HDFS FEDERATION

In HDFS Federation there can be multiple name nodes.

The name nodes contain subdirectories which store the metadata and the block mapping of files in particular sub directories.

The checkpointing is done by the secondary name node.

The sub directories of the name nodes are called as namespace volumes.

The files in a namespace volume is called a block pool.

The root directoy consists of two sub directories, /usr and /share

Here, even if a name node fails, the system does not fail as there are other name nodes.

The mapping can be handled by a name node while the other name node stores the meta data.

The data nodes can contain data blocks from different name nodes.

Let us consider name node 1 consists of namespace volume 1 and name node 2 consists of namespace volume 2.

The name node 1 is mapped to data node 1 and data node 2.

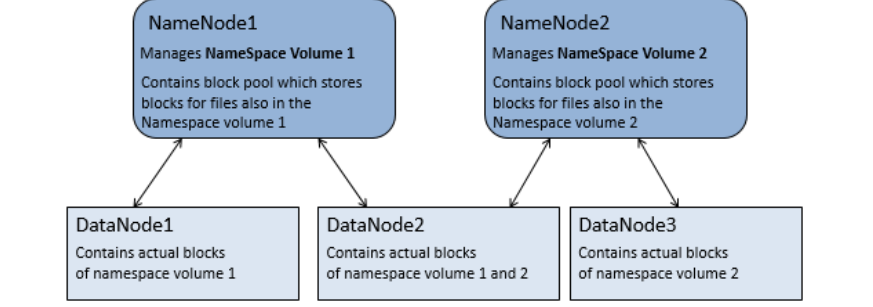
The name node 2 is mapped to data node 2 and data node 3.

Hence, the data node 2 contain blocks from both the name nodes 1 and 2.

Hence the namespace volumes are divided among the name nodes but not among the data nodes.

A drawback is that name node 1 cannot access data node 3 and name node 2 cannot access data node 1.

So if name node 1 fails the data in the data node 1 will be lost.



Hdfs Federation is used in case of increasing the cluster size in HDFS if a single name node is not able to hold the entire mapping of HDFS,so that they are distributed among multiple name nodes.

HDFS Availability

In the HDFS availability there a two name nodes, one is the active name node and the other is a standby name node.

The standby name node takes the control if the active name node fails.

Here, the name node is capable of storing the entire mapping of the HDFS.

All the data nodes are connected to both the active and the standby name nodes.

Here the check pointing is done by the standby name node.

The standby name node also checks the active name node periodically.

Data nodes send the blocked reports to both the name nodes because of block mappings.

The active name node must use its shared storage to share the edit logs.

The edit log is read by the standby name node when it takes the responsibility.

The data nodes notify the name nodes of the blocks it contains.

