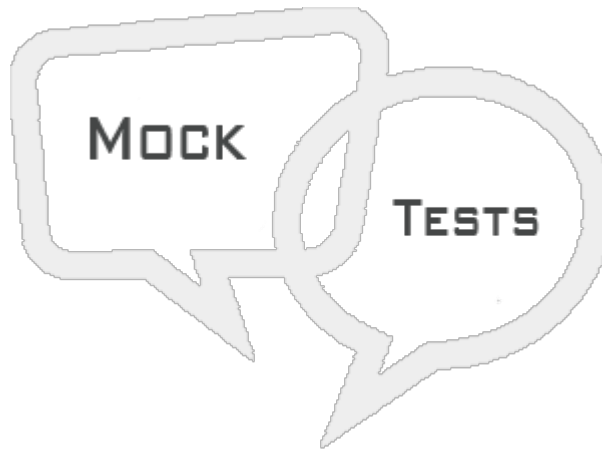


HADOOP MOCK TEST

<http://www.tutorialspoint.com>

Copyright © tutorialspoint.com

This section presents you various set of Mock Tests related to **Hadoop Framework**. You can download these sample mock tests at your local machine and solve offline at your convenience. Every mock test is supplied with a mock test key to let you verify the final score and grade yourself.



HADOOP MOCK TEST II

Q 1 - HDFS can be accessed over HTTP using

- A - viewfs URI scheme
- B - webhdfs URI scheme
- C - wasb URI scheme
- D - HDFS ftp

Q 2 - What is are true about HDFS?

- A - HDFS filesystem can be mounted on a local client's Filesystem using NFS.
- B - HDFS filesystem can never be mounted on a local client's Filesystem.
- C - You can edit a existing record in HDFS file which is already mounted using NFS.
- D - You cannot append to a HDFS file which is mounted using NFS.

Q 3 - The client reading the data from HDFS filesystem in Hadoop

- A - gets the data from the namenode
- B - gets the block location from the datanode
- C - gets only the block locations form the namenode
- D - gets both the data and block location from the namenode

Q 4 - Which scenario demands highest bandwidth for data transfer between nodes in Hadoop?

- A - Different nodes on the same rack

- B - Nodes on different racks in the same data center.
- C - Nodes in different data centers
- D - Data on the same node.

Q 5 - The current block location of HDFS where data is being written to,

- A - is visible to the client requesting for it.
- B - Block locations are never visible to client requests.
- C - May or may not be visible to the reader.
- D - becomes visible only after the buffered data is committed.

Q 6 - Which of this is not a scheduler options available with YARN?

- A - Optimal Scheduler
- B - FIFO scheduler
- C - Capacity scheduler
- D - Fair scheduler

Q 7 - Which of the following is not a Hadoop operation mode?

- A - Pseudo distributed mode
- B - Globally distributed mode
- C - Stand alone mode
- D - Fully-Distributed mode

Q 8 - The difference between standalone and pseudo-distributed mode is

- A - Stand alone cannot use map reduce
- B - Stand alone has a single java process running in it.
- C - Pseudo distributed mode does not use HDFS
- D - Pseudo distributed mode needs two or more physical machines.

Q 9 - The hadoop frame work is written in

- A - C++
- B - Python
- C - Java
- D - GO

Q 10 - The hdfs command to create the copy of a file from a local system is

- A - CopyFromLocal
- B - copyfromlocal
- C - CopyLocal
- D - copyFromLocal

Q 11 - The hadfs command put is used to

- A - Copy files from local file system to HDFS.
- B - Copy files or directories from local file system to HDFS.
- C - Copy files from from HDFS to local filesystem.
- D - Copy files or directories from HDFS to local filesystem.

Q 12 - Underreplication in HDFS means-

- A - No replication is happening in the data nodes.
- B - Replication process is very slow in the data nodes.
- C - The frequency of replication in data nodes is very low.
- D - The number of replicated copies is less than as specified by the replication factor.

Q 13 - When the namenode finds that some blocks are over replicated, it

- A - Stops the replication job in the entire hdfs file system.
- B - It slows down the replication process for those blocks
- C - It deletes the extra blocks.
- D - It leaves the extra blocks as it is.

Q 14 - Which of the below property gets configured on core-site.xml ?

- A - Replication factor
- B - Directory names to store hdfs files.
- C - Host and port where MapReduce task runs.
- D - Java Environment variables.

Q 15 - Which of the below property gets configured on hdfs-site.xml ?

- A - Replication factor
- B - Directory names to store hdfs files.
- C - Host and port where MapReduce task runs.
- D - Java Environment variables.

Q 16 - Which of the below property gets configured on mapred-site.xml ?

- A - Replication factor
- B - Directory names to store hdfs files.
- C - Host and port where MapReduce task runs.
- D - Java Environment variables.

Q 17 - Which of the below property gets configured on hadoop-env.sh?

- A - Replication factor
- B - Directory names to store hdfs files
- C - Host and port where MapReduce task runs
- D - Java Environment variables.

Q 18 - The command to check if Hadoop is up and running is –

- A - Jsp
- B - Jps
- C - Hadoop fs -test
- D - None

Q 19 - The information mapping data blocks with their corresponding files is stored in

- A - Data node
- B - Job Tracker
- C - Task Tracker
- D - Namenode

Q 20 - The file in Namenode which stores the information mapping the data block location with file name is –

- A - dfsimage
- B - nameimage
- C - fsimage
- D - image

Q 21 - The namenode knows that the datanode is active using a mechanism known as

- A - heartbeats
- B - datapulse
- C - h-signal

D - Active-pulse

Q 22 - The nature of hardware for the namenode should be

- A - Superior than commodity grade
- B - Commodity grade
- C - Does not matter
- D - Just have more Ram than each of the data nodes

Q 23 - In Hadoop, Snappy and LZO are examples of

- A - Mechanisms of file transport between data nodes
- B - Mechanisms of data compression
- C - Mechanisms of data Replication
- D - Mechanisms of Data synchronization

Q 24 - Which of the below apache system deals with ingesting streaming data to hadoop

- A - Ozie
- B - Kafka
- C - Flume
- D - Hive

Q 25 - The input split used in MapReduce indicates

- A - The average size of the data blocks used as input for the program
- B - The location details of where the first whole record in a block begins and the last whole record in the block ends.
- C - Splitting the input data to a MapReduce program into a size already configured in the mapred-site.xml
- D - None of these

Q 26 - The output of a mapper task is

- A - The Key-value pair of all the records of the dataset.
- B - The Key-value pair of all the records from the input split processed by the mapper
- C - Only the sorted Keys from the input split
- D - The number of rows processed by the mapper task.

Q 27 - The role of a Journal node is to

- A - Report the location of the blocks in a data node
- B - Report the edit log information of the blocks in the data node.
- C - Report the Schedules when the jobs are going to run
- D - Report the activity of various components handled by resource manager

Q 28 - The Zookeeper

- A - Detects the failure of the namenode and elects a new namenode.
- B - Detects the failure of datanodes and elects a new datanode.
- C - Prevents the hardware from overheating by shutting them down.
- D - Maintains a list of all the components IP address of the Hadoop cluster.

Q 29 - If the IP address or hostname of a datanode changes

- A - The namenode updates the mapping between file name and block name
- B - The namenode need not update mapping between file name and block name
- C - The data in that data node is lost forever
- D - There namenode has to be restarted

Q 30 - When a client contacts the namenode for accessing a file, the namenode responds with

- A - Size of the file requested.
- B - Block ID of the file requested.
- C - Block ID and hostname of any one of the data nodes containing that block.
- D - Block ID and hostname of all the data nodes containing that block.

Q 31 - HDFS stands for

- A - Highly distributed file system.
- B - Hadoop directed file system
- C - Highly distributed file shell
- D - Hadoop distributed file system.

Q 32 - The Hadoop tool used for uniformly spreading the data across the data nodes is named –

- A - Scheduler
- B - Balancer
- C - Spreader

D - Reporter

Q 33 - In the secondary namenode the amount of memory needed is

A - Similar to that of primary node

B - Should be at least half of the primary node

C - Must be double of that of primary node

D - Depends only on the number of data nodes it is going to handle

ANSWER SHEET

Question Number	Answer Key
-----------------	------------

1	B
2	A
3	C
4	C
5	D
6	A
7	B
8	B
9	C
10	D
11	B
12	D
13	C
14	B
15	A
16	C
17	D
18	B
19	D
20	C
21	A
22	A
23	B
24	C

25	B
26	B
27	B
28	A
29	B
30	D
31	D
32	B
33	A