

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.leads4pass.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

Which Hadoop component is responsible for managing the distributed file system metadata?

- A. NameNode
- B. Metanode
- C. DataNode
- D. NameSpaceManager

Correct Answer: A

QUESTION 2

You have written a Mapper which invokes the following five calls to the OutputCollector.collect method: `output.collect(new Text ("Apple"), new Text ("Red")) ;`

```
output.collect (new Text ("Banana"), new Text ("Yellow") ) ;
```

```
output.collect (new Text ("Apple"), new Text ("Yellow") ) ;
```

```
output.collect (new Text ("Cherry"), new Text ("Red") ) ;
```

```
output.collect (new Text ("Apple"), new Text ("Green") ) ;
```

How many times will the Reducer's reduce method be invoked?

- A. 6
- B. 3
- C. 1
- D. 0
- E. 5

Correct Answer: B

Explanation: reduce() gets called once for each [key, (list of values)] pair. To explain, let's say you called:

```
out.collect(new Text("Car"),new Text("Subaru");
```

```
out.collect(new Text("Car"),new Text("Honda");
```

```
out.collect(new Text("Car"),new Text("Ford");
```

```
out.collect(new Text("Truck"),new Text("Dodge");
```

```
out.collect(new Text("Truck"),new Text("Chevy");
```

Then reduce() would be called twice with the pairs reduce(Car,)

reduce(Truck,)

Reference: Mapper output.collect()?

QUESTION 3

Which project gives you a distributed, Scalable, data store that allows you random, realtime read/write access to hundreds of terabytes of data?

- A. HBase
- B. Hue
- C. Pig
- D. Hive
- E. Oozie
- F. Flume
- G. Sqoop

Correct Answer: A

Explanation: Use Apache HBase when you need random, realtime read/write access to your Big Data.

Note: This project's goal is the hosting of very large tables -- billions of rows X millions of columns -- atop clusters of commodity hardware. Apache HBase is an open-source, distributed, versioned, column-oriented store modeled after Google's Bigtable: A Distributed Storage System for Structured Data by Chang et al. Just as Bigtable leverages the distributed data storage provided by the Google File System, Apache HBase provides Bigtable-like capabilities on top of Hadoop and HDFS.

Features

Linear and modular scalability.

Strictly consistent reads and writes.

Automatic and configurable sharding of tables

Automatic failover support between RegionServers.

Convenient base classes for backing Hadoop MapReduce jobs with Apache HBase tables.

Easy to use Java API for client access.

Block cache and Bloom Filters for real-time queries. Query predicate push down via server side Filters Thrift gateway and a REST-ful Web service that supports XML, Protobuf, and binary data encoding options Extensible jruby-based

(JIRB) shell Support for exporting metrics via the Hadoop metrics subsystem to files or Ganglia; or via JMX

Reference: <http://hbase.apache.org/> (when would I use HBase? First sentence)

QUESTION 4

Which one of the following statements describes the relationship between the NodeManager and the ApplicationMaster?

- A. The ApplicationMaster starts the NodeManager in a Container
- B. The NodeManager requests resources from the ApplicationMaster
- C. The ApplicationMaster starts the NodeManager outside of a Container
- D. The NodeManager creates an instance of the ApplicationMaster

Correct Answer: D

QUESTION 5

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.

[HADOOP-PR000007 VCE Dumps](#)

[HADOOP-PR000007 Study Guide](#)

[HADOOP-PR000007 Exam Questions](#)