HDPCD^{Q&As}

Hortonworks Data Platform Certified Developer

Pass Hortonworks HDPCD Exam with 100% Guarantee

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

https://www.pass4itsure.com/hdpcd.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Hortonworks Official Exam Center

Instant Download After Purchase
100% Money Back Guarantee
365 Days Free Update
800,000+ Satisfied Customers

QUESTION 1

What is 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 same type.

Correct Answer: D

Explanation: 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

QUESTION 2

Consider the following two relations, A and B.

```
A = LOAD 'data1' AS (a1:int,a2:chararray);
DUMP A;
(1,apple)
(3,orange)
(4,peach)
(2,cherry)

B = LOAD 'data2' AS (b1:chararray,b2:int);
DUMP B;
(Jim,2)
(Brian,4)
(Kim,0)
(Terry,3)
(Chris,2)
```

A Pig JOIN statement that combined relations A by its first field and B by its second field would produce what output?

A. 2 Jim Chris 2 3 Terry 3 4 Brian 4

B. 2 cherry 2 cherry 3 orange 4 peach

C. 2 cherry Jim, Chris 3 orange Terry 4 peach Brian

D. 2 cherry Jim 2 2 cherry Chris 2 3 orange Terry 3 4 peach Brian 4

Correct Answer: D

QUESTION 3

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.

Correct Answer: A

Reference: How Map and Reduce operations are actually carried out

QUESTION 4

Which one of the following statements is true regarding a MapReduce job?

A. The job\\'s Partitioner shuffles and sorts all (key.value) pairs and sends the output to all reducers

B. The default Hash Partitioner sends key value pairs with the same key to the same Reducer

C. The reduce method is invoked once for each unique value

D. The Mapper must sort its output of (key.value) pairs in descending order based on value

Correct Answer: A

QUESTION 5

Which Two of the following statements are true about hdfs? Choose 2 answers

A. An HDFS file that is larger than dfs.block.size is split into blocks

B. Blocks are replicated to multiple datanodes

C. HDFS works best when storing a large number of relatively small files

D. Block sizes for all files must be the same size

Correct Answer: AB

HDPCD Practice Test   HDPCD Study Guide   HDPCD Exam Questions