

Configuring Cluster Monitors

Objectives

After completing this section, you should be able to describe the purpose of cluster monitors and the quorum procedures, query the monitor map, manage the configuration database, and describe Cephx.

Configuring Ceph Monitors

Ceph monitors (MONs) store and maintain the cluster map that clients use to find the MON and OSD nodes. Ceph clients must connect to a MON to retrieve the cluster map before they can read or write any data to OSDs. For this reason, the proper configuration of cluster MONs is critical.

MONs form a quorum and elect a *leader* by using a variation of the Paxos algorithm, to achieve consensus among a distributed set of computers.

MONs each have one of the following roles:

- *Leader*: the first MON to obtain the most recent version of the cluster map.
- *Provider*: a MON that has the most recent version of the cluster map, but is not the leader.
- *Requester*: a MON that does not have the most recent version of the cluster map and must synchronize with a provider before it can rejoin the quorum.

Synchronization always occurs when a new MON joins the cluster. Each MON periodically checks whether a neighboring monitor has a more recent version of the cluster map. If a MON does not have the most recent version of the cluster map, then it must synchronize and obtain it.

A majority of the MONs in a cluster must be running to establish a quorum. For example, if five MONs are deployed, then three must be running to establish a quorum. Deploy at least three MON nodes in your production Ceph cluster to ensure high availability. You can add or remove MONs in a running cluster.

The cluster configuration file defines the MON host IP addresses and ports for the cluster to operate. The `mon_host` setting can contain IP addresses or DNS names. The `cephadm` tool does not update the cluster configuration file. Define a strategy to keep the cluster configuration files synchronized across cluster nodes, such as with `r sync`.

```
[global]
mon_host = [v2:172.25.250.12:3300,v1:172.25.250.12:6789],
[v2:172.25.250.13:3300,v1:172.25.250.13:6789],
[v2:172.25.250.14:3300,v1:172.25.250.14:6789]
```



Important

Changing MON node IP addresses is not recommended after the cluster is deployed and running.