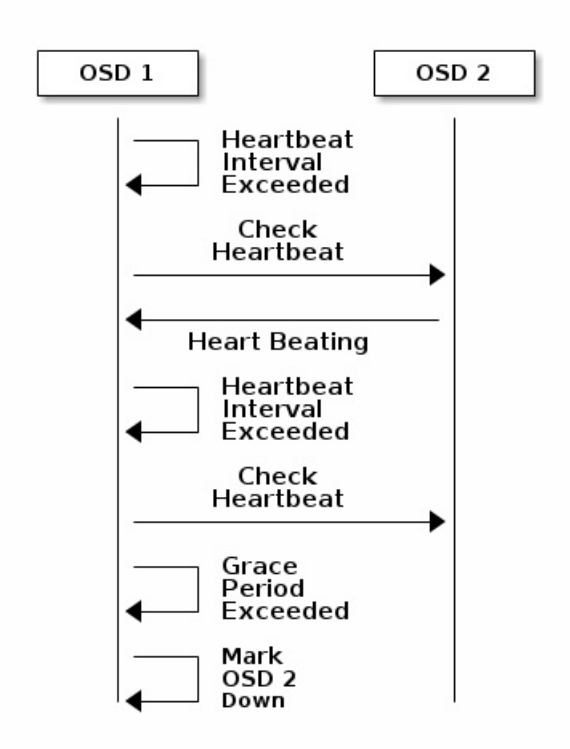
CONFIGURING MONITOR/OSD INTERACTION

After you have completed your initial Ceph configuration, you may deploy and run Ceph. When you execute a command such as `ceph health` or `ceph -s`, the monitor reports on the current state of the cluster. The monitor knows about the cluster by requiring reports from each OSD, and by receiving reports from OSDs about the status of their neighboring OSDs. If the monitor doesn't receive reports, or if it receives reports of changes in the cluster, the monitor updates the status of the cluster.

Ceph provides reasonable default settings for monitor/OSD interaction. However, you may override the defaults. The following sections describe how Ceph monitors and OSDs interact for the purposes of monitoring.

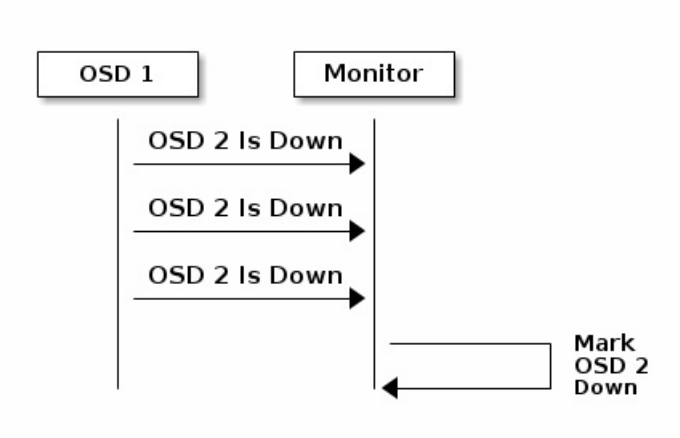
OSDS CHECK HEARTBEATS

Each OSD checks the heartbeat of other OSDs every 6 seconds. You can change the heartbeat interval by adding an `osd heartbeat interval` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime. If an OSD doesn't show a heartbeat within a 20 second grace period, the cluster may consider the OSD down. You may change this grace period by adding an `osd heartbeat grace` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime.



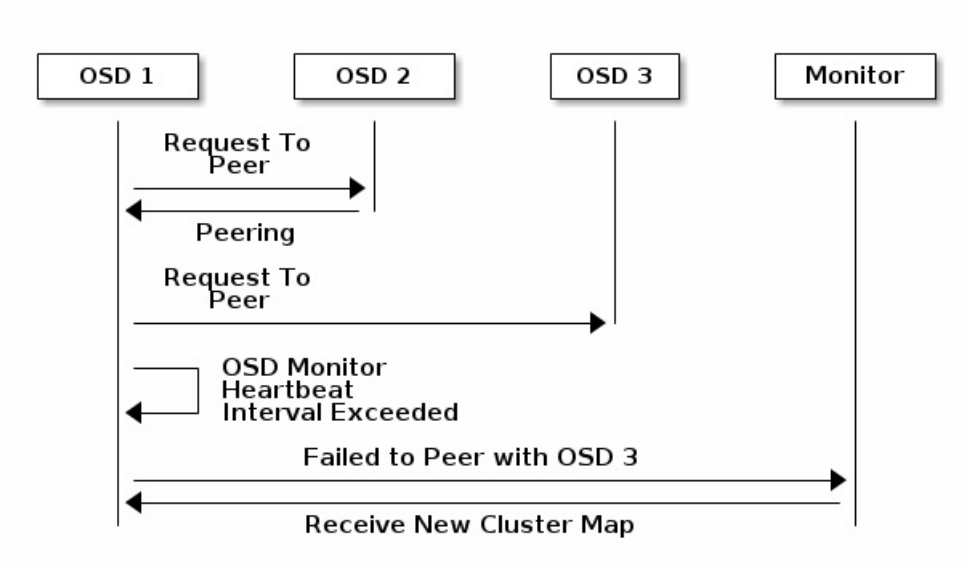
OSDS REPORT DOWN OSDS

By default, an OSD must report to the monitors that another OSD is down three times before the monitors acknowledge that the reported OSD is down. You can change the minimum number of `osd down` reports by adding an `osd min down reports` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime. By default, only one OSD is required to report another OSD down. You can change the number of OSDs required to report a monitor down by adding an `osd min down reporters` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime.



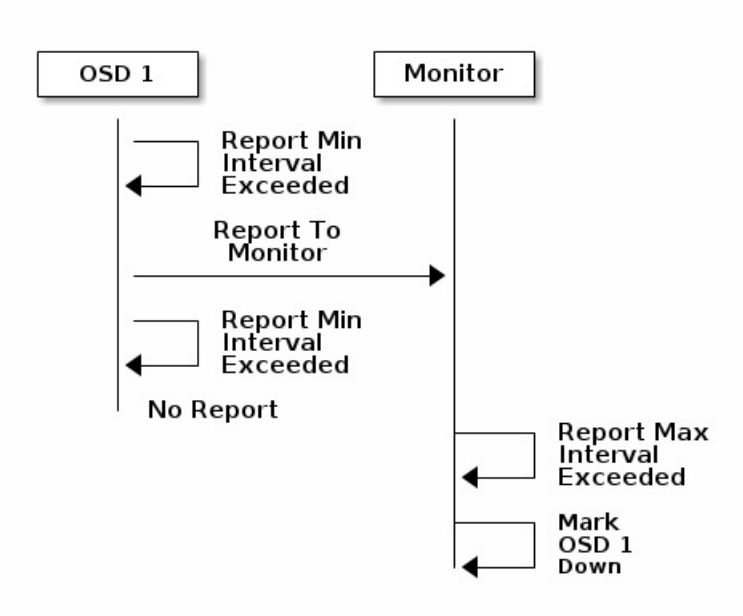
OSDS REPORT PEERING FAILURE

If an OSD cannot peer with any of the OSDs defined in its Ceph configuration file, it will ping the monitor for the most recent copy of the cluster map every 30 seconds. You can change the monitor heartbeat interval by adding an `osd mon heartbeat interval` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime.



OSDS REPORT THEIR STATUS

If an OSD doesn't report to the monitor once at least every 120 seconds, the monitor will consider the OSD down. You can change the monitor report interval by adding an `osd mon report interval max` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime. The OSD attempts to report on its status every 30 seconds. You can change the OSD report interval by adding an `osd mon report interval min` setting under the `[osd]` section of your Ceph configuration file, or by setting the value at runtime.



CONFIGURATION SETTINGS

When modifying heartbeat settings, you should include them in the `[global]` section of your configuration file.

MONITOR SETTINGS

`mon osd min up ratio`

Description: The minimum ratio of `up` OSDs before Ceph will mark OSDs `down`.
Type: Double
Default: `.3`

`mon osd min in ratio`

Description: The minimum ratio of `in` OSDs before Ceph will mark OSDs `out`.
Type: Double
Default: `.3`

`mon osd laggy halflife`

Description: The number of seconds laggy estimates will decay.
Type: Integer
Default: `60*60`

`mon osd laggy weight`

Description: The weight for new samples in laggy estimation decay.
Type: Double
Default: `0.3`

`mon osd adjust heartbeat grace`

Description: If set to `true`, Ceph will scale based on laggy estimations.
Type: Boolean
Default: `true`

`mon osd adjust down out interval`

Description: If set to `true`, Ceph will scaled based on laggy estimations.
Type: Boolean
Default: `true`

`mon osd auto mark in`

Description: Ceph will mark any booting OSDs as `in` the cluster.
Type: Boolean
Default: `false`

`mon osd auto mark auto out in`

Description: Ceph will mark booting OSDs auto marked `out` of the cluster as `in` the cluster.
Type: Boolean
Default: `true`

`mon osd auto mark new in`

Description: Ceph will mark booting new OSDs as `in` the cluster.
Type: Boolean
Default: `true`

`mon osd down out interval`

Description: The number of seconds Ceph waits before marking an OSD `down` and `out` if it doesn't respond.
Type: 32-bit Integer
Default: `300`

`mon osd downout subtree limit`

Description: The largest CRUSH unit type that Ceph will automatically mark `out`.
Type: String
Default: `rack`

`mon osd report timeout`

Description: The grace period in seconds before declaring unresponsive OSDs `down`.
Type: 32-bit Integer
Default: `900`

OSD SETTINGS

`osd heartbeat address`

Description: An OSD's network address for heartbeats.
Type: Address
Default: The host address.

`osd heartbeat interval`

Description: How often an OSD pings its peers (in seconds).
Type: 32-bit Integer
Default: 6

`osd heartbeat grace`

Description: The elapsed time when an OSD hasn't shown a heartbeat that the cluster considers it `down`.
Type: 32-bit Integer
Default: `20`

`osd mon heartbeat interval`

Description: How often the OSD pings a monitor if it has no OSD peers.
Type: 32-bit Integer
Default: `30`

`osd mon report interval max`

Description: The maximum time in seconds for an OSD to report to a monitor before the monitor considers the OSD `down`.

Type: 32-bit Integer
Default: `120`

`osd mon report interval min`

Description: The minimum number of seconds for an OSD to report to a monitor to avoid the monitor considering the OSD down.
Type: 32-bit Integer
Default: `5`
Valid Range: Should be less than `osd mon report interval max`

`osd mon ack timeout`

Description: The number of seconds to wait for a monitor to acknowledge a request for statistics.
Type: 32-bit Integer
Default: `30`

`osd min down reporters`

Description: The minimum number of OSDs required to report a down OSD.
Type: 32-bit Integer
Default: `1`

`osd min down reports`

Description: The minimum number of times an OSD must report that another is down.
Type: 32-bit Integer
Default: `3`