

CUDB System Split Partial Recovery Procedure

OPERATION INSTRUCTION

Copyright

© Ericsson AB 2016, 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.



Contents

1	Overview	1
1.1	Description	1
1.2	Prerequisites	1
1.3	Typographic Conventions	3
2	Procedure in Case of Site Failure	4
2.1	Situations	4
2.2	Emergency Procedure	4
2.3	Recovery Procedure	6
3	Procedure in Case of IP Backbone Failure	8
3.1	Situation	8
3.2	Emergency Procedure	8
3.3	Recovery Procedure	13
	Glossary	15
	Reference List	16





1 Overview

The purpose of this document is to provide system administrators clear operating instructions for recovering from a node failure or IP network loss, and to make traffic or provisioning available until the CUDB site or IP network becomes operational again.

1.1 Description

This document describes the Partial Recovery procedure that can be used as an emergency work-around solution for both available redundancy configurations. These configurations are as follows:

- 1+1 double geographically redundant CUDB system, with two sites. For more information, refer to *CUDB High Availability*, Reference [1].
- 1+1+1 triple geographically redundant CUDB system, with three sites. For more information, refer to *CUDB High Availability*, Reference [1].

1.2 Prerequisites

The procedure can be applied in the following situations:

- Symmetrical split situation.
 - In case of 1+1 double redundancy, one site is down for a long time. A site is considered down if all the nodes of the sites are down. If the surviving site is not hosting the master instance for PLDB, update operations from any Provisioning user are rejected. For more information, refer to *CUDB High Availability*, Reference [1].
 - In case of 1+1 double redundancy, an IP backbone failure isolates the sites from each other. One of the sites is not hosting the master instance for PLDB. Update operations from any Provisioning user are rejected in that site. Apply procedure to recover provisioning traffic through that specific site.
 - In case of 1+1+1 triple redundancy, one site fails, leaving the rest of the system in majority. After that, an IP backbone failure isolates the remaining sites from each other. Same as previous situation.
- Minority situation.
 - In case of 1+1+1 triple redundancy, two of the sites are down for a long time.



- In case of 1+1+1 triple redundancy, a simultaneous IP backbone failure isolates the three sites from each other.

Note: Do not execute the procedures of this document in case of a majority situation.

To detect the detailed situations above, check the following alarms in the CUDB system:

- Minority situation:

In this situation, two *Control, Remote Site Unreachable*, Reference [2], alarms are raised, one for each unreachable site. In case two of the sites are down, these alarms are only raised on the surviving site. In case the three sites are isolated among them, these alarms are raised on all the sites. This alarm is raised in the node where the System Monitor (SM) leader is located.

- Symmetrical split situation:

In this situation, the *Control, Remote Site Unreachable*, Reference [2], and *Control, Potential Split Brain Detected*, Reference [3] alarms are raised on the surviving sites.

For more information, refer to *Control, Remote Site Unreachable*, Reference [2], and *Control, Potential Split Brain Detected*, Reference [3].

In any other situation, the system is in majority situation, and the Partial Recovery procedure cannot be applied.

Warning!

The system must satisfy all the conditions listed above: in case of any differences in the setup, the instructions of this document are not applicable.

As a result of this procedure, the PLDB and DSG masterships are moved to the surviving CUDB site to recover provisioning.

The partial recovery procedure consists of two procedures:

- An Emergency Procedure, which must be applied when the symmetrical split or minority situation occurs.
- A Recovery Procedure, which must be applied when the split situation ends, so that all sites are operational.

The Selective Replica Check and Data Repair processes are part of the Automatic Handling of Network Isolation function. The aim of this automatic process in CUDB is to attempt to handle and repair data loss that happened



due to network split or unexpected PLDB or DSG mastership change. For more information on the Selective Replica Check and Data Repair processes, refer to *CUDB Data Storage Handling*, Reference [5].

1.3 Typographic Conventions

Typographic conventions can be found in the following document:

- *Typographic Conventions*



2 Procedure in Case of Site Failure

This section describes the procedure to follow in case of site failure.

2.1 Situations

This procedure can be applied in the following situations:

- Situation 1: In a 1+1 double geographical redundant CUDB system, one site fails (symmetrical split situation).
- Situation 2: In a 1+1+1 triple geographical redundant CUDB system, two sites fail (minority situation).

In *Situation 2*, the site in minority cannot handle any kind of traffic. To allow traffic operations only over DSGs, go to Section 2.2.2 on page 5.

In both cases, the surviving site is not hosting the master instance for PLDB. Update operations from any Provisioning user are rejected.

2.2 Emergency Procedure

This section describes the emergency procedure for provisioning recovery and traffic recovery.

2.2.1 Enabling Provisioning

In the situations listed in Section 2.1 on page 4, only one site is alive in the CUDB system. This site is called the “surviving site”. Perform the following steps in case of site failure:

1. Make sure that the failing CUDB sites cannot join to the CUDB system in an uncontrolled manner during the execution of the procedure.

Warning!

This preparatory step must be performed to avoid any potential conflicts and interference from the failed sites during the execution of the procedure.

2. Execute the `cudbTakeAllMasters` command.
 - Only applicable in *Situation 1*.



Only if the surviving site does not host the PLDB master, execute the `cudbTakeAllMasters` command on any CUDB node of the surviving site. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].

- Only applicable in *Situation 2*.

Execute the `cudbTakeAllMasters` command on any CUDB node of the surviving site. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].

Warning!

This command cannot be undone: once it is executed, the PLDB and DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.

3. After the above command has been executed, provisioning must be available.

Warning!

After successful command execution, provisioning continues. However, the system is not geographically redundant, so the missing sites and the system backbone must be restored immediately. The system recovers geographical redundancy after the recovery procedure (described in Section 2.3 on page 6) is executed.

2.2.2 Enabling Traffic

Only applicable in *Situation 2*.

To allow the traffic operations in the DSGs, execute the `cudbServiceContinuity` command in the surviving site by performing the following steps.

Note: The `cudbServiceContinuity` command does not avoid that this site could be chosen later as “surviving site” to recover the provisioning (see Section 2.2.1 on page 4).

If the `automaticServiceContinuity` parameter is enabled, this is done automatically (for more information, refer to *CUDB High Availability*, Reference [1]).



1. Execute the `cudbServiceContinuity` command on one node (and only one node) of the surviving site. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].

Warning!

This command cannot be undone: once it is executed, the DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.

2. After the above command has been executed, traffic must be available in all DSGs.
3. To recover the provisioning, execute the specific emergency procedure (see Section 2.2.1 on page 4).

2.3 Recovery Procedure

When the failed sites are working again, perform the following steps to make sure the original geographical redundancy configuration (either 1+1 double, or 1+1+1 triple geographical redundancy) is recovered:

1. Before connecting the CUDB nodes in the failed sites to the CUDB system, set the PLDB and all the DS Units in each node to maintenance mode executing the following command in each of the nodes of the failed sites:

```
cudbManageStore -a -o maintenance
```

2. Use the following command to check which node runs the SM leader:

```
cudbSystemStatus -B
```

3. Check that the *Control, Remote Site Unreachable*, Reference [2], alarm is not raised. To do so, login to the node where the SM leader is located, and check for the alarms raised in the node with the following command:

```
cudbSystemStatus -a
```

If the *Control, Remote Site Unreachable*, Reference [2], alarm is raised, then the original 1+1 double geographical redundant or 1+1+1 triple geographical redundant CUDB system is not recovered yet. Check that the nodes in the failed sites are properly connected to the CUDB system.

4. To set all the PLDB and DS Units to ready, execute the following command in each node of the failed sites:

```
cudbManageStore -a -o ready
```



5. At this point, the PLDB and the DS Units in the nodes of the failed sites become slaves of the master PLDB and master DS Units hosted in the surviving site, and start to gather the differences from the masters. Check the replication status of the nodes by executing the following command:

```
cudbSystemStatus -R
```

6. If a slave database cluster is unable to sync with the current master, check if *Storage Engine, Unable to Synchronize Cluster in DS, Major*, Reference [7] or *Storage Engine, Unable to Synchronize Cluster in PLDB, Major*, Reference [8] is raised. In those cases, execute the following command to restore the slave DS Unit or PLDB where the replication is not running:

```
cudbUnitDataBackupAndRestore
```

For more information, refer to *CUDB Backup and Restore Procedures*, Reference [9].

Note: If Self-Ordered Backup and Restore function is enabled, it does not start if all PLDB replicas in the site are slaves that are unable to synchronize with the current master. In case the Automatic Handling of Network Isolation or the Self-Ordered Backup and Restore functions or both are enabled, alarms related to these functions appear before the *Storage Engine, Unable to Synchronize Cluster in DS, Major*, Reference [7] or *Storage Engine, Unable to Synchronize Cluster in PLDB, Major*, Reference [8].

If Self-Ordered Backup and Restore process restores the replication automatically, these alarms are not raised at all.

Once the replication is working, if needed, move the DSG mastership to the desired node by using the **cudbDsgMastershipChange** command, taking into account if Automatic Mastership Change function is active.

For more information about the commands used in this section, refer to *CUDB Node Commands and Parameters*, Reference [4]. For more information on the Automatic Handling of Network Isolation procedure, refer to *CUDB High Availability*, Reference [1].



3 Procedure in Case of IP Backbone Failure

This section describes the procedure to follow in case of an IP backbone failure.

3.1 Situation

This procedure can be applied in the following situations:

- Situation 1: In a 1+1 double geographical redundant CUDB system, an IP backbone failure isolates the sites from each other (symmetrical split situation).
- Situation 2: In a 1+1+1 triple geographical redundant CUDB system, one site fails, leaving the rest of the system in majority. After that, an IP backbone failure isolates the remaining sites from each other (symmetrical split situation).
- Situation 3: In a 1+1+1 triple geographical redundant CUDB system, an IP backbone failure isolates the sites from each other causing all nodes to be in minority.
- Situation 4: In a 1+1+1 triple geographical redundant CUDB system, an IP backbone failure isolates one site from the others, being two sites in majority and the other one in minority.

In *Situations 1* and *2*, one of the sites is not hosting the master instance for PLDB. Update operations from any Provisioning user are rejected in that site. Apply procedure to recover provisioning traffic through that specific site.

In *Situation 3*, there are no master instances either for PLDB or any DSG in any of the sites.

In *Situation 4*, there are no master instances either for PLDB or any DSG in the site in minority.

In *Situations 3* and *4*, the site in minority cannot handle any kind of traffic. To allow traffic operations only over DSGs, go to Section 3.2.2 on page 12.

3.2 Emergency Procedure

In case of the situations described in Section 3.1 on page 8, the sites in the system are alive, but are isolated from each other (except *Situation 2*, where two sites are alive, and one is down). Therefore, the site where full service is to be granted (traffic and provisioning) must be chosen as the “surviving site”. Execute the below procedure on the surviving site.



3.2.1 Enabling Provisioning

This procedure is only applicable to *Situations 1* and *2*.

If Selective Replica Check and Data Repair are enabled, after the backbone problem is fixed and the sites are rejoined, Automatic Handling of Network Isolation fixes data inconsistencies caused by the split and re-establishes replication in the slave replicas, so there is no need to select a surviving site. For more information on Automatic Handling of Network Isolation, refer to *CUDB Automatic Handling of Network Isolation Output Description*, Reference [6].

If Selective Replica Check and Data Repair are disabled, perform the following:

Warning!

Execute the below procedure only on one node of the selected surviving site. Disable the rest of the sites, so that they do not join the system during the execution of the procedure. The execution of this procedure on the remote nodes leads to data inconsistency and data loss without any rollback possibility.

Do not select an unstable node in the surviving site. The node selected to execute the below steps in the surviving site must be free of Operating System (OS) alarms and SAF alarms.

1. Set the PLDB and DS Units to maintenance mode in all nodes on the sites not selected as the surviving site. Use the following command to do so:

```
cudbManageStore -a -o maintenance
```

Warning!

As the IP backbone is down, this step must be performed locally on-site by the operator (no remote configuration is available). Therefore, operators must be prepared to execute the below commands on remote nodes as well in time. Ericsson takes no responsibility for any data inconsistency if the commands in this step are not executed in time, or at all.

2. Only if the PLDB master is not in the surviving site, execute the **cudbTakeAllMasters** command on one node (and only one node) of the surviving site. For more information refer to *CUDB Node Commands and Parameters*, Reference [4].



Warning!

This command cannot be undone: once it is executed, the PLDB and DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.

3. After the above command has been executed, provisioning must be available.
-
-

Warning!

After successful command execution, provisioning continues. However, the system is not geographically redundant, so the missing sites and the system backbone must be restored immediately. The system recovers geographical redundancy after the recovery procedure (described in Section 3.3 on page 13) is executed.

This procedure is only applicable to *Situation 3*.

If Selective Replica Check and Data Repair are enabled, perform the following:

Warning!

Do not select an unstable node. The node selected to execute the below steps must be free of OS and SAF alarms.

1. Execute the `cudbTakeAllMasters` command on one node (and only one node) of the site where provisioning is to be enabled. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].
-
-

Warning!

This command cannot be undone: once it is executed, the PLDB and DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.



2. After the above command has been executed, provisioning must be available.

Warning!

After successful command execution, provisioning continues. However, the system is not geographically redundant, so the missing sites and the system backbone must be restored immediately. The system recovers geographical redundancy after the recovery procedure (described in Section 3.3 on page 13) is executed.

If Selective Replica Check and Data Repair are disabled, perform the following:

Warning!

Execute the below procedure only on one node of the selected surviving site. Disable the rest of the sites, so that they do not join the system during the execution of the procedure. The execution of this procedure on the remote nodes leads to data inconsistency and data loss without any rollback possibility.

Do not select an unstable node in the surviving site. The node selected to execute the below steps in the surviving site must be free of OS and SAF alarms.

1. Set the PLDB and DS Units to maintenance mode in all nodes on the sites not selected as the surviving site. Use the following command to do so:

```
cudbManageStore -a -o maintenance
```

Warning!

As the IP backbone is down, this step must be performed locally on-site by the operator (no remote configuration is available). Therefore, operators must be prepared to execute the below commands on remote nodes as well in time. Ericsson takes no responsibility for any data inconsistency if the commands in this step are not executed in time, or at all.

2. Execute the **cudbTakeAllMasters** command on one node (and only one node) of the surviving site. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].



Warning!

This command cannot be undone: once it is executed, the PLDB and DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.

3. After the above command has been executed, provisioning must be available.
-
-

Warning!

After successful command execution, provisioning continues. However, the system is not geographically redundant, so the missing sites and the system backbone must be restored immediately. The system recovers geographical redundancy after the recovery procedure (described in Section 3.3 on page 13) is executed.

3.2.2 Enabling Traffic

This procedure is only applicable in *Situations 3* and *4*.

To allow the traffic operations in the DSGs, execute the `cudbServiceContinuity` command in the site in minority by performing the following steps.

Note: The `cudbServiceContinuity` command does not avoid that this site could be chosen later as “surviving site” to recover the provisioning (see Section 3.2.1 on page 8).

If the `automaticServiceContinuity` parameter is enabled, this is done automatically (for more information, refer to *CUDB High Availability*, Reference [1]).

Warning!

As the IP backbone is down, this step must be performed locally on-site by the operator (no remote configuration is available).

1. Execute the `cudbServiceContinuity` command on one node (and only one node) of the surviving site. For more information, refer to *CUDB Node Commands and Parameters*, Reference [4].



Warning!

This command cannot be undone: once it is executed, the DSG masterships are moved to the surviving site. Killing the command during waiting does not stop the process of taking the mastership.

2. After the above command has been executed, traffic must be available in all DSGs.
3. To recover the provisioning in this site, execute the specific emergency procedure (see Section 3.2.1 on page 8).

3.3 Recovery Procedure

When the IP Backbone is working again, follow the steps below to restore the original redundancy configuration:

Note: In case the Selective Replica Check and Data Repair were enabled while performing the procedures in Section 3.2 on page 8, skip until Step 5.

1. Use the following command to check which node runs the SM leader:

```
cudbSystemStatus -B
```

2. Check that the *Control, Remote Site Unreachable*, Reference [2], alarm is not raised. To do so, login to the node where the SM leader is located, and check for the alarms raised in the node with the following command:

```
cudbSystemStatus -a
```

If the *Control, Remote Site Unreachable*, Reference [2], alarm is raised, then the original 1+1 double geographical redundant or 1+1+1 triple geographical redundant CUDB system is not recovered yet. Check that the nodes in the failed sites are properly connected to the CUDB system.

3. To set all the PLDB and DS Units to ready, execute the following command in all the nodes of the sites not selected as the surviving site:

```
cudbManageStore -a -o ready
```

4. At this point in time, the PLDB and the DS Units in the nodes of the failed sites become slaves of the master PLDB and master DS Units hosted in the surviving site, and start to gather the differences from the masters. Check the replication status of the nodes by executing the following command:

```
cudbSystemStatus -R
```



5. If Self-Ordered Backup and Restore function is enabled, it does not start if all PLDB replicas in the site are slaves that are unable to synchronize with the current master. If a slave database cluster is unable to sync with the current master, check if *Storage Engine, Unable to Synchronize Cluster in DS, Major*, Reference [7] or *Storage Engine, Unable to Synchronize Cluster in PLDB, Major*, Reference [8] is raised. In those cases, execute the following command to restore the slave DS Unit or PLDB where the replication is not running:

`cudbUnitDataBackupAndRestore`

For more information, refer to *CUDB Backup and Restore Procedures*, Reference [9].

Note: In case the Automatic Handling of Network Isolation or the Self-Ordered Backup and Restore functions or both are enabled, alarms related to these functions appear before the *Storage Engine, Unable to Synchronize Cluster in DS, Major*, Reference [7] or *Storage Engine, Unable to Synchronize Cluster in PLDB, Major*, Reference [8].

If Self-Ordered Backup and Restore process restores the replication automatically, these alarms are not raised at all.

Once the replication is working, if needed, move the DSG mastership to the desired node by using the **`cudbDsgMastershipChange`** command, taking into account if Automatic Mastership Change function is active.

For more information about the commands used in this section, refer to *CUDB Node Commands and Parameters*, Reference [4]. For more information on the Automatic Handling of Network Isolation procedure, refer to *CUDB High Availability*, Reference [1].



Glossary

For the terms, definitions, acronyms and abbreviations used in this document, refer to *CUDB Glossary of Terms and Acronyms*, Reference [10].



Reference List

Ericsson Documents

- [1] *CUDB High Availability*
- [2] *Control, Remote Site Unreachable*
- [3] *Control, Potential Split Brain Detected*
- [4] *CUDB Node Commands and Parameters*
- [5] *CUDB Data Storage Handling*
- [6] *CUDB Automatic Handling of Network Isolation Output Description*
- [7] *Storage Engine, Unable to Synchronize Cluster in DS, Major*
- [8] *Storage Engine, Unable to Synchronize Cluster in PLDB, Major*
- [9] *CUDB Backup and Restore Procedures*
- [10] *CUDB Glossary of Terms and Acronyms*