ASSIGNMENT-8.2

1. Explain the core changes made in Hadoop 2.x

> The main changes made in Hadoop 2.x is that high availability by increasing the name spaces such that the availability of namenode remains high which is the single point failure of Hadoop 1.x

>HDFS federation is the another point where the multiple namespaces are created so that multiple jobs can be performed simultaneously.

>By architecture ,map reduce has gone through a lot of change where the whole work is done by job tracker but here it is classified into resource manager and application manager to perform the functions of job tracker.

So heavy work done by job tracker has been reduced.

>Scheduler and application manager is also introduced that helps in better utilization of resources and for performing tasks.

2. Explain the difference between MapReduce 1 and MapReduce 2 / Yarn

|  |  |
| --- | --- |
| MAPREDUCE 1.X | MAPREDUCE 2.X (YARN) |
| >Job tracker and task tracker are the two components of mapreduce1.x  >Job tracker performs the resource allocation, resource utilization and distribution.  >Task tracker performs the map reduce work which is given by the job tracker  >it can handle only 4000 nodes | >Resource Manager and application manager are the two components of mapreduce2.x.  >Resource manager manages resources and coordinates the scheduler and application master  >Application manager manages various application and informs the performance to the resource manager  >it can handle nearly 10000 nodes |