

MAX_FAILURES

MAX_FAILURES is the number of failures that can occur **simultaneously** while ensuring that no data is lost. By default for (non-zone) configurations, a single node can become unavailable and ClustrixDB will resume operations without data loss. When zones are configured, a node or a zone (with any number of nodes per zone) can become unavailable with no loss of data.

The value of MAX_FAILURES is also used to determine the default number of replicas created for a table or index. For the default of MAX_FAILURES = 1, new tables are created with REPLICAS = 2.

Prerequisites for MAX_FAILURES

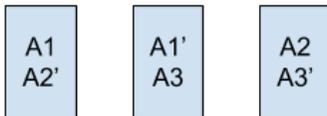
For a cluster to tolerate the configured value for MAX_FAILURES:

- 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](#)

MAX_FAILURES = 1

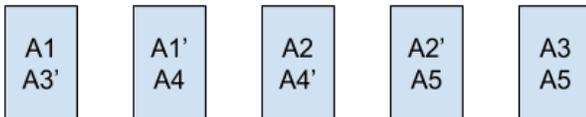
The default configuration for ClustrixDB is MAX_FAILURES = 1, which indicates REPLICAS = 2.

In the following example, Table A is configured with SLICES = 2 and REPLICAS = 2 (default for MAX_FAILURES = 1). Slices are labeled A1, and A2, and A3 and prime notation is used to denote replicas. A1 and A1' are different replicas of the same slice.



In this configuration any node can be lost with no data loss as the remaining nodes will contain a copy of the slice. After the loss of a node, ClustrixDB will continue to operate using the remaining nodes and work to create additional copies for replicas that were lost.

The default configuration of ClustrixDB is to have slices equal to the number of nodes (hash_dist_min_slices = 0). In this example, A1-A5 are used to label slices, and A1 and A1' are replicas of the same slice.



Any node can be lost with no data loss as the remaining nodes will contain a copy of the slice. Increasing the number of slices and cluster size increases parallelism and improves performance but does not change the degree of fault tolerance.

MAX_FAILURES = 2

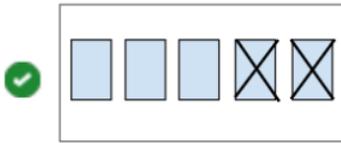
To configure ClustrixDB to tolerate additional failures, you can increase the value of MAX_FAILURES.

Increasing the value of MAX_FAILURES increases the number of replicas required, which can have a significant performance impact to writes and requires additional disk space.

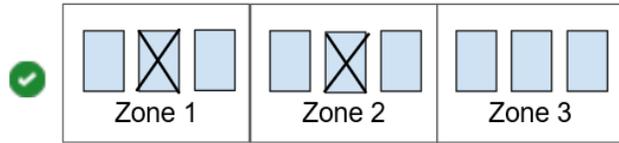
If additional failure protection is desired MAX_FAILURES can be set to 2. Setting MAX_FAILURES=2 will increase the default number of replicas maintained by the cluster and allow it to survive two simultaneous failures as long as the cluster is able to maintain a [quorum](#) of nodes ($N/2 + 1$). This means two nodes can fail simultaneously, or if Zones are configured, a single zone failure and an additional node can fail.

The following examples illustrate fault tolerance for MAX_FAILURES = 2.

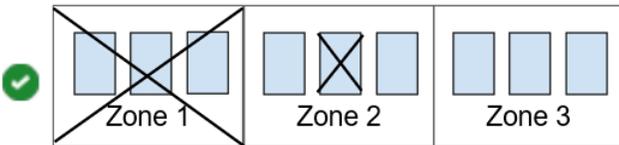
A 5 node cluster can sustain 2 simultaneous node failures with no loss of data.



A 9 node cluster deployed across 3 zones can sustain 2 simultaneous node failures (in any zone or zones) with no data loss.

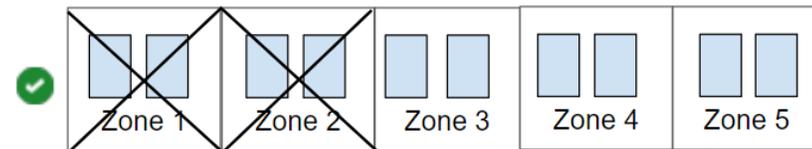


A 9 node cluster deployed across 3 zones can sustain a zone failure, plus one additional node failure with no data loss:

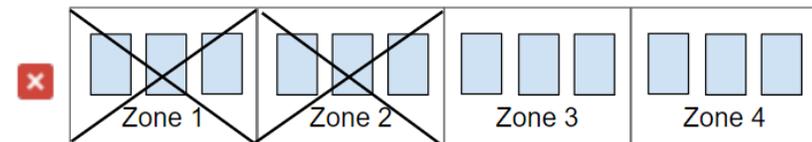


However, if a 9 node cluster deployed across 3 zones sustains 2 simultaneous zone failures, the remaining cluster does not meet the quorum requirement ($N/2 + 1$):

A 10 node cluster deployed across 5 zones can sustain 2 simultaneous zone failures:



When `MAX_FAILURES = 2`, there must be a minimum of 5 zones to sustain 2 zone failures in order to meet the quorum requirement ($N/2 + 1$):



Due to the high performance overhead, Clustrix does not recommend exceeding `MAX_FAILURES = 2`.

Changing the value of `MAX_FAILURES`

To change this value, follow the procedure outlined in [ALTER CLUSTER SET MAX_FAILURES](#). Clustrix supports a maximum of 5 zones.