



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

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 2

Which HDFS command copies an HDFS file named foo to the local filesystem as localFoo?

- A. `hadoop fs -get foo LocalFoo`
- B. `hadoop -cp foo LocalFoo`
- C. `hadoop fs -ls foo`
- D. `hadoop fs -put foo LocalFoo`

Correct Answer: A

QUESTION 3

In a large MapReduce job with m mappers and n reducers, how many distinct copy operations will there be in the sort/shuffle phase?

- A. $m \times n$ (i.e., m multiplied by n)
- B. n
- C. m
- D. $m+n$ (i.e., m plus n)
- E. m^n (i.e., m to the power of n)



Correct Answer: A

Explanation: A MapReduce job with m mappers and r reducers involves up to $m * r$ distinct copy operations, since each mapper may have intermediate output going to every reducer.

QUESTION 4

You want to count the number of occurrences for each unique word in the supplied input data. You've decided to implement this by having your mapper tokenize each word and emit a literal value 1, and then have your reducer increment a counter for each literal 1 it receives. After successfully implementing this, it occurs to you that you could optimize this by specifying a combiner. Will you be able to reuse your existing Reduces as your combiner in this case and why or why not?

- A. Yes, because the sum operation is both associative and commutative and the input and output types to the reduce method match.
- B. No, because the sum operation in the reducer is incompatible with the operation of a Combiner.
- C. No, because the Reducer and Combiner are separate interfaces.
- D. No, because the Combiner is incompatible with a mapper which doesn't use the same data type for both the key and value.
- E. Yes, because Java is a polymorphic object-oriented language and thus reducer code can be reused as a combiner.

Correct Answer: A

Explanation: Combiners are used to increase the efficiency of a MapReduce program. They are used to aggregate intermediate map output locally on individual mapper outputs. Combiners can help you reduce the amount of data that needs to be transferred across to the reducers. You can use your reducer code as a combiner if the operation performed is commutative and associative. The execution of combiner is not guaranteed, Hadoop may or may not execute a combiner. Also, if required it may execute it more than 1 times. Therefore your MapReduce jobs should not depend on the combiners execution.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What are combiners? When should I use a combiner in my MapReduce Job?

QUESTION 5

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

[HADOOP-PR000007 Study Guide](#)

[HADOOP-PR000007 Exam Questions](#)

[HADOOP-PR000007 Braindumps](#)