Vendor: Cloudera

Exam Code: CCD-470

Exam Name: Cloudera Certified Developer for Apache Hadoop CDH4 Upgrade Exam (CCDH)

Version: DEMO

QUESTION 1
You need to run the same job many times with minor variations. Rather than hardcoding all job configuration options in your drive code, you've decided to have your Driver subclass org.apache.hadoop.conf.Configured and implement the org.apache.hadoop.util.Tool interface. Indentify which invocation correctly passes.mapred.job.name with a value of Example to Hadoop?

A. hadoop "mapred.job.name=Example" MyDriver input output
B. hadoop MyDriver mapred.job.name=Example input output
C. hadoop MyDrive -D mapred.job.name=Example input output
D. hadoop setproperty mapred.job.name=Example MyDriver input output
E. hadoop setproperty ("mapred.job.name=Example") MyDriver input output

Answer: C

QUESTION 2
You are developing a MapReduce job for sales reporting. The mapper will process input keys representing the year (IntWritable) and input values representing product indentifies (Text).
Indentify what determines the data types used by the Mapper for a given job.

A. The key and value types specified in the JobConf.setMapInputKeyClass and JobConf.setMapInputValuesClass methods
B. The data types specified in HADOOP_MAP_DATATYPES environment variable
C. The mapper-specification.xml file submitted with the job determine the mapper's input key and value types.
D. The InputFormat used by the job determines the mapper's input key and value types.

Answer: D

QUESTION 3
Identify the MapReduce v2 (MRv2 / YARN) daemon responsible for launching application containers and monitoring application resource usage?

A. ResourceManager
B. NodeManager
C. ApplicationMaster
D. ApplicationMasterService
E. TaskTracker
F. JobTracker

Answer: C

QUESTION 4
Which best describes how TextInputFormat processes input files and line breaks?

A. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReader of the split that contains the beginning of the broken line.
B. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReaders of both splits containing the broken line.
C. The input file is split exactly at the line breaks, so each RecordReader will read a series of

complete
lines.
D. Input file splits may cross line breaks. A line that crosses file splits is ignored.
E. Input file splits may cross line breaks. A line that crosses file splits is read by the RecordReader
    of
    the split that contains the end of the broken line.

Answer: E

QUESTION 5
For each input key-value pair, mappers can emit:

A. As many intermediate key-value pairs as designed. There are no restrictions on the types of
    those
    key-value pairs (i.e., they can be heterogeneous).
B. As many intermediate key-value pairs as designed, but they cannot be of the same type as the
    input key-value pair.
C. One intermediate key-value pair, of a different type.
D. One intermediate key-value pair, but of the same type.
E. As many intermediate key-value pairs as designed, as long as all the keys have the same types
    and
    all the values have the same type.

Answer: E

QUESTION 6
You have user profile records in your OLPT database, that you want to join with web logs you
have already ingested into the Hadoop file system. How will you obtain these user records?

A. HDFS command
B. Pig LOAD command
C. Sqoop import
D. Hive LOAD DATA command
E. Ingest with Flume agents
F. Ingest with Hadoop Streaming

Answer: B

QUESTION 7
What is the disadvantage of using multiple reducers with the default HashPartitioner and
distributing your workload across you cluster?

A. You will not be able to compress the intermediate data.
B. You will longer be able to take advantage of a Combiner.
C. By using multiple reducers with the default HashPartitioner, output files may not be in globally
    sorted order.
D. There are no concerns with this approach. It is always advisable to use multiple reduces.

Answer: C

QUESTION 8

You need to perform statistical analysis in your MapReduce job and would like to call methods in the Apache Commons Math library, which is distributed as a 1.3 megabyte Java archive (JAR) file. Which is the best way to make this library available to your MapReducer job at runtime?

A. Have your system administrator copy the JAR to all nodes in the cluster and set its location in the HADOOP_CLASSPATH environment variable before you submit your job.
B. Have your system administrator place the JAR file on a Web server accessible to all cluster nodes
   and then set the HTTP_JAR_URL environment variable to its location.
C. When submitting the job on the command line, specify the 璀ibjars option followed by the JAR file path.
D. Package your code and the Apache Commands Math library into a zip file named JobJar.zip

Answer: C

QUESTION 9

The Hadoop framework provides a mechanism for coping with machine issues such as faulty configuration or impending hardware failure. MapReduce detects that one or a number of machines are performing poorly and starts more copies of a map or reduce task. All the tasks run simultaneously and the task finish first are used. This is called:

A. Combine
B. IdentityMapper
C. IdentityReducer
D. Default Partitioner
E. Speculative Execution

Answer: E

QUESTION 10

For each intermediate key, each reducer task can emit:

A. As many final key-value pairs as desired. There are no restrictions on the types of those key-value
   pairs (i.e., they can be heterogeneous).
B. As many final key-value pairs as desired, but they must have the same type as the intermediate key-value pairs.
C. As many final key-value pairs as desired, as long as all the keys have the same type and all the values have the same type.
D. One final key-value pair per value associated with the key; no restrictions on the type.
E. One final key-value pair per key; no restrictions on the type.

Answer: E

Thank You for Trying Our Product

Braindump2go Certification Exam Features:

★ More than 99,900 Satisfied Customers Worldwide.

★ Average 99.9% Success Rate.

★ Free Update to match latest and real exam scenarios.

★ Instant Download Access! No Setup required.

★ Questions & Answers are downloadable in PDF format and VCE test engine format.

★ Multi-Platform capabilities - Windows, Laptop, Mac, Android, iPhone, iPod, iPad.

★ 100% Guaranteed Success or 100% Money Back Guarantee.

★ Fast, helpful support 24x7.

View list of all certification exams: http://www.braindump2go.com/all-products.html

10% Discount Coupon Code: BDNT2014