



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.lead4pass.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

Your company stores user profile records in an OLTP databases. You want to join these records with web server logs you have already ingested into the Hadoop file system. What is the best way to obtain and ingest these user records?

- A. Ingest with Hadoop streaming
- B. Ingest using Hive's IQAD DATA command
- C. Ingest with sqoop import
- D. Ingest with Pig's LOAD command
- E. Ingest using the HDFS put command

Correct Answer: C

QUESTION 2

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 3

You need to analyze 60,000,000 images stored in JPEG format, each of which is approximately 25 KB. Because you 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 4

You suspect that your NameNode is incorrectly configured, and is swapping memory to disk. Which Linux commands help you to identify whether swapping is occurring? (Select all that apply)

- A. free
- B. df
- C. memcat
- D. top
- E. jps
- F. vmstat
- G. swapinfo

Correct Answer: ADF

QUESTION 5

Which YARN daemon or service negotiations map and reduce Containers from the Scheduler, tracking their status and monitoring progress?

- A. NodeManager
- B. ApplicationMaster
- C. ApplicationManager
- D. ResourceManager

Correct Answer: B

QUESTION 6

Table schemas in Hive are:

- A. Stored as metadata on the NameNode



- B. Stored along with the data in HDFS
- C. Stored in the Metadata
- D. Stored in ZooKeeper

Correct Answer: B

QUESTION 7

In CDH4 and later, which file contains a serialized form of all the directory and files inodes in the filesystem, giving the NameNode a persistent checkpoint of the filesystem metadata?

- A. fstime
- B. VERSION
- C. Fsimage_N (where N reflects transactions up to transaction ID N)
- D. Edits_N-M (where N-M transactions between transaction ID N and transaction ID N)

Correct Answer: C

QUESTION 8

You are configuring your cluster to run HDFS and MapReducer v2 (MRv2) on YARN. Which two daemons need to be installed on your cluster's master nodes? (Choose two)

- A. HMaster
- B. ResourceManager
- C. TaskManager
- D. JobTracker
- E. NameNode
- F. DataNode

Correct Answer: BE

QUESTION 9

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 10

You have a cluster running with a FIFO scheduler enabled. You submit a large job A to the cluster, which you expect to run for one hour. Then, you submit job B to the cluster, which you expect to run a couple of minutes only.

You submit both jobs with the same priority.

Which two best describes how FIFO Scheduler arbitrates the cluster resources for job and its tasks?

(Choose two)

- A. Because there is a more than a single job on the cluster, the FIFO Scheduler will enforce a limit on the percentage of resources allocated to a particular job at any given time
- B. Tasks are scheduled on the order of their job submission
- C. The order of execution of job may vary
- D. Given job A and submitted in that order, all tasks from job A are guaranteed to finish before all tasks from job B
- E. The FIFO Scheduler will give, on average, and equal share of the cluster resources over the job lifecycle
- F. The FIFO Scheduler will pass an exception back to the client when Job B is submitted, since all slots on the cluster are use

Correct Answer: AD

QUESTION 11

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 12

You have a 20 node Hadoop cluster, with 18 slave nodes and 2 master nodes running HDFS High Availability (HA). You want to minimize the chance of data loss in your cluster. What should you do?

- A. Add another master node to increase the number of nodes running the JournalNode which increases the number of machines available to HA to create a quorum
- B. Set an HDFS replication factor that provides data redundancy, protecting against node failure
- C. Run a Secondary NameNode on a different master from the NameNode in order to provide automatic recovery from a NameNode failure.
- D. Run the ResourceManager on a different master from the NameNode in order to load-share HDFS metadata processing
- E. Configure the cluster's disk drives with an appropriate fault tolerant RAID level

Correct Answer: D

QUESTION 13

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 14

You're upgrading a Hadoop cluster from HDFS and MapReduce version 1 (MRv1) to one running HDFS and MapReduce version 2 (MRv2) on YARN. You want to set and enforce version 1 (MRv1) to one running HDFS and MapReduce version 2 (MRv2) on YARN. You want to set and enforce a block size of 128MB for all new files written to the cluster after upgrade. What should you do?

- A. You cannot enforce this, since client code can always override this value
- B. Set dfs.block.size to 128M on all the worker nodes, on all client machines, and on the NameNode, and set the parameter to final
- C. Set dfs.block.size to 128 M on all the worker nodes and client machines, and set the parameter to final. You do not



need to set this value on the NameNode

D. Set dfs.block.size to 134217728 on all the worker nodes, on all client machines, and on the NameNode, and set the parameter to final

E. Set dfs.block.size to 134217728 on all the worker nodes and client machines, and set the parameter to final. You do not need to set this value on the NameNode

Correct Answer: C

QUESTION 15

You decide to create a cluster which runs HDFS in High Availability mode with automatic failover, using Quorum Storage. What is the purpose of ZooKeeper in such a configuration?

A. It only keeps track of which NameNode is Active at any given time

B. It monitors an NFS mount point and reports if the mount point disappears

C. It both keeps track of which NameNode is Active at any given time, and manages the Edits file. Which is a log of changes to the HDFS filesystem

D. If only manages the Edits file, which is log of changes to the HDFS filesystem

E. Clients connect to ZooKeeper to determine which NameNode is Active

Correct Answer: A

[Latest CCA-500 Dumps](#)

[CCA-500 Study Guide](#)

[CCA-500 Exam Questions](#)



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

Try our product !

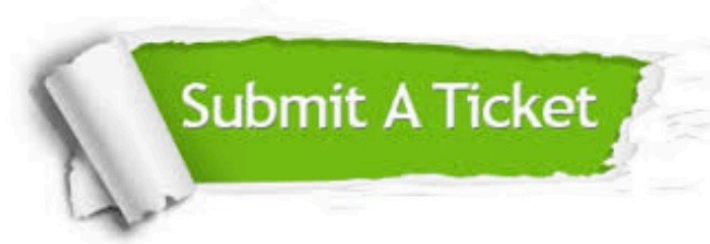
100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [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.lead4pass.com/allproducts>

Need Help

Please provide as much detail as possible so we can best assist you.
To update a previously submitted ticket:



 <p>One Year Free Update 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>Money Back Guarantee 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>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & 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 © lead4pass, All Rights Reserved.