



# HADOOP-PR000007<sup>Q&As</sup>

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.passapply.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

Identify the tool best suited to import a portion of a relational database every day as files into HDFS, and generate Java classes to interact with that imported data?

- A. Oozie
- B. Flume
- C. Pig
- D. Hue
- E. Hive
- F. Sqoop
- G. fuse-dfs

Correct Answer: F

Sqoop ("SQL-to-Hadoop") is a straightforward command-line tool with the following capabilities:

Imports individual tables or entire databases to files in HDFS  
Generates Java classes to allow you to interact with your imported data  
Provides the ability to import from SQL databases straight into your Hive data warehouse

Note:

Data Movement Between Hadoop and Relational Databases  
Data can be moved between Hadoop and a relational database as a bulk data transfer, or relational tables can be accessed from within a MapReduce map function.

Note:

\* Cloudera's Distribution for Hadoop provides a bulk data transfer tool (i.e., Sqoop) that imports individual tables or entire databases into HDFS files. The tool also generates Java classes that support interaction with the imported data. Sqoop supports all relational databases over JDBC, and Quest Software provides a connector (i.e., OraOop) that has been optimized for access to data residing in Oracle databases.

Reference: <http://log.medcl.net/item/2011/08/hadoop-and-mapreduce-big-data-analytics-gartner/> (Data Movement between hadoop and relational databases, second paragraph)

---

### QUESTION 2

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 3

You want to Ingest log files Into HDFS, which tool would you use?

- A. HCatalog
- B. Flume
- C. Sqoop
- D. Ambari

Correct Answer: B

### QUESTION 4

You need to perform statistical analysis in your MapReduce job and would like to call methods in the Apache Commons Math library, which is distributed as a 1.3 megabyte Java archive (JAR) file. Which is the best way to make this library available to your MapReducer job at runtime?

- A. Have your system administrator copy the JAR to all nodes in the cluster and set its location in the HADOOP\_CLASSPATH environment variable before you submit your job.
- B. Have your system administrator place the JAR file on a Web server accessible to all cluster nodes and then set the HTTP\_JAR\_URL environment variable to its location.
- C. When submitting the job on the command line, specify the ?ibjars option followed by the JAR file path.
- D. Package your code and the Apache Commands Math library into a zip file named JobJar.zip

Correct Answer: C

Explanation: The usage of the jar command is like this,

Usage: `hadoop jar [mainClass] args...`

If you want the commons-math3.jar to be available for all the tasks you can do any one of these



1. Copy the jar file in \$HADOOP\_HOME/lib dir

2.

or

Use the generic option -libjars.

---

#### QUESTION 5

Which one of the following statements is FALSE regarding the communication between DataNodes and a federation of NameNodes in Hadoop 2.0?

- A. Each DataNode receives commands from one designated master NameNode.
- B. DataNodes send periodic heartbeats to all the NameNodes.
- C. Each DataNode registers with all the NameNodes.
- D. DataNodes send periodic block reports to all the NameNodes.

Correct Answer: A

[HADOOP-PR000007 PDF Dumps](#)

[HADOOP-PR000007 Exam Questions](#)

[HADOOP-PR000007 Braindumps](#)