



# CCA-500<sup>Q&As</sup>

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 need to analyze 60,000,000 images stored in JPEG format, each of which is approximately 25 KB. Because your Hadoop cluster isn't optimized for storing and processing many small files, you decide to do the following actions:

1. Group the individual images into a set of larger files
2. Use the set of larger files as input for a MapReduce job that processes them directly with python using Hadoop streaming.

Which data serialization system gives the flexibility to do this?

- A. CSV
- B. XML
- C. HTML
- D. Avro
- E. SequenceFiles
- F. JSON

Correct Answer: E

---

**QUESTION 2**

You have recently converted your Hadoop cluster from a MapReduce 1 (MRv1) architecture to 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 to specify the number of reduce tasks when a specific job runs. Which method should you tell that developers to implement?

- A. 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.
- B. In YARN, resource allocation is a function of megabytes of memory in multiples of 1024mb. Thus, they should specify the amount of memory resource they need by executing `D mapreducereduces.memory-mb-2048`
- C. In YARN, the ApplicationMaster is responsible for requesting the resource required for a specific launch. Thus, executing `D yarn.applicationmaster.reduce.tasks=2` will specify that the ApplicationMaster launch two task containers on the worker nodes.
- D. Developers specify reduce tasks in the exact same way for both MapReduce version 1 (MRv1) and MapReduce version 2 (MRv2) on YARN. Thus, executing `D mapreduce.job.reduces=2` will specify reduce tasks.
- E. In YARN, resource allocation is a function of virtual cores specified by the ApplicationManager making requests to the NodeManager where a reduce task is handled by a single container (and thus 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`

Correct Answer: D

---

### QUESTION 3

Which scheduler would you deploy to ensure that your cluster allows short jobs to finish within a reasonable time without starting long-running jobs?

- A. Complexity Fair Scheduler (CFS)
- B. Capacity Scheduler
- C. Fair Scheduler
- D. FIFO Scheduler

Correct Answer: C

---

### 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

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 6

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 7

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, the statestored daemon, 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 8**

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 9**

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 10**

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

- A. 25GB on each hard drive may not be used to store HDFS blocks
- B. 100GB on each hard drive may not be used to store HDFS blocks
- C. All hard drives may be used to store HDFS blocks as long as at least 100 GB in total is available on the node
- D. A maximum of 100 GB on each hard drive may be used to store HDFS blocks

Correct Answer: C

---

**QUESTION 11**

What two processes must you do if you are running a Hadoop cluster with a single NameNode and six DataNodes, and



you want to change a configuration parameter so that it affects all six DataNodes. (Choose two)

- A. You must modify the configuration files on the NameNode only. DataNodes read their configuration from the master nodes
- B. You must modify the configuration files on each of the DataNodes machines
- C. You don't need to restart any daemon, as they will pick up changes automatically
- D. You must restart the NameNode daemon to apply the changes to the cluster
- E. You must restart all six DataNode daemon to apply the changes to the cluster

Correct Answer: AD

---

### QUESTION 12

You are working on a project where you need to chain together MapReduce, Pig jobs. You also need the ability to use forks, decision points, and path joins. Which ecosystem project should you use to perform these actions?

- A. Oozie
- B. ZooKeeper
- C. HBase
- D. Sqoop
- E. HUE

Correct Answer: A

[Latest CCA-500 Dumps](#)

[CCA-500 PDF Dumps](#)

[CCA-500 VCE Dumps](#)



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

## Try our product !

100% Guaranteed Success

100% Money Back Guarantee

365 Days Free Update

Instant Download After Purchase

24x7 Customer Support

Average 99.9% Success Rate

More than 800,000 Satisfied Customers Worldwide

Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.pass4itsure.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p><b>One Year Free Update</b> Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p><b>Money Back Guarantee</b> To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.

All trademarks are the property of their respective owners.

Copyright © pass4itsure, All Rights Reserved.