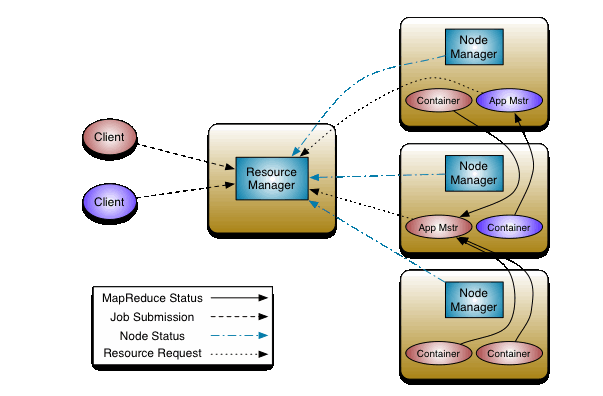
Q1- Explain in brief the architecture of Apache Hadoop Yarn.

Answer:



Components of YARN:

1. Global Resource Manager:

• Assigns resources among applications for optimal resource utilization.

• 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.

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.