**1.Importance of Name Node in Hadoop Cluster**

The Name node is one of the component of an HDFS file system. It is the root tree of all files in the system, and retrieves where the file data is kept. It does not store the data, only contains where the data is located.

Client sends the request to the name node whenever they want to locate a file. When the cluster or name node starts, data nodes join the cluster.

In this process, they send the information about the blocks stored by them to the name node. Name node keeps this information in memory.

Data nodes keep on sending the same information periodically to name node, so that name node remains in sync with latest mapping.

The Name node responds for that requests by returning relevant data node servers where the data is presented.

Name node contains two important files:

Fs image

Edit log

File System Image:

Fs image contains the HDFS metadata which means the data about data node.

Edit log:

Edits contain the changes in the HDFS metadata.

Name node is a single point of failure in HDFS system. When name node fails then automatically system goes down. For those case, HDFS has optional which is called secondary node. Secondary node is in the separate machine. It is used to merging the edit file with fs image and gives the exact view of HDFS structure.

Edits can become very large with time, the Name node start up time can become very lengthy. So, it is good to have them merged periodically.