1. **Namenode**
   1. NameNode is the centerpiece of  HDFS.
   2. NameNode is also known as the Master
   3. NameNode only stores the metadata of HDFS – the directory tree of all files in the file system, and tracks the files across the cluster.
   4. NameNode does not store the actual data. The data itself is actually stored in the DataNodes.
   5. when the NameNode is down, HDFS is inaccessible and it will get down.
2. **Datanode**
   1. DataNode is responsible for storing the actual data in HDFS.
   2. DataNode is also known as the Slave
   3. NameNode and DataNode are in constant communication.
   4. When a DataNode starts up it announce itself to the NameNode along with the list of blocks it is responsible for.
   5. When a DataNode is down, it does not affect the availability of data. NameNode will arrange for replication for the blocks managed by the DataNode that is not available.
3. **Resource manager**
   1. The ResourceManager (RM) is responsible for tracking the resources.
   2. scheduling applications.
   3. the ResourceManager is the single point of failure in a YARN cluster.
   4. RM is the master that arbitrates all the available cluster resources.
   5. The Resource Manager is the core component in the Hadoop 2.x framework (YARN)
4. **Node manager**
   1. takes care of the individual nodes in a Hadoop cluster.
   2. Keeping up-to date details to the RM like node health, logs.
   3. He will initiate the service and start the service.
   4. For each node there will be an node status, which used to give update about the node to RM.
   5. Nm contains a containers which used to run a program.