Hadoop 2.x has the following three Major Components:

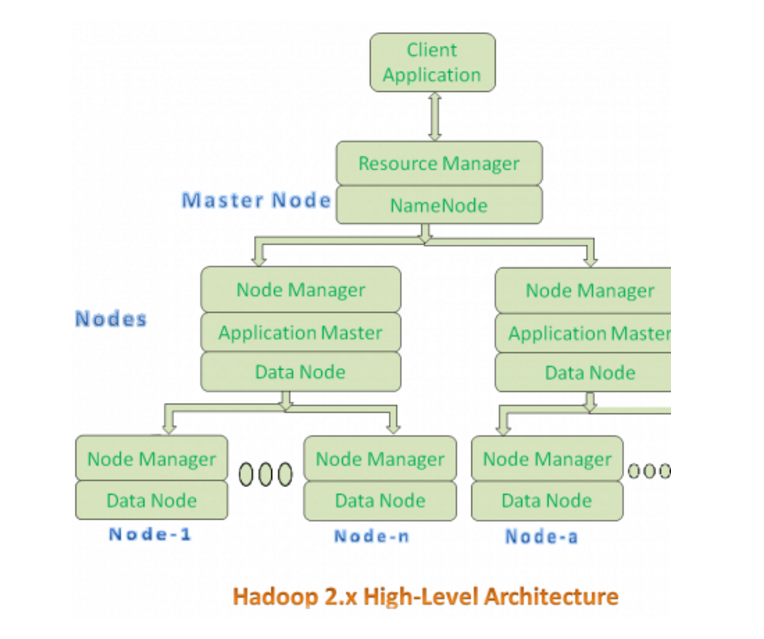
HDFS.

YARN.

MapReduce.

In Hadoop 2.x major key component change is YARN.

**Hadoop 2.x Components High-Level Architecture**



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

It’s HDFS component is also knows as NameNode.

NameNode is used to store Meta Data.

**Resource Manager:**

* 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.

**Application Manager**

* Managing assigned Application Life cycle.
* It interacts with both Resource Manager’s Scheduler and Node Manager
* It interacts with Scheduler to acquire required resources.
* It interacts with Node Manager to execute assigned tasks and monitor those task’s status.

**Node Manager:**

It is responsible for:

* 1. Managing the life-cycle of the Container.
  2. Monitoring each Container’s Resources utilization.

**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).