

100% Money Back Guarantee

Vendor: Cloudera

Exam Code: CCA-500

Exam Name: Cloudera Certified Administrator for Apache Hadoop (CCA-H)

Version: Demo

QUESTION 1

Your cluster's mapred-start.xml includes the following parameters

```
<name>mapreduce.map.memory.mb</name>
<value>4096</value>
<name>mapreduce.reduce.memory.mb</name>
<value>8192</value>
```

And any cluster's yarn-site.xml includes the following parameters

```
<name>yarn.nodemanager.vmem-pmem-ratio</name>
<value>2.1</value>
```

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 2

Assuming you're not running HDFS Federation, what is the maximum number of NameNode daemons you should run on your cluster in order to avoid a "split-brain" scenario with your NameNode when running HDFS High Availability (HA) using Quorum-based storage?

- A. Two active NameNodes and two Standby NameNodes
- B. One active NameNode and one Standby NameNode
- C. Two active NameNodes and one Standby NameNode
- D. Unlimited. HDFS High Availability (HA) is designed to overcome limitations on the number of NameNodes you can deploy

Correct Answer: B

QUESTION 3

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 4

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 5

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 6

Each node in your Hadoop cluster, running YARN, has 64GB memory and 24 cores. Your yarn.site.xml has the following configuration:

```
<property>
<name>yarn.nodemanager.resource.memory-mb</name>
<value>32768</value>
</property>
<property>
<name>yarn.nodemanager.resource.cpu-vcores</name>
<value>12</value>
</property>
```

You want YARN to launch no more than 16 containers per node. What should you do?

- A. Modify yarn-site.xml with the following property:
<name>yarn.scheduler.minimum-allocation-mb</name>
<value>2048</value>
- B. Modify yarn-sites.xml with the following property:
<name>yarn.scheduler.minimum-allocation-mb</name>
<value>4096</value>
- C. Modify yarn-site.xml with the following property:
<name>yarn.nodemanager.resource.cpu-vcores</name>
- D. No action is needed: YARN's dynamic resource allocation automatically optimizes the node memory and cores

Correct Answer: A

QUESTION 7

You want to node to only swap Hadoop daemon data from RAM to disk when absolutely necessary. What should you do?

- A. Delete the /dev/vmswap file on the node
- B. Delete the /etc/swap file on the node
- C. Set the ram.swap parameter to 0 in core-site.xml
- D. Set vm.swapfile file on the node
- E. Delete the /swapfile file on the node

Correct Answer: D

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

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 10

You are running a Hadoop cluster with a NameNode on host mynamenode, a secondary NameNode on host mysecondarynamenode and several DataNodes.

Which best describes how you determine when the last checkpoint happened?

- A. Execute hdfs namenode report on the command line and look at the Last Checkpoint information
- B. Execute hdfs dfsadmin saveNamespace on the command line which returns to you the last checkpoint value in fstime file
- C. Connect to the web UI of the Secondary NameNode (<http://mysecondary:50090/>) and look at the "Last Checkpoint" information
- D. Connect to the web UI of the NameNode (<http://mynamenode:50070/>) and look at the "Last Checkpoint" information

Correct Answer: C

QUESTION 11

What does CDH packaging do on install to facilitate Kerberos security setup?

- A. Automatically configures permissions for log files at & MAPRED_LOG_DIR/userlogs
- B. Creates users for hdfs and mapreduce to facilitate role assignment
- C. Creates directories for temp, hdfs, and mapreduce with the correct permissions
- D. Creates a set of pre-configured Kerberos keytab files and their permissions
- E. Creates and configures your kdc with default cluster values

Correct Answer: B

QUESTION 12

You want to understand more about how users browse your 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 across logs into your Hadoop cluster 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. Channel these clickstreams into Hadoop using Hadoop Streaming
- D. Import all user clicks from your OLTP databases into Hadoop using Sqoop

- E. Write a MapReduce job with the web servers for mappers and the Hadoop cluster nodes for reducers

Correct Answer: B

QUESTION 13

Which three basic configuration parameters must you set to migrate your cluster from MapReduce 1 (MRv1) to MapReduce V2 (MRv2)? (Choose three)

- A. Configure the NodeManager to enable MapReduce services on YARN by setting the following property in yarn-site.xml:
<name>yarn.nodemanager.hostname</name>
<value>your_nodeManager_shuffle</value>
- B. Configure the NodeManager hostname and enable node services on YARN by setting the following property in yarn-site.xml:
<name>yarn.nodemanager.hostname</name>
<value>your_nodeManager_hostname</value>
- C. Configure a default scheduler to run on YARN by setting the following property in mapred-site.xml:
<name>mapreduce.jobtracker.taskScheduler</name>
<value>org.apache.hadoop.mapred.JobQueueTaskScheduler</value>
- D. Configure the number of map tasks per job on YARN by setting the following property in mapred-site.xml:
<name>mapreduce.job.maps</name>
<value>2</value>
- E. Configure the ResourceManager hostname and enable node services on YARN by setting the following property in yarn-site.xml:
<name>yarn.resourcemanager.hostname</name>
<value>your_resourceManager_hostname</value>
- F. Configure MapReduce as a Framework running on YARN by setting the following property in mapred-site.xml:
<name>mapreduce.framework.name</name>
<value>yarn</value>

Correct Answer: AEF

QUESTION 14

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 15

Identify two features/issues that YARN is designed to address: (Choose two)

- A. Standardize on a single MapReduce API
B. Single point of failure in the NameNode
C. Reduce complexity of the MapReduce APIs

- D. Resource pressure on the JobTracker
- E. Ability to run framework other than MapReduce, such as MPI
- F. HDFS latency

Correct Answer: DE

QUESTION 16

Which YARN daemon or service monitors a Controller's per-application resource using (e.g., memory CPU)?

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

Correct Answer: A

QUESTION 17

Which is the default scheduler in YARN?

- A. YARN doesn't configure a default scheduler, you must first assign an appropriate scheduler class in yarn-site.xml
- B. Capacity Scheduler
- C. Fair Scheduler
- D. FIFO Scheduler

Correct Answer: B

QUESTION 18

Which YARN process run as "container 0" of a submitted job and is responsible for resource requests?

- A. ApplicationManager
- B. JobTracker
- C. ApplicationMaster
- D. JobHistoryServer
- E. ResoureManager
- F. NodeManager

Correct Answer: C

QUESTION 19

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 20

Your cluster is configured with HDFS and MapReduce version 2 (MRv2) on YARN. What is the result when you execute: `hadoop jar SampleJar MyClass` on a client machine?

- A. SampleJar.Jar is sent to the ApplicationMaster which allocates a container for SampleJar.Jar
- B. Sample.jar is placed in a temporary directory in HDFS

- C. SampleJar.jar is sent directly to the ResourceManager
- D. SampleJar.jar is serialized into an XML file which is submitted to the ApplicationMaster

Correct Answer: A

QUESTION 21

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 22

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 23

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 one of your monitoring UIs displays the details shown in the exhibit.

What does this tell you?

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

Trying 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 **69,000** Satisfied Customers Worldwide
- ★ Multi-Platform capabilities - **Windows, Mac, Android, iPhone, iPod, iPad, Kindle**

Need Help

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

To update a previously submitted ticket:



 One Year Free Update <p>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>	 Money Back Guarantee <p>To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 Security & Privacy <p>We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p>
---	---	--

Guarantee & Policy | Privacy & Policy | Terms & Conditions

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

All trademarks are the property of their respective owners.