**Assignment 3.1**

Hadoop 2.x has the following three Major Components:

* HDFS
* YARN
* MapReduce

These three are also known as Three Pillars of Hadoop 2. Here major key component change is YARN.Hadoop 2.x components follows below architecture to interact each other and to work parallel in a reliable, highly available and fault-tolerant manner.

Client Application

Name Node

Resource Manager

Master Node

Data Node

Application Master

Node Manager

Node Manager

Application Master

Data Node

Nodes

Node 1 Node 2 Node 3 Node 4

Data Node

Node Manager

Node Manager

Node Manager

Node Manager

Data Node

Data Node

Data Node

* All Master Nodes and Slave Nodes contains both MapReduce and HDFS Components.
* One Master Node has two components:
  1. Resource Manager(YARN or MapReduce v2)
  2. HDFS
* One Master Node has two components:
  1. Resource Manager(YARN or MapReduce v2)
  2. HDFS

It’s HDFS component is also knows as NameNode. It’s NameNode is used to store Meta Data.

* Slave Nodes have two components:
  1. Node Manager
  2. HDFS

It’s HDFS component is also knows as Data Node. It’s Data Node component is used to store actual our application Big Data. These nodes does not contain Application Master component.

Hadoop2 Architecture has mainly 2 set of daemons

* **HDFS 2.x Daemons:**  Name Node, Secondary Name Node (not required in HA) and Data Nodes
* **MapReduce 2.x Daemons (YARN):**  Resource Manager, Node Manager

**Resource Manager:**

* Resource Manager is a Per-Cluster Level Component.
* Resource Manager is again divided into two components:
  1. Scheduler
  2. Application Manager
* Resource Manager’s Scheduler is :
  1. Responsible to schedule required resources to Applications (that is Per-Application Master).
  2. It does only scheduling.
  3. It does care about monitoring or tracking of those Applications.
* 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.

**Application Master:**

* Application Master is a per-application level component. It is responsible for:
  1. Managing assigned Application Life cycle.
  2. It interacts with both Resource Manager’s Scheduler and Node Manager
  3. It interacts with Scheduler to acquire required resources.
  4. It interacts with Node Manager to execute assigned tasks and monitor those task’s status.
  5. This is the component where the job actually resides and the Application Master component is responsible for managing each and every Map Reduce job and is concluded once the job completes processing.

**Node Manager:**

* Node Manager is a Per-Node Level component.
* It is responsible for:
  1. Managing the life-cycle of the Container.
  2. Monitoring each Container’s Resources utilization.
* 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.This is the job history server component of [YARN](https://www.dezyre.com/Big-Data-and-Hadoop/19)which will furnish the information about all the completed jobs. The NM keeps a track of all the users’ jobs and their workflow on any particular given node.

**Container:**

* Each Master Node or Slave Node contains set of Containers. In this diagram, Main Node’s Name Node is not showing the Containers. However, it also contains a set of Containers.
* Container is a portion of Memory in HDFS (Either Name Node or Data Node).