



# 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.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

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 2

MapReduce v2 (MRv2/YARN) is designed to address which two issues?

- A. Single point of failure in the NameNode.
- B. Resource pressure on the JobTracker.
- C. HDFS latency.
- D. Ability to run frameworks other than MapReduce, such as MPI.
- E. Reduce complexity of the MapReduce APIs.
- F. Standardize on a single MapReduce API.

Correct Answer: AB

Reference: Apache Hadoop YARN ?Conceptsand; Applications

---

### QUESTION 3

What does Pig provide to the overall Hadoop solution?

- A. Legacy language Integration with MapReduce framework
- B. Simple scripting language for writing MapReduce programs
- C. Database table and storage management services



D. C++ interface to MapReduce and data warehouse infrastructure

Correct Answer: B

---

#### QUESTION 4

Which one of the following statements is true about a Hive-managed table?

- A. Records can only be added to the table using the Hive INSERT command.
- B. When the table is dropped, the underlying folder in HDFS is deleted.
- C. Hive dynamically defines the schema of the table based on the FROM clause of a SELECT query.
- D. Hive dynamically defines the schema of the table based on the format of the underlying data.

Correct Answer: B

---

#### QUESTION 5

On a cluster running MapReduce v1 (MRv1), a TaskTracker heartbeats into the JobTracker on your cluster, and alerts the JobTracker it has an open map task slot.

What determines how the JobTracker assigns each map task to a TaskTracker?

- A. The amount of RAM installed on the TaskTracker node.
- B. The amount of free disk space on the TaskTracker node.
- C. The number and speed of CPU cores on the TaskTracker node.
- D. The average system load on the TaskTracker node over the past fifteen (15) minutes.
- E. The location of the InputSplit to be processed in relation to the location of the node.

Correct Answer: E

Explanation: The TaskTrackers send out heartbeat messages to the JobTracker, usually every few minutes, to reassure the JobTracker that it is still alive. These message also inform the JobTracker of the number of available slots, so the JobTracker can stay up to date with where in the cluster work can be delegated. When the JobTracker tries to find somewhere to schedule a task within the MapReduce operations, it first looks for an empty slot on the same server that hosts the DataNode containing the data, and if not, it looks for an empty slot on a machine in the same rack.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, How JobTracker schedules a task?

[Latest HADOOP-PR000007 Dumps](#)

[HADOOP-PR000007 Study Guide](#)

[HADOOP-PR000007 Braindumps](#)