

FailureScenarios

How ZooKeeper Handles Failure Scenarios

Network partitions

When there is a network partition that separates members of the ZooKeeper ensemble and there exists a quorum side and a non-quorum side of the partition:

- if the leader is in the quorum side of the partition, all read and write requests to the quorum side continue normally. On the non-quorum side of the partition, there is no leader, so requests to read or write to those leaderless ZooKeeper servers result in an [error](#).
- if the leader is in the non-quorum side of the partition, that side of the partition will recognize that it no longer has a quorum of the ensemble. For writes, you need a quorum in order to commit a change so all writes will fail. For reads, there may be a narrow window where reads are returned with stale data - note that no consistency guarantees are violated here because it's legal to serve a stale value as long as you yourself haven't overwritten it. On the quorum-side of the partition, the nodes will recognize that there is not a leader among them and will elect a new leader. When the partition is resolved, the non-quorum side will join the quorum side and update their data.