

QUESTIONS & ANSWERS

Kill your exam at first Attempt



Cloudera

CCA-500

*Cloudera Certified Administrator for Apache Hadoop
(CCA-H)*

QUESTION: 54

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`

Answer: D

QUESTION: 55

Your Hadoop cluster contains nodes in three racks. You have not configured the `dfs.hosts` property in the NameNode’s configuration file. What results?

- A. The NameNode will update the `dfs.hosts` property to include machines running the DataNode daemon on the next NameNode reboot or with the command `dfsadmin -refreshNodes`

- B. No new nodes can be added to the cluster until you specify them in the dfs.hosts file
- C. Any machine running the DataNode daemon can immediately join the cluster
- D. Presented with a blank dfs.hosts property, the NameNode will permit DataNodes specified in mapred.hosts to join the cluster

Answer: C

QUESTION: 56

You are running a Hadoop cluster with MapReduce version 2 (MRv2) on YARN. You consistently see that MapReduce map tasks on your cluster are running slowly because of excessive garbage collection of JVM, how do you increase JVM heap size property to 3GB to optimize performance?

- A. yarn.application.child.java.opts=-Xsx3072m
- B. yarn.application.child.java.opts=-Xmx3072m
- C. mapreduce.map.java.opts=-Xms3072m
- D. mapreduce.map.java.opts=-Xmx3072m

Answer: C

Reference:

<http://hortonworks.com/blog/how-to-plan-and-configure-yarn-in-hdp-2-0/>

QUESTION: 57

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?

- 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, an 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

Answer: A, D

QUESTION: 58

A user comes to you, complaining that when she attempts to submit a Hadoop job, it fails. There is a Directory in HDFS named /data/input. The Jar is named j.jar, and the driver class is named DriverClass. She runs the command: `Hadoop jar j.jar DriverClass /data/input/data/output` The error message returned includes the line: `PrivilegedActionException as:training (auth:SIMPLE) cause:org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: file:/data/input` What is the cause of the error?

- A. The user is not authorized to run the job on the cluster
- B. The output directory already exists
- C. The name of the driver has been spelled incorrectly on the command line
- D. The directory name is misspelled in HDFS
- E. The Hadoop configuration files on the client do not point to the cluster

Answer: A

QUESTION: 59

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

Answer: C

QUESTION: 60

Which two are features of Hadoop's rack topology?

- A. Configuration of rack awareness is accomplished using a configuration file. You cannot use a rack topology script.
- B. Hadoop gives preference to intra-rack data transfer in order to conserve bandwidth
- C. Rack location is considered in the HDFS block placement policy
- D. HDFS is rack aware but MapReduce daemon are not
- E. Even for small clusters on a single rack, configuring rack awareness will improve performance

Answer: B, C

KILL EXAMS

KILLEXAMS.COM

For More exams visit <https://killexams.com> -



[KILLEXAMS.COM](https://killexams.com)

Kill your exam at First Attempt....Guaranteed!