RECOVERING FROM MONITOR FAILURES

In production clusters, we recommend running the cluster with a minimum of three monitors. The failure of a single monitor should not take down the entire monitor cluster, provided a majority of the monitors remain available. If the majority of nodes are available, the remaining nodes will be able to form a quorum.

When you check your cluster's health, you may notice that a monitor has failed. For example:

```
ceph health
HEALTH_WARN 1 mons down, quorum 0,2
```

For additional detail, you may check the cluster status:

```
ceph status
HEALTH_WARN 1 mons down, quorum 0,2
mon.b (rank 1) addr 192.168.106.220:6790/0 is down (out of quorum)
```

In most cases, you can simply restart the affected node. For example:

```
service ceph -a restart {failed-mon}
```

If there are not enough monitors to form a quorum, the ceph command will block trying to reach the cluster. In this situation, you need to get enough `ceph-mon` daemons running to form a quorum before doing anything else with the cluster.

CLIENT CAN'T CONNECT/MOUNT

Check your IP tables. Some OS install utilities add a REJECT rule to `iptables`. The rule rejects all clients trying to connect to the host except for `ssh`. If your monitor host's IP tables have such a REJECT rule in place, clients connecting from a separate node will fail to mount with a timeout error. You need to address `iptables` rules that reject clients trying to connect to Ceph daemons. For example, you would need to address rules that look like this appropriately:

```
REJECT all -- anywhere anywhere reject-with icmp-host-prohibited
```

You may also need to add rules to IP tables on your Ceph hosts to ensure that clients can access the ports associated with your Ceph monitors (i.e., port 6789 by default) and Ceph OSDs (i.e., 6800 et. seq. by default). For example:

```
iptables -A INPUT -m multiport -p tcp -s {ip-address}/{netmask} --dports 6789,6800:6810 -j AC
```

LATENCY WITH DOWN MONITORS

When you have a monitor that is down, you may experience some latency as clients will try to connect to a monitor in the configuration even though it is down. If the client fails to connect to the monitor within a timeout window, the client will try another monitor in the cluster.

You can also specify the `-m` option to point to a monitor that is up and in the quorum to avoid latency.

=