



CCA-500^{Q&As}

Cloudera Certified Administrator for Apache Hadoop (CCA-H)

Pass Cloudera CCA-500 Exam with 100% Guarantee

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

<https://www.passapply.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 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

QUESTION 2

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.local-DIFS

Correct Answer: E

QUESTION 3

Your Hadoop cluster is configuring with HDFS and MapReduce version 2 (MRv2) on YARN. Can you configure a worker node to run a NodeManager daemon but not a DataNode daemon and still have a functional cluster?

- A. Yes. The daemon will receive data from the NameNode to run Map tasks
- B. Yes. The daemon will get data from another (non-local) DataNode to run Map tasks
- C. Yes. The daemon will receive Map tasks only
- D. Yes. The daemon will receive Reducer tasks only



Correct Answer: B

QUESTION 4

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.vmem-pmem-ratio 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 5

Your Hadoop cluster contains nodes in three racks. You have not configured the dfs.hosts property in the NameNode's configuration file. What results?

- A. The NameNode will update the dfs.hosts property to include machines running the DataNode daemon on the next NameNode reboot or with the command dfsadmin refreshNodes
- B. No new nodes can be added to the cluster until you specify them in the dfs.hosts file
- C. Any machine running the DataNode daemon can immediately join the cluster
- D. Presented with a blank dfs.hosts property, the NameNode will permit DataNodes specified in mapred.hosts to join the cluster

Correct Answer: C

[CCA-500 PDF Dumps](#)

[CCA-500 Practice Test](#)

[CCA-500 Braindumps](#)