Performing Cluster Maintenance Operations

Objectives

After completing this section, you should be able to perform common cluster maintenance tasks, such as adding or removing MONs and OSDs, and recovering from various component failures.

Adding or Removing OSD Nodes

The cluster can operate and serve clients in a degraded state during cluster maintenance activities. However, adding or removing OSDs can affect cluster performance. Backfilling operations can generate large data transfers between OSDs, causing cluster performance to degrade.

Evaluate the potential performance impact before performing cluster maintenance activities. The following factors typically affect cluster performance when adding or removing OSD nodes:

- Client load

  If an OSD node has a pool that is experiencing high client loads, then performance and recovery time could be negatively affected. Because write operations require data replication for resiliency, write-intensive client loads increase cluster recovery time.

- Node capacity

  The capacity of the node being added or removed affects the cluster recovery time. The node's storage density also affects recovery times. For example, a node with 36 OSDs takes longer to recover than a node with 12 OSDs.

- Spare cluster capacity

  When removing nodes, verify that you have sufficient spare capacity to avoid reaching the full or near full ratios. When a cluster reaches the full ratio, Ceph suspends write operations to prevent data loss.

- CRUSH rules

  A Ceph OSD node maps to at least one CRUSH hierarchy, and that hierarchy maps to at least one pool via a CRUSH rule. Each pool using a specific CRUSH hierarchy experiences a performance impact when adding and removing OSDs.

- Pool types

  Replication pools use more network bandwidth to replicate data copies, while erasure-coded pools use more CPU to calculate data and coding chunks.

  The more data copies that exist, the longer it takes for the cluster to recover. For example, an erasure-coded pool with many chunks takes longer to recover than a replicated pool with fewer copies of the same data.

- Node hardware