

Active-Active Geographical Redundancy

Ericsson Service-Aware Policy Controller

FACILITY DESCRIPTION

Copyright

© Ericsson España, S.A. 2017. 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.

Abstract

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



Contents

1	Active-Active Geographical Redundancy Introduction	1
1.1	Active-Active Geographical Redundancy Concepts	1
2	Active-Active Geographical Redundancy Function	2
2.1	Active-Active Geographical Redundancy Overview	2
2.2	Active-Active Geographical Redundancy Data Mirroring	3
2.3	Active-Active Geographical Redundancy Network traffic handling	6
2.4	Active-Active Geographical Redundancy States	9
2.5	Active-Active Geographical Redundancy Supervision and Control Functions	13
3	Active-Active Geographical Redundancy Traffic Cases	18
3.1	SAPC in Distributed state restarts	18
3.2	Replication Channel unavailable but Application Channel available	20
3.3	Both Replication Channel and Application Channel unavailable	22
4	Active-Active Geographical Redundancy Capabilities	24
	Reference List	25





1 Active-Active Geographical Redundancy Introduction

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

1.1 Active-Active Geographical Redundancy Concepts

Active SAPC	A SAPC that is processing traffic and can handle provisioning operations.
Application Channel	Connection between SAPC peers in a geographical redundant configuration used for geographical redundancy supervision and control functions.
Asynchronous replication	The data is committed in an active SAPC and then it is replicated to the mated peer.
Database ownership	When opening a geographical redundant object, a lock is held by the local SAPC, using a mechanism called “ownership” of the relevant object. This means that the next access to the same data from the SAPC mated peer, is possible and consistent, but requires more processing and adds latency.
Failover	Mechanism that allows to SAPC neighbour peers to switch to a redundant SAPC upon the failure of the previously active SAPC.
Mated peer	For a SAPC, the mated peer is the other SAPC that is part of the geographical redundancy function.
Replication channel	Connection between SAPC peers in a geographical redundant configuration, used for data replication and also for geographical redundancy supervision and control functions.
Subscriber and Session Stickiness	Feature provided by Diameter Clients to bind all the session requests for the same subscriber, to a specific SAPC peer instance. This prevents the database ownership latency.



2 Active-Active Geographical Redundancy Function

2.1 Active-Active 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-Active Geographical Redundancy, if one of the SAPC peers fails down, the neighbour nodes shall failover to the other SAPC (see Figure 1). The failover is not transparent for the SAPC clients (see Section 2.3.1 on page 6).

Both SAPC peers are interconnected through two different network connections, called the Replication Channel and the Application Channel links. The Replication Channel link is used to transfer changes in database information done in one SAPC to the mated peer. The Replication Channel link is also used for geographical redundancy supervision and control, mainly to monitor the redundancy state of the mated peer. The Application Channel is only used for geographical redundancy supervision and control, to detect the availability of the mated peer, even when the Replication Channel fails.

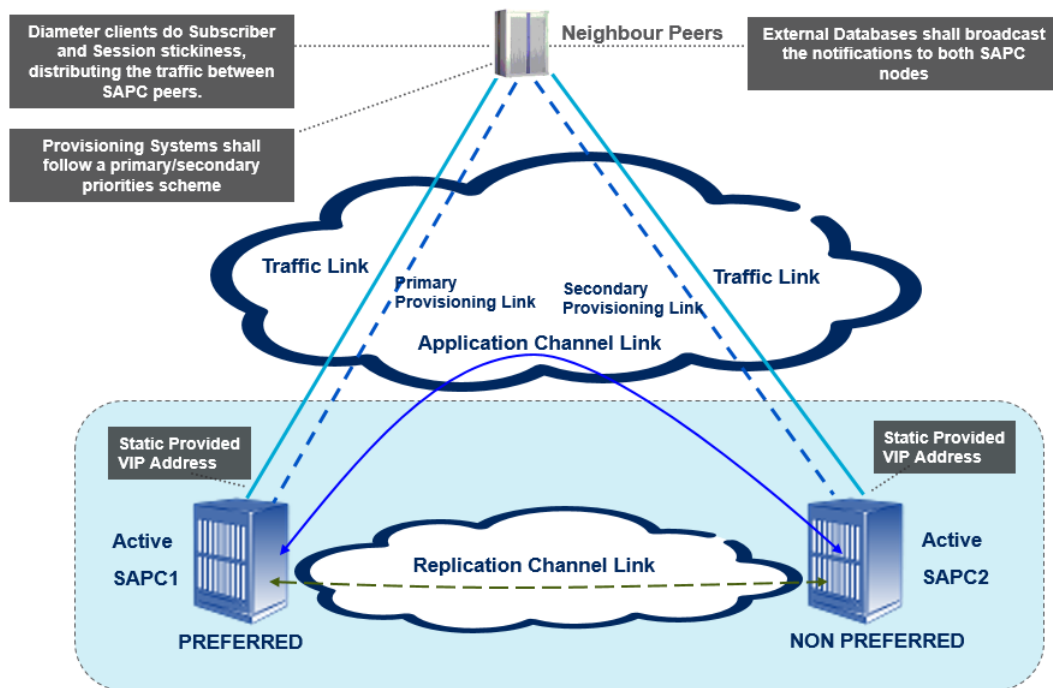


Figure 1 SAPC Active-Active Geographical Redundancy

The function consists of the following main parts:

- Data mirroring, which keeps each active SAPC up-to-date regarding the mated peer data.
- Geographical redundancy supervision and control, which maintains the correct redundancy state in each SAPC.

To prevent time discrepancies in session management, for example when each SAPC belongs to different time zones, the SAPC in geographical redundancy uses the UTC time standard regardless of the configured time zone.

2.2 Active-Active 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 each SAPC, are mirrored and applied in the mated peer.

Data mirroring between the SAPC peers, keeps data in each SAPC synchronized with the mated peer. The SAPC that commits data, forwards data transaction updates to the mated 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 one SAPC, towards a traffic processor in the mated peer.

When opening a geographical redundant object, a lock is held by the local SAPC, using a mechanism called “ownership” of the relevant object. This means that the next access to the same data from the SAPC mated peer, is possible and consistent, but requires more processing and adds latency.

Note: It is recommended to always have both SAPC peers equally sized in terms of processing capacity and memory capacity. Each SAPC shall also be sized to support the processing of maximum amount of traffic (dimensioned traffic for the whole system), when one SAPC is down or unavailable.

2.2.1 Redundant Data

To allow neighbour peers a transparent failover, between one SAPC to the mated peer, 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). It is responsibility of the operator, to align the non replicated data between both SAPC peers. This comprises the following data:

- The SAPC configuration data that is provided through COM, and through configuration (cfg) files. Therefore, the configuration data has to be provided to each SAPC individually.
- Licensing information has to be managed in each SAPC individually. For capacity licenses based on SAPC capacity measures, it shall be considered that both SAPC peers participate in each measure. For example, the SAPC mobileActiveSessions capacity measure, counts the total number of simultaneous active mobile sessions in both SAPC peers.
- Alarms, logging events and performance measurements are local to each SAPC.
- User management information and authentication credentials.
- Networking configuration.

2.2.2 Backlog

Data mirroring is performed asynchronously for performance reasons. Hence, database changes in one SAPC are temporarily stored in memory while they are sent to the mated peer. Next, the received changes are saved and a confirmation



is sent back to the mated peer. The SAPC can then apply the received changes in the local database, and the mated peer can release the transmitted data. These memory buffers in both SAPC peers, 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 SAPC 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, because of overload in either the sender or the receiver side, or disturbances in the Replication Channel. To handle this case in terms of resource use, it is possible to configure the maximum amount of memory used by the backlog.

2.2.3 SAPC DBS Synchronization

Data mirroring synchronizes database changes between SAPC 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. In the following cases, a complete synchronization of the database is needed:

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

Synchronization itself, transmits a consistent view of the database from one SAPC to the mated peer. Any changes done in the SAPC transmitting the data in the meantime, are also transferred as normal database changes, which are applied in the mated peer once 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 mated 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 preferred SAPC is maintained, and the non-preferred SAPC synchronizes from the preferred one.



The SAPC that starts a complete synchronization of the database with the mated peer data, can not neither handle traffic nor provisioning operation until the synchronization is finished.

2.3 Active-Active Geographical Redundancy Network traffic handling

2.3.1 Active-Active Geographical Redundancy Requirements

An Active-Active solution, means that any of the SAPC mated peers can handle traffic and provisioning requests from the client peers, maintaining a SAPC synchronization between them (see Section 2.2.3 on page 5). It is required in the neighboring traffic and provisioning plane the following:

- diameter clients/DRA nodes shall support subscriber and session stickiness (preventing the latency from the database ownership), distributing homogeneously the traffic between both active SAPC peers. One of the possible solutions is that the diameter client/DRA distribute traffic based on subscribers range. Each SAPC should be assigned with a range, ensuring the traffic from each range is balanced.
- diameter clients/DRA nodes, shall be able to do failover to the SAPC that remains active in case of SAPC peer failure. Once the failed SAPC is again available, it shall be able to do failback and continue the homogenous traffic distribution between both SAPC peers.
- provisioning systems, should follow a primary/secondary priorities scheme, where all the provisioning orders have to be sent to the SAPC configured as preferred (primary), if it is available. In case of preferred SAPC failure, the provisioning system shall redirect the provisioning orders to the non-preferred SAPC (secondary). The provisioning priorities scheme, prevents the loss of data after split-brain scenarios, where the non-preferred SAPC synchronizes from the preferred one (all the provisioned data in the non-preferred node during the split-brain is not maintained).
- in network deployments with the CUDB or an external database function, both SAPC peers shall be configured in the CUDB or external database node, to broadcast the SOAP notifications..

The SAPC configuration, operation and maintenance procedures are performed in each SAPC individually.

One of the SAPC peers must be configured as the preferred node in geographical redundancy. This is used when resolving some faulty situations where is not possible to know which of the SAPC peers holds the most up-to-date database. It is also used to prevent loss of data in some Replication Channel failure scenarios. These are the possible situations where the preferred configuration is used:

- After recovery of the network connectivity (neither the Replication Channel nor the Application Channel are available) between both SAPC peers, the



database from the preferred SAPC is maintained, and the database from the non-preferred 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.
- When the Replication Channel is down but the Application Channel is available, the preferred SAPC takes the role of active and the non-preferred goes to standby, to prevent loss of data after recovery of the Replication Channel.

2.3.2 Active-Active Geographical Redundancy Network connectivity

Clients need to maintain one connection towards each SAPC in a redundant mated active-active pair. Seen from the external network, the two SAPC peers are as if they were two different standalone deployments, both of them with the capacity of handling traffic and provisioning. The redundant SAPC solution exposes two VIP addresses for traffic, another two VIP addresses for provisioning and two diameter host names.

Each SAPC handles the following mandatory VIP addresses:

Traffic VIP	This is the VIP address that the SAPC clients use to send diameter traffic to the SAPC. Each SAPC has its own traffic VIP. It is recommended the use of a DRA (Diameter Routing Agent), for traffic separation based on subscribers range, providing session stickiness, to ensure that all Diameter sessions established over the Gx and Rx, for a certain IP-CAN session, reach the same SAPC. In network deployments with Online Charging System, this is also the VIP address where the SAPC handles the Sy interface. This is also the VIP address that the SAPC exposes for the Application Channel to the mated peer.
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.

And the following optional VIP addresses:



Provisioning VIP

This is an optional VIP address (if not configured, the O&M Local VIP can be used instead for provisioning), that the provisioning SAPC clients use to send provisioning orders. Each SAPC has its own provisioning VIP. It is recommended to perform the provisioning in the preferred SAPC. The data is replicated in the mated peer, transparent to the provisioning server.

ExtDB VIP

Only required 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. Each SAPC has its own external database VIP. Both SAPC peers shall be configured in the CUDB or external database node, to broadcast the SOAP notifications. Only the SAPC with the database ownership of the affected subscriber's IP-CAN session/s, process the SOAP notification, being ignored by the SAPC mated peer.

The following picture shows all the IP addresses involved in the geographical redundancy scenario with details about which IP addresses are available for each SAPC. As in standalone deployments, additional VIPs can be added if traffic separation between interfaces is required.

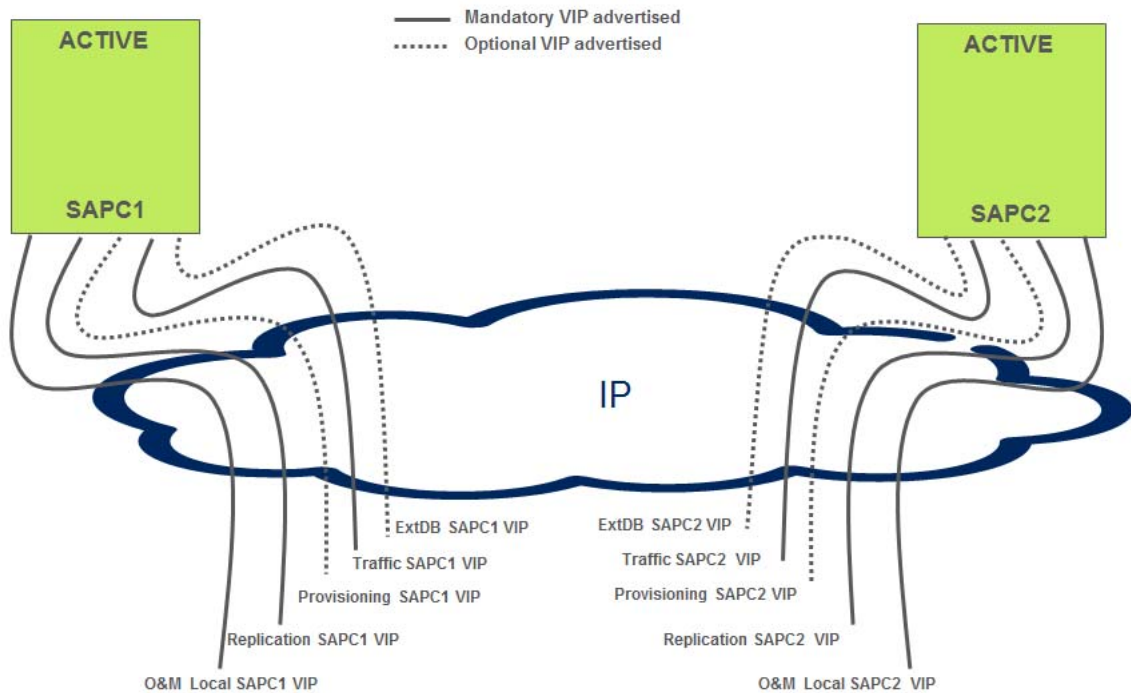


Figure 2 VIP addresses in the SAPC active-active geographical redundancy solution

2.4 Active-Active Geographical Redundancy States

The Active-Active Geographical Redundancy function is initiated in the SAPC by performing an operational procedure. The SAPC geographical redundancy state, reflects the traffic handling ability. The states are the following:

- Initial: This is the initial state when the SAPC is installed. The SAPC does not handle traffic, neither external database nor provisioning operations. The SAPC can only transition from this state when ordered by operational procedure.
- Synchronizing: Temporal transition state when data mirroring starts. The SAPC does not handle traffic, neither external database nor provisioning operations.
- Distributed: The SAPC is handling traffic, provisioning and replicating the changes from the mated peer, that it is also in distributed state.
- Active: The SAPC is handling traffic and provisioning, but there is no data mirroring (the Replication Channel is down). If the Application Channel with the mated peer is up, the mated peer is in standby state. If the Application



Channel with the mated peer is also down (split-brain situation), the mated peer is also in active state.

- Standby: Only the SAPC configured as non-preferred, can be in standby state. It happens when the Replication Channel is down but the Application Channel is up. The non-preferred SAPC does not handle traffic, neither external database nor provisioning operations, and it is not replicating the changes from the active SAPC (to prevent the loss of session information data on the Replication Channel recovery, in case the Application Channel also failed before this recovery). It is ready to transition to active state, if the mated peer fails or if the Application Channel also fails (split-brain situation).
- 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 handle traffic, neither external database nor provisioning operations, and does not replicate database changes from the mated peer. However, the O&M Local VIP can be used for SAPC configuration through COM.

2.4.1 Transition between States

The SAPC executes transitions from one state to another depending on the information provided by the supervision functions explained in Section 2.5 on page 13.

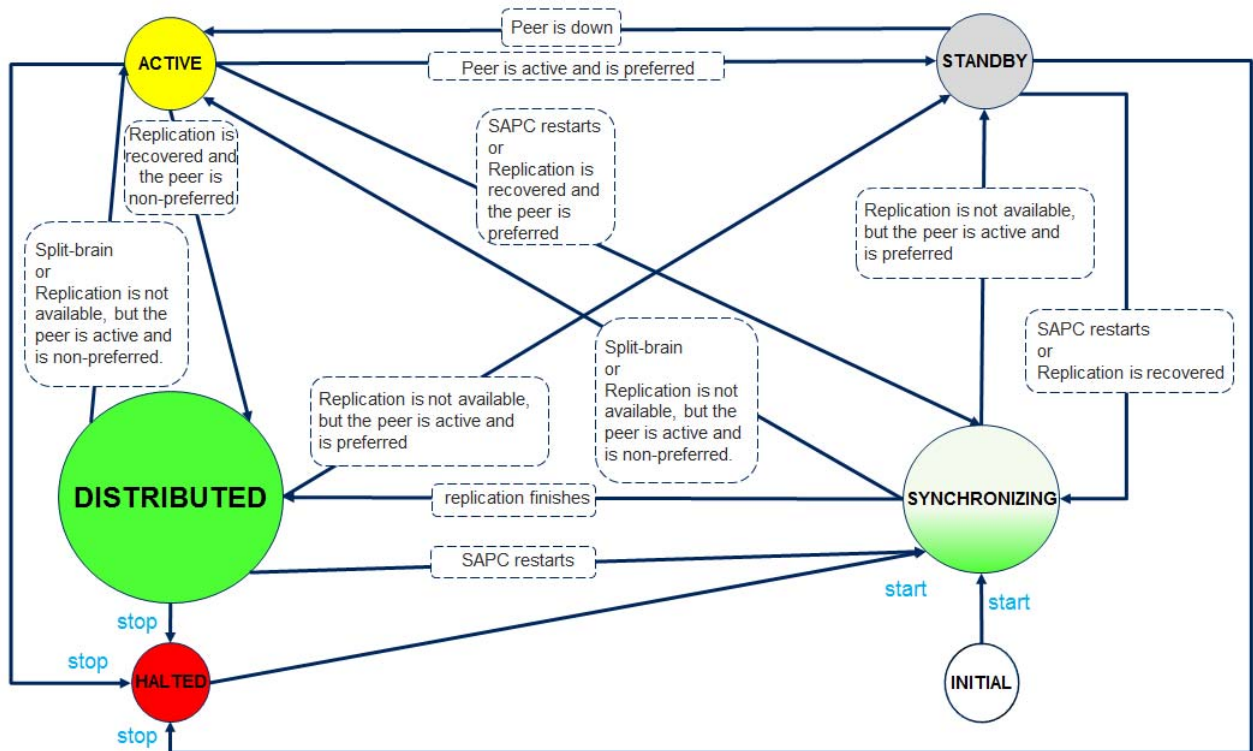


Figure 3 SAPC Active-Active geographical redundancy state machine

2.4.1.1 Transitions from Initial State

In this state, the SAPC can only make one transition when ordered by operational procedure:

- Initial to Synchronizing. This is done when the SAPC is ordered to start active-active geographical redundancy.

2.4.1.2 Transitions from Synchronizing State

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

- Synchronizing to Distributed. This is done when the data mirroring finishes the replication from the mated peer.
- Synchronizing to Active. This is done in two situations:
 - When the Replication Channel fails but the Application Channel is up, and the mated peer is the non-preferred SAPC.



- On split-brain (see Section 2.5.2.1 on page 16). The preferred setting in the SAPC is not considered in this case.
- Synchronizing to Standby. This is done when the Replication Channel fails but the Application Channel is up, and the mated peer is the preferred SAPC.

2.4.1.3 Transitions from Distributed State

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

- Distributed to Synchronizing. This is done in two situations:
 - The SAPC restarts.
 - The SAPC is set as non-preferred ordered by operational procedure. The SAPC discards the local session and provisioning data, starting the replication from the mated peer.
- Distributed to Active. This is done in two situations:
 - When the Replication Channel fails but the Application Channel is up, and the mated peer is the non-preferred SAPC.
 - On split-brain situation.
- Distributed to Standby. This is done when the Replication Channel fails but the Application Channel is up, and the mated peer is the preferred SAPC.
- Distributed to Halted. This is done when ordered by operational procedure. This transition is allowed on any mated peer state.

2.4.1.4 Transitions from Active State

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

- Active to Distributed. This is done when the Replication Channel is recovered and the peer is the non-preferred SAPC.
- Active to Synchronizing. This is done in two situations:
 - The SAPC restarts.
 - When the Replication Channel is recovered and the peer is the preferred SAPC.
- Active to Standby. This is done when the Application Channel recovers but the Replication Channel stills down, and the mated peer is the preferred SAPC.
- Active to Halted. This is done when ordered by operational procedure. This transition is allowed on any mated peer state.



2.4.1.5 Transitions from Standby State

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

- Standby to Active. This is done when the SAPC determines that the mated peer is unavailable or in halted state.
- Standby to Synchronizing. This is done in two situations:
 - The SAPC restarts.
 - When the Replication Channel is recovered. The SAPC is the non-preferred one and discards the local session and provisioning data, starting the replication from the mated peer.
- Standby to Halted. This is done when ordered by operational procedure. This transition is allowed on any mated peer state.

It is not allowed in the Standby state, to order a change in the configuration from non-preferred SAPC to preferred one, to avoid loss of session and provisioning data from the mated peer when the Replication Channel recovers.

2.4.1.6 Transitions from Halted State

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

- Halted to Synchronizing. This happens when the active-active geographical redundancy function, is restarted by operational procedure. This transition is not allowed if the mated peer is in Active state and it is configured as non-preferred SAPC, to avoid the loss of session and provisioning data on the replication procedure.

2.5 Active-Active 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, like Database Service (DBS) temporal unavailability.
- Monitor the redundancy state of the mated peer, to determine when it is unreachable or unavailable and whether it is in distributed, active, standby or halted state.
- Manage the geographical redundancy state machine, and take the appropriate actions in each state or transition. This includes:
 - Handle of VIP 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.
- Set the correct value for the SAPC `Origin-Host` upon failure or when a diameter traffic operation is managed by the SAPC peer different than the one who managed the diameter session establishment.

2.5.1 Mated Peer Supervision

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

The SAPC uses a two simultaneous heartbeat mechanism:

- A Replication Channel heartbeat, to monitor both the availability and redundancy state of the mated peer through the IP network. The heartbeat is sent to the replication VIP address.
- A Application Channel heartbeat, to monitor the availability of the mated peer through the IP network. The heartbeat is sent to the traffic VIP address.

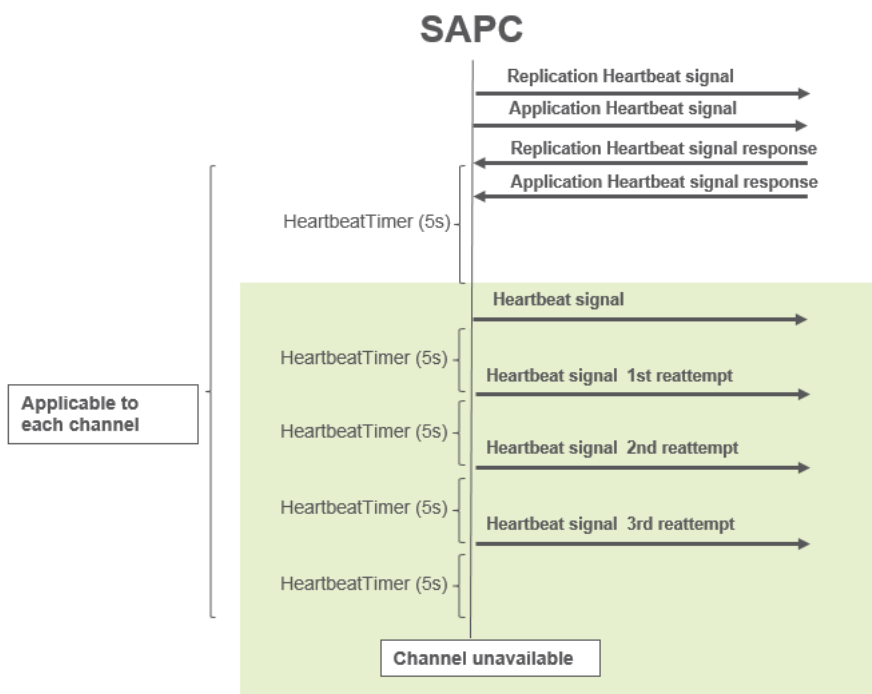


Figure 4 SAPC mated peer supervision

- Each SAPC monitors the redundancy and availability of the mated peer at regular time intervals, by sending two heartbeat signals. This allows to detect



a problem in the Replication Channel or in the availability of the mated peer. Both heartbeat intervals have a default value of 5 seconds.

- The mated peer answers each heartbeat signal by sending an acknowledgment.
- If the acknowledgment of each heartbeat signal is not received, after several attempts, the corresponding channel is declared unavailable. If both channels are unavailable an alarm is raised. The number of reattempts for each channel, has a default value of 3 retransmissions.
- The maximum time period before assuming that any channel 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 one SAPC transitions to halted state by through operational procedure, the mated peer is notified immediately and transitions to active state. The SAPC in halted state, neither send nor answer any heartbeat.

Replication Heartbeat / Application Heartbeat	SAPC1 state	SAPC2 state
Available / Available	DISTRIBUTED	DISTRIBUTED
Available / Unavailable	DISTRIBUTED	DISTRIBUTED
Unavailable / Available	ACTIVE (preferred)	STANDBY (non-preferred)
Unavailable / Unavailable	ACTIVE	ACTIVE

Figure 5 SAPC Active-Active geographical redundancy states based on heartbeat signals

The Application Channel is only considered when the Replication Channel is down. There are three different supervision situations:

1. When the Replication Channel heartbeat response is received, the redundancy and availability of the mated peer are considered available. The Application Channel heartbeat responses, are not considered. Both SAPC peers are in Distributed state.
2. When the Replication Channel heartbeat response is not received (exceeding the maximum time period to consider that the channel is unavailable), but the Application Channel heartbeat response is received. The SAPC configured as preferred, makes the transition to Active state. The SAPC configured as non-preferred makes the transition to Standby state.
3. When neither Replication Channel heartbeat nor Application Channel heartbeat responses are received. This is the split-brain situation, both SAPC peers make the transition to Active state.



2.5.2 Fault Detection and Recovery

The basic principle for the Active-Active Geographical Redundancy function, is that both SAPC peers process all incoming traffic and provisioning operations. Each SAPC keeps the state of the mated peer. However, there are situations that might cause one SAPC not to be synchronized with the peer. This chapter describes those scenarios and how they are handled.

2.5.2.1 Split Brain Scenario

This situation happens when the SAPC detects that neither Replication Channel nor Application Channel are available. If each SAPC is healthy, that is, only the heartbeats are lost (owing to loss of connectivity between the SAPC peers), it switches from distributed state to active state and continues serving the traffic and provisioning operations. In this scenario, the end result consists of two SAPC peers in active state serving traffic and provisioning but without data mirroring. This situation is known as an split brain scenario.

The split brain has the following consequences:

- 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, a complete database synchronization is performed. Then, the non-preferred SAPC is automatically restarted. Once restarted, it recovers the most current database information from the mated peer. Once the complete synchronization is finished, both SAPC peers set the replication state to distributed.

2.5.2.2 Simultaneous SAPC Restart

If both SAPC peers fail, they probably do not reload at the same time. Therefore, the first one that restarts, observes loss of network connectivity to the mated peer and takes the necessary actions, as described in Section 2.5.2.1 on page 16.

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

2.5.2.3 Temporary Differences between Databases

The data mirroring functionality makes use of backlogs while transferring database changes from each SAPC to the mated peer as described in Section 2.2.2 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 of the SAPC peers. 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 preferred SAPC, provides data to the non-preferred SAPC by using the procedure described in Section 2.2.3 on page 5.

2.5.3 Handling of the SAPC Origin State Id

In a standalone 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 an Active-Active Geographical Redundancy deployment, the `Origin State Id` information is replicated between both SAPC peers. Upon a SAPC restart, the `Origin State Id` is not incremented, it is obtained from the most up-to-date database information during synchronization with the mated peer. This enables to each SAPC to send the same `Origin State Id` value. Transitions between SAPC redundancy states do not increase the value of the `SAPC Origin State Id`, as those are transparent to the peer diameter nodes in the external network.

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 because of loss of network connectivity in the Replication Channel), the SAPC increments its own `Origin State Id`.

When one 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 Replication Channel is re-established, the preferred SAPC continues providing service, but the `Origin State Id` is not increased (the non-preferred SAPC makes a complete synchronization with the mated peer and then, replicates the same `Origin State Id`).

If the SAPC restarts in a split-brain situation, the SAPC 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 Replication Channel is recovered, the `Origin State Id` in the preferred SAPC is maintained (the non-preferred SAPC makes a complete synchronization with the mated peer and then replicates the same `Origin State Id`).



2.5.4 Handling of the SAPC Origin-Host

When a diameter client requests for a new diameter session establishment, the SAPC `Origin-Host` included in the diameter answer, is the one from the SAPC peer that manages the session establishment. Each SAPC peer have a different `Origin-Host` value. In case of failure in the SAPC that managed the session establishment, the SAPC mated peer (that handles all the sessions while the failed SAPC recovers), always include as origin-host, the one used in the session establishment instead of its own, to avoid any problem in the diameter client.

3 Active-Active Geographical Redundancy Traffic Cases

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

- SAPC in Distributed state restarts.
- Replication Channel unavailable but Application Channel available.
 - SAPC in Active state restarts.
- Both Replication Channel and Application Channel unavailable.

3.1 SAPC in Distributed state restarts

The following figure shows the high level flow that takes place when one SAPC in Distributed state restarts, and the main actions are taken by the peer SAPC, to perform the Active-Active Geographical Redundancy functionality.

A failure in one SAPC, makes the mated peer transition from distributed to active state. Diameter clients/DRA nodes shall be able to do failover to the SAPC that remains active. Once the SAPC recovers from the restart and completes synchronization from the mated peer, diameter clients/DRA nodes, shall be able to do failback to the recovered SAPC and continue the homogenous traffic distribution between both SAPC peers. During the transition, ongoing traffic events may fail. The duration of the traffic switch depends on several factors, see Section 4 on page 23.

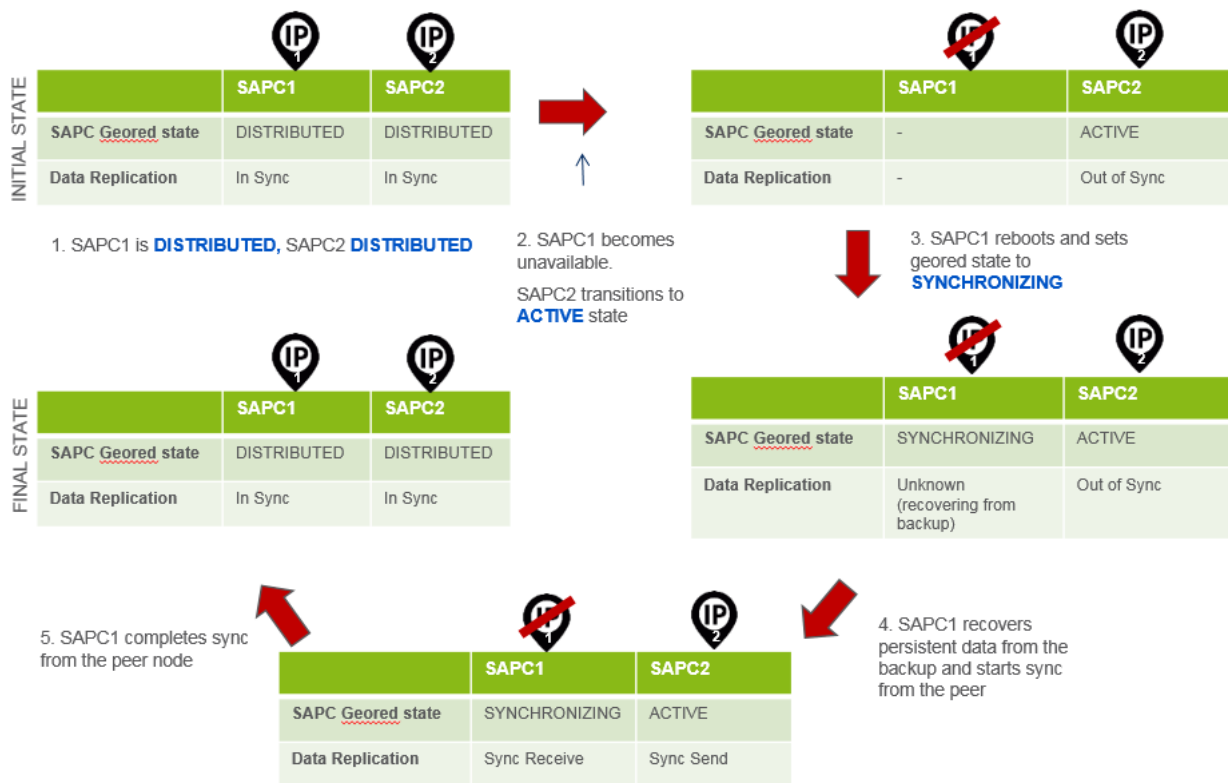


Figure 6 SAPC in Distributed state restarts

1. This is the initial working condition for Active-Active Geographical Redundancy. Both SAPC peers are in distributed state, processing traffic. Data mirroring is fully operational and keeps data in both SAPC peers synchronized.
2. SAPC1 fails, SAPC2 detects that SAPC1 is not available (because of heartbeat time-outs), rises two alarms (Unable to Reach Peer, one for the Replication Channel and one for the Application Channel), and transitions to active state. 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 synchronizing. 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 from the active SAPC to the recovered one, where it is imported. Any changes



done on the active SAPC in the meantime, are also transferred as normal database changes, which are applied in the recovered SAPC 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, both SAPC peers are in distributed state, processing traffic.

3.2 Replication Channel unavailable but Application Channel available

The following figure shows the high level flow that takes place when the Replication Channel becomes temporarily unavailable, but the Application Channel keeps available.

A failure in the Replication Channel, makes each SAPC to check the availability of the mated peer using the Application Channel heartbeat. Once each SAPC determines that the peer is alive, the preferred SAPC transitions from distributed to active state. The non-preferred SAPC transitions from distributed to standby, setting the traffic and provisioning VIP addresses unavailable. Diameter clients/DRA nodes shall be able to do failover to the SAPC that remains active. Once the SAPC recovers the Replication Channel, the non-preferred SAPC makes a complete synchronization from the mated peer, and makes available again the traffic and provisioning VIP addresses. Diameter clients/DRA nodes, shall be able to do failback to the recovered SAPC and continue the homogenous traffic distribution between both SAPC peers. During the transition, ongoing traffic events may fail. The duration of the traffic switch depends on several factors, see Section 4 on page 23.

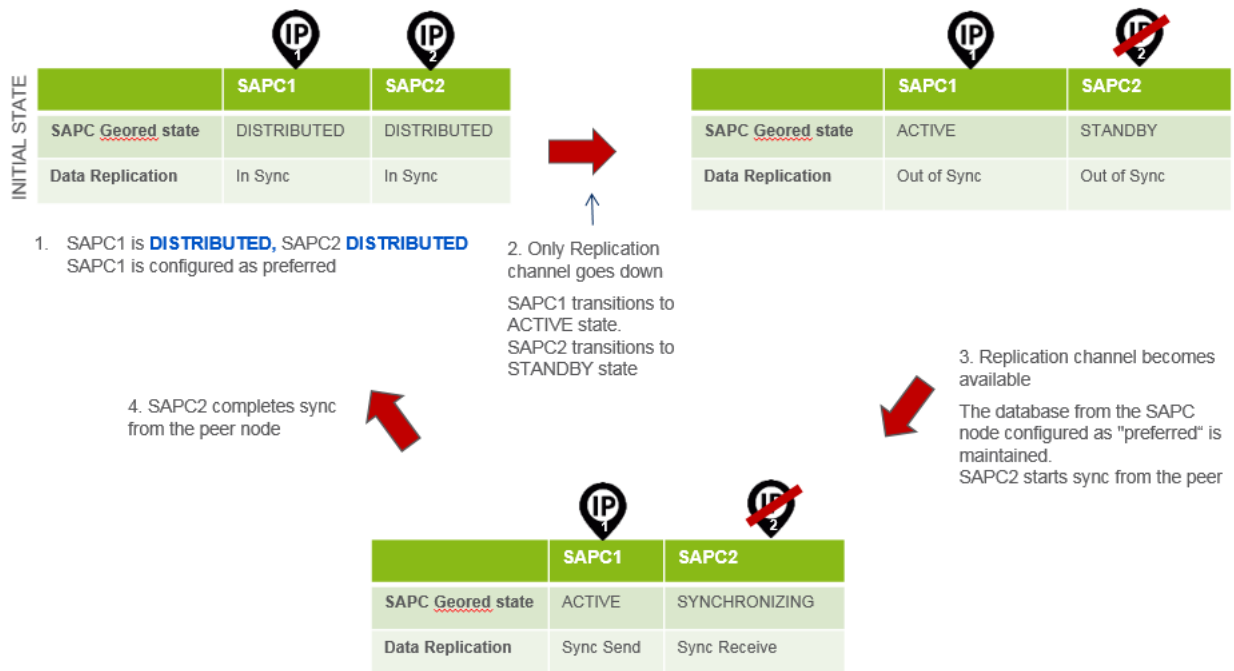


Figure 7 Replication channel temporarily unavailable

1. This is the initial working condition for Active-Active Geographical Redundancy. Both SAPC peers are in distributed state, processing traffic. Data mirroring is fully operational and keeps data in both SAPC peers synchronized. The SAPC1 is configured as the preferred node for geographical redundancy.
2. The Replication Channel fails, but SAPC1 detects that SAPC2 is available through the Application Channel. SAPC1 rises an alarm (Unable to Reach Peer) for the Replication Channel and transitions to Active state. SAPC2 also detects the same situation, transitions to Standby state and makes unavailable the traffic and provisioning VIP addresses. Data mirroring is interrupted and the SAPC DBS component also rises an alarm (Connection Loss) in both SAPC peers. Only SAPC1 handles traffic, and database changes are applied locally. This handle of traffic in the SAPC1, makes the SAPC DBS component to rise another alarm in the SAPC1 (Synchronization Needed).
3. The Replication Channel becomes available and both SAPC peers clear the corresponding alarms (Unable to Reach Peer, Connection Loss and Synchronization Needed). Data mirroring restarts and detects that a complete synchronization is needed. The SAPC2 connects to the mated peer (which is in active state) and sets the redundancy state to synchronizing. 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 handling traffic.
4. SAPC2 completes successfully synchronization from the active SAPC1 and the system goes back to the initial state.



3.2.1 SAPC in Active state restarts

The following actions are done if before the Replication Channel recovers, the Active SAPC1 fails:

1. SAPC2 detects that the SAPC1 is down through the Application Channel. It transitions from Standby to Active state, makes available the traffic and provisioning VIP addresses and rises two instance alarms (Unable to Reach Peer), one for the Application Channel and another for the Replication Channel. Only SAPC2 handles traffic and provisioning operations, applying database changes locally. The SAPC DBS component also rises two alarms (Connection Loss and Synchronization Needed) in SAPC2.
2. Once SAPC1 is recovered, detects that the non-preferred SAPC2 is available through the Application Channel, but the Replication Channel continues unavailable, so it transitions to Active state again and makes available the traffic and provisioning VIP addresses. SAPC1 rises an alarm (Unable to Reach Peer) for the Replication Channel. The SAPC DBS component also rises two alarms (Connection Loss and Synchronization Needed) in SAPC1.
3. The non-preferred SAPC2 detects that SAPC1 is available again, so it transitions to Standby state. SAPC2 makes unavailable the traffic and provisioning VIP addresses. Only SAPC1 handles traffic and provisioning operations, applying database changes locally. All the session data managed in the SAPC2 until the SAPC1 recovered, is lost, because the Replication Channel continues unavailable.

3.3 Both Replication Channel and Application Channel unavailable

The following figure shows the high level flow that takes place when both the Replication Channel and Application Channel become temporarily unavailable, and the main actions taken by the SAPC to perform the Active-Active Geographical Redundancy functionality.

A failure in both channels, results in both SAPC peers taking the role of active SAPC, handling traffic but without data replication. In this situation, the databases become inconsistent. When the connectivity is re-established, the database from the preferred SAPC is maintained, and the database from the non-preferred SAPC is discarded.

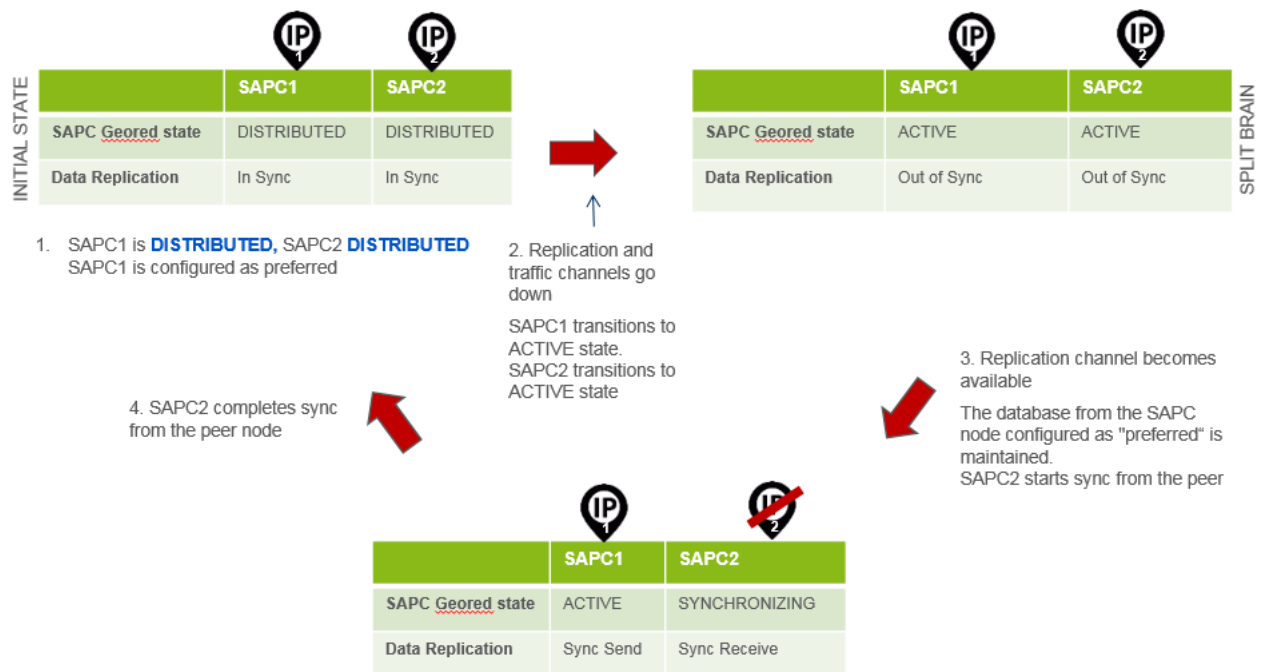


Figure 8 Replication and Application Channels temporarily unavailable

1. This is the initial working condition for Active-Active Geographical Redundancy. Both SAPC peers are in distributed state, processing traffic. Data mirroring is fully operational and keeps data in both SAPC peers synchronized. The SAPC1 is configured as the preferred node for geographical redundancy.
2. Both, the Replication Channel and Application Channel fail, SAPC1 detects that SAPC2 is not available (because of heartbeat time-outs), rises two alarms (Unable to Reach Peer, one for the Replication Channel and one for the Application Channel), and transitions to active state. SAPC2 also detects that SAPC1 is not available, and make the same actions. Data mirroring is interrupted and the SAPC DBS component also rises an alarm (Connection Loss) in both SAPC peers. This is an 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 SAPC peers clear the corresponding alarms (Unable to Reach Peer, Connection Loss). Data mirroring restarts and detects that a complete synchronization is needed. The SAPC2 connects to the mated peer (which is in active state) and sets the redundancy state to synchronizing. Then the SAPC2 starts synchronization from the preferred SAPC, discarding the local data. In the meantime, the SAPC1 continues handling traffic.
4. SAPC2 completes successfully synchronization from the active SAPC1 and the system goes back to the initial state.

4 Active-Active Geographical Redundancy Capabilities

For Active-Active Geographical Redundancy the following capabilities must be considered:

- Enough bandwidth. The Replication Channel and Application 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 fulfill 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 each SAPC can replicate from the mated peer without lagging behind.
- Node capacity. Each SAPC, must be dimensioned to support the traffic from both zones on temporal Replication Channel unavailability or mated peer failure.
- 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 (Replication Channel, Application Channel or both), 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 distributed state to active state, may vary from 5 seconds to 20 seconds.



Reference List

Ericsson Documents

- [1] Availability and Scalability