



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.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

Which one of the following statements is true regarding a MapReduce job?

- A. The job's Partitioner shuffles and sorts all (key.value) pairs and sends the output to all reducers
- B. The default Hash Partitioner sends key value pairs with the same key to the same Reducer
- C. The reduce method is invoked once for each unique value
- D. The Mapper must sort its output of (key.value) pairs in descending order based on value

Correct Answer: A

QUESTION 2

Which one of the following classes would a Pig command use to store data in a table defined in HCatalog?

- A. org.apache.hcatalog.pig.HCatOutputFormat
- B. org.apache.hcatalog.pig.HCatStorer
- C. No special class is needed for a Pig script to store data in an HCatalog table
- D. Pig scripts cannot use an HCatalog table

Correct Answer: B

QUESTION 3

You want to perform analysis on a large collection of images. You want to store this data in HDFS and process it with MapReduce but you also want to give your data analysts and data scientists the ability to process the data directly from HDFS with an interpreted high-level programming language like Python. Which format should you use to store this data in HDFS?

- A. SequenceFiles
- B. Avro
- C. JSON
- D. HTML
- E. XML
- F. CSV

Correct Answer: B

Reference: Hadoop binary files processing introduced by image duplicates finder



QUESTION 4

In Hadoop 2.0, which one of the following statements is true about a standby NameNode? The Standby NameNode:

- A. Communicates directly with the active NameNode to maintain the state of the active NameNode.
- B. Receives the same block reports as the active NameNode.
- C. Runs on the same machine and shares the memory of the active NameNode.
- D. Processes all client requests and block reports from the appropriate DataNodes.

Correct Answer: B

QUESTION 5

You need to create a job that does frequency analysis on input data. You will do this by writing a Mapper that uses TextInputFormat and splits each value (a line of text from an input file) into individual characters. For each one of these characters, you will emit the character as a key and an InputWritable as the value. As this will produce proportionally more intermediate data than input data, which two resources should you expect to be bottlenecks?

- A. Processor and network I/O
- B. Disk I/O and network I/O
- C. Processor and RAM
- D. Processor and disk I/O

Correct Answer: B

[Latest HADOOP-PR000007 Dumps](#)

[HADOOP-PR000007 PDF Dumps](#)

[HADOOP-PR000007 VCE Dumps](#)