

[By OnlineInterviewQuestions.com](http://OnlineInterviewQuestions.com)

[Yarn Interview Questions](#)

Are you looking to bag a dream job as a **Hadoop YARN developer**? If yes, then you must buck up your efforts and start preparing for all the competition out there. In this article, you shall come across questions that may be asked during an interview and answers, which shall be most appropriate.

Being an aspirer and looking to get into a big corporate set up demands a lot of sincerity and preparation. Your subject knowledge must be good. All the important sections of Hadoop YARN are covered in the questions given below. These can be helpful for both fresher's and experienced.

Find latest interview questions on Yarn

Hadoop Yarn can provide a great exposure to any developer due to unlimited opportunities associated with it. One can work its way hard to get into the some of the top organizations.

So here is a list of some of the top Yarn Interview questions that you can expect at your interview:

Q1. [What Is Yarn?](#)

Full form of YARN is 'Yet Another Resource Negotiator.' YARN is a great and productive feature rolled out as a part of Hadoop 2.0. YARN is an extensive scale circulated system for running big information applications. YARN gives APIs for requesting and working with Hadoop's bunch assets. These APIs are generally utilized by parts of Hadoop's distributed systems, for example, Map Reduce, Spark, and Tez and much more which are on top of YARN.

Q2. [What Is Resource Manager In YARN?](#)

The Resource Manager is the rack-aware master node in YARN. It is in charge of taking stock of accessible assets and runs several critical services, the most imperative of which is the Scheduler. Resource Manager is the master that referees all the accessible cluster assets and thus assists in managing the dispersed applications running on the YARN system.

Resource Manager has two main parts:

- Scheduler
- Application supervisor

Q3. What Are The Core Changes In Hadoop 2.x?

Numerous changes, the particular single point of failure and Decentralize Job Tracker power to information nodes are the main changes. Whole job tracker design changed. Some of the principal difference between Hadoop 1.x and 2.x provided below:

- One point of failure – Rectified
- Limitations of nodes (4000-to boundless) – Rectified.
- Job Tracker bottleneck – Rectified
- High accessibility – Available
- Support both Interactive, diagram iterative algorithms.
- Allows different applications additionally to coordinate with HDFS.

Q4. Is the yarn Replacement of Map-reduce?

NO, Yarn isn't the replacement of map-reduce. Map Reduce and YARN are unquestionably unique. Map Reduce is Programming Model; YARN is the architecture for the allocation cluster. Hadoop 2 utilizing YARN for asset management. On the other hand, Hadoop supports a programming model which supports parallel handling that we know as Map Reduce

Q5. What Is Apache Hadoop Yarn?

Apache Hadoop YARN is the job scheduling, and resource management innovation in the open source Hadoop distributes preparing structure. One of Apache Hadoop's center segments, YARN is in charge of designating system assets to the different applications running in a Hadoop cluster and scheduling tasks to be executed on various cluster nodes.

Q6. What Are The Additional Benefits Yarn Brings In To Hadoop?

The YARN structure, presented in Hadoop, is intended to share the responsibilities of Map Reduce and deal with the cluster administration task. This enables Map Reduce to execute information preparing and consequently, streamline the procedure. In Hadoop Map Reduce there are different openings for Map and Reduce errands while in YARN there is no fixed space. A similar container can be utilized for Map and Reduce undertakings prompting better usage.

Q7. How Is The Distance Between Two Nodes Defined In Hadoop?

Measuring bandwidth is quite challenging in Hadoop, so the network is signified as a tree in Hadoop. The space between two nodes in the tree plays a crucial part in shaping a Hadoop cluster and is characterized by the system topology and Java interface DNS to Switch Mapping. The distance is equivalent to the sum of the distance to the nearest basic predecessor of both the nodes. The technique gets Distance (Node node1, Node node2) is utilized to ascertain the distance between two nodes with the expectations that the distance from a node to its parent node is dependable

Q8. What Is the Difference between MapReduce 1 And MapReduce 2/yarn?

In Map Reduce 1, Hadoop concentrated all tasks to the Job Tracker. It dispenses assets and scheduling the jobs over the cluster. Whereas in YARN, de-centralized this to facilitate the work pressure at job Tracker. The responsibility of Resource Manager is to allocate assets to the specific nodes and Node administrators schedule the jobs on the application Master. YARN permits parallel execution and Application Master overseeing and execute the activity. This approach can ease numerous Job Tracker issues and enhances to scale up capacity and advance the job execution. Moreover, YARN can permit to make numerous applications to scale up on the disseminated condition.

Q9. What Are The Scheduling Policies Available In Yarn?

The YARN design has pluggable scheduling policies that rely upon the application's prerequisites and the utilization case characterized for the running application. You can discover the YARN scheduling confirmations in the yarn-site.xml file. You can also locate the running application scheduling data in the Resource Manager UI.

As there is three kind of scheduling policies that the YARN scheduler follows:

- FIFO(First In First Out) scheduler
- Capacity Scheduler
- Fair scheduler

Q10. What Is Difference between Hadoop Name node Federation, NFS And Journal Node?

Here are some of the considerable differences:

- HDFS is considered as a write-once file system so a client can't update the files once they exist it is possible that they can read or write to it. But, under specific situations in the enterprise condition like uploading, downloading, file browsing or information streaming – it isn't conceivable to accomplish such an excess of utilizing the standard HDFS. This is the place an appropriated file framework protocol Network File System (NFS) is utilized. NFS enables access to files on remote machines only like how the

local file system is pervaded to by applications.

- Name node is the core of the HDFS file framework that keeps up the metadata and tracks where the record data is kept over the Hadoop group.
- Reserve Nodes and Active Nodes communicate with a group of lightweight nodes to keep their state synchronized. These are called Journal Nodes

Please Visit OnlineInterviewquestions.com to download more pdfs