



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.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

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



**QUESTION 1**

You have just run a MapReduce job to filter user messages to only those of a selected geographical region. The output for this job is in a directory named westUsers, located just below your home directory in HDFS. Which command gathers these into a single file on your local file system?

- A. Hadoop fs getmerge R westUsers.txt
- B. Hadoop fs getemerge westUsers westUsers.txt
- C. Hadoop fs cp westUsers/* westUsers.txt
- D. Hadoop fs get westUsers westUsers.txt

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

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 4**

You have a cluster running with the fair Scheduler enabled. There are currently no jobs running on the cluster, and you submit a job A, so that only job A is running on the cluster. A while later, you submit Job B. now Job A and Job B are running on the cluster at the same time. How will the Fair Scheduler handle these two jobs? (Choose two)

- A. When Job B gets submitted, it will get assigned tasks, while job A continues to run with fewer tasks.
- B. When Job B gets submitted, Job A has to finish first, before job B can get scheduled.
- C. When Job A gets submitted, it doesn't consume all the task slots.
- D. When Job A gets submitted, it consumes all the task slots.

Correct Answer: B

QUESTION 5

You observed that the number of spilled records from Map tasks far exceeds the number of map output records. Your child heap size is 1GB and your io.sort.mb value is set to 1000MB. How would you tune your io.sort.mb value to achieve maximum memory to disk I/O ratio?

- A. For a 1GB child heap size an io.sort.mb of 128 MB will always maximize memory to disk I/O
- B. Increase the io.sort.mb to 1GB
- C. Decrease the io.sort.mb value to 0
- D. Tune the io.sort.mb value until you observe that the number of spilled records equals (or is as close to equals) the number of map output records.

Correct Answer: D

[Latest CCA-500 Dumps](#)

[CCA-500 Practice Test](#)

[CCA-500 Study Guide](#)