



HADOOP-PR000007^{Q&As}

Hortonworks Certified Apache Hadoop 2.0 Developer (Pig and Hive Developer)

Pass Hortonworks HADOOP-PR000007 Exam with 100% Guarantee

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

<https://www.geekcert.com/hadoop-pr000007.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

Given the following Hive command:

```
CREATE EXTERNAL TABLE mytable (name string, age int) ROW FORMAT DELIMITED FIELDS TERMINATED BY  
STORED AS TEXTFILE LOCATION '/home/user/mydata/';
```

Which one of the following statements is true?

- A. The files in the mydata folder are copied to a subfolder of /apps/hlve/warehouse
- B. The files in the mydata folder are moved to a subfolder of /apps/hive/wa re house
- C. The files in the mydata folder are copied into Hive\'s underlying relational database
- D. The files in the mydata folder do not move from their current location In HDFS

Correct Answer: D

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

Which one of the following statements is false about HCatalog?

- A. Provides a shared schema mechanism
- B. Designed to be used by other programs such as Pig, Hive and MapReduce
- C. Stores HDFS data in a database for performing SQL-like ad-hoc queries
- D. Exists as a subproject of Hive

Correct Answer: C

QUESTION 4

On a cluster running MapReduce v1 (MRv1), a TaskTracker heartbeats into the JobTracker on your cluster, and alerts the JobTracker it has an open map task slot.

What determines how the JobTracker assigns each map task to a TaskTracker?

- A. The amount of RAM installed on the TaskTracker node.
- B. The amount of free disk space on the TaskTracker node.
- C. The number and speed of CPU cores on the TaskTracker node.
- D. The average system load on the TaskTracker node over the past fifteen (15) minutes.
- E. The location of the InputSplit to be processed in relation to the location of the node.

Correct Answer: E

Explanation: The TaskTrackers send out heartbeat messages to the JobTracker, usually every few minutes, to reassure the JobTracker that it is still alive. These message also inform the JobTracker of the number of available slots, so the JobTracker can stay up to date with where in the cluster work can be delegated. When the JobTracker tries to find somewhere to schedule a task within the MapReduce operations, it first looks for an empty slot on the same server that hosts the DataNode containing the data, and if not, it looks for an empty slot on a machine in the same rack.



Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, How JobTracker schedules a task?

QUESTION 5

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

[HADOOP-PR000007 Study Guide](#)

[HADOOP-PR000007 Exam Questions](#)

[HADOOP-PR000007 Braindumps](#)