

# APACHE-HADOOP-DEVELOPER<sup>Q&As</sup>

Hadoop 2.0 Certification exam for Pig and Hive Developer

## Pass Hortonworks APACHE-HADOOP-DEVELOPER Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.pass2lead.com/apache-hadoop-developer.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 best describes what the map method accepts and emits?

- A. It accepts a single key-value pair as input and emits a single key and list of corresponding values as output.
- B. It accepts a single key-value pairs as input and can emit only one key-value pair as output.
- C. It accepts a list key-value pairs as input and can emit only one key-value pair as output.
- D. It accepts a single key-value pairs as input and can emit any number of key-value pair as output, including zero.

Correct Answer: D

Explanation: public class Mapper extends Object Maps input key/value pairs to a set of intermediate key/value pairs.

Maps are the individual tasks which transform input records into a intermediate records. The transformed intermediate records need not be of the same type as the input records. A given input pair may map to zero or many output pairs.

Reference: org.apache.hadoop.mapreduce

Class Mapper

---

### QUESTION 2

Examine the following Pig commands:

```
customers = LOAD 'data' USING PigStorage(',')
AS (name:chararray,age:int,zip:int,receipts:double);
s = SAMPLE customers 0.2;
DUMP s;
```

Which one of the following statements is true?

- A. The SAMPLE command generates an "unexpected symbol" error
- B. Each MapReduce task will terminate after executing for 0.2 minutes
- C. The reducers will only output the first 20% of the data passed from the mappers
- D. A random sample of approximately 20% of the data will be output

Correct Answer: D

---

### QUESTION 3

How are keys and values presented and passed to the reducers during a standard sort and shuffle phase of MapReduce?

- A. Keys are presented to reducer in sorted order; values for a given key are not sorted.

- B. Keys are presented to reducer in sorted order; values for a given key are sorted in ascending order.
- C. Keys are presented to a reducer in random order; values for a given key are not sorted.
- D. Keys are presented to a reducer in random order; values for a given key are sorted in ascending order.

Correct Answer: A

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 4

You use the `hadoop fs -put` command to write a 300 MB file using and HDFS block size of 64 MB. Just after this command has finished writing 200 MB of this file, what would another user see when trying to access this file?

- A. They would see Hadoop throw an `ConcurrentFileAccessException` when they try to access this file.
- B. They would see the current state of the file, up to the last bit written by the command.
- C. They would see the current of the file through the last completed block.
- D. They would see no content until the whole file written and closed.

Correct Answer: C

---

#### QUESTION 5

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

[Latest APACHE-HADOOP-DEVELOPER Dumps](#)

[APACHE-HADOOP-DEVELOPER Exam Questions](#)

[APACHE-HADOOP-DEVELOPER Braindumps](#)