



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 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

QUESTION 2

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 3

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 4

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 5

For each YARN job, the Hadoop framework generates task log file. Where are Hadoop task log files stored?

- A. Cached by the NodeManager managing the job containers, then written to a log directory on the NameNode
- B. Cached in the YARN container running the task, then copied into HDFS on job completion
- C. In HDFS, in the directory of the user who generates the job
- D. On the local disk of the slave node running the task

Correct Answer: D

QUESTION 6

Which process instantiates user code, and executes map and reduce tasks on a cluster running MapReduce v2 (MRv2) on YARN?

- A. NodeManager
- B. ApplicationMaster
- C. TaskTracker



- D. JobTracker
- E. NameNode
- F. DataNode
- G. ResourceManager

Correct Answer: A

QUESTION 7

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

QUESTION 8

Which YARN daemon or service negotiates 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 9

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 specify to 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 allocations 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 contains 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 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 10

Cluster Summary:

45 files and directories, 12 blocks = 57 total. Heap size is 15.31 MB/193.38MB(7%)

Configured capacity	:	17.33GB
DFS Used	:	144KB
Non DFS Used	:	5.49GB
DFS Remaining	:	11.84GB
DFS Used %	:	0%
DFS Remaining %	:	68.32GB
Live Nodes	:	6
Dead Nodes	:	1
Decommissioning Nodes	:	0
Number of Under-Replicated Blocks	:	6

Refer to the above screenshot.

You configure a Hadoop cluster with seven DataNodes and on of your monitoring UIs displays the details shown in the exhibit.

What does the this tell you?

A. The DataNode JVM on one host is not active

B. Because your under-replicated blocks count matches the Live Nodes, one node is dead, and your DFS Used % equals 0%, you can't be certain that your cluster has all the data you've written it.

C. Your cluster has lost all HDFS data which had bocks stored on the dead DatNode



D. The HDFS cluster is in safe mode

Correct Answer: A

[CCA-500 VCE Dumps](#)

[CCA-500 Study Guide](#)

[CCA-500 Exam Questions](#)