

Hadoop admin interview questions

Which operating system(s) are supported for production Hadoop deployment?

The main supported operating system is Linux. However, with some additional software [Hadoop](#) can be deployed on Windows.

What is the role of the namenode?

The namenode is the “brain” of the [Hadoop cluster](#) and responsible for managing the distribution blocks on the system based on the replication policy. The namenode also supplies the specific addresses for the data based on the client requests.

What happen on the namenode when a client tries to read a data file?

The namenode will look up the information about file in the edit file and then retrieve the remaining information from filesystem memory snapshot. Since the namenode needs to support a large number of the clients, the primary namenode will only send information back for the data location. The datanode itself is responsible for the retrieval.

What are the hardware requirements for a Hadoop cluster (primary and secondary namenodes and datanodes)?

There are no requirements for datanodes. However, the namenodes require a specified amount of RAM to store filesystem image in memory. Based on the design of the primary namenode and secondary namenode, entire filesystem information will be stored in memory. Therefore, both namenodes need to have enough memory to contain the entire filesystem image.

What mode(s) can Hadoop code be run in?

[Hadoop](#) can be deployed in stand alone mode, pseudo-distributed mode or fully-distributed mode. Hadoop was specifically designed to be deployed on multi-node cluster. However, it also can be deployed on single machine and as a single process for testing purposes

How would an Hadoop administrator deploy various components of Hadoop in production?

Deploy namenode and jobtracker on the master node, and deploy datanodes and tasktrackers on multiple slave nodes. There is a need for only one namenode and jobtracker on the system. The number of datanodes depends on the available hardware

What is the best practice to deploy the secondary namenode

Deploy secondary namenode on a separate standalone machine. The secondary namenode needs to be deployed on a separate machine. It will not interfere with primary namenode operations in this way. The secondary namenode must have the same memory requirements as the main namenode.

Is there a standard procedure to deploy Hadoop?

No, there are some differences between various distributions. However, they all require that Hadoop jars be installed on the machine. There are some common requirements for all [Hadoop](#) distributions but the specific procedures will be different for different vendors since they all have some degree of proprietary software

What is the role of the secondary namenode?

Secondary namenode performs CPU intensive operation of combining edit logs and current filesystem snapshots. The secondary namenode was separated out as a process due to having CPU intensive operations and additional requirements for metadata back-up

What are the side effects of not running a secondary name node?

The cluster performance will degrade over time since edit log will grow bigger and bigger. If the secondary namenode is not running at all, the edit log will grow significantly and it will slow the system down. Also, the system will go into safemode for an extended time since the namenode needs to combine the edit log and the current filesystem checkpoint image.

What happen if a datanode loses network connection for a few minutes?

The namenode will detect that a datanode is not responsive and will start replication of the data from remaining replicas. When datanode comes back online, the extra replicas will be The replication factor is actively maintained by the namenode. The namenode monitors the status of all datanodes and keeps track which blocks are located on that node. The moment the datanode is not available it will trigger replication of the data from the existing replicas. However, if the datanode comes back up, overreplicated data will be deleted. Note: the data might be deleted from the original datanode.

What happen if one of the datanodes has much slower CPU?

The task execution will be as fast as the slowest worker. However, if speculative execution is enabled, the slowest worker will not have such big impact Hadoop was specifically designed to work with commodity hardware. The speculative execution helps to offset the slow workers. The multiple instances of the same task will be created and job tracker will take the first result into consideration and the second instance of the task will be killed.

What is speculative execution?

If speculative execution is enabled, the job tracker will issue multiple instances of the same task on multiple nodes and it will take the result of the task that finished first. The other instances of the task will be killed.

The speculative execution is used to offset the impact of the slow workers in the cluster. The jobtracker creates multiple instances of the same task and takes the result of the first successful task. The rest of the tasks will be discarded.

How many racks do you need to create an [Hadoop](#) cluster in order to make sure that the cluster operates reliably?

In order to ensure a reliable operation it is recommended to have at least 2 racks with rack placement configured Hadoop has a built-in rack awareness mechanism that allows data distribution between different racks based on the configuration.

Are there any special requirements for namenode?

Yes, the namenode holds information about all files in the system and needs to be extra reliable. The namenode is a single point of failure. It needs to be extra reliable and metadata need to be replicated in multiple places. Note that the community is working on solving the single point of failure issue with the namenode.

If you have a file 128M size and replication factor is set to 3, how many blocks can you find on the cluster that will correspond to that file (assuming the default apache and cloudera configuration)?

6

Based on the configuration settings the file will be divided into multiple blocks according to the default block size of 64M. $128M / 64M = 2$. Each block will be replicated according to replication factor settings (default 3). $2 * 3 = 6$.

What is distributed copy (distcp)?

Distcp is a Hadoop utility for launching MapReduce jobs to copy data. The primary usage is for copying a large amount of data. One of the major challenges in the Hadoop environment is copying data across multiple clusters and distcp will allow multiple datanodes to be leveraged for parallel copying of the data.

What is replication factor?

Replication factor controls how many times each individual block can be replicated –

Data is replicated in the Hadoop cluster based on the replication factor. The high replication factor guarantees data availability in the event of failure.

What daemons run on Master nodes?

NameNode, Secondary NameNode and JobTracker

[Hadoop](#) is comprised of five separate daemons and each of these daemons run in its own JVM. NameNode, Secondary NameNode and JobTracker run on Master nodes. DataNode and TaskTracker run on each Slave nodes.

What is rack awareness?

Rack awareness is the way in which the namenode decides how to place blocks based on the rack definitions. Hadoop will try to minimize the network traffic between datanodes within the same rack and will only contact remote racks if it has to. The namenode is able to control this due to rack awareness

What is the role of the jobtracker in an Hadoop cluster?

The jobtracker is responsible for scheduling tasks on slave nodes, collecting results, retrying failed tasks. The job tracker is the main component of the map-reduce execution. It controls the division of the job into smaller tasks, submits tasks to individual tasktrackers, tracks the progress of the jobs and reports results back to calling code.

How does the Hadoop cluster tolerate datanode failures?

Since Hadoop is designed to run on commodity hardware, the datanode failures are expected. Namenode keeps track of all available datanodes and actively maintains replication factor on all data.

The namenode actively tracks the status of all datanodes and acts immediately if the datanodes become non-responsive. The namenode is the central “brain” of the [HDFS](#) and starts replication of the data the moment a disconnect is detected.

What is the procedure for namenode recovery?

A namenode can be recovered in two ways: starting new namenode from backup metadata or promoting secondary namenode to primary namenode.

The namenode recovery procedure is very important to ensure the reliability of the data. It can be accomplished by starting a new namenode using backup data or by promoting the secondary namenode to primary.

Web-UI shows that half of the datanodes are in decommissioning mode. What does that mean? Is it safe to remove those nodes from the network?

This means that namenode is trying to retrieve data from those datanodes by moving replicas to remaining datanodes. There is a possibility that data can be lost if administrator removes those datanodes before decommissioning finished .

Due to replication strategy it is possible to lose some data due to datanodes removal en masse prior to completing the decommissioning process. Decommissioning refers to namenode trying to retrieve data from datanodes by moving replicas to remaining datanodes.

What does the Hadoop administrator have to do after adding new datanodes to the Hadoop cluster?

Since the new nodes will not have any data on them, the administrator needs to start the balancer to redistribute data evenly between all nodes.

Hadoop cluster will detect new datanodes automatically. However, in order to optimize the cluster performance it is recommended to start rebalancer to redistribute the data between datanodes evenly.

If the Hadoop administrator needs to make a change, which configuration file does he need to change?

Each node in the Hadoop cluster has its own configuration files and the changes need to be made in every file. One of the reasons for this is that configuration can be different for every node.

Map Reduce jobs are failing on a cluster that was just restarted. They worked before restart. What could be wrong?

The cluster is in a safe mode. The administrator needs to wait for namenode to exit the safe mode before restarting the jobs again

This is a very common mistake by Hadoop administrators when there is no secondary namenode on the cluster and the cluster has not been restarted in a long time. The namenode will go into safemode and combine the edit log and current file system timestamp

Map Reduce jobs take too long. What can be done to improve the performance of the cluster?

One of the most common reasons for performance problems on Hadoop cluster is uneven distribution of the tasks. The number of tasks has to match the number of available slots on the cluster

Hadoop is not a hardware aware system. It is the responsibility of the developers and the administrators to make sure that the resource supply and demand match.

How often do you need to reformat the namenode?

Never. The namenode needs to be formatted only once in the beginning. Reformatting of the namenode will lead to loss of the data on entire

The namenode is the only system that needs to be formatted only once. It will create the directory structure for file system metadata and create namespaceID for the entire file system.

After increasing the replication level, I still see that data is under replicated. What could be wrong?

Data replication takes time due to large quantities of data. The [Hadoop administrator](#) should allow sufficient time for data replication

Depending on the data size the data replication will take some time. Hadoop cluster still needs to copy data around and if data size is big enough it is not uncommon that replication will take from a few minutes to a few hours.

Hadoop admin interview question

1. Which operating system(s) are supported for production Hadoop deployment?
2. What is the role of the namenode?
3. What happen on the namenode when a client tries to read a data file?
4. What are the hardware requirements for a Hadoop cluster (primary and secondary namenodes and datanodes)?
5. What mode(s) can Hadoop code be run in?
6. How would an Hadoop administrator deploy various components of Hadoop in production?
7. What is the best practice to deploy the secondary namenode
8. Is there a standard procedure to deploy Hadoop?
9. What is the role of the secondary namenode?
10. What are the side effects of not running a secondary name node?
11. What happen if a datanode loses network connection for a few minutes?
12. What happen if one of the datanodes has much slower CPU?
13. What is speculative execution?
14. After increasing the replication level, I still see that data is under replicated. What could be wrong?
15. How many racks do you need to create an Hadoop cluster in order to make sure that the cluster operates reliably?
16. Are there any special requirements for namenode?
17. If you have a file 128M size and replication factor is set to 3, how many blocks can you find on the cluster that will correspond to that file (assuming the default apache and cloudera configuration)?
18. What is distributed copy (distcp)?
19. What is replication factor?
20. What daemons run on Master nodes?
21. What is rack awareness?
22. What is the role of the jobtracker in an Hadoop cluster?
23. How does the Hadoop cluster tolerate datanode failures?
24. What is the procedure for namenode recovery?
25. Web-UI shows that half of the datanodes are in decommissioning mode. What does that mean? Is it safe to remove those nodes from the network?
26. What does the Hadoop administrator have to do after adding new datanodes to the Hadoop cluster?
27. If the Hadoop administrator needs to make a change, which configuration file does he need to change?
28. Map Reduce jobs are failing on a cluster that was just restarted. They worked before restart. What could be wrong?
29. Map Reduce jobs take too long. What can be done to improve the performance of the cluster?
30. How often do you need to reformat the namenode?