Administering Failure and Recovery

This section describes how ClustrixDB handles various types of failures and tells you how best to recover from them if they occur.

Front-end Network Failure

If a node cannot communicate on its front end Ethernet network port, for example, due to an inadvertent cable pull, switch misconfiguration, or NIC failure, no manual intervention is required. The cluster takes the following actions:

- Additional connections are not assigned to the failed instance.
- If the failed node was handling replication slave connections, the connections are moved to another node.

Back-end Network Failure

Backend network failures look like Node Failures (see below).

Disk Failure

ClustrixDB maintains two copies of your data (replicas) by default. When the cluster detects a disk failure, the system will automatically schedule reprotect operations to generate new copies of data. The administrator does not need to take any action to reprotect the dataset. The cluster will also issue Database Alerts when any tables fall below the specified number of copies and following the completion of the resulting reprotect task.

In some situations, the system may detect errors on the disk. However, if the errors are below the threshold at which we mark the drive as failed, some user queries may get occasional errors attempting to read from the failed device. In such cases, the administrator may manually deactivate the node without reducing the number of available copies. The system will attempt to safely move all data on such devices to another device in the system. To do this, follow the steps in Reducing Your Cluster's Capacity - Flex Down.

Node Failure

This section describes two types of node failure: transient, where the node is offline briefly (for example, due to a crash or power failure), and permanent, where a node has failed completely and is not expected to return (for example, due to a hardware failure).

Transient Node Failure and Rebalancer Reprotect

When the cluster loses contact with an individual node for any reason, surviving nodes in the cluster form a new group without this node and continue serving clients. All services, such as replication slaves, are reassigned across the surviving nodes. Clients that were distributed to the failed node must reconnect. Clients that were directly connected to the failed node are unable to query the database. You will receive an email alert and a message like following will appear in the clustrix.log for one of the nodes:

> *ALERT PROTECTION_LOST WARNING Full protection lost for some data; queueing writes for down node; reprotection will begin in 600 seconds if node has not recovered*

This message simply indicates that not all data has a full set of replicas available. The global variable rebalancer_reprotect_queue_interval_s specifies how long the Rebalancer should wait for a node to re-join the cluster before starting to create additional replicas.

If a node re-joins within *rebalancer_reprotect_queue_interval_s*	**If a node joins after *rebalancer_reprotect_queue_interval_s* has passed**

• ClustrixDB replays the changes that were made since the last time the node was in quorum, thereby enabling the node to rejoin the cluster quickly. • The node rejoins the cluster and begins accepting work. No further action is necessary.	• The Rebalancer begins copying under-protected slices to create new replicas throughout the surviving nodes. • The cluster performs a group change. • Assuming there is sufficient disk space, ClustrixDB will automatically handle the reprotect process and no manual intervention is required.

Use this sql to view the Rebalancer reprotection tasks that have not been finished.

```
sql> SELECT * FROM system.rebalancer_activity_log where finished is null;
```

Once there are sufficient copies of all replicas (either because a node was recovered or the Rebalancer is done making copies), you will receive an alert and a message like the following will appear in the clustrix.log for one of the nodes:

> *ALERT PROTECTION_RESTORED WARNING Full protection restored for all data after 20 minutes and 40 seconds*

Softfailing a Node

If a node becomes unreliable and you would like to remove it from the cluster, Clustrix recommends marking it as softfailed (using the Flex Down) procedure. You can simultaneously incorporate a replacmement using the Flex Up procedure. The high level steps are:

1. Provision replacement node(s) by installing ClustrixDB and adding them to the cluster using ALTER CLUSTER ADD
2. Mark the node(s) in question as softfailed using ALTER CLUSTER SOFTFAIL
3. Once softfail operations complete, execute ALTER CLUSTER REFORM to remove the softfailed node(s)

Permanent Node Failure

If a node has failed permanently, the Rebalancer will automatically create additional replicas as described above. The lost node is still considered to be a quorum participant until it is removed explicitly.

Manually drop a permanently failed node
`ALTER CLUSTER DROP nodeid;`

This command results in a group change.

Dropping a node before reprotect has completed can leave the cluster vulnerable to data loss.

To incorporate a replacement node, follow the instructions for Expanding Your Cluster's Capacity - Flex Up.

Multiple Node Failures

ClustrixDB can be configured to withstand multiple failures by setting the value of MAX_FAILURES.

For a cluster to tolerate the configured value for MAX_FAILURES:

- All representations must have sufficient replicas. If MAX_FAILURES is updated, all tables created previously must have their replicas updated manually.
- There must be a quorum (at least N/2+1) of nodes available
- Clustrix recommends provisioning enough disk space so that the cluster has enough space to reprotect after an unexpected failure. See Allocating Disk Space for Fault Tolerance and Availability

Zone Failure

When zones are configured, a failure of an entire zone is analgous to a node failure. ClustrixDB will automatically resume operation with the nodes from available zones and automatically reprotect. To remove a zone from the cluster, mark all nodes in the zone as softfailed.

When Reprotect Cannot Complete

Insufficient Disk Space

If there is insufficient disk space for all replicas, the Reprotect process will be unable to complete. Consider adding additional capacity by Expanding Your Cluster's Capacity - Flex Up. See Managing File Space and Database Capacity.

Missing Replicas

If the cluster has lost more nodes and/or zones than specified for MAX_FAILURES , the cluster will be unable to reprotect. The Rebalancer activity log (system.rebalancer_activity_log) will show "Representation not found" errors.

To see tables with replicas that are not on any of the currently available nodes:

```
sql> SELECT `Database`, `Table`, `Index`, slice, status
     FROM (SELECT `Database`, `Table`, `Index`, slice, MIN(status)
     AS status FROM system.table_replicas
     GROUP BY slice)
     AS x
     WHERE x.status > 1;
```

If the unavailable nodes cannot be recovered, these tables must be restored from backup.