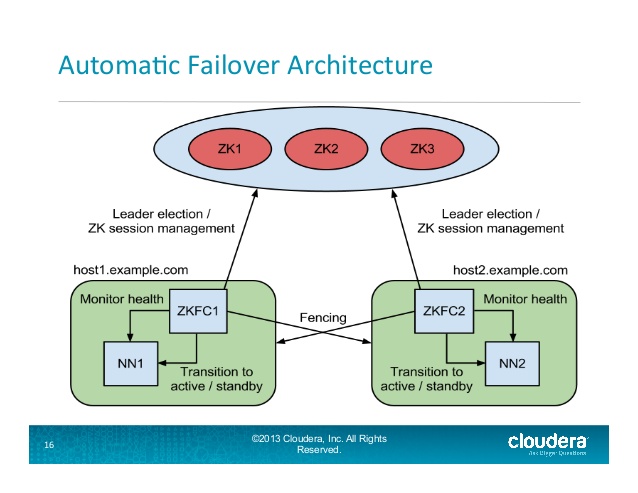
How does Hadoop Namenode failover process works?

The ZKFailoverController (ZKFC) is a new component which is a ZooKeeper client which also monitors and manages the state of the NameNode. Each of the machines which runs a NameNode also runs a ZKFC, and that ZKFC is responsible for:

* Health monitoring - the ZKFC pings its local NameNode on a periodic basis with a health-check command. So long as the NameNode responds in a timely fashion with a healthy status, the ZKFC considers the node healthy. If the node has crashed, frozen, or otherwise entered an unhealthy state, the health monitor will mark it as unhealthy.
* ZooKeeper session management - when the local NameNode is healthy, the ZKFC holds a session open in ZooKeeper. If the local NameNode is active, it also holds a special "lock" znode. This lock uses ZooKeeper's support for "ephemeral" nodes; if the session expires, the lock node will be automatically deleted.
* ZooKeeper-based election - if the local NameNode is healthy, and the ZKFC sees that no other node currently holds the lock znode, it will itself try to acquire the lock. If it succeeds, then it has "won the election", and is responsible for running a failover to make its local NameNode active.

Automatic Namenode failover process in Hadoop:

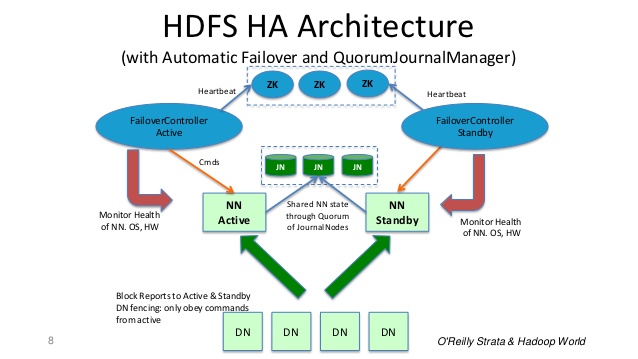
In a typical HA cluster, two separate machines are configured as NameNodes. At any point in time, exactly one of the NameNodes is in an Active state, and the other is in a Standby state. The Active NameNode is responsible for all client operations in the cluster, while the Standby is simply acting as a slave, maintaining enough state to provide a fast failover if necessary.

In order for the Standby Namenode to keep its state synchronized with the Active Namenode, both nodes communicate with a group of separate daemons called JournalNodes (JNs).

When any namespace modification is performed by the Active node, it durably logs a record of the modification to a majority of these JNs. The Standby node is reads these edits from the JNs and apply to its own name space.

In the event of a failover, the Standby will ensure that it has read all of the edits from the JounalNodes before promoting itself to the Active state. This ensures that the namespace state is fully synchronized before a failover occurs.

It is vital for an HA cluster that only one of the NameNodes is Active at a time. ZooKeeper has been used to avoid split brain scenario so that name node state is not getting diverged due to failover.

In Summary: Name Node is Daemon & Failover controller is a Daemon. If Name Node Daemon fails, Failover controller Daemon detects and takes corrective action. Even if entire machine crashes, ZooKeeper server detects it and lock will be expired and other Standby name node will be elected as Active Name node.