**Hadoop 2.x is the high level Architecture**

In hadoop 1.x the major components are HDFS and Map Reduce. For some additional purpose hadoop 2.x has three components.

* HDFS
* YARN
* Map Reduce

**HDFS** stands for Hadoop Distributed File System. It is also called as HDFS V2. It is the part of hadoop 2.x with some features. HDFS is used for storing the files into nodes.

**YARN** stands for Yet Another Resource Negotiator. It is a new component in Hadoop 2.x. It is also known as MR V2.

**Map Reduce** is a distributed processing. It is mainly used for analysing purpose.

All master nodes and slave nodes contains HDFS as well as Map Reduce.

**Master node** has two components:

* Resource Manager which may be YARN or Map Reduce.
* HDFS

Master node’s HDFS component is known as name node. Name node is used to store the meta data.

In hadoop 2.x some more node acts as master node. This particular master nodes has three components.

* Node Manager
* Application Master
* Data Node

These second level master node has one or more their slave nodes.

Each **slave node** contains two components.

* Node Manager
* HDFS

This HDFS is also known as Data Nodes. Data Node is used to store the actual data.

**Resource Manager:**

It is a cluster level component.

Resource Manager is used to getting the request from the client and find the data details which is required by the client . These task is accomplished using second level master node. It sends the request to application master. This application master interacts with slave nodes by sending comments. If required data is available then the details is collected from application master.

Resource manager has divided into two components

* Scheduler
* Application Manager

**Scheduler** is mainly for scheduling process. It does not care about monitoring the application.

**Application Manager** is a application level component. It is for Managing assigned task. It interacts between the scheduler and node manager.

It interacts with scheduler to get the required resource and it also interacts with node manager for getting assigned task’s status.

**Node Manager:**

It is a node level component.

It is used to monitor the each container level which means how much amount of data is available, which data is available.

Each node contains set of containers. Name node also has some container. Container is just like the data slots like blocks. It is mainly used for memory purpose.