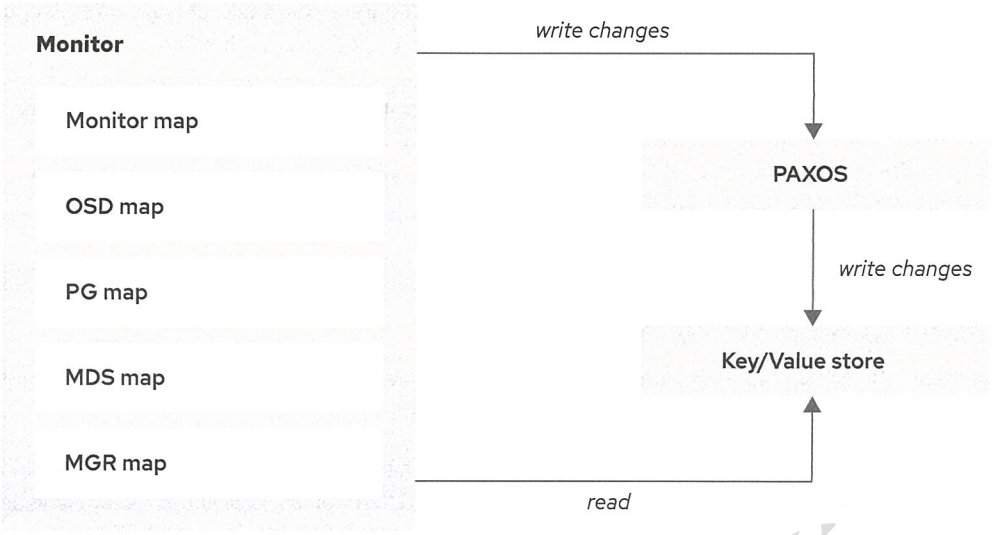

Figure 5.2: Cluster map consistency using Paxos

Propagating the OSD Map

OSDs regularly report their status to the monitors. In addition, OSDs exchange heartbeats so that an OSD can detect the failure of a peer and report that event to the monitors.

When a leader monitor learns of an OSD failure, it updates the map, increments the epoch, and uses the Paxos update protocol to notify the other monitors, at the same time revoking their leases. After a majority of monitors acknowledge the update, and the cluster has a quorum, the leader monitor issues a new lease so that the monitors can distribute the updated OSD map. This method avoids the map epoch ever going backwards anywhere in the cluster, and finding previous leases that are still valid.

OSD Map Commands

Use the following commands to manage the OSD map as an administrator:

Command	Action
`ceph osd dump`	Dump the OSD map to standard output.
`ceph osd getmap -o` *`binfile`*	Export a binary copy of the current map.
`osdmaptool --print` *`binfile`*	Display a human-readable copy of the map to standard output.
`osdmaptool --export-crush` *`crushbinfile binfile`*	Extract the CRUSH map from the OSD map.
`osdmaptool --import-crush` *`crushbinfile binfile`*	Embed a new CRUSH map.
`osdmaptool --test-map-pg` *`pgid binfile`*	Verify the mapping of a given PG.