CCA-500 Q&As

Cloudera Certified Administrator for Apache Hadoop (CCAH)

Pass Cloudera CCA-500 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass4itsure.com/cca-500.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Cloudera Official Exam Center

QUESTION 1

Your cluster has the following characteristics:

A rack aware topology is configured and on

Replication is set to 3

Cluster block size is set to 64MB

Which describes the file read process when a client application connects into the cluster and requests a 50MB file?

A. The client queries the NameNode for the locations of the block, and reads all three copies. The first copy to complete transfer to the client is the one the client reads as part of hadoop\\'s speculative execution framework.

B. The client queries the NameNode for the locations of the block, and reads from the first location in the list it receives.

C. The client queries the NameNode for the locations of the block, and reads from a random location in the list it receives to eliminate network I/O loads by balancing which nodes it retrieves data from any given time.

D. The client queries the NameNode which retrieves the block from the nearest DataNode to the client then passes that block back to the client.

Correct Answer: B

QUESTION 2

For each YARN job, the Hadoop framework generates task log file. Where are Hadoop task log files stored?

A. Cached by the NodeManager managing the job containers, then written to a log directory on the NameNode

B. Cached in the YARN container running the task, then copied into HDFS on job completion

C. In HDFS, in the directory of the user who generates the job

D. On the local disk of the slave mode running the task

Correct Answer: D

QUESTION 3

You are planning a Hadoop cluster and considering implementing 10 Gigabit Ethernet as the network fabric. Which workloads benefit the most from faster network fabric?

A. When your workload generates a large amount of output data, significantly larger than the amount of intermediate data

B. When your workload consumes a large amount of input data, relative to the entire capacity if HDFS

C. When your workload consists of processor-intensive tasks

D. When your workload generates a large amount of intermediate data, on the order of the input data itself

Correct Answer: A

QUESTION 4

Which YARN daemon or service negotiations map and reduce Containers from the Scheduler, tracking their status and monitoring progress?

A. NodeManager

B. ApplicationMaster

C. ApplicationManager

D. ResourceManager

Correct Answer: B

QUESTION 5

You have installed a cluster HDFS and MapReduce version 2 (MRv2) on YARN. You have no dfs.hosts entry(ies) in your hdfs-site.xml configuration file. You configure a new worker node by setting fs.default.name in its configuration files to point to the NameNode on your cluster, and you start the DataNode daemon on that worker node. What do you have to do on the cluster to allow the worker node to join, and start sorting HDFS blocks?

A. Without creating a dfs.hosts file or making any entries, run the commands hadoop.dfsadminrefreshModes on the NameNode

B. Restart the NameNode

C. Creating a dfs.hosts file on the NameNode, add the worker Node\\'s name to it, then issue the command hadoop dfsadmin refresh Nodes = on the Namenode

D. Nothing; the worker node will automatically join the cluster when NameNode daemon is started

Correct Answer: A

QUESTION 6

During the execution of a MapReduce v2 (MRv2) job on YARN, where does the Mapper place the intermediate data of each Map Task?

A. The Mapper stores the intermediate data on the node running the Job\\'s ApplicationMaster so that it is available to YARN ShuffleService before the data is presented to the Reducer

B. The Mapper stores the intermediate data in HDFS on the node where the Map tasks ran in the HDFS /usercache/and(user)/apache/application_and(appid) directory for the user who ran the job

C. The Mapper transfers the intermediate data immediately to the reducers as it is generated by the Map Task

D. YARN holds the intermediate data in the NodeManager\\'s memory (a container) until it is transferred to the Reducer

E. The Mapper stores the intermediate data on the underlying filesystem of the local disk in the directories yarn.nodemanager.locak-DIFS

Correct Answer: E

QUESTION 7

Assuming a cluster running HDFS, MapReduce version 2 (MRv2) on YARN with all settings at their default, what do you need to do when adding a new slave node to cluster?

A. Nothing, other than ensuring that the DNS (or/etc/hosts files on all machines) contains any entry for the new node.

B. Restart the NameNode and ResourceManager daemons and resubmit any running jobs.

C. Add a new entry to /etc/nodes on the NameNode host.

D. Restart the NameNode of dfs.number.of.nodes in hdfs-site.xml

Correct Answer: A

QUESTION 8

What does CDH packaging do on install to facilitate Kerberos security setup?

A. Automatically configures permissions for log files at and MAPRED_LOG_DIR/userlogs

B. Creates users for hdfs and mapreduce to facilitate role assignment

C. Creates directories for temp, hdfs, and mapreduce with the correct permissions

D. Creates a set of pre-configured Kerberos keytab files and their permissions

E. Creates and configures your kdc with default cluster values

Correct Answer: B

QUESTION 9

Your cluster\\'s mapred-start.xml includes the following parameters

mapreduce.map.memory.mb 4096 mapreduce.reduce.memory.mb 8192

And any cluster\\'s yarn-site.xml includes the following parameters

yarn.nodemanager.vmen-pmen-ration 2.1

What is the maximum amount of virtual memory allocated for each map task before YARN will kill its Container?

A. 4 GB

B. 17.2 GB

C. 8.9 GB

D. 8.2 GB

E. 24.6 GB

Correct Answer: D

QUESTION 10

You are migrating a cluster from MApReduce version 1 (MRv1) to MapReduce version 2 (MRv2) on YARN. You want to maintain your MRv1 TaskTracker slot capacities when you migrate. What should you do/

A. Configure yarn.applicationmaster.resource.memory-mb and yarn.applicationmaster.resource.cpu-vcores so that ApplicationMaster container allocations match the capacity you require.

B. You don\\'t need to configure or balance these properties in YARN as YARN dynamically balances resource management capabilities on your cluster

C. Configure mapred.tasktracker.map.tasks.maximum and mapred.tasktracker.reduce.tasks.maximum ub yarn-site.xml to match your cluster\\'s capacity set by the yarn-scheduler.minimum-allocation

D. Configure yarn.nodemanager.resource.memory-mb and yarn.nodemanager.resource.cpu- vcores to match the capacity you require under YARN for each NodeManager

Correct Answer: D