

# 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.pass2lead.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

In a MapReduce job with 500 map tasks, how many map task attempts will there be?

- A. It depends on the number of reduces in the job.
- B. Between 500 and 1000.
- C. At most 500.
- D. At least 500.
- E. Exactly 500.

Correct Answer: D

From Cloudera Training Course: Task attempt is a particular instance of an attempt to execute a task ?There will be at least as many task attempts as there are tasks ?If a task attempt fails, another will be started by the JobTracker ?Speculative execution can also result in more task attempts than completed tasks

---

### QUESTION 2

Assuming default settings, which best describes the order of data provided to a reducer's reduce method:

- A. The keys given to a reducer aren't in a predictable order, but the values associated with those keys always are.
- B. Both the keys and values passed to a reducer always appear in sorted order.
- C. Neither keys nor values are in any predictable order.
- D. The keys given to a reducer are in sorted order but the values associated with each key are in no predictable order

Correct Answer: D

Explanation: Reducer has 3 primary phases:

1.

Shuffle

The Reducer copies the sorted output from each Mapper using HTTP across the network.

2.

Sort

The framework merge sorts Reducer inputs by keys (since different Mappers may have output the same key).

The shuffle and sort phases occur simultaneously i.e. while outputs are being fetched they are merged.

SecondarySort

To achieve a secondary sort on the values returned by the value iterator, the application should extend the key with the

secondary key and define a grouping comparator. The keys will be sorted using the entire key, but will be grouped using the grouping comparator to decide which keys and values are sent in the same call to reduce.

### 3. Reduce

In this phase the reduce(Object, Iterable, Context) method is called for each in the sorted inputs. The output of the reduce task is typically written to a RecordWriter via TaskInputOutputContext.write

(Object, Object).

The output of the Reducer is not re-sorted.

Reference: org.apache.hadoop.mapreduce, Class

Reducer

---

### QUESTION 3

Can you use MapReduce to perform a relational join on two large tables sharing a key? Assume that the two tables are formatted as comma-separated files in HDFS.

- A. Yes.
- B. Yes, but only if one of the tables fits into memory
- C. Yes, so long as both tables fit into memory.
- D. No, MapReduce cannot perform relational operations.
- E. No, but it can be done with either Pig or Hive.

Correct Answer: A

Explanation: Note:

\*

Join Algorithms in MapReduce A) Reduce-side join B) Map-side join

C) In-memory join / Striped Striped variant variant / Memcached variant

\*

Which join to use? / In-memory join > map-side join > reduce-side join / Limitations of each? In-memory join: memory  
Map-side join: sort order and partitioning Reduce-side join: general purpose

---

### QUESTION 4

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. mn (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 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

[Latest HADOOP-PR000007 Dumps](#)

[HADOOP-PR000007 Practice Test](#)

[HADOOP-PR000007 Exam Questions](#)