|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  | Ques 1) List the components of Hadoop 2.x and explain them in detail.  Ans 1) |
|  |  | Hadoop 2.x has the following three Major Components: |
|  |  |  |
|  |  | 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: |
|  |  | Scheduler |
|  |  | Application Manager |
|  |  | Resource Manager’s Scheduler is : |
|  |  | Responsible to scheduler required resources to Applications (that is Per-Application Master). |
|  |  | It does only scheduling. |
|  |  | It does care about monitoring or tracking of those Applications.  This is a relentless YARN service that is designed for receiving and running the applications on Hadoop cluster. |
|  |  |  |
|  |  | Application Master: |
|  |  | 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.  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: |
|  |  | Managing the life-cycle of the Container. |
|  |  | Monitoring each Container’s Resources utilization.  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). |
|  |  | In Hadoop 2.x, Container is similar to Data Slots in Hadoop 1.x. We will see the major differences between |
|  |  | these two Components: Slots Vs Containers in my coming posts. |

Diagram of HDFS 2.X