HDFS Federation and High availability

In hadoop, The large amount of data is stored in distributed environment. It contains more than 3000 nodes on a single cluster. Name node must gather the data details of those many nodes and also scale up the details for each updating. It is quite difficult to handle many details on a single name node. To avoid this problem we have to federate the name node. Using multiple name nodes scale up process will be easy.

In this system, there are several nodes are connected. If failure occurs in slave system it will not affect the whole cluster. Because the data is replicated into more systems. So if one data node is failure it will manage by the other node which holds replicated data.

But this will not suitable for name node. Because we have one name node. To avoid this problem we have 2 name nodes in hadoop cluster.

Hadoop 2, has 2 name node

Active name node

Stand by name node

Using standby name the high availability is achieved. Because when the name node failure the standby name node is used to recover the details of data node so we can easily rebuild the active name node. Active name node is used to contain the metadata using fs image and edit log. Fs image is used to store the details of data and edit log is used to store the update the details whenever the data is uploaded. Standby name node is mainly for merging the fsimage and editlog to get the meta data. Standby name node doesn’t act as the active name node. It only used for recovery ang merge process.

The data won’t be loss even the particular node gets fail. so the hadoop is achieved high availability and federation concept.

**Handling Failures During Writing a File**

The pipeline is closed and any packets in the ack queue are added to the front of the data queue

The current block on the good DataNodes is given a new identity, which is communicated to the NameNode

The failed DataNode is removed from the pipeline, and a new pipeline is constructed from the two good DataNodes

The remainder of the block’s data is written to the good DataNodes in the pipeline

The NameNode notices that the block is under-replicated, and it arranges for a further replica to be created on another node

As long as dfs.namenode.replication.min replicas are written, the write will succeed

The block will be asynchronously replicated across the cluster until its target replication factor is reached.