Hadoop 2.x follows Master slave architecture same as Hadoop 1.x.Core components in Hadoop 2.x

* 1. HDFS(Hadoop Distributed File system)- Distributed storage system
  2. YARN(Yet Another resource Negotiator) It is also know as “MR V2”.

Hadoop 2.x consists of the following daemons:

* **Name Node**
  1. Name node is the master of HDFS that directs the slave DataNode daemons to perform the low-level I/O tasks same as Hadoop 1.x
  2. Hadoop 2.x is featured with Name Node HA which is referred as HDFS High Availability (HA).
  3. Hadoop 2.x supports two Name Nodes at a time one node is active and another is standby node
  4. Active Name Node handles the client operations in the cluster, StandBy Name Node manages metadata same as Secondary Name Node in Hadoop 1.x
  5. When Active Name Node is down, Standby Name Node takes over and will handle the client operations then after
* **Secondary Name Node**
  1. Hadoop 2.x Non HA mode has same Name Node and Secondary Name Node working same as in Hadoop 1.x architecture
* **DateNode (**Hadoop FileSystem Component):Stores Blocks from files
  1. A DataNode stores the actual data in the HDFS.
  2. A functional filesystem typically have more than one DataNode in the cluster, with data replicated across them. On startup, a DataNode connects to the NameNode; spinning until that service comes up.
  3. It then responds to requests from the NameNode for filesystem operations.

MapReduce2 has replace old daemon process Job Tracker and Task Tracker with YARN components Resource Manager and Node Manager respectively. These two components are responsible for executing distributed data computation jobs in Hadoop 2

* **Resource Manager (YARN Component)**
  1. The function of the Resource Manager is simple: Keeping track of available resources. One per cluster.
  2. It contains two main components: Scheduler and ApplicationsManager.  
     The Scheduler is responsible for allocating resources to the various running applications.  
     The ApplicationsManager is responsible for accepting job-submissions, negotiating the first container for ApplicationMaster and provides the service for restarting the ApplicationMaster container on failure
* **Node Manager (YARN Component)**
  1. This daemon process runs on slave nodes (normally on HDFS Data node machines)
  2. It is responsible for coordinating with Resource Manager for task scheduling and tracking the resource utilization on the slave node
  3. It also reports the resource utilization back to the Resource Manager
  4. It uses other daemon process like Application Master and Container for MapReduce task scheduling and execution on the slave node