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.leads4pass.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 does Pig provide to the overall Hadoop solution?

A. Legacy language Integration with MapReduce framework

B. Simple scripting language for writing MapReduce programs

C. Database table and storage management services

D. C++ interface to MapReduce and data warehouse infrastructure

Correct Answer: B

QUESTION 2

You want to count the number of occurrences for each unique word in the supplied input data. You\\'ve decided to implement this by having your mapper tokenize each word and emit a literal value 1, and then have your reducer increment a counter for each literal 1 it receives. After successful implementing this, it occurs to you that you could optimize this by specifying a combiner. Will you be able to reuse your existing Reduces as your combiner in this case and why or why not?

A. Yes, because the sum operation is both associative and commutative and the input and output types to the reduce method match.

B. No, because the sum operation in the reducer is incompatible with the operation of a Combiner.

C. No, because the Reducer and Combiner are separate interfaces.

D. No, because the Combiner is incompatible with a mapper which doesn\\'t use the same data type for both the key and value.

E. Yes, because Java is a polymorphic object-oriented language and thus reducer code can be reused as a combiner.

Correct Answer: A

Explanation: Combiners are used to increase the efficiency of a MapReduce program. They are used to aggregate intermediate map output locally on individual mapper outputs. Combiners can help you reduce the amount of data that needs to be transferred across to the reducers. You can use your reducer code as a combiner if the operation performed is commutative and associative. The execution of combiner is not guaranteed, Hadoop may or may not execute a combiner. Also, if required it may execute it more then 1 times. Therefore your MapReduce jobs should not depend on the combiners execution.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What are combiners? When should I use a combiner in my MapReduce Job?

QUESTION 3

Table metadata in Hive is:

A. Stored as metadata on the NameNode.

B. Stored along with the data in HDFS.

C. Stored in the Metastore.

D. Stored in ZooKeeper.

Correct Answer: C

Explanation: By default, hive use an embedded Derby database to store metadata information. The metastore is the "glue" between Hive and HDFS. It tells Hive where your data files live in HDFS, what type of data they contain, what tables they belong to, etc.

The Metastore is an application that runs on an RDBMS and uses an open source ORM layer called DataNucleus, to convert object representations into a relational schema and vice versa. They chose this approach as opposed to storing this information in hdfs as they need the Metastore to be very low latency. The DataNucleus layer allows them to plugin many different RDBMS technologies.

Note:

*

By default, Hive stores metadata in an embedded Apache Derby database, and other client/server

databases like MySQL can optionally be used.

*

features of Hive include:

Metadata storage in an RDBMS, significantly reducing the time to perform semantic checks during query

execution.

Reference: Store Hive Metadata into RDBMS

QUESTION 4

Which Hadoop component is responsible for managing the distributed file system metadata?

A. NameNode

B. Metanode

C. DataNode

D. NameSpaceManager

Correct Answer: A

QUESTION 5

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)
```

What is the output of the following Pig commands?

X = GROUP A BY S1;

DUMP X;

A. (group,{(apple,peach,cherry,orange)})

B. {apple,peach,cherry,orange}

C. {1,4,2,3}

D.
```
(apple,{(1,apple)})
(peach,{(4,peach)})
(cherry,{(2,cherry)})
(orange,{(3,orange)})
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: D