

APACHE-HADOOP-DEVELOPER^{Q&As}

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

You want to count the number of occurrences for each unique word in the supplied input data. You've decided to implement this by having your mapper tokenize each word and emit a literal value 1, and then have your reducer increment a counter for each literal 1 it receives. After successfully implementing this, it occurs to you that you could optimize this by specifying a combiner. Will you be able to reuse your existing Reduces as your combiner in this case and why or why not?

- A. Yes, because the sum operation is both associative and commutative and the input and output types to the reduce method match.
- B. No, because the sum operation in the reducer is incompatible with the operation of a Combiner.
- C. No, because the Reducer and Combiner are separate interfaces.
- D. No, because the Combiner is incompatible with a mapper which doesn't use the same data type for both the key and value.
- E. Yes, because Java is a polymorphic object-oriented language and thus reducer code can be reused as a combiner.

Correct Answer: A

Explanation: Combiners are used to increase the efficiency of a MapReduce program. They are used to aggregate intermediate map output locally on individual mapper outputs. Combiners can help you reduce the amount of data that needs to be transferred across to the reducers. You can use your reducer code as a combiner if the operation performed is commutative and associative. The execution of combiner is not guaranteed, Hadoop may or may not execute a combiner. Also, if required it may execute it more than 1 times. Therefore your MapReduce jobs should not depend on the combiners execution.

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, What are combiners? When should I use a combiner in my MapReduce Job?

QUESTION 2

Given the following Hive command:

```
INSERT OVERWRITE TABLE mytable SELECT * FROM myothertable;
```

Which one of the following statements is true?

- A. The contents of myothertable are appended to mytable
- B. Any existing data in mytable will be overwritten
- C. A new table named mytable is created, and the contents of myothertable are copied into mytable
- D. The statement is not a valid Hive command

Correct Answer: B

QUESTION 3

You wrote a map function that throws a runtime exception when it encounters a control character in input data. The input supplied to your mapper contains twelve such characters total, spread across five file splits. The first four file splits each have two control characters and the last split has four control characters.

Identify the number of failed task attempts you can expect when you run the job with `mapred.max.map.attempts` set to 4:

- A. You will have forty-eight failed task attempts
- B. You will have seventeen failed task attempts
- C. You will have five failed task attempts
- D. You will have twelve failed task attempts
- E. You will have twenty failed task attempts

Correct Answer: E

Explanation: There will be four failed task attempts for each of the five file splits.

Note:

When the jobtracker is notified of a task attempt that has failed (by the tasktracker's heartbeat call), it will reschedule execution of the task. The jobtracker will try to avoid rescheduling the task on a tasktracker where it has previously failed. Furthermore, if a task fails four times (or more), it will not be retried further. This value is configurable: the maximum number of attempts to run a task is controlled by the `mapred.map.max.attempts` property for map tasks and `mapred.reduce.max.attempts` for reduce tasks. By default, if any task fails four times (or whatever the maximum number of attempts is configured to), the whole job fails.

QUESTION 4

Which HDFS command uploads a local file X into an existing HDFS directory Y?

- A. `hadoop scp X Y`
- B. `hadoop fs -localPut X Y`
- C. `hadoop fs-put X Y`
- D. `hadoop fs -get X Y`

Correct Answer: C

QUESTION 5

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

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

[APACHE-HADOOP-DEVELOPER Study Guide](#)

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