

Active-Standby Geographical Redundancy

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	Active-Standby Geographical Redundancy Introduction	1
2	Active-Standby Geographical Redundancy Concepts	1
3	Active-Standby Geographical Redundancy Function	2
3.1	Active-Standby Geographical Redundancy Overview	2
3.2	Active-Standby Geographical Redundancy Data Mirroring	4
3.2.1	Redundant Data	4
3.2.2	Backlog	5
3.2.3	SAPC DBS Synchronization	5
3.3	Active-Standby Geographical Redundancy Network traffic handling	6
3.3.1	Hot Standby Concepts	6
3.3.1.1	Active SAPC	6
3.3.1.2	Standby SAPC	6
3.3.1.3	Preferred SAPC	6
3.3.2	Active-Standby Geographical Redundancy Network connectivity	7
3.4	Active-Standby Geographical Redundancy States	8
3.4.1	Transition between States	9
3.4.1.1	Transitions from Initial State	9
3.4.1.2	Transitions from Active State	10
3.4.1.3	Transitions from Standby State	10
3.4.1.4	Transitions from Halted State	10
3.5	Active-Standby Geographical Redundancy Supervision and Control Functions	11
3.5.1	Mated Peer Supervision	11
3.5.2	Fault Detection and Recovery	13
3.5.2.1	Split Brain Scenario	13
3.5.2.2	Simultaneous SAPC Restart	13
3.5.2.3	Temporary Differences between Databases	14
3.5.3	Handling of the SAPC Origin State Id	14
4	Active-Standby Geographical Redundancy Traffic Cases	15
4.1	Active SAPC Restarts	15
4.2	Standby SAPC Restarts	17
4.3	Replication Channel Unavailable	19
5	Active-Standby Geographical Redundancy Capabilities	20
	Reference List	23





1 Active-Standby Geographical Redundancy Introduction

This document describes the Active-Standby Geographical Redundancy function provided by the SAPC.

2 Active-Standby Geographical Redundancy Concepts

Active SAPC

The SAPC that is processing traffic and provisioning operations.

Asynchronous replication

The data is committed in the active SAPC and then it is replicated to the standby SAPC. Therefore, the standby SAPC lags behind the active SAPC until the data is replicated.

Mated peer

For a SAPC, the mated peer is the other SAPC that is part of the geographical redundancy function.

Replication channel

Connection between the active and standby SAPC peers in a geographical redundant configuration used for data replication and also for geographical redundancy supervision and control functions.

Standby SAPC

The SAPC that is replicating data from the active. This SAPC does not process traffic nor provisioning operations but is ready to take over in case of failure in the active SAPC.



3 Active-Standby Geographical Redundancy Function

3.1 Active-Standby Geographical Redundancy Overview

The SAPC, as a network element, provides high availability as explained in the document *Availability and Scalability*. However, this level of availability does not help in the case of complete power failure, natural disasters, such as fire or earthquakes, or deliberately destructive human behavior, such as bombings or terrorist attacks. Operators may also require the possibility to shut down clusters completely for planned or unplanned maintenance (for example, hardware or software change). The geographical network redundancy function provides this extra level of redundancy at network level. The SAPC with this feature offers a system availability target figure of 99.999%. The SAPC provides two geographical redundancy solutions which enhances the In-Service Performance (ISP) for traffic and O&M interfaces:

— Active-Standby Geographical Redundancy

The solution is based on a hot-standby system (1+1 redundancy), composed of two SAPC peers and network connection allowing communication between them. The active SAPC, processes the incoming traffic and provisioning operations. The standby SAPC, replicates the state from the active one, and it is ready to process the incoming traffic and provisioning when the active SAPC cannot handle it

— Active-Active Geographical Redundancy

The solution is based on a 1+1 redundancy, where both SAPC peers are active. Both SAPC peers, can process the incoming traffic and provisioning operations. Each SAPC keeps the state of the mated peer, and it is ready to process the incoming traffic and provisioning when the mated peer cannot handle it.

In Active-Standby Geographical Redundancy normal operation, both peers work in an active/standby model. The active SAPC processes the incoming traffic and provisioning operations. The standby keeps the state of the active, and it is ready to process the incoming traffic and provisioning when the active SAPC cannot handle it. If the peer being active fails down, the control is automatically taken by the other peer (see Figure 1); this procedure is known as switchover. This switchover is transparent for the neighboring traffic and provisioning plane nodes, as well as for the nodes sending XML/SOAP requests/notifications.

Both SAPC peers are interconnected through the network by a connection, called the replication channel link. The replication channel link is used to transfer changes in database information done in the active SAPC to the standby peer. The replication channel link is also used for geographical redundancy supervision



and control, mainly to monitor the redundancy state of the mated peer and detect when the active SAPC is no longer operational.

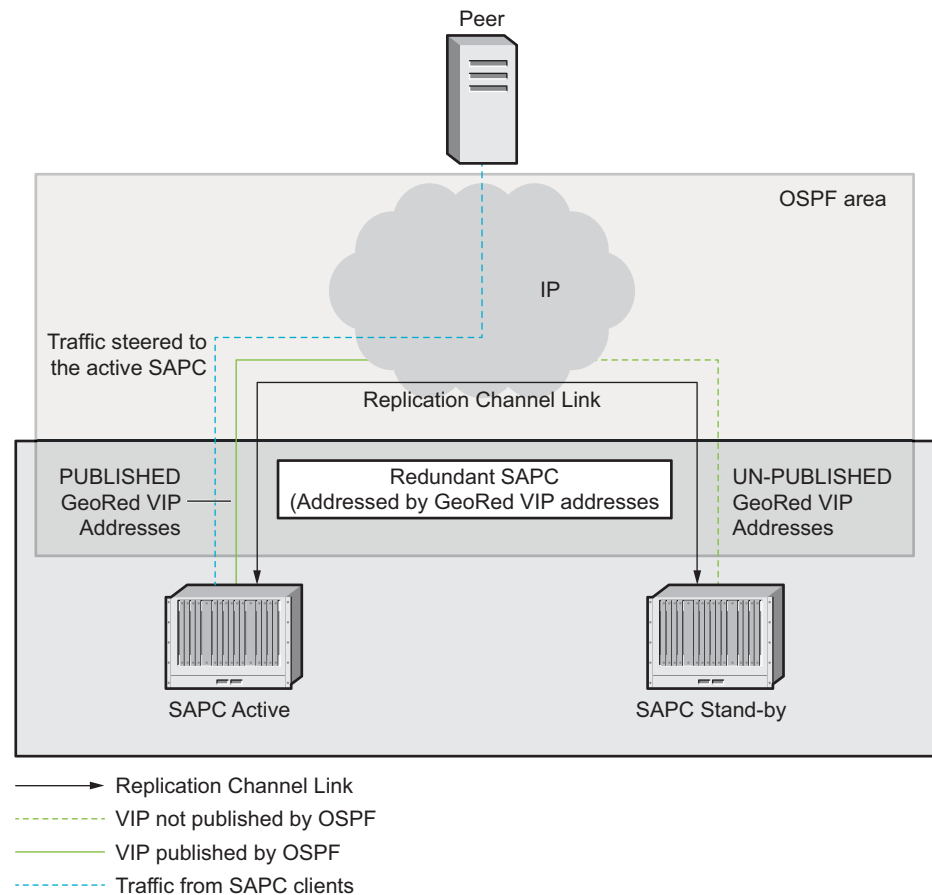


Figure 1 SAPC Geographical Redundancy.

The function consists of the following main parts:

- Data mirroring, which keeps the standby peer up-to-date regarding the SAPC data that is replicated from the active peer.
- Network traffic handling, which routes requests from the network to the active SAPC.
- Geographical redundancy supervision and control, which maintains the correct redundancy state in each SAPC.

To prevent time discrepancies, for example when active and standby peers belong to different time zones, the SAPC in geographical redundancy uses the UTC time standard regardless of the configured time zone.



3.2 Active-Standby Geographical Redundancy Data Mirroring

The SAPC geographical redundancy solution is based on the replication capability provided by the Database Service (DBS) of the Ericsson Component Based Architecture (CBA) platform. This is an asynchronous replication function, that guarantees that database changes done in one SAPC are mirrored and applied in the mated peer.

Data mirroring (replication) between the peers keeps data in the standby peer synchronized with the active peer.. The active peer forwards data transaction updates to the standby peer over the replication channel.

The data mirroring functionality is distributed over all payloads in the SAPC. There is one TCP/IP connection originating from each traffic processor in the active peer, towards a traffic processor in the standby peer.

Note: It is recommended to always have the active and standby SAPC peers equally sized in terms of processing capacity and memory capacity. However, the function does not depend on an exact match.

3.2.1 Redundant Data

To be able to perform a transparent switchover in case of failure between the active and the standby SAPC peers, the following information is replicated:

- Subscriber data (including Subscriber Fair usage accumulated data).
- Provisioning data (policies, subscriber groups data, services, and profiles).
- Session information, notification data and time trigger data.

The SAPC does not replicate any other data which is not stored in the Database Service (DBS). This comprises the following data:

- The SAPC configuration data that is provided through COM, and through configuration (cfg) files is not replicated. Therefore, the configuration data has to be provided to each SAPC individually.
- Licensing information is not replicated. Therefore license information has to be managed in each SAPC individually.
- Alarms, logging events and performance measurements are local to each SAPC.
- User management information and authentication credentials are not replicated.
- Networking configuration is not replicated.



3.2.2 Backlog

Data mirroring is performed asynchronously for performance reasons. Hence, database changes in the active peer are temporarily stored in memory while they are sent to the standby peer. Next, the received changes are saved in the standby peer and a confirmation is sent back to the active peer. The standby peer can then apply the received changes in the local database, and the active peer can release the transmitted data. These memory buffers can be regarded as a backlog.

This procedure allows the SAPC to handle failures during data mirroring and ensure database consistency by forcing the same order of changes in both the active and standby peers. Thus, having a backlog is a normal condition. However, there are situations when some transactions in the backlog cannot be processed immediately, for example, due to overload in either the sender or the receiver side, or disturbances in the replication channel. To handle this case in terms of resource utilization, it is possible to configure the maximum amount of memory used by the backlog.

3.2.3 SAPC DBS Synchronization

Data mirroring synchronizes database changes between the active and standby peers. However, there are a number of cases where the original contents of the databases are different, therefore synchronizing the changes does not make the contents identical either. In the following cases, a complete synchronization of the database is needed:

- After initial startup of a SAPC.
- When transactions have to be discarded because of memory and network capacity limitations.
- After split-brain situations (loss of network connectivity between active and standby peers).
- When data mirroring functionality is resumed after being disabled as a result of an operational procedure, for example when one of the peers is set to Halted state or after a geographical redundancy configuration change.

Synchronization itself transmits a consistent view of the database of the active SAPC to the standby peer, where it is imported. Any changes done on the active peer in the meantime are also transferred as normal database changes, which will then be applied in the standby peer when the initial database is fully imported. The synchronization process is automatically triggered by the SAPC when it is required, no manual intervention is needed. The SAPC DBS component sends notifications about the start and the end (successful or unsuccessful) of the synchronization process.

When a SAPC reloads, it starts synchronization from the still-running peer, that is, the SAPC that reboots copies the database from the active (running) peer. However, in other situations where a complete synchronization is needed, it is not possible to know which of the SAPC peers holds the most up-to-date database.



In those cases, for example after a split-brain situation, the database from the SAPC configured as "preferred" is maintained, and the database from the other peer is discarded.

3.3 Active-Standby Geographical Redundancy Network traffic handling

3.3.1 Hot Standby Concepts

A hot standby solution means that the standby peer is always able to take over traffic in the event of a failure in the active peer. In addition, SAPC traffic handling is adapted to follow the hot standby principles, by always attempting to route traffic to the correct peer. On the other hand, node configuration, operation and maintenance procedures are performed in each SAPC individually.

3.3.1.1 Active SAPC

The traffic is always routed to the node that is considered as active, which means that it processes all traffic. The active node is automatically switched if a fault is detected, see Section 3.4.1 on page 9. The active node can also be switched manually, for instance, to do planned maintenance of the node.

All provisioning operations (including subscription data, services, profiles and policies) are performed in the active node and replicated in the standby node. This is transparent to the provisioning server. Incoming and outgoing traffic messages, IP-CAN session reauthorizations, AF session events, access to external database, interactions with the online charging system, end user notifications, are also handled by the active node.

3.3.1.2 Standby SAPC

The standby node is not allowed to handle traffic nor provisioning operations, but can adopt the role of the active node, when needed.

3.3.1.3 Preferred SAPC

One of the SAPC peers must be configured as the preferred SAPC in geographical redundancy, and this is used when resolving some fault situations where is not possible to know which of the SAPC peers holds the most up-to-date database.

- After recovery of the network connectivity between both SAPC peers (split-brain situation), the database from the preferred SAPC is maintained, and the database from the other peer is discarded.
- When both SAPC peers reload and come up nearly at the same time, the preferred SAPC takes the role of active and provides data to the not preferred SAPC.



3.3.2 Active-Standby Geographical Redundancy Network connectivity

There is only one connection point to the redundant SAPC pair seen from the external network independently of the peer that is handling traffic and provisioning. The redundant SAPC solution exposes by default one VIP address for traffic and another VIP address for provisioning. These are called redundant virtual IP addresses.

The active SAPC notifies, through Open Shortest Path First (OSPF) mechanism, that it is the one that is processing traffic and provisioning, while the standby SAPC does not. If the standby peer has to take the control, it notifies, through OSPF mechanism, that it is going to process the incoming traffic and provisioning. The other SAPC peer removes that notification. The advertisement is done with Link-State Advertisements (LSA) messages. For more information, refer to RFC 2328 – OSPF Version 2, Reference [1].

Each SAPC handles the following VIP addresses:

Traffic GeoVIP	This is the VIP address that the SAPC clients use to send diameter traffic to the SAPC. In network deployments with Online Charging System, this is also the VIP address where the SAPC handles the Sy interface. This VIP is handled by both SAPC peers.
Provisioning GeoVIP	This is the VIP address that the SAPC clients use to send provisioning orders to the SAPC. This VIP is handled by both SAPC peers.
ExtDB GeoVIP	In network deployments with the CUDB or an external database function, this is the VIP address used to provide access to an external database system and receive SOAP notifications. This VIP is handled by both SAPC peers.
Replication VIP	This is the VIP address that the SAPC exposes for the replication channel to the mated peer. Each SAPC has its own replication VIP.
O&M Local VIP	There is an extra local VIP associated to each SAPC. This is the VIP used to manage the SAPC information model through COM.

The following picture shows all the IP addresses involved in the geographical redundancy scenario with details about which IP address is advertised by each SAPC.

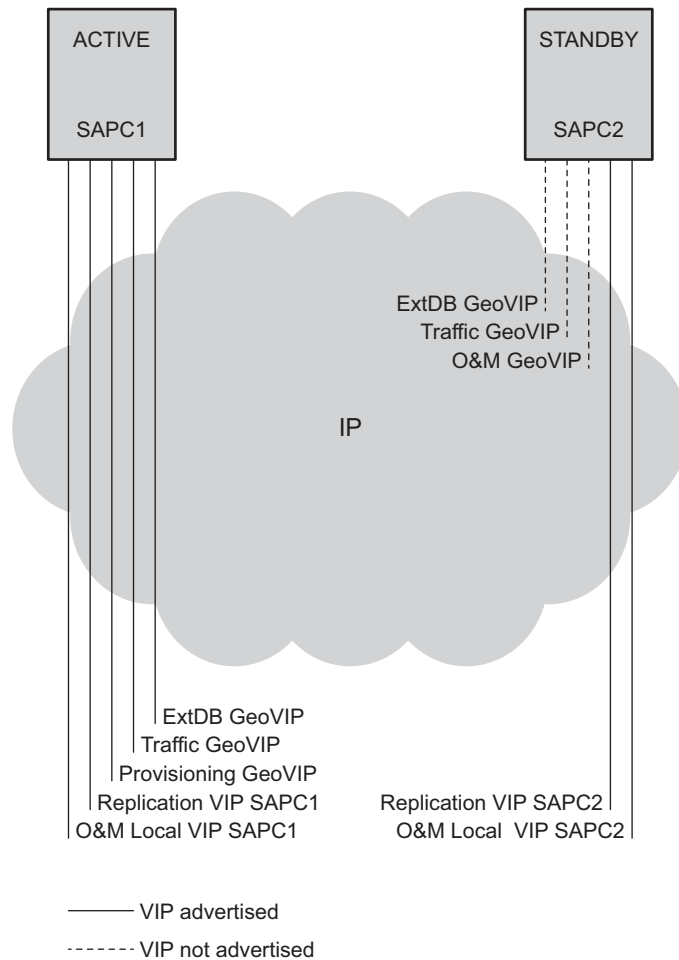


Figure 2 VIP addresses in the SAPC geographical redundancy solution.

3.4 Active-Standby Geographical Redundancy States

The geographical redundancy function is initiated in the SAPC by performing an operational procedure to promote one of the SAPC peers to active and another to standby. The SAPC state reflects the traffic handling ability of each of the peers. The valid states are the following:

- **Initial:** This is the initial state when the SAPC is installed. The SAPC does not handle traffic as it does not publish the GeoVIP addresses for traffic, external database and provisioning. The SAPC can only transition from this state when ordered by operational procedure.
- **Active:** The SAPC is handling traffic and provisioning while the mated peer is replicating the changes. In this state, the SAPC publishes the GeoVIP addresses for traffic, external database and provisioning.



- Standby: The SAPC is not handling traffic nor provisioning but is replicating the changes from the active SAPC. Therefore, it is ready to take over traffic if the active fails. In this state, the SAPC does not publish the GeoVIP addresses for traffic external database and provisioning.
- Halted: The SAPC is in this state when the geographical redundancy function is stopped as a result of an operational procedure. In this state, the SAPC does not handle traffic nor provisioning. The SAPC does not publish the GeoVIP addresses for traffic, external database and provisioning and does not replicate database changes from the mated peer. However, the O&M Local VIP can be used for SAPC configuration through COM.

3.4.1 Transition between States

The SAPC executes transitions from one state to another depending on the information provided by the supervision functions explained in Section 3.5 on page 10.

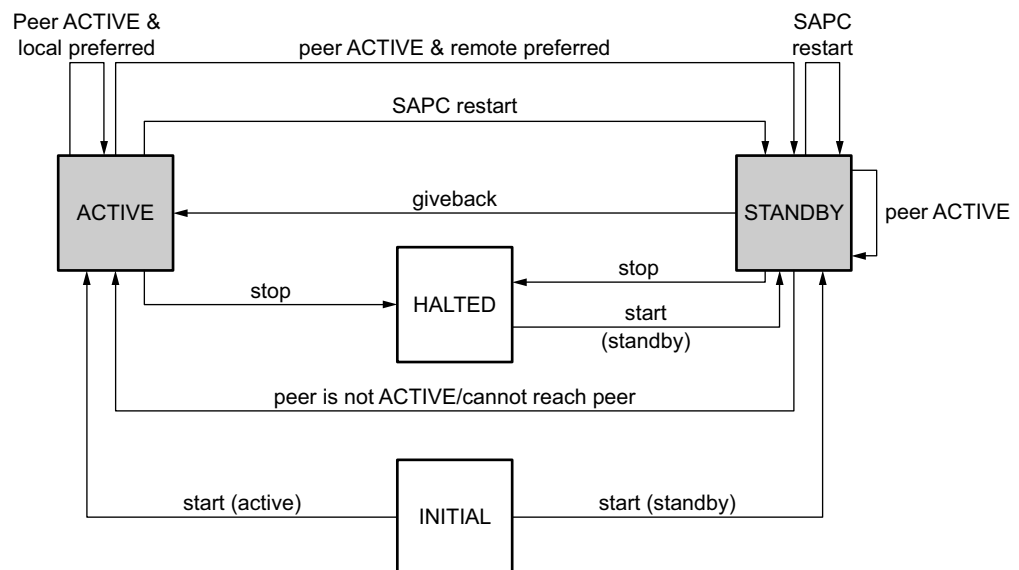


Figure 3 SAPC geographical redundancy state machine

3.4.1.1 Transitions from Initial State

In this state, a SAPC can make the following transitions when ordered by operational procedure:

- Initial to Active. This is done when the SAPC is ordered to start geographical redundancy with the role of active peer.
- Initial to Standby. This is done when the SAPC is ordered to start geographical redundancy with the role of standby peer.



3.4.1.2 Transitions from Active State

In this state, a SAPC can make the following transitions:

- Active to Standby. This is done in the following situations:
 - The SAPC restarts and determines that the mated peer is available and in active state.
 - The SAPC detects that the Database Service (DBS) is temporarily unavailable.
 - The SAPC determines that the mated peer is the preferred active peer and is currently in active state. This can happen after a split brain situation as explained in Section 3.5.2.1 on page 13.
- Active to Halted. This is done when ordered by operational procedure. This transition is not allowed if the mated peer is in Initial or Halted state.

The SAPC remains in active state when it is the preferred active peer and detects that the mated peer is also in active state.

3.4.1.3 Transitions from Standby State

In this state, a SAPC can make the following transitions:

- Standby to Active. This is done in the following situations:
 - The SAPC determines that the mated peer is unavailable, in standby state or in halted state.
 - The SAPC configured as the preferred node is ordered by operational procedure to return to active state, typically after a takeover.
- Standby to Halted. This is done when ordered by operational procedure. This transition is only allowed if the mated peer is in active state or unavailable.

The SAPC remains in standby state when the SAPC restarts and determines that the mated peer is available and in active state.

3.4.1.4 Transitions from Halted State

In this state, a SAPC can make the following transitions:

- Halted to Standby. This happens when the geographical redundancy function is restarted by operational procedure.



3.5 Active-Standby Geographical Redundancy Supervision and Control Functions

The Geographical Redundancy control function has the following responsibilities:

- Supervise the own state of the SAPC to recognize when there is an internal failure, such as if the Database Service (DBS) is temporarily unavailable.
- Monitor the redundancy state of the mated peer, to determine when it is unreachable or unavailable and whether it is in active, standby or halted state.
- Manage the geographical redundancy state machine, and take the appropriate actions in each state or transition. This includes:
 - To publish or not publish the GeoVIP addresses for traffic, access to external database and provisioning.
 - Set the correct value for the SAPC `Origin State Id` upon failure.
- Handle the operational procedures to perform transitions from the Initial state and transitions to the Halted state, and provide configuration, logs and alarms related to the geographical redundancy functionality.

If the geographical redundancy control function detects that the Database Service (DBS) is temporarily unavailable in the active SAPC, a switchover is executed and the standby SAPC takes the role of active SAPC.

If the two System Controllers in the SAPC cluster fail in geographical redundancy configuration, the SAPC restarts. If the failure occurs in the active SAPC, a switchover is executed and the standby SAPC takes the role of active SAPC. This is to ensure that the O&M and provisioning functions are available in the SAPC when the two System Controllers fail.

3.5.1 Mated Peer Supervision

The SAPC stores its own replication state in a specific Managed Object Class (MOC) that can be managed via the NETCONF interface. This MO also stores the previous state and a time stamp when the transition between states happened.

The SAPC uses a heartbeat mechanism to monitor the availability and redundancy state of the mated peer through the IP network. This heartbeat is sent to the replication VIP address of the mated peer.

- The active SAPC monitors the availability of the standby peer at regular time intervals, by sending a heartbeat signal. This allows the active SAPC to detect a problem in the replication channel or in the mated peer. The heartbeat interval has a default value of 5 seconds.
- The standby peer answers the heartbeat signal by sending an acknowledgment to the active SAPC.



- If the active SAPC does not receive the acknowledgement to the heartbeat signal after several attempts, the mated peer is declared unreachable and an alarm is raised. The number of reattempts has a default value of 3 retransmissions.
- If the standby peer does not receive a heartbeat signal during more than a predefined time period, it transitions to active state. The maximum time period before assuming that the mated peer is unavailable is derived from the heartbeat interval and the number of reattempts as $\text{Heartbeat_Timeout} = \text{Heartbeat_Interval} * (1 + \text{Number_of_Reattempts})$
- If the active SAPC transitions to halted state by means of operational procedure, it notifies the standby peer, and the standby peer transitions to active state. This allows the mated peer to react immediately upon the state transition. The SAPC in halted state also answers the heartbeat signal by sending an acknowledgment to the active SAPC.
- If the standby peer transitions to halted state by means of operational procedure, the standby peer signals the active SAPC and waits for the answer. This allows the SAPC to determine if the mated peer is in active state, and so permit the state transition.

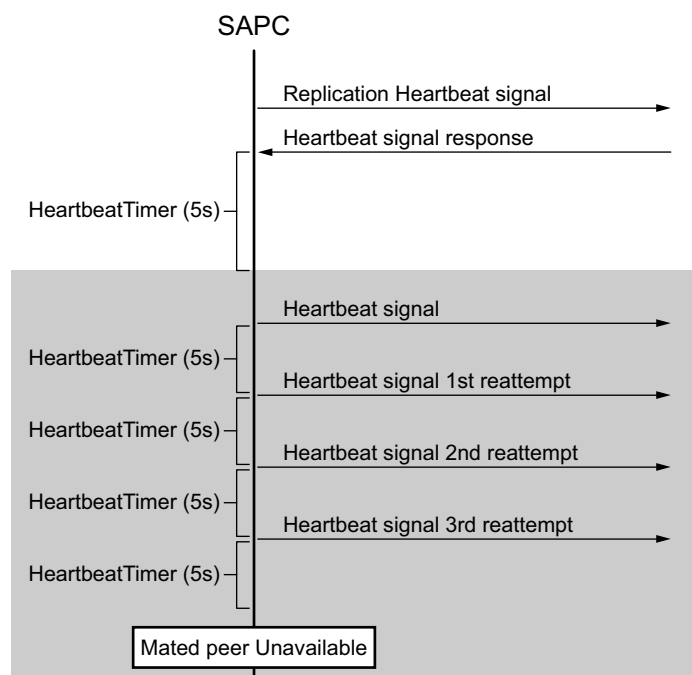


Figure 4 SAPC mated peer supervision



3.5.2 Fault Detection and Recovery

The basic principle for the geographical redundancy function is that the active SAPC processes all incoming traffic and provisioning operations, while the standby peer keeps the state of the active, and it is ready to take over when the active SAPC fails. However, there are situations that might cause the standby peer not to be synchronized with the active SAPC. This chapter describes those scenarios and how they are handled.

3.5.2.1 Split Brain Scenario

This situation happens when the standby SAPC detects a failure in the mated peer. The standby SAPC switches to active state and starts to announce the GeoVIP addresses to serve the traffic. If the active SAPC is healthy, that is, only the heartbeat is lost (owing to loss of connectivity between the peers), the active SAPC continues announcing the GeoVIP address. In this scenario, the end result consists of two SAPC peers in active state announcing the GeoVIP addresses to serve traffic and provisioning. This situation is known as a split brain scenario.

The split brain has the following consequences:

- The standby peer switches to active state, so it starts announcing the GeoVIP addresses and serving traffic.
- Both SAPC peers may commit traffic and provisioning operations so the databases may become inconsistent because changes cannot be replicated.
- Traffic operations may fail if the messages related to a session are steered to the SAPC that does not hold the correct session information.

When the connectivity between the SAPC peers is re-established, the peers will not communicate database changes until a complete synchronization is performed. Then, the SAPC that is not configured as the preferred node is automatically restarted. When this SAPC reloads, it recovers the most current database information from the active SAPC and sets the replication state to standby.

3.5.2.2 Simultaneous SAPC Restart

If both SAPC peers (active and standby) fail, they probably do not reload at the same time. Therefore, the first one that restarts will observe loss of network connectivity to the mated peer and take the necessary actions, as described in Section 3.5.2.1 on page 13.

If both SAPC peers come up nearly at the same time and both observe that its peer is already running, both of them will try to synchronize data from the peer. Resolving this situation is automatic. The SAPC that is configured as preferred, provides data to the not preferred SAPC. The procedure requires a complete synchronization of the non-preferred SAPC.



3.5.2.3 Temporary Differences between Databases

The data mirroring functionality makes use of backlogs while transferring database changes from the active to the standby SAPC is described in Section 3.2.2 on page 4. Normally, any specific transaction should be processed soon, and, therefore, be removed from the backlog. If a transaction remains in a queue for more than a minute, an alarm is raised to report that redundancy is compromised. If the database differences turn out to be temporary, the alarm is automatically cleared.

However, it can happen that SAPC overload or network connectivity problems persist, and the backlog reaches the configured memory limit and transactions have to be dropped by either the active or the standby SAPC. In this case, the only way to reach a state where the database contents are the same is to fully synchronize the databases. Then, the SAPC that is configured as preferred, provides data to the not preferred SAPC by using the procedure described in Section 3.2.3 on page 5.

3.5.3 Handling of the SAPC Origin State Id

In a stand-alone deployment, when the SAPC recovers from a restart, the database information recovered from the backup may not be fully up to date. Hence the SAPC increments its own `Origin State Id` and includes the new value in every response message alerting the peer diameter nodes about the loss of previous session state.

In a deployment with geographical redundancy, the `Origin State Id` information is replicated between the active and standby SAPC peers. Upon a node restart, the SAPC does not increment the `Origin State Id`, but obtain the `Origin State Id` value together with the most up-to-date database information during synchronization with the mated peer. This enables the standby SAPC to send the same `Origin State Id` value as the active SAPC, upon a failure in the active SAPC. Transitions between SAPC redundancy states do not increase the value of the `Origin State Id`, as those are transparent to the peer diameter nodes in the external network.

In a deployment with geographical redundancy, the SAPC only increments the `Origin State Id` if both SAPC peers restart. This is, when the SAPC recovers from a restart and cannot replicate the `Origin State Id` information (cannot sync with the mated peer, for example due to loss of network connectivity in the replication channel), the SAPC increments its own `Origin State Id`.

When the active or standby SAPC detects a split-brain situation, the own `Origin State Id` is not increased in order to provide service availability, as the session information is available in both SAPC peers. When the network connectivity is re-established, the SAPC that is configured as preferred continues providing service, but the `Origin State Id` is not increased.

If a SAPC restarts in a split-brain situation, it autonomously increments its own `Origin State Id`, and this information cannot be replicated in the mated peer. As a result each SAPC may send a different `Origin State Id` value as long as the



replication channel is unavailable. When the network connectivity is recovered the `Origin State Id` in the SAPC that has been configured as preferred is maintained.

4 Active-Standby Geographical Redundancy Traffic Cases

This chapter explains the high level interactions that occur in the most common use cases for the Active-Standby Geographical Redundancy functionality:

- Active SAPC restarts.
- Standby SAPC restarts.
- Replication channel unavailable.

4.1 Active SAPC Restarts

The following figure shows the high level flow that takes place when the active SAPC restarts, and the main actions taken by the SAPC to perform the Active-Standby Geographical Redundancy functionality.

A failure in the active SAPC makes the standby peer transition to active state and start publishing the GeoVIP addresses and processing traffic. When the SAPC recovers from the restart and completes synchronization from the mated peer, it takes the role of the standby peer. During the transition, ongoing traffic events may fail and diameter connections need to be reestablished, in a similar way to when a temporary connectivity problem occurs in the diameter link. The duration of the traffic switch depends on several factors, see Section 5 on page 20.

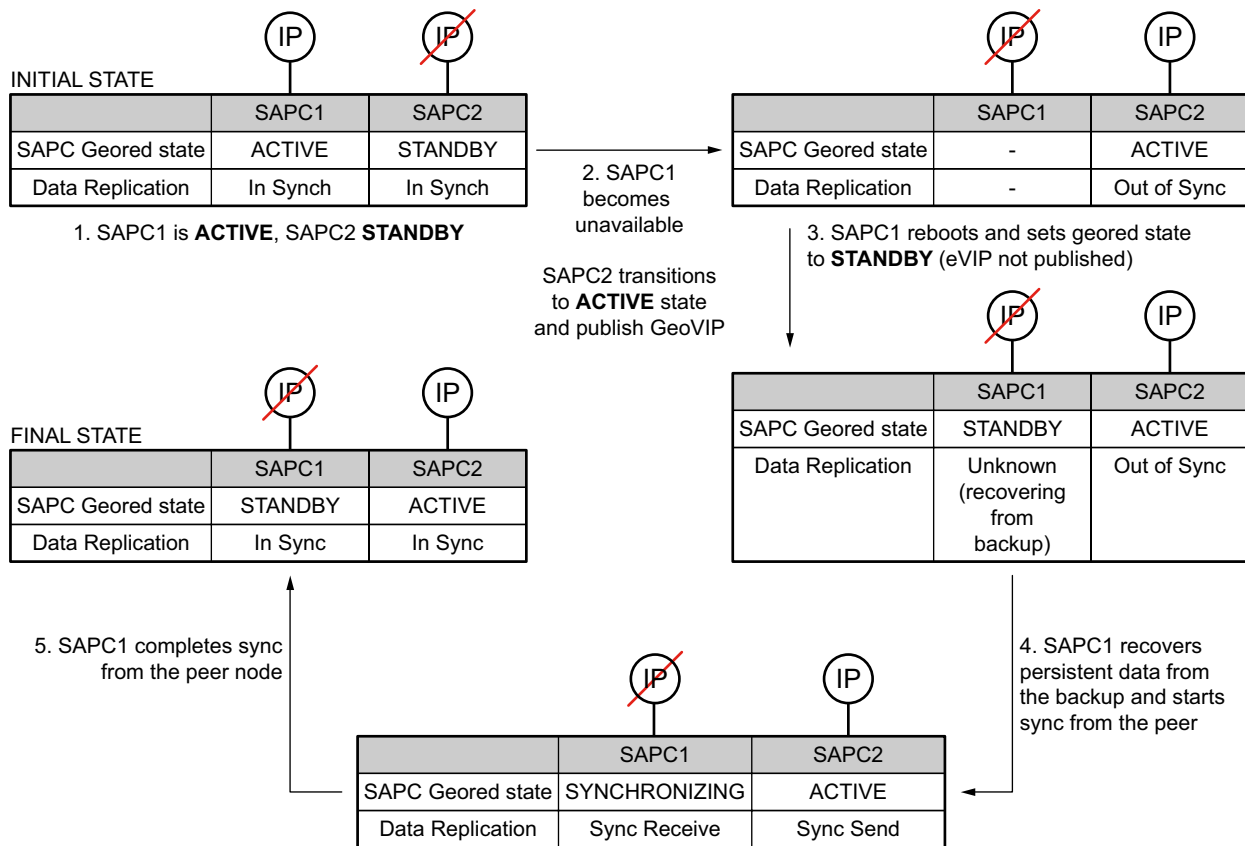


Figure 5 Active SAPC restarts

1. This is the initial working condition for geographical redundancy. The SAPC1 is in active state publishing the GeoVIP addresses and processing traffic, whilst the SAPC2 is in standby state, not publishing the GeoVIP addresses, not processing traffic and replicating the changes from the active SAPC. Data mirroring (replication) is fully operational and keeps data in the standby peer synchronized with the active SAPC.
2. SAPC1 fails, SAPC2 detects that SAPC1 is not available (due to heartbeat timeout), rises an alarm (Unable to Reach Peer), transitions to active state and publish the GeoVIP addresses. Data mirroring is interrupted and the SAPC DBS component also rises an alarm (Connection Loss). Database transactions that were pending to be replicated, are dropped. As the SAPC2 starts to handle traffic, database changes are applied but can no longer be replicated in the mated peer. This makes the SAPC DBS component to rise another alarm in the SAPC2 (Synchronization Needed).
3. SAPC1 completes the software reload, connects to the mated peer (which is in active state) and sets the redundancy state to standby (GeoVIP addresses not published). The SAPC2 clears the corresponding alarms (Connection Loss, Unable to Reach Peer).



4. Simultaneously with the previous step, the SAPC1 recovers all persistent database data from the latest backup and detects that the local database is out of sync. Then the SAPC DBS component rises an alarm (*Initial Synchronization Needed*) and starts synchronization from the active SAPC. The synchronization process transmits a snapshot of the database of the active SAPC to the standby peer, where it is imported. Any changes done on the active SAPC in the meantime are also transferred as normal database changes, which will then be applied in the standby peer when the base view is fully imported. The SAPC DBS component in the SAPC2 also clears the corresponding alarm (*Synchronization Needed*).
5. SAPC1 completes successfully synchronization from the active SAPC2 and clears the corresponding alarm (*Initial Synchronization Needed*). In the final state, the SAPC2 is in active state publishing the GeoVIP addresses and processing traffic, whilst the SAPC1 is in standby state replicating the changes from the active SAPC.

4.2 Standby SAPC Restarts

The following figure shows the high level flow that takes place when the standby SAPC restarts, and the main actions taken by the SAPC to perform the Active-Standby Geographical Redundancy functionality.

A failure in the standby SAPC does not affect the SAPC traffic handling capability and is not detected by the external network nodes. However, there is a time period during which the standby SAPC is not ready to take over with up-to-date database information, in the event of a failure of the active peer.

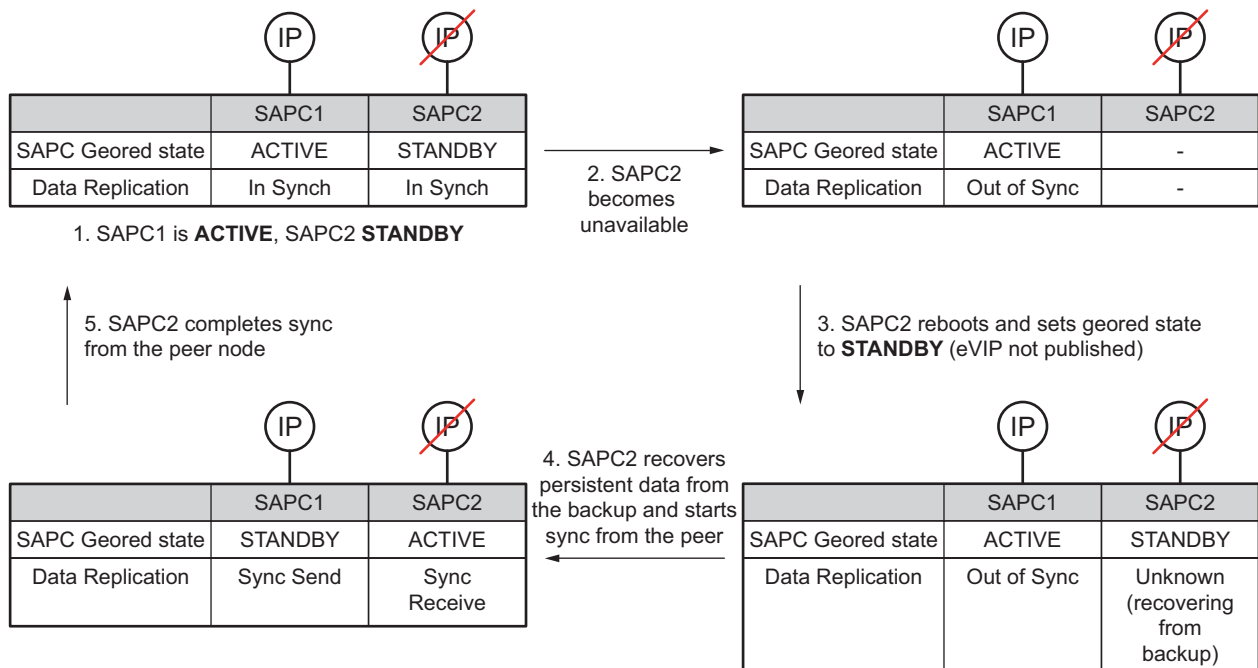


Figure 6 Standby SAPC restarts

1. This is the initial working condition for geographical redundancy. The SAPC1 is in active state publishing the GeoVIP addresses and processing traffic, whilst the SAPC2 is in standby state, not publishing the GeoVIP addresses, not processing traffic and replicating the changes from the active peer. Data mirroring (replication) is fully operational and keeps data in the standby SAPC synchronized with the active peer.
2. SAPC2 fails, SAPC1 detects that SAPC2 is not available (due to heartbeat timeout) and rises an alarm (Unable to Reach Peer). Data mirroring is interrupted and the SAPC DBS component also rises an alarm (Connection Loss). Database transactions in the active SAPC that are pending to be sent to standby SAPC and those that arrived from the active SAPC but are still not applied in the standby SAPC, are dropped. As the SAPC1 continues to handle traffic, database changes are applied but can no longer be replicated in the standby SAPC. This makes the SAPC DBS component to rise another alarm in the SAPC1 (Synchronization Needed).
3. SAPC2 completes the software reload, connects to the mated peer (which is in active state) and sets the redundancy state to standby (GeoVIP addresses not published). The SAPC1 clears the corresponding alarms (Connection Loss, Unable to Reach Peer).
4. Simultaneously with the previous step, the SAPC2 recovers all persistent database data from the latest backup and detects that the local database is out of sync. Then the SAPC DBS component rises an alarm (Initial Synchronization Needed) and starts synchronization from the active SAPC. The synchronization process transmits a snapshot of the database of the



active SAPC to the standby SAPC, where it is imported. Any changes done on the active SAPC in the meantime are also transferred as normal database changes, which will then be applied in the standby SAPC when the base view is fully imported. The SAPC DBS component also clears the corresponding alarm in the SAPC1 (Synchronization Needed).

- SAPC2 completes successfully synchronization from the active SAPC1 and clears the corresponding alarm (Initial Synchronization Needed). The system goes back to the initial state, where SAPC1 is active and SAPC2 standby.

4.3 Replication Channel Unavailable

The following figure shows the high level flow that takes place when the replication channel becomes temporarily unavailable, and the main actions taken by the SAPC to perform the Active-Standby Geographical Redundancy functionality.

A failure in the replication channel results in both SAPC peers taking the role of active SAPC, publish the GeoVIP addresses and start handling traffic. In this situation, the databases become inconsistent. When the connectivity is re-established, the database from the SAPC configured as preferred is maintained, and the database from the peer is discarded.

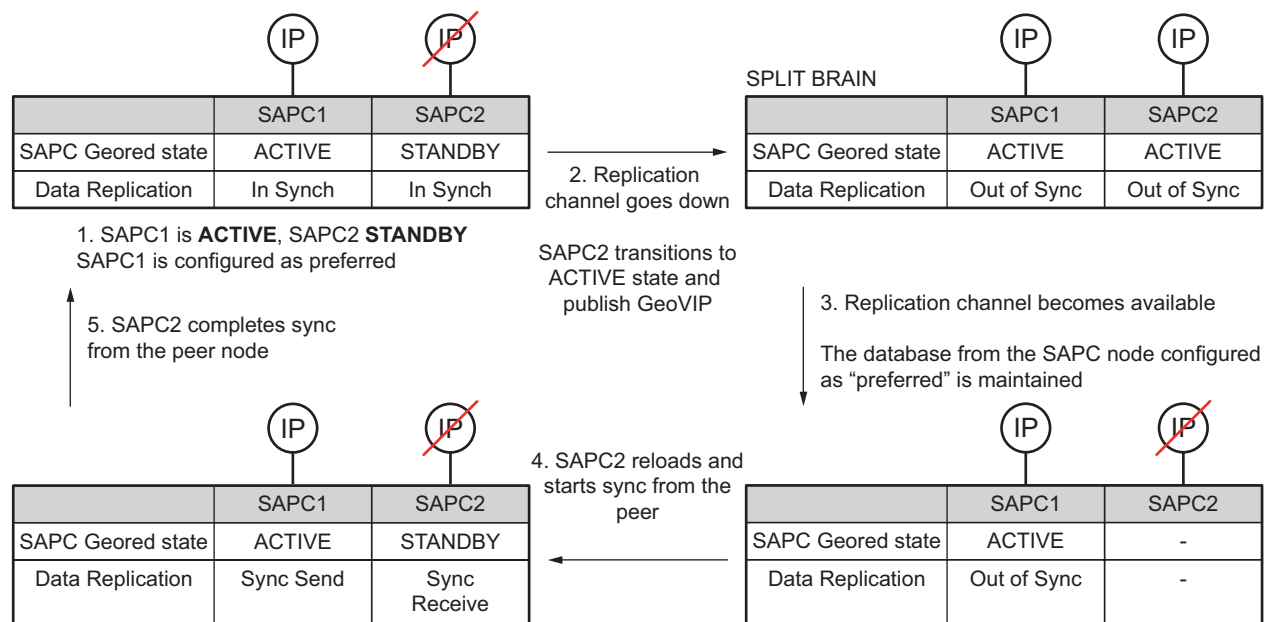


Figure 7 Replication channel temporarily unavailable

- This is the initial working condition for geographical redundancy. The SAPC1 is in active state publishing the GeoVIP addresses and processing traffic, whilst the SAPC2 is in standby state, not publishing the GeoVIP addresses,



not processing traffic and replicating the changes from the active SAPC. Data mirroring (replication) is fully operational and keeps data in the standby SAPC synchronized with the active SAPC. The SAPC1 is configured as the preferred SAPC for geographical redundancy.

2. The replication channel fails, SAPC1 detects that SAPC2 is not available and rises an alarm (`Unable to Reach Peer`). SAPC2 detects that SAPC1 is not available, also rises an alarm, transitions to active state and publish the GeoVIP addresses. Data mirroring is interrupted and the SAPC DBS component also rises an alarm (`Connection Loss`) in both peers. This is a split brain scenario. Both SAPC peers start handling traffic, database changes are applied locally but can no longer be replicated, so they become inconsistent.
3. The replication channel becomes available and both peers clear the corresponding alarms (`Unable to Reach Peer`, `Connection Loss`). Data mirroring restarts and detects that a complete synchronization is needed. Then, the SAPC that is not configured as the preferred SAPC is automatically restarted.
4. SAPC2 completes the software reload, connects to the mated peer (which is in active state) and sets the redundancy state to standby (GeoVIP addresses not published). Then the SAPC2 recovers all persistent database data from the latest backup and starts synchronization from the active SAPC. In the meantime the SAPC1 continues to handle traffic.
5. SAPC2 completes successfully synchronization from the active SAPC1 and the system goes back to the initial state.

5 Active-Standby Geographical Redundancy Capabilities

For Geographical Redundancy the following capabilities must be considered:

- Enough bandwidth. The replication channel must be dimensioned to be able to handle the required bandwidth according to the traffic scenario and hardware configuration.
- Link quality. The characteristics of the link (latency and error rate) set a limit in the maximum throughput that can be achieved in the replication channel, and this throughput must fulfil the bandwidth requirements. To avoid the effects of poor performance in the replication channel link, the maximum One-Way Delay (OWD) of the replication channel, must be no more than 20 ms and the packet loss rate must be no more than 0,0001.



- System dimensioning. The system limit is imposed by the maximum sustained load at which the standby SAPC can replicate from the active SAPC without lagging behind.
- Backlog size. The maximum size of the backlog must be correctly dimensioned in order to cope with temporary overload in either the sender or the receiver side, or disturbances in the replication channel.

Regarding response times, the following capabilities must be considered for SAPC transitions:

- The time to detect loss of connectivity to the mated peer depends on the configured values for the heartbeat interval and number of re-attempts, according to the reliability of the operator transport network. Typical values to perform a transition from standby to active, may vary from 5 seconds to 20 seconds.
- Time to stop advertising the GeoVIP addresses at transition from active to other state: a Link State Update is sent after one second (controlled by the transmit_delay OSPF parameter) without the VIP address. The LSA is (the same OSPF parameter) flooded on all the interfaces from the SAPC router, and throughout the network after one second. The next step is that the SAPC routers and the other routers in the network recalculate their routing tables. The exact time depends on the time needed to recalculate the routes. The time required to run the algorithm depends on a combination of the size of the area and the number of routes in the database. It can take up to 10 seconds. After that, the routers stop sending packets to the former active SAPC.
- Time to start advertising the GeoVIP at transition from standby to active state: a Link State Update packet is sent after one second (controlled by the transmit_delay OSPF parameter provided by eVIP component) with the VIP address. The LSA is (the same OSPF parameter) flooded on all the interfaces from the SAPC router and throughout the network after one second. The next step is that the SAPC routers and the other routers in the network recalculate their routing tables. The exact time depends on the time needed to recalculate the routes. The time required to run the algorithm depends on a combination of the size of the area and the number of routes in the database. It can take up to 10 seconds. After that, the routers start sending packets to the active SAPC.





Reference List

Ericsson Documents

- [1] Availability and Scalability
- [2] Active-Active Geographical Redundancy

Standards

- [3] RFC 2328 – OSPF Version 2