filter or AQP)
- Action (allow/drop)
- Session-filter specific
  - → Session-filter name
  - → Session-filter entry
- AQP specific
  - → Drop reason
  - → Fragment Offset (if applicable)
  - → Fragment ID (if applicable)
- If an out of order fragment triggers the log, then whatever 5-tuple information is available is included.

Note that for AQPs, only **drop** events are captured in the log. The logs do not capture drops due to flow policers.

The operator can configure up to one log per partition, with a maximum configurable log size of 100,000 events per log.

SeGW Firewall Protection

Application Assurance SeGW FW deployed in 3G/4G/Femto access networks provides the operator with back-end core network security protection. AA Firewall provides protection for:

1. S1-MME (SCTP) traffic
2. S1- U (GTP-U) traffic
3. OAM traffic

SAPs on the private side of Tunnel ISA are diverted to AA for firewall protection. If per eNB/Femto Access Point (FAP) control is desired, then AA auto-configures /instantiate subscribers based on the "seen-ip" transit-AA subscriber model (no RADIUS interaction is required). This auto-creates a subscriber per eNB/FAP. Alternatively, AA applies firewall rules to the diverted SAP (for all eNBs/FAPs) at the aggregate level (for all eNBs/FAPs).

One AA ISA is supported per Tunnel-ISA group. Therefore, all private side SAPs that are diverted to AA for Firewalling service go to the same AA ISA module with no need to load balance the traffic into different AA ISAs. If the capacity of one AA ISA is not sufficient, then the IPSec tunnel group is split into two (or more) groups. Each group is served by an AA ISA.

OAM traffic Firewall

For details on OAM Traffic protection, refer to the Stateful /Stateless Packet Filtering and Inspection with Application-Level Gateway (ALG) Support on page 151 and Denial Of Service (DOS) Protection on page 154 sections.

S1- MME (SCTP) Firewall

Network flooding attacks, malformed packets and port scans are examples of DoS attacks that can be carried out using a compromised eNB/FAP. AA FW provides inspection of SCTP (the protocol used to communicate to MME). Such inspection includes checking for SCTP protocol Id, source/destination ports, PPID, SCTP chunk checking and malformed SCTP packet (such as checksum validation).

SCTP chunk checking includes checking for:

- Invalid length values. Frames with invalid length value are dropped regardless of the chunk type .
- Data chunks with length value that is too small to accommodate PPID. Such frames are dropped as invalid/badly formed.
- Data chunks with length value that is too large for chunk. Such frames are dropped as invalid/badly-formed.

For S1-MME traffic, the operator can configure various AA actions:

- Drop packets with invalid checksum, src/dest IP and/or port numbers (malformed Packet protection) by appropriately configuring session filters and /or **error-drop** [**event-log** <*event-log-name*>] AQP action command.
- PPID Filtering, using SCTP-Filter command
- Rate limit the amount of S1-MME traffic (flooding protection) in terms of Bandwidth (bits/sec), using AA bandwidth policers.
- Limit the number of concurrent SCTP flows (flooding protection) using AA flow count policers.
- Limit the SCTP flow setup rate (flows/sec) to protect against DoS flooding using AA flow rate policers.
- Drop fragmented packets or drop out of order fragmented packets using the **fragment-drop** {**all** | **out-of-**order} AQP action command.

The actions above can be applied per eNB/FAP IP address and/or per MME (to control aggregate traffic per MME).

SCTP PPID Filtering

AA allows the operator to configure PPID filters that contain a list of PPIDs to allow or deny.

```
configure>application-assurance>group <aa-group-id>[:<partition>]
         sctp-filter <sctp-filter-name> [create]
                description <description-string>
                no description
                event-log <event-log-name>
                no event-log
                ppid-range min <min-ppid> max <max-ppid>   //[0..4294967295]
                no ppid-range
                ppid
                    default-action {permit | deny}
                    entry <entry-id> value <ppid-value> action {permit|deny}
                             //<entry-id>: [1..255]
                                 <ppid-value> : [0..4294967295]D | [256 chars max]
                                 <permit|deny>: permit|deny
                    no value <entry-id>
         no sctp-filter <sctp-filter-name>
```

The filter can then be used within an AQP action.

AA identifies DATA chunks within SCTP payloads (for example, as first, nth or last chunk) and filters according to the configure PPID filter. If any chunk PPID matches a PPID on the configured blocked PPID list, the whole SCTP packet is dropped.

SCTP packets without DATA chunks are not impacted or accounted for by an SCTP Filter.

For IP fragmentation, and in the case where the operator did not configure AA ISA to drop “all fragmented traffic”, only the first IP fragment is inspected and subject to the PPID filtering. Any action applied to the first fragment is also applied to the remaining fragments. Out-of-order fragments appearing before the first fragment receive the default action (for example, drop action of “out-of-order-Frag”).

S1-U GTP Traffic Protection

7750 SeGW with AA FW provides protection of SGW/SGSN infrastructure against an attack from a compromised eNB/FAP. AA FW supports:

1. Protection against malformed GTP packets attack:

   For GTP-v1 traffic carried over UDP port number port 2152, AA performs various packet sanity checks, such as:

   → UDP destination port is 2152

   → Version: GTP-U should always be version 1.

   → Protocol Type bit should be 1

   → Invalid/Missing Mandatory Header Fields

   → Invalid Optional/Spare Header Fields

   → Invalid/Missing Header Extensions

   → Invalid Length

   For S1-U interface, only GTP-v1 is supported. No support for GTP-v2 (as there is no signaling on S1-U interface).

   Details of the various GTP sanity checks that are performed for different GTP-U message types are shown in Table 14:

Table 14: GTP-U Message Types

Payload Size	Encapsulated Data Checks	IE Checks	Header Extension Checks	Optional HEADER Check			GTP Mandatory Header Checks			
				If E, S or PN =1	Length	TEID	Spare	PT	version	
>0	PayloadSize is assumed to be the size of the remainder of the packet, unless the packet is fragmented No checking of the encapsulated data	No checks	Valid types = Service Class Indicator & • PDCP PDU Number Extension size= 4*# of extensions	OptionalSize = 8 IF E= 0, ExtSize = 0	OptionalSize + ExtensionSize + PayloadSize	<>0	0	1	1	G-PDU (Encapsulated Data Delivery) – Message Type 255
	No payload after the IEs	Only private extensions are allowed.	No external header allowed.	No option headers allowed.	IE Size	0	0	1		Echo Request – Message Type 1

Table 14: GTP-U Message Types (Continued)

Payload Size	Encapsulated Data Checks	IE Checks	Header Extension Checks	Optional HEADER Check	GTP Mandatory Header Checks					
				If E, S or PN =1	Length	TEID	Spare	PT	version	
	No payload after the IEs	Recovery ID is present Private extensions allowed.	No external header allowed.	No option headers allowed.	IE Size	0	0	1	1	Echo Response – Message Type 2
	No payload after the IEs	Extension Header Type List IE is present Private extensions allowed No checking on the extension header value	No external header allowed.	No option headers allowed.	IE Size	0	0	0	1	Supported Extension Headers Notification - Message Type 31

Table 14: GTP-U Message Types (Continued)

Payload Size	Encapsulated Data Checks	IE Checks	Header Extension Checks	Optional HEADER Check			GTP Mandatory Header Checks			
				If E, S or PN =1	Length	TEID	Spare	PT	version	
	No payload after the IEs	TEID IE & GTP-U Peer Address IE are present IE type and length are verified Private extensions allowed	Only the UDP Port Extension Header is valid	Option Size = 8	OptionalSize + ExtensionSize +IESize	<>0	0	1	1	Error Indication – Message Type 26
	No payload after the IEs	Only Private extensions are allowed	no valid external header allowed.	OptionalSize = 8 IF E = 0, ExtSize = 0	IE Size	<>0	0	1	1	End Marker – Message Type 254

To enable GTP packet sanity checks, the operator must configure:

```
configure>application-assurance>group <aa-group-id>[:<partition>]
```

Note that once the **gtp** command is issued for a partition, AA treats traffic with UDP destination port number 2152 as GTP. It applies the different GTP level firewall functions as configured by the operator. However, it does not look beyond the GTP header for further inner L3-L7 packet classifications and actions. For example, Ipfix record for GTP traffic contains the 5 tuples of the GTP-u tunnel (eNB, SGW IPs and port numbers, etc., no TIED).

2. Protection against un-supported GTP messages

   AA allows the operator to configure a GTP filter to indicate which GTP message types are to be allowed/denied as well as the maximum allowed GTP message length:

```
configure>application-assurance>group <aa-group-id>[:<partition>]>gtp
        gtp-filter <gtp-filter-name> [create]
            max-payload-length <bytes>           //[0..65535]
            message-type
                default-action {permit | deny}
                entry <entry-id> value <gtp-message-value> action {permit|deny}
```

   There are approximately 67 valid message names to enter in the above GTP filter:

   echo-request, echo-response, error-indication, g-pdu, end-marker and supported-extension-headers-notification.

   Once a GTP filter is configured, it can then be included as an AQP action:

```
configure>application-assurance>group <aa-group-id>[:<partition>]> policy
        app-qos-policy
            entry <entry-id> [create]
                action
                    gtp-filter <gtp-filter-name>
```

   Note: Extensive GTP header sanity checks (included in Table 14) that are based on different GTP message types are only performed when these GTP messages are permitted by the GTP filter. If no GTP filter is configured, then no extensive GTP-U header checks are performed. In other words, if the operator wants to allow all GTP-U packets and perform all GTP header sanity checks, then the operator needs to configure a GTP filter with default action of **permit** and no values, such as:

```
configure>application-assurance>group 1:100> gtp
     gtp-filter “allow-all” create
          message-type
             default-action permit
```

3. Protection against flooding attack:

   AA can be configured to drop all fragments and/or out of order fragments, using AQP action: **fragment-drop** {**all** | **out-of-order**}

   In the case that the IP **fragment-drop** command is not set, then the following conditions apply to the way AA inspects GTP traffic:

- → Permit/deny decisions are entirely based on the first fragment. The first fragment contains the entire GTP header in almost all of the cases.
- → Max packet length check is not done across fragments. Only the first fragment length is checked. In other words, AA ISA may permit a packet that is larger than the max packet allowed if it is fragmented, with the first fragment smaller than the configured maximum packet size allowed.
- → First fragmented packet is discarded (and logged), as well as subsequent fragments:
  - – If the first packet is too small to contain the mandatory header (12 bytes, ending with the TEID).
  - – If the mandatory header indicates there should be an optional header, and the fragment is too small to contain the optional header (mandatory + optional = 16 bytes).

Service Monitoring and Debugging

Operators can use AA-specific tools in addition to system tools that allow them to monitor, adjust, debug AA services. The following are examples of some of the available functions:

1. Display and monitor AA ISA group status and statistics (AA ISA status and capacity planning/monitoring).
2. Clear AA ISA group statistics (clears all system and per-AA-subscriber statistics).
3. Special study mode for real-time monitoring of AA-subscriber traffic (ESM or transit subscriber, SAP or spoke SDP).
4. Display per AQP entry statistics for number of hits (flow matching the entry) and conflicts (actions ignored due to per-flow-action-limit exceeded).
5. Mirror (all or any subset of traffic seen by an AA ISA group).
6. Display all the per-ISA statistics from the aa-performance record, for examining resource loading of each ISA
7. Display the top active AA-subscribers per ISA by bytes, packets or flows, for traffic in each direction

CPU Utilization

The ISA show status command displays per ISA CPU utilization by main tasks, to provide insight into what aspects of load may be loading the ISA. These are split into 2 main areas:

- Management CPU, which includes all tasks related to communication between the CPM and the ISA, with the following usage percentage reported:
  - → System — Various infrastructure and overhead work
  - → Management — Managing AA policy, AA subscriber and trap configurations and handling tools requests
  - → Statistics — Collecting and reporting statistics and Cflowd reporting
  - → Idle
- Datapath CPUs, which includes all tasks related to datapath packet and flow processing on the ISA, with the following usage percentage reported:
  - → System — Various infrastructure and overhead work
  - → Packet processing — Receiving, associating with flows, applying application QoS policy and transmitting
  - → Application ID — Using protocol signatures and other techniques to identify application/app-group and determine the application QoS policy

CLI Batch: Begin, Commit and Abort Commands

The Application Assurance uses CLI batch capability in policy definition. To start editing a policy, a begin command must be executed. To finish editing either abort (discard all changes) or commit (accept all changes) needs to be executed. CLI batch state is preserved on an HA activity switch.

To enter the mode to create or edit policies, the **begin** keyword must be entered at the prompt. Other editing commands include:

- The **commit** command saves changes made to policies during a session. The newly committed policy takes affect immediately for all new flows. Existing flows will transition onto the new policy shortly after the commit.
- The **abort** command discards changes that have been made to policies during a session.

To allow flexible order for policy editing, the **policy>commit** function cross references policy components to verify, among others:

- Whether all ASO characteristics have a default value and are defined in the app-profile.
- Checks whether limits are adhered.