Braindump2go Cloudera CCA-500 Practice Exam Questions Guarantee Your Cloudera CCA-500 Real Test Totally Success (1-11)

QUESTION 1   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 Answer: ABQUESTION 2   Identify two features/issues that YARN is designated to address: 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 Answer: BD   Explanation: http://www.revelytix.com/?q=content/hadoop-ecosystem(YARN, first para) QUESTION 3   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 Answer: A QUESTION 4   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 Answer: B Explanation: http://hadoop.apache.org/docs/r2.4.1/hadoop-yarn/hadoop-yarn-site/CapacityScheduler.html QUESTION 5   Which YARN process run as "container 0" of a submitted job and is responsible for resource qrequests? A.   ApplicationManager   B.   JobTracker   C.   ApplicationMaster D.   JobHistoryServer   E.   ResoureManager   F.   NodeManager Answer: C QUESTION 6   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 Answer: C   Explanation: http://hadoop.apache.org/docs/r1.2.1/fair_scheduler.html QUESTION 7   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 Answer: A QUESTION 8   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 Answer: A QUESTION 9 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 Answer: D QUESTION 10   Which two features does Kerberos security add to a Hadoop cluster? A.   User authentication on all remote procedure calls (RPCs)   B.   Encryption for data during transfer between the Mappers and Reducers   C.   Encryption for data on disk ("at rest")   D.   Authentication for user access to the cluster against a central server   E.   Root access to the cluster for users hdfs and mapred but non-root access for clients Answer: BD QUESTION 11   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 Offe 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 Answer: B   Braindump2go New Released Cloudera CCA-500 Dump PDF Free Download, 71 Questions in all, Passing Your Exam 100% Easily! http://www.braindump2go.com/cca-500.html