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.geekcert.com/cca-500.html

100% Passing Guarantee
100% Money Back Assurance

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

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

You have a Hadoop cluster HDFS, and a gateway machine external to the cluster from which clients submit jobs. What do you need to do in order to run Impala on the cluster and submit jobs from the command line of the gateway machine?

A. Install the impalad daemon statestored daemon, and daemon on each machine in the cluster, and the impala shell on your gateway machine

B. Install the impalad daemon, the statestored daemon, the catalogd daemon, and the impala shell on your gateway machine

C. Install the impalad daemon and the impala shell on your gateway machine, and the statestored daemon and catalogd daemon on one of the nodes in the cluster

D. Install the impalad daemon on each machine in the cluster, the statestored daemon and catalogd daemon on one machine in the cluster, and the impala shell on your gateway machine

E. Install the impalad daemon, statestored daemon, and catalogd daemon on each machine in the cluster and on the gateway node

Correct Answer: D

QUESTION 2

You want to node to only swap Hadoop daemon data from RAM to disk when absolutely necessary. What should you do?

A. Delete the /dev/vmswap file on the node

B. Delete the /etc/swap file on the node

C. Set the ram.swap parameter to 0 in core-site.xml

D. Set vm.swapfile file on the node

E. Delete the /swapfile file on the node

Correct Answer: D

QUESTION 3

Assuming you\\'re not running HDFS Federation, what is the maximum number of NameNode daemons you should run on your cluster in order to avoid a "split-brain" scenario with your NameNode when running HDFS High Availability (HA) using Quorum-based storage?

A. Two active NameNodes and two Standby NameNodes

B. One active NameNode and one Standby NameNode

C. Two active NameNodes and on Standby NameNode

D. Unlimited. HDFS High Availability (HA) is designed to overcome limitations on the number of NameNodes you can deploy

Correct Answer: B

QUESTION 4

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 5

You want to understand more about how users browse your public website. For example, you want to know which pages they visit prior to placing an order. You have a server farm of 200 web servers hosting your website. Which is the most efficient process to gather these web server across logs into your Hadoop cluster analysis?

A. Sample the web server logs web servers and copy them into HDFS using curl

B. Ingest the server web logs into HDFS using Flume

C. Channel these clickstreams into Hadoop using Hadoop Streaming

D. Import all user clicks from your OLTP databases into Hadoop using Sqoop

E. Write a MapReeeduce job with the web servers for mappers and the Hadoop cluster nodes for reducers

Correct Answer: B