Zones

ClustrixDB allows clusters to be deployed across multiple zones to provide fault tolerance during unplanned zone failure. A zone in ClustrixDB is grouping of nodes, such as AWS Availability Zones within the same Region, different server racks, different network switches, different power sources, or even separate servers in different data centers. ClustrixDB requires that network latency between zones not exceed 2ms.

When ClustrixDB is configured for zones, the Rebalancer becomes zone-aware and ensures replicas are placed in different zones. Additionally, Paxos accepters are distributed across multiple zones. The result is that ClustrixDB is fault tolerant in the face of unexpected zone failure, regardless of how many nodes are in the zone.

Prerequisites

Use of zones with ClustrixDB has the following prerequisites:

- Clustrix supports configurations with a minimum of 3 and a maximum of 5 zones.
- There must be a sufficient number of zones to maintain a quorum of nodes following a failure
- Zones must be in the same geographical region with <2ms network latency between each zone.
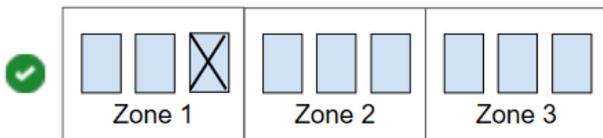- A cluster deployed across zones must have the same number of nodes in each zone.

Fault Tolerance and Zones

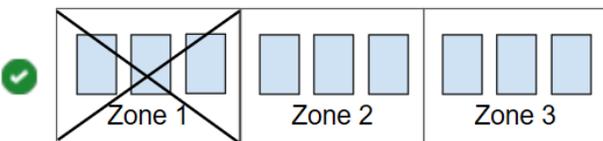The following examples illustrate how zones work with the default MAX_FAILURES = 1.

There must be the same number of nodes in each zone. Clustrix recommends deploying a 9 node cluster across 3 zones.
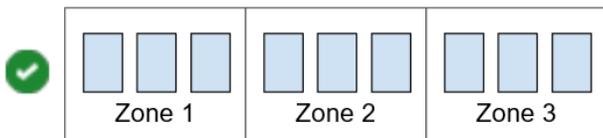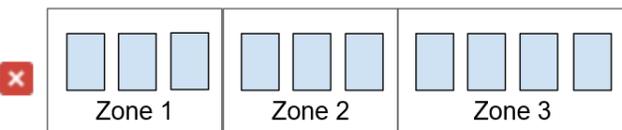

This cluster can tolerate one node failure:

Or one zone failure:

There must be the same number of nodes in each zone. The following configuration is not supported.

If you require a larger cluster, simply add the same number of nodes to each zone.

See MAX_FAILURES for additional information.

Configuring Zones

To configure zones, use ALTER CLUSTER ZONE to assign nodes to a zone. Zone ids must be greater than 0 and there must be an equal number of nodes in each zone. For example, a 9 node cluster deployed across 3 zones should have 3 nodes each in zones 1, 2, and 3.

Clustrix recommends provisioning enough free disk space to reprotect the data after a zone failure. See Allocating Disk Space for Fault Tolerance and Availability.

> Clustrix currently supports a minimum of 3 and a maximum of 5 zones.

Creating a New Cluster in Multiple Zones

To create a new cluster, follow these steps:

1. Provision nodes in different zones, where zones are separate server racks, separate servers in different data centers, separate network switches, independent power sources, or AWS Availability Zones.
2. Use ALTER CLUSTER ADD to build a cluster.
3. Use ALTER CLUSTER ZONE to assign nodes to zones.
4. Verify that zones have been configured using clx stat. All nodes should show a zone number of 1 - *nn*. Zone 0 should not appear in the results.

Migrating an Existing Cluster to Multiple Zones

If you would like to migrate your existing cluster, the high-level steps are:

1. Temporarily disable the Rebalancer.
2. Provision nodes in new zones and install ClustrixDB on each node.
3. Use the Flex Up procedure to add the new nodes to the cluster.
4. Configure zones using ALTER CLUSTER ZONE.
5. Use the Flex Down procedure to remove surplus nodes from the cluster.
6. Re-enable the Rebalancer.

Once this procedure is complete, you should have the same number of nodes in each zone and all nodes should have a non-zero zone assigned. Because the flex up and flex down operations can be time consuming, it is often simpler to migrate to a new cluster when moving from default to zone configurations.

Resources

For further information, see Understanding Fault Tolerance in ClustrixDB.