CCD-410^{Q&As}

Cloudera Certified Developer for Apache Hadoop (CCDH)

Pass Cloudera CCD-410 Exam with 100% Guarantee

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

https://www.geekcert.com/ccd-410.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

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.

Correct Answer: E

Reducer reduces a set of intermediate values which share a key to a smaller set of values.

Reducing lets you aggregate values together. A reducer function receives an iterator of input values from an input list. It then combines these values together, returning a single output value. Reference: Hadoop Map-Reduce Tutorial; Yahoo! Hadoop Tutorial, Module 4: MapReduce

QUESTION 2

You want to run Hadoop jobs on your development workstation for testing before you submit them to your production cluster. Which mode of operation in Hadoop allows you to most closely simulate a production cluster while using a single machine?

A. Run all the nodes in your production cluster as virtual machines on your development workstation.

B. Run the hadoop command with the jt local and the fs file:///options.

C. Run the DataNode, TaskTracker, NameNode and JobTracker daemons on a single machine.

D. Run simldooop, the Apache open-source software for simulating Hadoop clusters.

Correct Answer: C

QUESTION 3

In a large MapReduce job with m mappers and n reducers, how many distinct copy operations will there be in the sort/shuffle phase?

A. mXn (i.e., m multiplied by n)

B. n

C. m

D. m+n (i.e., m plus n)

E. mn (i.e., m to the power of n)

Correct Answer: A

A MapReduce job with m mappers and r reducers involves up to m * r distinct copy operations, since each mapper may have intermediate output going to every reducer.

QUESTION 4

Can you use MapReduce to perform a relational join on two large tables sharing a key? Assume that the two tables are formatted as comma-separated files in HDFS.

A. Yes.

B. Yes, but only if one of the tables fits into memory

C. Yes, so long as both tables fit into memory.

D. No, MapReduce cannot perform relational operations.

E. No, but it can be done with either Pig or Hive.

Correct Answer: A

Note:

*

 Join Algorithms in MapReduce A) Reduce-side join B) Map-side join C) In-memory join / Striped Striped variant variant / Memcached variant

*

 Which join to use? / In-memory join > map-side join > reduce-side join / Limitations of each? In-memory join: memory Map-side join: sort order and partitioning

Reduce-side join: general purpose

QUESTION 5

Indentify which best defines a SequenceFile?

A. A SequenceFile contains a binary encoding of an arbitrary number of homogeneous Writable objects

B. A SequenceFile contains a binary encoding of an arbitrary number of heterogeneous Writable objects

C. A SequenceFile contains a binary encoding of an arbitrary number of WritableComparable objects, in sorted order.

D. A SequenceFile contains a binary encoding of an arbitrary number key-value pairs. Each key must be the same type.

Each value must be the same type.

Correct Answer: D

SequenceFile is a flat file consisting of binary key/value pairs.

There are 3 different SequenceFile formats:

Uncompressed key/value records.

Record compressed key/value records - only \\'values\\' are compressed here.

Block compressed key/value records - both keys and values are collected in \\'blocks\\' separately and

compressed. The size of the \\'block\\' is configurable.

Reference: http://wiki.apache.org/hadoop/SequenceFile

Latest CCD-410 Dumps   CCD-410 Study Guide   CCD-410 Braindumps