

Cloudera Certified Administrator for Apache Hadoop Upgrade Exam -- CCA-505

Cloudera CCA-505 Dumps Available Here at:

<https://www.certification-questions.com/cloudera-exam/cca-505-dumps.html>

Enrolling now you will get access to 45 questions in a unique set of CCA-505 dumps

Question 1

You have installed a cluster running HDFS and MapReduce version 2 (MRv2) on YARN. You have no `afs.hosts` entry()ies in your `hdfs-alte.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 storing HDFS blocks?

Options:

- A. Nothing; the worker node will automatically join the cluster when the DataNode daemon is started.
- B. Without creating a `dfs.hosts` file or making any entries, run the command `hadoop dfsadmin -refreshHadoop` on the NameNode
- C. Create a `dfs.hosts` file on the NameNode, add the worker node's name to it, then issue the command `hadoop dfsadmin -refreshNodes` on the NameNode
- D. Restart the NameNode

Answer: D

Question 2

Given:

[user1@host1 ~] & yarn application -list
Total Applications: 3

Application-Id	Application-Name	Application-Type	User	Queue	State
Application_1374638600275_0109	Sleep job	MAPREDUCE	user1	default	Killed
Application_1374638600275_0121	sleep job	MAPREDUCE	user1	default	SUCCESS
Application_1374638600275_0020	sleep job	MAPREDUCE	user1	default	SUCCESS

You want to clean up this list by removing jobs where the state is KILLED. What command you enter?

Options:

- A. Yarn application -kill application_1374638600275_0109
- B. Yarn radmin -refreshQueue
- C. Yarn application -refreshJobHistory
- D. Yarn radmin -kill application_1374638600275_0109

Answer: A

Explanation:

Reference: http://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.1-latest/bk_using-apache-hadoop/content/common_mrv2_commands.html

Question 3

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 a cluster?

Options:

- A. Nothing, other than ensuring that DNS (or /etc/hosts files on all machines) contains an entry for the new node.
- B. Restart the NameNode and ResourceManager daemons and resubmit any running jobs
- C. Increase the value of dfs.number.of.nodes in hdfs-site.xml
- D. Add a new entry to /etc/nodes on the NameNode host.
- E. Restart the NameNode daemon.

Answer: B

Question 4

You have a 20 node Hadoop cluster, with 18 slave nodes and 2 master nodes running HDFS High Availability (HA). You want to minimize the chance of data loss in you cluster. What should you do?

Options:

A. Add another master node to increase the number of nodes running the JournalNode which increases

the number of machines available to HA to create a quorum

B. Configure the cluster's disk drives with an appropriate fault tolerant RAID level

C. Run the ResourceManager on a different master from the NameNode in the order to load share HDFS

metadata processing

D. Run a Secondary NameNode on a different master from the NameNode in order to load provide automatic recovery from a NameNode failure

E. Set an HDFS replication factor that provides data redundancy, protecting against failure

Answer: C

Question 5

You decide to create a cluster which runs HDFS in High Availability mode with automatic failover, using Quorum-based Storage. What is the purpose of ZooKeeper in such a configuration?

Options:

A. It manages the Edits file, which is a log changes to the HDFS filesystem.

B. It monitors an NFS mount point and reports if the mount point disappears

C. It both keeps track of which NameNode is Active at any given time, and manages the Edits file, which is

a log of changes to the HDFS filesystem

D. It only keeps track of which NameNode is Active at any given time

E. Clients connect to ZoneKeeper to determine which NameNode is Active

Answer: D

Explanation:

Reference: <http://www.cloudera.com/content/cloudera-content/cloudera-docs/CDH4/latest/PDF/CDH4-High-Availability-Guide.pdf> (page 15)

Question 6

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

Options:

A. The Mapper stores the intermediate data on the node running the job's ApplicationMaster so that is

available to YARN's 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/&[user]spcache/application_&(appid) directory for the user who ran the job

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

the Reducers

D. The Mapper stores the intermediate data on the underlying filesystem of the local disk in the directories

yarn.nodemanager.local-dirs

E. The Mapper transfers the intermediate data immediately to the Reducers as it generated by the Map

task

Answer: E

Question 7

Which Yarn daemon or service monitors a Container's per-application resource usage (e.g, memory, CPU)?

Options:

A. NodeManager

- B. ApplicationMaster
- C. ApplicationManagerService
- D. ResourceManager

Answer: A

Explanation:

Reference: http://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.0.0.2/bk_using-apache-hadoop/content/ch_using-apache-hadoop-4.html (4th para)

Question 8

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

Options:

- A. When your workload generates a large amount of output data, significantly larger than amount of intermediate data
- B. When your workload generates a large amount of intermediate data, on the order of the input data itself
- C. When workload consumers a large amount of input data, relative to the entire capacity of HDFS
- D. When your workload consists of processor-intensive tasks

Answer: B

Question 9

For each YARN Job, the Hadoop framework generates task log files. Where are Hadoop's files stored?

Options:

- A. In HDFS, In the directory of the user who generates the job
- B. On the local disk of the slave node running the task
- C. Cached In the YARN container running the task, then copied into HDFS on job completion
- D. Cached by the NodeManager managing the job containers, then written to a log directory on the

NameNode

Answer: B

Explanation:

Reference: <http://hortonworks.com/blog/simplifying-user-logs-management-and-access-in-yarn/>

Question 10

You are the hadoop fs -put command to add a file "sales.txt" to HDFS. This file is small enough that it fits into a single block, which is replicated to three nodes in your cluster (with a replication factor of 3). One of the nodes holding this file (a single block) fails. How will the cluster handle the replication of this file in this situation/

Options:

A. The cluster will re-replicate the file the next time the system administrator reboots the NameNode

daemon (as long as the file's replication doesn't fall two)

B. This file will be immediately re-replicated and all other HDFS operations on the cluster will halt until the

cluster's replication values are restored

C. The file will remain under-replicated until the administrator brings that nodes back online

D. The file will be re-replicated automatically after the NameNode determines it is under replicated based

on the block reports it receives from the DataNodes

Answer: B

Would you like to see more? Don't miss our CCA-505 PDF file at:

<https://www.certification-questions.com/cloudera-pdf/cca-505-pdf.html>