

Group Changes

[What is a Group Change?](#)

[When Does a Cluster Experience a Group Change?](#)

[When Node\(s\) are Added to a Cluster](#)

[When Node\(s\) Leave the Cluster](#)

[What Happens During a Group Change?](#)

[Details of a Group Change](#)

[Cluster Pauses Processing and Performs Internal Operations](#)

[Cluster Forms New Group](#)

[What Happens to Processes That Were Running?](#)

[Special Considerations](#)

What is a Group Change?

Xpand uses a distributed group membership protocol to maintain the static set of all nodes known to the cluster and checks that the nodes maintain active communication between each other. Xpand refers to this as a Group.

When the set of nodes changes, there is a change in the group and a Group Change occurs. During a Group Change, Xpand performs tasks to ensure:

- Data consistency
- Data availability
- Effective query distribution

When Does a Cluster Experience a Group Change?

When Node(s) are Added to a Cluster

A Group Change occurs in conjunction with [Expanding Your Cluster's Capacity - Flex Up](#). Following the Group Change, the [Rebalancer](#) will work in the background to move slices to the new node(s). You may notice a slight degradation of performance during that time.

When Node(s) Leave the Cluster

When reducing a cluster's capacity using the [Flex Down](#) procedure, the `ALTER CLUSTER REFORM` command, used to remove nodes, will invoke a Group Change.

A cluster will also Group Change if node(s) are dropped using the emergency procedure of [ALTER CLUSTER DROP](#).

Additionally, there are several unscheduled events that can cause a cluster to Group Change:

- A cluster experiences unexpected node failure(s) due to hardware failure, network failure, or kernel panic.
- A node or node(s) are unable to be reached during a regular heartbeat check of the cluster.

Following a node loss, if the node was not previously soft-failed, the Rebalancer will automatically work to reprotect all data and ensure all data has sufficient copies throughout the cluster. You may notice performance degradation during the reprotect process.

What Happens During a Group Change?

If Xpand detects a change in its group, it will recover automatically as long as a [quorum](#) of nodes is available. Your cluster will experience a brief period during which the group is being reformed and the consistency of the database is ensured. Connections from applications to surviving nodes will remain but transactions and queries for those connections will be temporarily paused.

The cluster can recover from multiple simultaneous node failures if the total number of failed nodes does not exceed the value configured for `max_failures`.

Details of a Group Change

Group Changes are relatively short (generally measured in seconds), though the duration of each Group Change depends on factors such as the number of containers, workload, and cluster size. The underlying steps of a Group Change are the same, regardless of cluster size and workload.

Cluster Pauses Processing and Performs Internal Operations

When there is a Group Change, the cluster pauses all processing and determines whether a [quorum](#) of nodes is available. If true, Xpand performs a series of internal operations in preparation for the new group. Together these operations may take a few seconds, or 10s or seconds, depending on how large the cluster is, how large the database is, and how many transactions were in-process when the Group Change occurred. These steps ensure that the consistency of the database is guaranteed despite having lost a member of the cluster.

- Initializing subsystems such as flow control and the Rebalancer.
- Synchronizing global cluster state, including internal system catalogs and global variables.
- Resolving (or re-resolving) in-process transactions, including rolling back transactions that were interrupted by the Group Change.
- Invalidating or rebuilding internal caches, such as the [Query Plan Cache](#).
- Creating [recovery queues](#) for downed replicas, or "flipping" queues that are no longer needed.
- Performing checks for [licensing](#) and [nResiliency](#).
- [Resizing device files](#) if necessary.

Cluster Forms New Group

Once the cluster is ready to resume operations, a new group is formed and the clustrix.log will contain an informational message that includes details of the new group:

```
[INFO] Node 1 has new group effff: { 1-4 down: 5 }
```

This example shows a cluster that has re-grouped without node 5. The database then resumes its operations.

What Happens to Processes That Were Running?

If any of the following processes were running when a Group Change occurred, they will be impacted as shown:

Process	Result
Queries (DML and DDL)	If a transaction or statement is interrupted by a Group Change before it has a chance to commit, it will receive an error. If the global autoretry is true and a transaction was submitted with autocommit enabled, the database will automatically retry. If the retried statements cannot be executed successfully, the application will receive another error.
Replication	Replication processes will automatically restart at the proper binlog location following a Group Change.
Other Connections	Connections to nodes that are still in quorum will be maintained and will not experience any errors. Connections to non-communicative nodes will be lost.

Special Considerations

In-Memory tables may be impacted by a Group Change. See [In-Memory Tables](#) for more information.