

Availability and Scalability

Ericsson Service-Aware Policy Controller

FACILITY DESCRIPTION

Copyright

© Ericsson España, S.A. 2018. All rights reserved. No part of this document may be reproduced in any form without the written permission of the copyright owner.

Disclaimer

The contents of this document are subject to revision without notice due to continued progress in methodology, design and manufacturing. Ericsson shall have no liability for any error or damage of any kind resulting from the use of this document.

Trademark List

All trademarks mentioned herein are the property of their respective owners. These are shown in the document Trademark Information.



Contents

1	Availability and Scalability Introduction	1
2	Availability and Scalability Function	1
2.1	Availability	1
2.2	Scalability	9
3	Availability and Scalability Operational Conditions	12
3.1	Availability and Scalability External Conditions	12
3.2	Availability and Scalability Function Administration	13
3.3	Availability and Scalability Security	13
	Reference List	15





1 Availability and Scalability Introduction

This document provides a description of the Availability and Scalability function provided by the Ericsson Service-Aware Policy Controller (SAPC).

2 Availability and Scalability Function

2.1 Availability

The SAPC is built on a high available architecture where a single failure does not stop the operation of the cluster. It is built over a cluster of nodes of three types:

- System Controllers (SC): There are always two SCs in the cluster. They provide the OAM and provisioning services through the OAM virtual IP address.
- Traffic Processors (TP): The number of traffic processors is determined depending on the operator needs (the minimum value is two TPs). TPs provide the SAPC traffic services through virtual IP addresses: they provide the traffic interface, balance the traffic processing among all TPs in the cluster and also store provisioning and dynamic data in a replicated way:

- Primary copy and one replica for dynamic data (sessions, accumulators) and subscribers.
- Totally replicated for remaining provisioning data.

Note: Diameter traffic processing is distributed among all Traffic Payloads in the cluster based on the `session-id` AVP.

- Virtual Routers (VRs): Only provided, optionally, for Cloud Data Center deployments. There are two VRs providing access to the SAPC OAM Virtual IP Address and two VRs providing access to the SAPC Traffic Virtual IP Address.

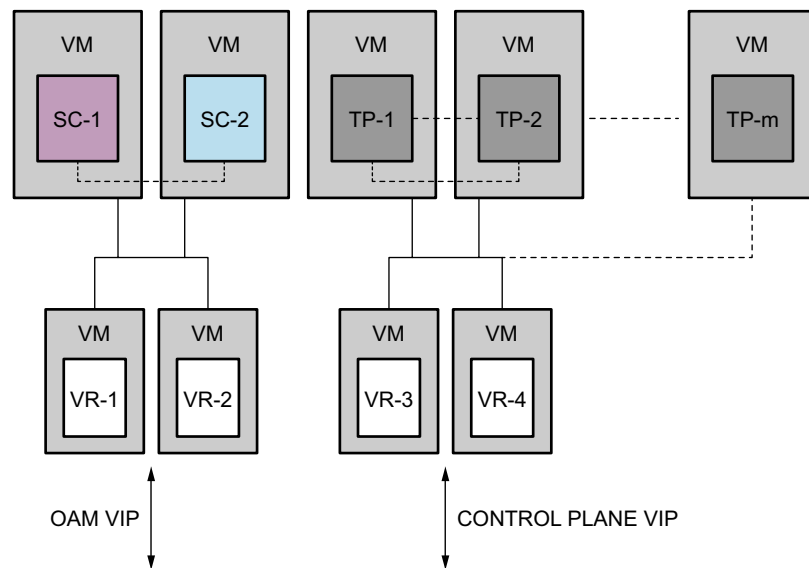


Figure 1 SAPC Cluster Architecture

The two SCs provide the OAM and provisioning services in an active-standby mode, which means that if an SC goes down, all services considering it as the active one, are managed by the other SC. The rest of the node types work in an active-active mode. The incoming traffic is distributed by a maximum of six TPs (usually, the first six TPs) among all the available traffic processors in the cluster. These TPs are also the ones publishing the traffic virtual IP address to the external network. If one of these TPs goes down, the publishing of the virtual IP address and the traffic distribution functions are moved to another available TP. Also, this TP is not considered to receive traffic in the distribution until it is up again.

The following situations, in which multiple failures were produced simultaneously, would affect the SAPC service availability:

- In case the two SC processors were restarted simultaneously, if, after a period of 15 minutes, none of the SC processors are recovered, then a cluster restart is produced until, at least, one of the SCs is recovered. During the absence of the SCs, Diameter traffic is processed, but the OAM traffic cannot be handled because the OAM VIP is not available. Performance counters are stored in memory during the absence of SCs and the result is dumped in the performance files when the SCs are up again. Alarms and logs are not available during the absence of SCs and recovered when they are up again. For the Geographical Redundancy scenario, the cluster reboot is produced immediately after the restart of the SC processors.
- In case two TP processors containing both replicas of the dynamic data or subscribers were restarted, the SAPC database would be restarted to assure consistency. In this case, the Gx session data (also including, the time-related session events and the internal session usage accumulator information) is not recovered and Diameter connections with PCEF are closed.



- In case the two VRs providing access to the OAM virtual IP address are restarted simultaneously, the OAM incoming traffic would be affected.
- In case the two VRs providing access to the Traffic virtual IP address are restarted simultaneously, the incoming traffic would be affected.

To increase the reliability and availability of the system, the SAPC includes several control mechanisms, such as restoration procedures, overload control, session clean up procedure, and mechanisms to overcome connectivity loss.

2.1.1 Restart and Restore Procedures

The SAPC provides mechanisms to handle restart situations both for the SAPC itself and for the peer traffic plane nodes, and also procedures to restore.

2.1.1.1 SAPC Restart

Even when the SAPC provides a high level of availability, in case both SCs fail simultaneously during more than 15 minutes, the SAPC is restarted. Once the SAPC recovers from a restart, the last database information is recovered from the stored backups. The information recovered may not be fully up to date and, for this reason, some actions are performed by the SAPC to consolidate this information.

The next sections describe the actions performed by the SAPC , in PCC deployment scenario, after a cluster restart.

The SAPC increments its own `Origin-State-Id` and includes the new value in every response message alerting the peer nodes about the loss of previous session state.

Note: The `Origin-State-Id` is a monotonously increasing value that is increased whenever a Diameter entity restarts with the loss of the previous state.

The sessions available before the restart are not recovered from a backup. Therefore, all dynamic data related to sessions are not recovered either:

- Time Trigger reauthorization events are lost.
- For IP-CAN sessions with end-user notifications, the same notification message can be sent again after the restart.
- For Fair Usage, the usage accumulators may not be accurate.

In case the fair usage feature is active and usage accumulator data was stored before the SAPC restart, the absolute usage accumulator is recovered from the stored backups. After the restart, the recovered data only contains usage activity until the time the last backup was performed. This usage information is used for the quota calculation for the new and ongoing sessions after the restart.



The Gx, Sd, Sy, and Rx sessions are identified by the Diameter `session_id`. A session is considered unknown if the SAPC does not find a session in its internal database with the same `session_id`. After a SAPC restart, requests sent from the PCEFs, AFs, TDFs or Online Charging Systems for an unknown session will be answered by the SAPC with the `DIAMETER_UNKNOWN_SESSION_ID` error code.

All subscriber-related data are recovered from the stored backups.

2.1.1.2 Peer Restart

The SAPC is able to detect diameter peer node restarts based on the standard mechanism described for Diameter nodes in RFC 6733 (refer to [Diameter Base Protocol, IETF RFC 6733](#)).

The SAPC provides the following mechanisms to handle restart situations for the peer traffic plane nodes:

— PCEF Restart:

The SAPC detects that a Gx Diameter client restarted after comparing the `Origin-State-Id` information received in every Gx client's activation request with its locally stored state information for that client. When a received `Origin-State-Id` is different from the stored one, the SAPC considers that this peer has restarted and, after a period of time of 2 hours, identifies all the sessions established by restarted client and removes them. See Section 2.1.2.2 for a more detailed procedure.

The Gx massive clean up mechanism starts 2 hours after the PCEF restart is detected. During this period of time, the SAPC removes the obsolete sessions by the Basic Clean up mechanism. See Section 2.1.2.1 for further information. This delay in starting the massive clean up avoids or at least minimizes collisions between both clean up mechanisms being executed at the same time.

A PCEF can be a cluster of several master nodes and a backup node. Each master and backup can contain several Diameter client applications, each one identified by its `origin-host-id`. In this scenario, both master and backup nodes should be configured with the same `diameter_node_id` that is used in the SAPC to identify the session. In case there is a failover from one of the master nodes to the backup node, the massive clean up mechanism is not used, since in that case, the `origin-host-id` changes. The SAPC performs the basic clean up mechanism when the backup node sends a session establishment message again for each session.

— AF Restart:

The SAPC detects that an AF restarted after comparing the `Origin-State-Id` information received in every Rx client's request (initial or update) with its locally stored state information for that client. When a received `Origin-State-Id` is different from the stored one, the SAPC considers that this peer has restarted, identifies immediately all the AF sessions established



by the restarted client and removes them. See Section 2.1.2.2 for a more detailed procedure.

- OCS Restart:

The SAPC does not detect when the OCS restarted. In case the OCS restarts losing charging and Sy session information, the network must be configured to terminate the affected IP-CAN sessions. The SAPC removes the Sy sessions at IP-CAN session termination and creates new Sy sessions at IP-CAN session establishment.

- Diameter Routing Agent (DRA) Restart

The SAPC does not keep track of the Origin-State-Id received in the diameter base protocol messages. The SAPC does not act or clean-up sessions in the event of a DRA restart.

- TDF Restart

The SAPC does not keep track of the Origin-State-Id received in the Sd diameter protocol messages, so it does not detect when the TDF has restarted. The SAPC only removes Sd sessions at IP-CAN session termination.

2.1.1.3

SAPC Restore

The SAPC provides the System Data type of restore.

System Data backup is used to do a system data fallback to recover to a former version of the whole system with consistency. After restoring a System Data backup, the SAPC reestablishes the following information:

- Software installed.
- Node configuration.
- Static Subscriber profile and provisioning information stored in the SAPC database.
- Dynamic Subscriber information (that is, usage accumulators) stored in the SAPC database.

And the SAPC loses the following data:

- Sessions.
- Time Trigger events.

2.1.2

Session Cleanup Mechanisms

The following mechanisms are implemented in the SAPC to remove obsolete information.



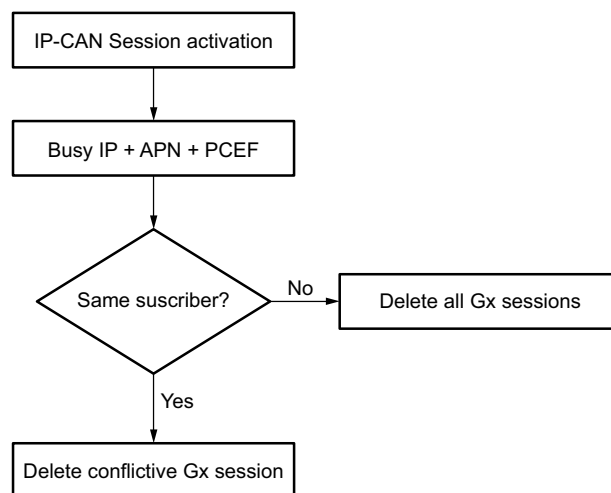
2.1.2.1 Basic Session Cleanup Mechanism

The following mechanism is related to the removal of specific obsolete sessions:

— Gx Mechanism

When the SAPC receives an IP-CAN session activation request from a given PCEF including an APN and an IP address, the SAPC checks if any IP-CAN session exist that was established from the same PCEF, with the same IP address, and the same APN. If any IP-CAN sessions are found:

- a If the subscriber received in the activation request is different, the SAPC removes the IP-CAN session with all its Gx sessions and a new one is created with the new IP-CAN session request information.
- b If the subscriber received in the activation request is the same, only the Gx session matching the received PCEF is removed and a new Gx session is created with the new IP-CAN session activation request information.



The Session usage accumulator is reset.

As a consequence of the removal of each Gx session, the AF, Sd, and Sy sessions are removed as it is done when an IP-CAN session is terminated. The SAPC requests the AF to revoke all the identified AF sessions, the TDF to terminate the associated Sd session, and the Online Charging System to terminate the indicated Sy sessions.

2.1.2.2 Massive Cleanup Mechanism

Massive Gx Session Clean up at PCEF Restart

This clean up mechanism consists of deleting all the obsolete IP-CAN sessions existing in the SAPC for a restarted PCEF considering also:



- The AF sessions corresponding to the IP-CAN sessions are deleted and a request is sent to the corresponding AF to revoke the identified AF sessions.
- In a multiple Gx scenario, the SAPC does not remove associated sessions of the other PCEF Diameter clients.
- Since the sessions are lost in the PCEF, the session usage accumulator is removed.
- The Sd sessions associated with the removed IP-CAN sessions are deleted and a request is sent to the corresponding TDF to remove the identified Sd sessions.
- The Sy sessions corresponding to the removed IP-CAN sessions are also deleted if no more IP-CAN sessions are bound to the subscriber. When the Sy session is deleted, a session termination message is sent to the Online Charging System.
- The Gx massive clean up mechanism is load regulated as explained in [Overload Control](#).

Massive Gx Session Clean up at PCEF Peer Removal

When a `diameterNode` peer is removed from the configuration data, the SAPC removes all the IP-CAN sessions established by that peer, using the same PCEF restart mechanism.

Massive Rx Session Clean up at AF Restart

This clean up mechanism consists of deleting all the obsolete AF sessions existing in the SAPC for a restarted AF considering also:

- The PCC rules from the IP-CAN session bound to the deleted AF sessions are removed, and corresponding PCEF is notified by means of a Gx RAR message.
- Rx massive clean up is load regulated as explained in [Overload Control](#).

Note: The SAPC provides a robust mechanism that allows to clean the obsolete sessions even in case of geored switchover or scaling scenarios.

Both massive Gx and Rx clean up processes continue scanning and removing sessions until all the obsolete IP-CAN or AF sessions of the restarted peer have been removed.

2.1.2.3

Session Inactivity Cleanup Mechanism

This clean up mechanism consists of deleting all the inactive Gx sessions (no request is received or sent for them in a configurable period of time) existing in the SAPC considering also:

- Depending on the configuration, the SAPC checks whether the Gx session is still alive in the PCEF sending a Gx RAR message. Only in the case when

the PCEF sends an RAA message with the `DIAMETER_UNKNOWN_SESSION_ID` error, the SAPC removes the session.

- The same behavior as in PCEF restart mechanism for the AF, Sd and Sy sessions and session usage accumulator corresponding to the removed Gx session.
- This mechanism is load regulated as explained in [Overload Control](#).

This mechanism is daily and enabled or disabled by configuration together with other parameters, as explained in [Configure Session Inactivity Cleanup Mechanism](#).

If there is a massive clean up running or detected at the same time with a session inactivity cleanup process, the SAPC stops the session inactivity cleanup process.

2.1.3 Virtual Machine Evacuation

Virtual Machine (VM) Evacuation is a feature provided by the Network Function Virtualization Infrastructure (NFVI).

The evacuation of SAPC VMs permits that when the physical host where any of them is allocated is down, the Infrastructure re-creates the VM in a different host, restoring the VNF High Availability state. The evacuated VM is created from scratch, so runtime data is lost. To fully guarantee the HA of the SAPC, the evacuation of the VM should be done in such a way that the anti-affinity rules recommended for the VNF are applied during the re-creation of the new VM. See chapter for anti-affinity rules in [Virtual Service-Aware Policy Controller 1](#).

No specific configuration is needed in the SAPC to support the feature, except for CEE deployments, where an evacuation policy must be configured during the VNF deployment. See the `ha-policy` parameter of the Descriptor Generator Tool Configuration File in [SAPC VNF Descriptor Generator Tool](#).

For SAPC deployments on OpenStack based NFVIs, the VM Evacuation is only supported for Traffic Processors and Virtual Router VMs. In those deployments, when a host allocating an SC is down, the SC VM is started in this same host once it is up again.

2.1.4 EBM Server Connectivity Loss

The events generated by the SAPC are sent to the Event-Based Monitoring (EBM) server through a set of connections that the SAPC establishes. Connections towards the EBM server can be lost or suffer network disturbances. Meanwhile, if any of these connections are not working because of connectivity loss or server unavailability, events generated by the SAPC are stored in an internal buffer per each EBM server. For each EBM server there is a pool of connections as well.

This buffer is dimensioned to store the number of events generated to each EBM Server in a 10 second period, in which the SAPC handles the maximum incoming



traffic the node can withstand. When the maximum capacity of the buffer is reached, the new events will overwrite the older events, following the First In First Out (FIFO) criteria in the buffer. As a result, some events can be lost if the network disturbance lasts more than 10 seconds.

When the SAPC detects that a connection is broken, the following actions are performed:

- The SAPC tries to reestablish the connection up to 3 times, at 5 second intervals.
- Every time the timer elapses, the SAPC tries to send the events stored in the buffer. If the SAPC detects that the connection keeps being broken, a new timer of 5 second is set.
- If after 3 reattempts the SAPC fails to establish the connection with the EBM server, the reattempts will continue until the connection is successfully reestablished, but at intervals of 10 seconds. The `EbmCommunicationFailure` alarm is raised when the 3 reattempts fail, and it is cleared when the connection is successfully reestablished or the EBM feature is disabled.

2.1.5 Live Migration with VMware vMotion

Live Migration is a feature provided by the NFVI.

To ease physical host maintenance tasks (for example HW upgrade), Live Migration allows to move a running VM between different physical hosts without perceived downtime. Memory, storage, and network connectivity of the VM are transferred from the original guest machine to the destination.

No specific configuration is needed in the SAPC to support the VMware vMotion feature.

The migration may involve only the VM, or only the datastore for this VM, or both. During the procedure, the new host, where they are migrated, is selected.

Shared storage is required for VMware vMotion. Also, the same networking configuration (same port groups) must be guaranteed in the destination host.

2.2 Scalability

The SAPC is built on a scalable architecture providing the ability to, on runtime, increase the capacity for traffic processing adding additional processors (Scale-out) or reduce the capacity removing existing processors (Scale-in). The SAPC is able to keep performance levels with few seconds impact on the ongoing traffic when Scale-out or Scale-in functions are performed. The node types to be scaled are only the Traffic Processors (TPs).

Figure 2 shows a SAPC cluster initially installed with m TP nodes that have been scaled-out up to z TP nodes.

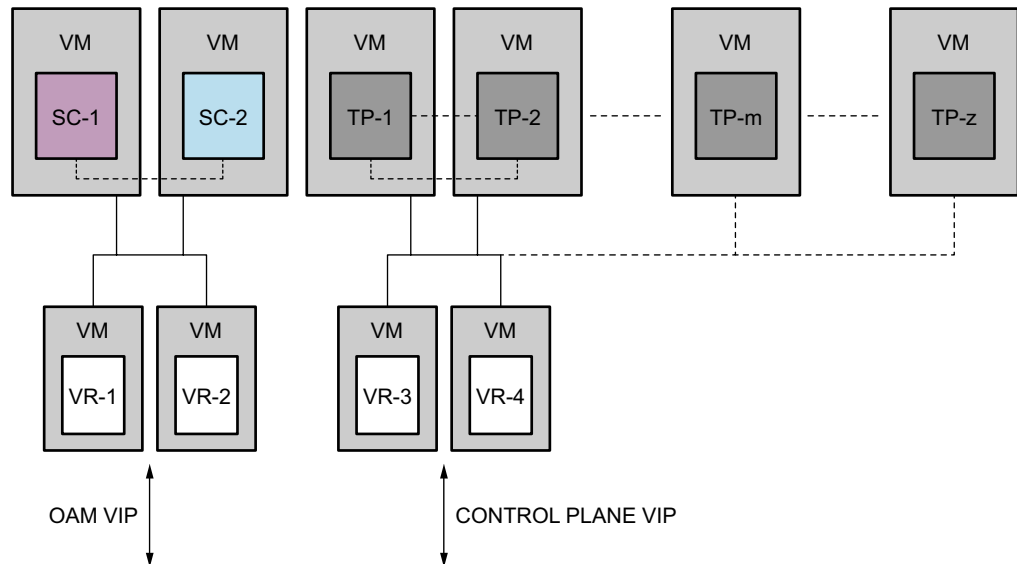


Figure 2 SAPC Cluster where New TP Nodes Are Added

When TPs are scaled, the traffic interface and traffic distribution functionalities are also included, up to six running instances in six different TPs. From this number onwards, the new scaled TPs provide these functions in a spare way (as standby to become ready if any other active instance gets down).

2.2.1 Multi-Site Support

The SAPC supports geographical distribution (multi-site) configurations when a single SAPC does not have enough capacity to handle all the subscribers' traffic in the following scenarios.

2.2.1.1 SAPC with Common Database

In this deployment, the operator has multiple SAPCs deployed in different sites and a common database to store the subscriber data. Hence, any SAPC can serve IP-CAN session from any subscriber. Fair Usage Accumulators must be stored in the common database, so that any of the SAPCs can access and modify the data at any time

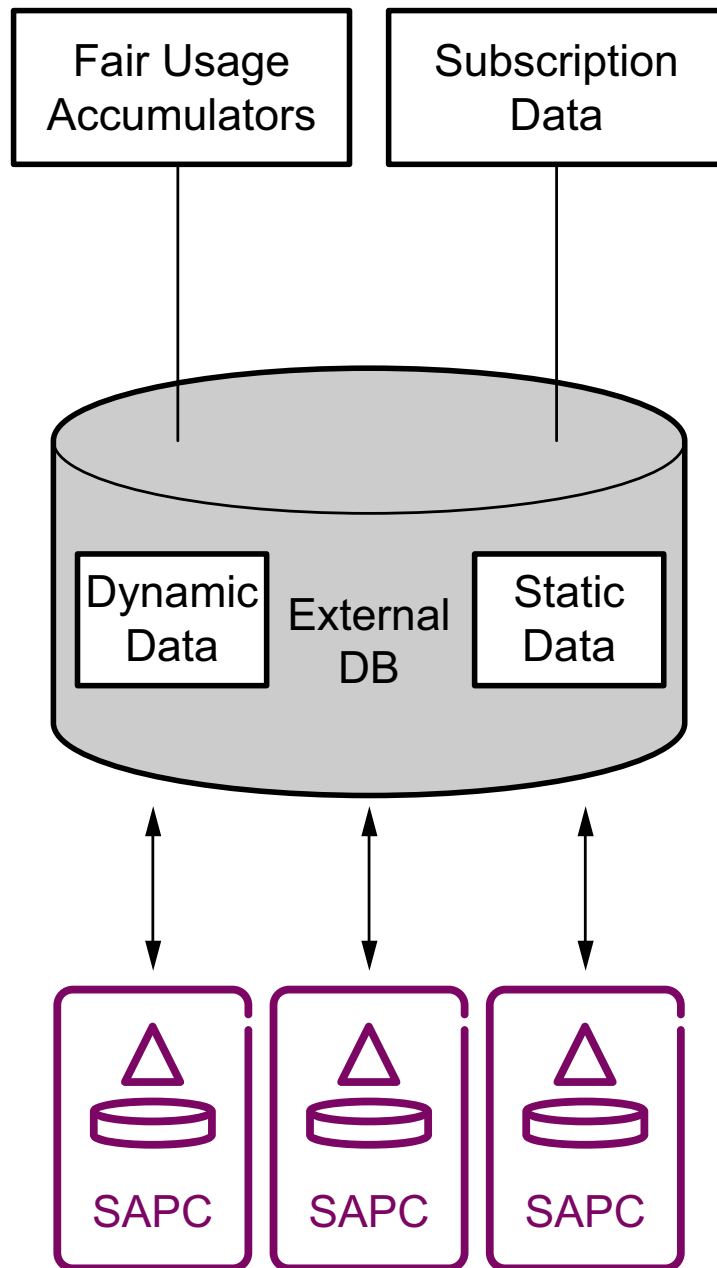


Figure 3 Multi-Site Deployment with Common Database

The subscribers' static data and Fair Usage Accumulators are centralized in the external database. The rest of the static data as Subscriber Groups, Services, and so on, together with operator defined policies, are provisioned in all the SAPCs.



2.2.1.2 Network Dependencies

The following considerations must be taken into account in deployments with multiple SAPCs:

- The SAPC keeps track of the state of the session, hence all the messages belonging to the same IP-CAN session must be routed to the same SAPC.
- In scenarios with dynamic policy control, successful session binding requires that all diameter sessions established over the Gx and Rx reference points for a certain IP-CAN session are routed to the same SAPC.
- Network deployments with multiple PCEFs require that Gx sessions from two or more PCEFs that belong to the same user IP-CAN session, are routed to the same SAPC.
- The Fair Usage with Multiple IP-CAN session functionality requires that all the IP-CAN sessions for the same subscriber are routed to the same SAPC. In external database scenarios, concurrent updates of the subscriber usage accumulators as a result of usage reporting received from different IP-CAN sessions can lead to some concurrent usage reporting is not accumulated as SAPC does not perform concurrency control.
- The Shared Device Plans functionality requires that IP-CAN sessions that share a subscription are routed to the same SAPC.
- The Shared Subscriber Plans functionality requires that IP-CAN sessions from all subscribers sharing usage quota are routed to the same SAPC.

Note: Shared Subscriber Plans are not supported with External Database.

- In scenarios where the SAPC is integrated with an Online Charging System for monetary spending limit reporting, it is recommended that all IP-CAN session establishments from the same subscriber are routed to the same SAPC. This allows the SAPC to establish a single session to the OCS per subscriber.

3 Availability and Scalability Operational Conditions

3.1 Availability and Scalability External Conditions

VIP Gateway routers are not part of a SAPC but are needed in all kinds of deployments of a SAPC.



3.2 Availability and Scalability Function Administration

The following sections list the relevant Operation and Maintenance related actions, alarms, logs, notifications, and statistics data related to the function.

3.2.1 Availability and Scalability Alarms

There are no specific SAPC alarms related to its availability, apart from the ones provided by the platform:

- “COM SA, AMF SI Unassigned” alarm is risen when a processor is restarted.
- “COM SA, CLM Cluster Node Unavailable” alarm is risen when the cluster cannot access a defined membership machine.
- “eVIP, Gateway Unavailable” alarm is risen when contact is lost with an external gateway. If contact is lost to all external gateways, all traffic is lost.

3.2.2 Availability and Scalability Logging

The following events are logged:

- Existing IP Session removed (Basic clean up)
- Non-persistent data, such as Time of Day, Gx session, and Subscriber notifications are empty
- Diameter peer restarted (Gx and Rx massive clean up)
- Start deleting old sessions (Gx and Rx massive clean up)
- End deleting old sessions (Gx and Rx massive clean up)
- Start deleting inactive sessions (session inactivity clean up)
- End deleting inactive sessions (session inactivity clean up)

3.2.3 Availability and Scalability Notifications

There are no specific SAPC notifications related to service availability, apart from the ones provided by the platform.

- "Link down" and "Link up" notifications are risen when a processor is restarted.

3.3 Availability and Scalability Security

Not applicable.





Reference List

Standards

- [1] [Diameter Base Protocol, IETF RFC 6733](#)