**1. Explain in brief the architecture of Apache Hadoop Yarn :**

Job Tracker in Hadoop 1.x has 3 major functions:

-- Resource Management

-- Job Scheduling

-- Job Monitoring

> The fundamental idea of YARN is to split up the major functions of the JobTracker i.e. resource management

and job scheduling/monitoring , into separate daemons:

-- a global Resource Manager and per-application Application Master

**Components of YARN :**

1. Global Resource Manager -

- It assigns resources among applications for optimal resource utilization or we can say that Resource Manager

is the ultimate authority that divides resources among all the applications in the system.

- One cluster has one instance of Resource Manager.

2. Node Manager -

- Runs on each node and communicates with Resource Manager about resource usage on the machine.

- It receives requests from resource manager about resource allocation to jobs and maintains life cycle of containers.

3. Application-specific Application Master

- It is the actual instance which does processing.

- It requests Resouce Manager for resources and works with NodeManager to get those resources for task execution.

Application Master could be MapReduce or any other processing framework.

4. Scheduler

- It is plugged with Resource Manager to help in resource allocation. Different schedulers allocate resources using

different algorithms.The Scheduler performs its scheduling function based on the resource requirements of the applications.

5. Container

- It is a set of allocated system resources (CPU Core and Memory). Containers are allocated and managed by NodeManager

and are used by tasks.