Split Brain Recovery

Revised June 30, 2011

Split brain mode refers to a state in which each database server does not know the high availability (HA) role of its redundant peer, and cannot determine which server currently has the primary HA role. In split brain mode, data modifications may have been made on either node, and those changes may not be replicated to the peer. Also, neither or both nodes may be functioning in the primary HA role.

Split brain mode occurs when there is a temporary failure of the network connections between the two database servers, for example, due to one of the following occurrences:

- Restart of either database server during synchronization.
- Physical disconnection of the Ethernet cables from a database server.
- Loss of power to one or both database servers.

**Note**

If the current primary database server loses power, or its integrated management module (IMM) becomes unreachable by the secondary database server (for example, due to network connectivity issues), the secondary database server cannot automatically take over as primary. If the current primary database server fails under these conditions, your system may or may not enter split brain mode. In this situation, take one of the following actions:

- If the primary database server comes back up, the system may enter split brain mode; proceed to the “Diagnosing Split Brain Mode” section on page 30-1.
- If the primary database server remains down, the split brain recovery procedure is not applicable; instead, see the “Recovering from a Failed Primary Database Server” section on page 33-1.

This chapter includes the following topics:

- Diagnosing Split Brain Mode, page 30-1
- Recovering from Split Brain Mode, page 30-3
- Verifying Synchronization of the Database Servers, page 30-4
- Diagnosing Corrupted DRBD Metadata, page 30-6
- Recovering from Corrupted DRBD Metadata, page 30-7

**Diagnosing Split Brain Mode**

Use this procedure to determine whether your database servers are in split brain mode.
Diagnosing Split Brain Mode

Before You Begin
Make sure that the database servers are correctly cabled. See the “Cabling Requirements for the Database Servers” section on page 4-3.

Procedure

Step 1 Log in to the CLI of each database server.

Step 2 On each database server, enter the `utils service database status` command.

Step 3 If the output indicates one or more of the following conditions, the database servers are in split brain mode:

- The connection state (cs) is “StandAlone.”
- The role (ro) values display one of the following combinations:
  - “Primary/Unknown” on one server and “Secondary/Unknown” on the other server.
  - “Secondary/Secondary” on both servers—In this particular case, if the connection state (cs) on both servers is “Connected,” then the MySQL database is corrupted, and the split brain recovery procedure will not help. Instead, see the “Corrupted MySQL Database Recovery” chapter.
  - “Secondary/Unknown” on both servers—In this particular case, if you know that one of the database servers had a reboot during the initial synchronization process, then your database system is functioning in a mode for which the split brain recovery procedure will not help. To recover, you need to reinstall both database servers. See the “Installing and Synchronizing the Cisco TelePresence Exchange System Database Servers” section on page 5-4.

Step 4 To recover from split brain mode, proceed to the “Recovering from Split Brain Mode” section on page 30-3.

Example
In the following example, the connection state (cs) of one of the database servers is StandAlone, which indicates that the nodes are in split brain mode:

```
admin: utils service database status
--------------------------------------------------------------------------------
The initial configured HA role of this node : secondary
The current HA role of this node : primary
The database vip address : 10.22.130.54
The database primary node name : ctx-db-1
The database primary node IP address : 10.22.130.49
The database secondary node name : ctx-db-2
The database secondary node IP address : 10.22.130.57
Mon status : Running pid 2820
MySQL status : Running pid 2810
Heartbeat status : Running pid 3752
--------------------------------------------------------------------------------
drbd driver loaded OK; device status:
version: 8.3.2 (api:88/proto:86-90)
m:res cs ro ds p mounted fstype
0:mysql StandAlone Primary/Unknown UpToDate/Unknown r---- ext3
--------------------------------------------------------------------------------
```

Related Topics
- Command Reference, page C-1
Recovering from Split Brain Mode

Use this procedure to recover your database servers from split brain mode.

Before You Begin

- Complete the “Diagnosing Split Brain Mode” section on page 30-1 to confirm that your system is in split brain mode.
- Decide which node has the data that you want to keep. In this procedure, you will give this node the primary HA role. All data on the other node will be lost during this procedure and will not be recoverable.

If you do not know which node has the most recent or most valuable data, follow these recommendations:

- If the `utils service database status` command output on both nodes indicates that one node currently has the primary HA role while the other node currently has the secondary HA role, you should choose the current primary node to keep as the primary database server.

  ```
  admin: utils service database status
  --------------------------------------------------------------------------------
  The initial configured HA role of this node : primary
  The current HA role of this node                 : primary
  The database vip address : 10.22.130.54
  ... 
  ```

- If the `utils service database status` command output on both nodes indicates that neither or both nodes have the primary HA role, choose the node that you initially installed as the primary server to keep as the primary database server.

  ```
  admin: utils service database status
  --------------------------------------------------------------------------------
  The initial configured HA role of this node : primary
  The current HA role of this node                 : secondary
  The database vip address : 10.22.130.54
  ... 
  ```

Procedure

**Step 1** Log in to the CLI of the database server which has the data that you want to keep.

**Step 2** Enter the `utils service database drbd keep-node` command to reset the server to currently function in the primary HA role.

  ```
  admin: utils service database drbd keep-node
  This command will make this node as Primary
  Trying to assume primary role......... [Done]
  Reconnecting to MySQL........... [Done]
  ```

**Step 3** Log in to the CLI of the other database server.

**Step 4** Enter the `utils service database drbd discard-node` command to reset the server to currently function in the secondary HA role.

  ```
  admin: utils service database drbd discard-node
  This command will make this node as Secondary
  Trying to assume secondary role......... [Done]
  Ensuring DRBD volume unmounted...
  Ensuring DRBD role is Secondary...
  Discarding local MySQL data..... [Done]
  ```
Verifying Synchronization of the Database Servers

Procedure

Step 1
Log in to the CLI of each database server.

Step 2
On each database server, enter the `utils service database status` command.

The following examples show that synchronization is in progress and proceeding successfully, because each node is aware of the HA role of its redundant peer, and the output displays the percentage of the synchronization progress. Also, the current primary database server identifies itself as the SyncSource, while the current secondary database server identifies itself as the SyncTarget.

Sample output from the current primary database server:

```
admin: utils service database status
---------------------------------------------------------------
The initial configured HA role of this node : primary
The **current HA role** of this node : primary
The database vip address : 10.22.130.54
The database primary node name : ctx-db-1
The database primary node IP address : 10.22.130.49
The database secondary node name : ctx-db-2
The database secondary node IP address : 10.22.130.57
Mon status : Running pid 10183
MySQL status : Running pid 10100
Heartbeat status : Running pid 20414
---------------------------------------------------------------

drbd driver loaded OK; device status:
version: 8.3.2 (api:88/proto:86-90)
m:res cs ro ds p mounted fstype
... sync'ed: 2.0% (44104/44980)M
0:mysql  SyncSource Primary/Secondary UpToDate/Inconsistent C /mnt/mysql ext3
---------------------------------------------------------------
```

Sample output from the current secondary database server:

```
admin: utils service database status
---------------------------------------------------------------
The initial configured HA role of this node : secondary
The **current HA role** of this node : secondary
The database vip address : 10.22.130.54
The database primary node name : ctx-db-1
The database primary node IP address : 10.22.130.49
The database secondary node name : ctx-db-2
The database secondary node IP address : 10.22.130.57
Mon status : Not running (only runs on primary)
MySQL status : Not running (only runs on primary)
Heartbeat status : Running pid 17842
```

Synchronization begins between the two database servers.

Step 5
Proceed to the “Verifying Synchronization of the Database Servers” section on page 30-4.

Related Topics

- Command Reference, page C-1
Step 3  
To confirm that the synchronization is complete, enter the **`utils service database status`** command on both the primary and secondary database servers.

The following examples show that synchronization is complete, because the disk state (ds) of the current secondary server is now up to date.

**Sample output from the primary database server:**

```
admin: **`utils service database status`**
--------------------------------------------------------------------------------
The initial configured HA role of this node : primary
The current HA role of this node             : primary
The database vip address                     : 10.22.130.54
The database primary node name               : ctx-db-1
The database primary node IP address         : 10.22.130.49
The database secondary node name             : ctx-db-2
The database secondary node IP address       : 10.22.130.57
Mon status                                   : Running pid 10183
MySQL status                                  : Running pid 10100
Heartbeat status                              : Running pid 20414
--------------------------------------------------------------------------------
```

**Sample output from the secondary database server:**

```
admin: **`utils service database status`**
--------------------------------------------------------------------------------
The initial configured HA role of this node : secondary
The current HA role of this node             : secondary
The database vip address                     : 10.22.130.54
The database primary node name               : ctx-db-1
The database primary node IP address         : 10.22.130.49
The database secondary node name             : ctx-db-2
The database secondary node IP address       : 10.22.130.57
Mon status                                   : Not running (only runs on primary)
MySQL status                                  : Not running (only runs on primary)
Heartbeat status                              : Running pid 17842
--------------------------------------------------------------------------------
```

---

**Note**  
The synchronization takes approximately 40 minutes. During this time, the disk state (ds) of the current secondary server is shown as **inconsistent**. An inconsistent disk state indicates that the synchronization between the database servers is not complete.

---

drbd driver loaded OK; device status:
version: 8.3.2 (api:88/proto:86-90)
 m:res  cs  ro  ds  p mounted  fstype
 ...  sync'ed: 2.1%  (44060/44980)M
0:mysql  SyncTarget  Secondary/Primary  Inconsistent/UpToDate  C
Diagnosing Corrupted DRBD Metadata

If, after you complete the split brain recovery procedure, the database servers still cannot connect to each other and complete synchronization, the metadata for the Distributed Replicated Block Device (DRBD) may be corrupted. The DRBD is what synchronizes the secondary database with changes that are made on the primary database.

**Before You Begin**
This procedure applies only after you attempt split brain recovery. (See the “Recovering from Split Brain Mode” section on page 30-3.)

**Procedure**

**Step 1** Log in to the CLI of each database server.

**Step 2** On each database server, enter the `utils service database status` command.

**Step 3** The DRBD metadata is corrupted if the disk state (ds) value is “Inconsistent/Inconsistent” while the connection state (cs) is “StandAlone” or “WFConnection” on one or both servers.

**Step 4** To recover from corrupted DRBD metadata, proceed to the “Recovering from Corrupted DRBD Metadata” section on page 30-7.

**Example**
In the following example, the status of one database server indicates that the nodes have corrupted DRBD metadata:

```
admin: utils service database status
--------------------------------------------------------------------------------
The initial configured HA role of this node      : secondary
The current HA role of this node                 : secondary
The database vip address                         : 10.22.130.54
The database primary node name                   : ctx-db-1
The database primary node IP address             : 10.22.130.49
The database secondary node name                 : ctx-db-2
The database secondary node IP address           : 10.22.130.57
Mon status                                       : Not running (only runs on primary)
MySQL status                                     : Not running (only runs on primary)
Heartbeat status                                 : Running pid 11459
--------------------------------------------------------------------------------
drbd driver loaded OK; device status:
version: 8.3.2 (api:88/proto:86-90)
m:res cs ro ds p mounted fstype
0:mysql WFConnection Secondary/Unknown Inconsistent/Inconsistent C
```

**Related Topics**
- Command Reference, page C-1
Recovering from Corrupted DRBD Metadata

Before You Begin

- Make sure that the database servers are correctly cabled. See the “Cabling Requirements for the Database Servers” section on page 4-3.
- Complete the “Diagnosing Corrupted DRBD Metadata” section on page 30-6 to confirm that your system has corrupted DRBD metadata.

Procedure

Step 1  Log in to the CLI of the database server which has the data that you want to keep.
This should be the same node whose data you decided to keep when you completed the procedure in the “Recovering from Split Brain Mode” section on page 30-3.

Step 2  Enter the `utils service database drbd force-keep-node` command to reset the DRBD metadata and set the server to currently function in the primary HA role.

```
admin: utils service database drbd force-keep-node
This command will make this node as Primary
Trying to assume primary role........... [Done]
Overwriting peer data... [Done]
```

Step 3  Log in to the CLI of the other database server.

Step 4  Enter the `utils service database drbd force-discard-node` command to reset the DRBD metadata and set the server to currently function in the secondary HA role.

```
admin: utils service database drbd force-discard-node
Shutting down Heartbeat...
Stopping High-Availability services:
[ OK ]
Ensuring DRBD volume unmounted...
unmount: /dev/drbd0: not mounted
Taking down DRBD Resource...
Recreating DRBD meta-data...
NOT initialized bitmap
Bringing up DRBD...
Starting Heartbeat...
Starting High-Availability services:
[ OK ]
[Done]
```

Synchronization begins between the two database servers.

Step 5  Proceed to the “Verifying Synchronization of the Database Servers” section on page 30-4.

Related Topics

- Command Reference, page C-1