# CCA-505 Q&As

Cloudera Certified Administrator for Apache Hadoop (CCAH) CDH5 Upgrade Exam

## Pass Cloudera CCA-505 Exam with 100% Guarantee

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

**https://www.passapply.com/cca-505.html**

100% Passing Guarantee
100% Money Back Assurance

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

**QUESTION 1**

A slave node in your cluster has four 2TB hard drives installed (4 x 2TB). The DataNode is configured to store HDFS blocks on the disks. You set the value of the dfs.datanode.du.reserved parameter to 100GB. How does this alter HDFS block storage?

A. A maximum of 100 GB on each hard drive may be used to store HDFS blocks

B. All hard drives may be used to store HDFS blocks as long as atleast 100 GB in total is available on the node

C. 100 GB on each hard drive may not be used to store HDFS blocks

D. 25 GB on each hard drive may not be used to store HDFS blocks

Correct Answer: B

**QUESTION 2**

You observe 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 100 MB. How would you tune your io.sort.mb value to achieve maximum memory to disk I/O ratio?

A. Decrease the io.sort.mb value to 0

B. Increase the io.sort.mb to 1GB

C. For 1GB child heap size an io.sort.mb of 128 MB will always maximize memory to disk I/O

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

**QUESTION 3**

You have converted your Hadoop cluster from a MapReduce 1 (MRv1) architecture to a MapReduce 2 (MRv2) on YARN architecture. Your developers are accustomed to specifying map and reduce tasks (resource allocation) tasks when they run jobs. A developer wants to know how specify to reduce tasks when a specific job runs. Which method should you tell that developer to implement?

A. Developers specify reduce tasks in the exact same way for both MapReduce version 1 (MRv1) and MapReduce version 2 (MRv2) on YARN. Thus, executing p mapreduce.job.reduce-2 will specify 2 reduce tasks.

B. In YARN, the ApplicationMaster is responsible for requesting the resources required for a specific job. Thus, executing p yarn.applicationmaster.reduce.tasks-2 will specify that the ApplicationMaster launch two task containers on the worker nodes.

C. In YARN, resource allocation is a function of megabytes of memory in multiple of 1024mb. Thus, they should specify the amount of memory resource they need by executing D mapreduce.reduce.memory-mp-2040

D. In YARN, resource allocation is a function of virtual cores specified by the ApplicationMaster making requests to the

NodeManager where a reduce task is handled by a single container (and this a single virtual core). Thus, the developer needs to specify the number of virtual cores to the NodeManager by executing p yarn.nodemanager.cpu-vcores=2

E. MapReduce version 2 (MRv2) on YARN abstracts resource allocation away from the idea of "tasks" into memory and virtual cores, thus eliminating the need for a developer to specify the number of reduce tasks, and indeed preventing the developer from specifying the number of reduce tasks.

Correct Answer: D

**QUESTION 4**

You want to understand more about how users browse you 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 logs into your Hadoop cluster for 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. Import all users clicks from your OLTP databases into Hadoop using Sqoop

D. Write a MApReduce job with the web servers from mappers and the Hadoop cluster nodes reducers

E. Channel these clickstream into Hadoop using Hadoop Streaming

Correct Answer: AB

**QUESTION 5**

Which three basic configuration parameters must you set to migrate your cluster from MapReduce1 (MRv1) to MapReduce v2 (MRv2)?

A. Configure the NodeManager hostname and enable services on YARN by setting the following property in yarn-site.xml: yarn.nodemanager.hostname your_nodeManager_hostname

B. Configure the number of map tasks per job on YARN by setting the following property in mapredsite.xml: mapreduce.job.maps 2

C. Configure MapReduce as a framework running on YARN by setting the following property in mapredsite.xml: mapreduce.framework.name yarn

D. Configure the ResourceManager hostname and enable node services on YARN by setting the following property in yarn-site.xml: yarn.resourcemanager.hostname your_responseManager_hostname

E. Configure a default scheduler to run on YARN by setting the following property in sapred- site.xml: mapreduce.jobtracker.taskScheduler org.apache.hadoop.mapred.JobQueueTaskScheduler

F. Configure the NodeManager to enable MapReduce services on YARN by adding following property in yarn-site.xml: yarn.nodemanager.aux-services mapreduce_shuffle

Correct Answer: ABD

CCA-505 PDF Dumps CCA-505 Practice Test CCA-505 Study Guide