



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

All keys used for intermediate output from mappers must:

- A. Implement a splittable compression algorithm.
- B. Be a subclass of FileInputFormat.
- C. Implement WritableComparable.
- D. Override isSplittable.
- E. Implement a comparator for speedy sorting.

Correct Answer: C

Explanation: The MapReduce framework operates exclusively on pairs, that is, the framework views the input to the job as a set of pairs and produces a set of pairs as the output of the job, conceivably of different types.

The key and value classes have to be serializable by the framework and hence need to implement the Writable interface. Additionally, the key classes have to implement the WritableComparable interface to facilitate sorting by the framework.

Reference: MapReduce Tutorial

QUESTION 2

Given the following Pig commands: Which one of the following statements is true?

```
logevents = LOAD 'input/my.log';  
severe = FILTER logevents BY ($1 == 'severe' AND $2 >= 500);  
grouped = GROUP severe BY $2;  
DUMP grouped;
```

- A. The \$1 variable represents the first column of data in 'my.log'
- B. The \$1 variable represents the second column of data in 'my.log'
- C. The severe relation is not valid
- D. The grouped relation is not valid

Correct Answer: B

QUESTION 3

You want to run Hadoop jobs on your development workstation for testing before you submit them to your production cluster. Which mode of operation in Hadoop allows you to most closely simulate a production cluster while using a



single machine?

- A. Run all the nodes in your production cluster as virtual machines on your development workstation.
- B. Run the hadoop command with the `-jt local` and the `-fs file:///options`.
- C. Run the DataNode, TaskTracker, NameNode and JobTracker daemons on a single machine.
- D. Run simldoop, the Apache open-source software for simulating Hadoop clusters.

Correct Answer: C

QUESTION 4

You want to understand more about how users browse your public website, such as which pages they visit prior to placing an order. You have a farm of 200 web servers hosting your website. How will you gather this data for your analysis?

- A. Ingest the server web logs into HDFS using Flume.
- B. Write a MapReduce job, with the web servers for mappers, and the Hadoop cluster nodes for reduces.
- C. Import all users' clicks from your OLTP databases into Hadoop, using Sqoop.
- D. Channel these clickstreams into Hadoop using Hadoop Streaming.
- E. Sample the weblogs from the web servers, copying them into Hadoop using curl.

Correct Answer: A

QUESTION 5

What is a SequenceFile?

- A. A SequenceFile contains a binary encoding of an arbitrary number of homogeneous writable objects.
- B. A SequenceFile contains a binary encoding of an arbitrary number of heterogeneous writable objects.
- C. A SequenceFile contains a binary encoding of an arbitrary number of WritableComparable objects, in sorted order.
- D. A SequenceFile contains a binary encoding of an arbitrary number key-value pairs. Each key must be the same type. Each value must be same type.

Correct Answer: D

Explanation: SequenceFile is a flat file consisting of binary key/value pairs.

There are 3 different SequenceFile formats:

Uncompressed key/value records.



Record compressed key/value records - only '\values\' are compressed here. Block compressed key/value records - both keys and values are collected in '\blocks\' separately and compressed. The size of the '\block\' is configurable.

Reference: <http://wiki.apache.org/hadoop/SequenceFile>

[Latest HADOOP-PR000007 Dumps](#)

[HADOOP-PR000007 Study Guide](#)

[HADOOP-PR000007 Exam Questions](#)