YARN

Apache Yarn – "Yet Another Resource Negotiator" is the resource management layer of Hadoop. The Yarn was introduced in Hadoop 2.x. Yarn allows different data processing engines like graph processing, interactive processing, stream processing as well as batch processing to run and process data stored in HDFS (Hadoop Distributed File System). Apart from resource management, Yarn also does job Scheduling. Yarn extends the power of Hadoop to other evolving technologies, so they can take the advantages of HDFS.

Apache Yarn Framework consists of a master daemon known as "Resource Manager", slave daemon called node manager (one per slave node) and Application Master (one per application).

Resource Manager (RM)

It is the master daemon of Yarn. RM manages the global assignments of resources (CPU and memory) among all the applications. It arbitrates system resources between competing applications.

 Resource Manager has two Main components
- Scheduler
- Application manager

a) Scheduler

The scheduler is responsible for allocating the resources to the running application. The scheduler is pure scheduler it means that it performs no monitoring no tracking for the application and even does not guarantees about restarting failed tasks either due to application failure or hardware failures.

b) Application Manager

It manages running Application Masters in the cluster, i.e., it is responsible for starting application masters and for monitoring and restarting them on different nodes in case of failures.

Node Manager (NM)

It is the slave daemon of Yarn. NM is responsible for containers monitoring their resource usage and reporting the same to the Resource Manager. Manage the user process on that machine. Yarn Node Manager also tracks the health of the node on which it is running. The design also allows plugging long-running auxiliary services to the NM; these are application-specific services, specified as part of the configurations and loaded by the NM during startup. A shuffle is a typical auxiliary service by the NMs for MapReduce applications on YARN.