Apache YARN

Functional Area: Cluster Resource Management

Overview:

Apache Hadoop YARN (Yet Another Resource Negotiator) came about as part of a major redesign of Hadoop in 2011 to address some of its limitations. The resource management and scheduling/monitoring functions, both of which had previously been performed by the Hadoop JobTracker, were implemented through the ResourceManager, ApplicationMaster and NodeManager processes within YARN.

YARN has made it possible to use Hadoop for non-MapReduce jobs. It also eliminates the acknowledged performance bottleneck associated with the JobTracker of the previous Hadoop releases, allowing more scalability.

YARN, which has been a sub-project of Apache Hadoop since 2012, represents a key point in Hadoop’s evolution, but it also adds additional complexity to a development environment which was already considered challenging to use.