

# 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.leads4pass.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**

Your client application submits a MapReduce job to your Hadoop cluster. Identify the Hadoop daemon on which the Hadoop framework will look for an available slot schedule a MapReduce operation.

- A. TaskTracker
- B. NameNode
- C. DataNode
- D. JobTracker
- E. Secondary NameNode

Correct Answer: D

Explanation: JobTracker is the daemon service for submitting and tracking MapReduce jobs in Hadoop. There is only One Job Tracker process run on any hadoop cluster. Job Tracker runs on its own JVM process. In a typical production cluster its run on a separate machine. Each slave node is configured with job tracker node location. The JobTracker is single point of failure for the Hadoop MapReduce service. If it goes down, all running jobs are halted. JobTracker in Hadoop performs following actions(from Hadoop Wiki:)

Client applications submit jobs to the Job tracker. The JobTracker talks to the NameNode to determine the location of the data The JobTracker locates TaskTracker nodes with available slots at or near the data The JobTracker submits the work to the chosen TaskTracker nodes. The TaskTracker nodes are monitored. If they do not submit heartbeat signals often enough, they are deemed to have failed and the work is scheduled on a different TaskTracker. A TaskTracker will notify the JobTracker when a task fails. The JobTracker decides what to do then: it may resubmit the job elsewhere, it may mark that specific record as something to avoid, and it may even blacklist the TaskTracker as unreliable. When the work is completed, the JobTracker updates its status.

Client applications can poll the JobTracker for information.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What is a JobTracker in Hadoop? How many instances of JobTracker run on a Hadoop Cluster?

---

**QUESTION 2**

To process input key-value pairs, your mapper needs to load a 512 MB data file in memory. What is the best way to accomplish this?

- A. Serialize the data file, insert in it the JobConf object, and read the data into memory in the configure method of the mapper.
- B. Place the data file in the DistributedCache and read the data into memory in the map method of the mapper.
- C. Place the data file in the DataCache and read the data into memory in the configure method of the mapper.
- D. Place the data file in the DistributedCache and read the data into memory in the configure method of the mapper.

Correct Answer: C

---

## QUESTION 3

Which YARN component is responsible for monitoring the success or failure of a Container?

- A. ResourceManager
- B. ApplicationMaster
- C. NodeManager
- D. JobTracker

Correct Answer: A

---

## QUESTION 4

For each intermediate key, each reducer task can emit:

- A. As many final key-value pairs as desired. There are no restrictions on the types of those key-value pairs (i.e., they can be heterogeneous).
- B. As many final key-value pairs as desired, but they must have the same type as the intermediate key-value pairs.
- C. As many final key-value pairs as desired, as long as all the keys have the same type and all the values have the same type.
- D. One final key-value pair per value associated with the key; no restrictions on the type.
- E. One final key-value pair per key; no restrictions on the type.

Correct Answer: C

Reference: Hadoop Map-Reduce Tutorial; Yahoo! Hadoop Tutorial, Module 4: MapReduce

---

## 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.

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

[APACHE-HADOOP-DEVELOPER Practice Test](#)

[APACHE-HADOOP-DEVELOPER Braindumps](#)