

APACHE-HADOOP-DEVELOPER^{Q&As}

Hadoop 2.0 Certification exam for Pig and Hive Developer

Pass Hortonworks APACHE-HADOOP-DEVELOPER Exam with 100% Guarantee

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

<https://www.leads4pass.com/apache-hadoop-developer.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

Which two of the following are true about this trivial Pig program\ (choose Two)

```
$ pig
grunt> ABC = LOAD 'myfile';
grunt> DUMP ABC;
```

- A. The contents of myfile appear on stdout
- B. Pig assumes the contents of myfile are comma delimited
- C. ABC has a schema associated with it
- D. myfile is read from the user\’s home directory in HDFS

Correct Answer: AD

QUESTION 2

You write MapReduce job to process 100 files in HDFS. Your MapReduce algorithm uses TextInputFormat: the mapper applies a regular expression over input values and emits key- values pairs with the key consisting of the matching text, and the value containing the filename and byte offset. Determine the difference between setting the number of reduces to one and settings the number of reducers to zero.

- A. There is no difference in output between the two settings.
- B. With zero reducers, no reducer runs and the job throws an exception. With one reducer, instances of matching patterns are stored in a single file on HDFS.
- C. With zero reducers, all instances of matching patterns are gathered together in one file on HDFS. With one reducer, instances of matching patterns are stored in multiple files on HDFS.
- D. With zero reducers, instances of matching patterns are stored in multiple files on HDFS. With one reducer, all instances of matching patterns are gathered together in one file on HDFS.

Correct Answer: D

Explanation: * It is legal to set the number of reduce-tasks to zero if no reduction is desired.

In this case the outputs of the map-tasks go directly to the FileSystem, into the output path set by setOutputPath(Path). The framework does not sort the map-outputs before writing them out to the FileSystem.

* Often, you may want to process input data using a map function only. To do this, simply set mapreduce.job.reduces to zero. The MapReduce framework will not create any reducer tasks. Rather, the outputs of the mapper tasks will be the final output of the job.

Note:

Reduce

In this phase the reduce(WritableComparable, Iterator, OutputCollector, Reporter) method is called for each pair in the grouped inputs.

The output of the reduce task is typically written to the FileSystem via OutputCollector.collect (WritableComparable, Writable).

Applications can use the Reporter to report progress, set application-level status messages and update Counters, or just indicate that they are alive.

The output of the Reducer is not sorted.

QUESTION 3

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

QUESTION 4

Examine the following Hive statements:

```
CREATE TABLE x (name STRING, age INT, zip INT, salary DOUBLE)
ROW FORMAT DELIMITED FIELDS TERMINATED BY
',' LOCATION '/user/joe/x';
LOAD DATA INPATH 'input/File1' OVERWRITE INTO TABLE x;
```

Assuming the statements above execute successfully, which one of the following statements is true?

- A. Hive reformats File1 into a structure that Hive can access and moves into to/user/joe/x/
- B. The file named File1 is moved to to/user/joe/x/
- C. The contents of File1 are parsed as comma-delimited rows and loaded into /user/joe/x/

D. The contents of File1 are parsed as comma-delimited rows and stored in a database

Correct Answer: B

QUESTION 5

You need to run the same job many times with minor variations. Rather than hardcoding all job configuration options in your driver code, you've decided to have your Driver subclass `org.apache.hadoop.conf.Configured` and implement the `org.apache.hadoop.util.Tool` interface.

Identify 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 MyDriver ? 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

Correct Answer: C

Explanation: Configure the property using the `-D key=value` notation:

```
-D mapred.job.name='My Job'
```

You can list a whole bunch of options by calling the streaming jar with just the `-info` argument

Reference: Python hadoop streaming : Setting a job name

[APACHE-HADOOP-DEVELOPER PDF Dumps](#)

[APACHE-HADOOP-DEVELOPER Study Guide](#)

[APACHE-HADOOP-DEVELOPER Braindumps](#)