![Leads4Pass]

CCA-500 ^{Q&As}

Cloudera Certified Administrator for Apache Hadoop (CCAH)

Pass Cloudera CCA-500 Exam with 100% Guarantee

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

https://www.leads4pass.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 cluster is running MapReduce version 2 (MRv2) on YARN. Your ResourceManager is configured to use the FairScheduler. Now you want to configure your scheduler such that a new user on the cluster can submit jobs into their own queue application submission. Which configuration should you set?

A. You can specify new queue name when user submits a job and new queue can be created dynamically if the property yarn.scheduler.fair.allow-undecleared-pools = true

B. Yarn.scheduler.fair.user.fair-as-default-queue = false and yarn.scheduler.fair.allow- undecleared-pools = true

C. You can specify new queue name when user submits a job and new queue can be created dynamically if yarn .schedule.fair.user-as-default-queue = false

D. You can specify new queue name per application in allocations.xml file and have new jobs automatically assigned to the application queue

Correct Answer: A

QUESTION 2

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 3

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 4

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 ApplicatoionMaster

Correct Answer: A

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 gets scheduled.

C. When Job A gets submitted, it doesn\\'t consumes all the task slots.

D. When Job A gets submitted, it consumes all the task slots.

Correct Answer: B

[Latest CCA-500 Dumps](#) [CCA-500 PDF Dumps](#) [CCA-500 VCE Dumps](#)