**NameNode**

It contains all the information of your DataNode,access permissions, locations of your Nodes.

However it does not contain the actual information of your data.

It is single point of failure of HDFS.

**DataNode –**

It contains your actual information or data.It is the slave component in hdfs

**SecondaryNameNode-**

Takes snapshot of fsi image and snapshot of the editlogs on frequent interval. Acts as passive namenode.

**Resource Manager**

* This daemon process runs on master node (may run on the same machine as name node for smaller clusters)
* It is responsible for getting job submitted from client and schedule it on cluster, monitoring running jobs on cluster and allocating proper resources on the slave node
* It communicates with Node Manager daemon process on the slave node to track the resource utilization
* It uses two other processes named *Application Manager*and *Scheduler*for MapReduce task and resource management

**Node Manager**

* This daemon process runs on slave nodes (normally on HDFS Data node machines)
* It is responsible for coordinating with Resource Manager for task scheduling and tracking the resource utilization on the slave node
* It also reports the resource utilization back to the Resource Manager
* It uses other daemon process like Application Master and Container for MapReduce task scheduling and execution on the slave node