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

QUESTION 1

What does the following WebHDFS command do?

Curl -1 -L "http://host:port/webhdfs/v1/foo/bar?op=OPEN"

A. Make a directory /foo/bar

B. Read a file /foo/bar

C. List a directory /foo

D. Delete a directory /foo/bar

Correct Answer: B

QUESTION 2

Which project gives you a distributed, Scalable, data store that allows you random, realtime read/write access to hundreds of terabytes of data?

A. HBase

B. Hue

C. Pig

D. Hive

E. Oozie

F. Flume

G. Sqoop

Correct Answer: A

Explanation: Use Apache HBase when you need random, realtime read/write access to your Big Data.

Note: This project\\'s goal is the hosting of very large tables -- billions of rows X millions of columns -- atop

clusters of commodity hardware. Apache HBase is an open-source, distributed, versioned, column-

oriented store modeled after Google\\'s Bigtable: A Distributed Storage System for Structured Data by

Chang et al. Just as Bigtable leverages the distributed data storage provided by the Google File System,

Apache HBase provides Bigtable-like capabilities on top of Hadoop and HDFS.

Features

Linear and modular scalability.

Strictly consistent reads and writes.

Automatic and configurable sharding of tables

Automatic failover support between RegionServers.

Convenient base classes for backing Hadoop MapReduce jobs with Apache HBase tables.

Easy to use Java API for client access.

Block cache and Bloom Filters for real-time queries.

Query predicate push down via server side Filters

Thrift gateway and a REST-ful Web service that supports XML, Protobuf, and binary data encoding options

Extensible jruby-based (JIRB) shell

Support for exporting metrics via the Hadoop metrics subsystem to files or Ganglia; or via JMX

Reference: http://hbase.apache.org/ (when would I use HBase? First sentence)

QUESTION 3

Review the following data and Pig code: What command to define B would produce the output (M,62,95l02) when invoking the DUMP operator on B?

```
M,38,95111
F,29,95060
F,45,95192
M,62,95102
F,56,95102

A = LOAD 'data' USING PigStorage(',')
AS (gender:chararray, age:int, zip:chararray);
```

A. B = FILTER A BY (zip = = \\'95102\\' AND gender = = M");

B. B= FOREACH A BY (gender = = \\'M\\' AND zip = = \\'95102\\');

C. B = JOIN A BY (gender = = \\'M\\' AND zip = = \\'95102\\');

D. B= GROUP A BY (zip = = \\'95102\\' AND gender = = \\'M\\');

Correct Answer: A

QUESTION 4

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

Correct Answer: C

Reference: Hadoop and Pig for Large-Scale Web Log Analysis

QUESTION 5

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: C

Reference: Hadoop Map-Reduce Tutorial; Yahoo! Hadoop Tutorial, Module 4: MapReduce