Chapter 11 | Managing a Red Hat Ceph Storage Cluster

Configuration option	Description
osd_mon_heartbeat_interval	Ceph monitor heartbeat interval.
mon_osd_report_timeout	The time-out (in seconds) before the MON marks an OSD as down if it does not report.

Monitoring OSD Capacity

Red Hat Ceph Storage provides configuration parameters to help prevent data loss due to a lack of storage space in the cluster. You can set these parameters to provide an alert when OSDs are low on storage space.

When the value of the mon_osd_full_ratio setting is reached or exceeded, the cluster stops accepting write requests from clients and enters the HEALTH_ERR state. The default full ratio is 0.95 (95%) of the available storage space in the cluster. Use the full ratio to reserve enough space so that if OSDs fail, there is enough space left that automatic recovery succeeds without running out of space.

The mon_osd_nearfull_ratio setting is a more conservative limit. When the value of the mon_osd_nearfull_ratio limit is reached or exceeded, the cluster enters the HEALTH_WARN state. This is intended to alert you to the need to add OSDs to the cluster or fix issues before you reach the full ratio. The default near full ratio is 0.85 (85%) of the available storage space in the cluster.

The mon_osd_backfillfull_ratio setting is the threshold at which cluster OSDs are considered too full to begin a backfill operation. The default backfill full ratio is 0.90 (90%) of the available storage space in the cluster.

Use the ceph osd set-full-ratio, ceph osd set-nearfull-ratio, and ceph osd set-backfillfull-ratio commands to configure these settings.

[ceph: root@node /]# ceph osd set-full-ratio .85
[ceph: root@node /]# ceph osd set-nearfull-ratio .75
[ceph: root@node /]# ceph osd set-backfillfull-ratio .80



Note

The default ratio settings are appropriate for small clusters, such as the one used in this lab environment. Production clusters typically require lower ratios.

Different OSDs might be at full or nearfull depending on exactly what objects are stored in which placement groups. If you have some OSDs full or nearfull and others with plenty of space remaining, analyze your placement group distribution and CRUSH map weights.

Monitoring Placement Groups

Every placement group (PG) has a status string assigned to it that indicates its health state. When all placement groups are in the active+clean state, the cluster is healthy. A PG status