

## CCD-410<sup>Q&As</sup>

Cloudera Certified Developer for Apache Hadoop (CCDH)

### Pass Cloudera CCD-410 Exam with 100% Guarantee

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

<https://www.leads4pass.com/ccd-410.html>

100% Passing Guarantee  
100% Money Back Assurance

Following Questions and Answers are all new published by Cloudera  
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



**QUESTION 1**

You have the following key-value pairs as output from your Map task:

(the, 1) (fox, 1) (faster, 1) (than, 1) (the, 1) (dog, 1)

How many keys will be passed to the Reducer's reduce method?

- A. Six
- B. Five
- C. Four
- D. Two
- E. One
- F. Three

Correct Answer: B

Only one key value pair will be passed from the two (the, 1) key value pairs.

---

**QUESTION 2**

You've written a MapReduce job that will process 500 million input records and generated 500 million keyvalue pairs. The data is not uniformly distributed. Your MapReduce job will create a significant amount of intermediate data that it needs to transfer between mappers and reduces which is a potential bottleneck. A custom implementation of which interface is most likely to reduce the amount of intermediate data transferred across the network?

- A. Partitioner
- B. OutputFormat
- C. WritableComparable
- D. Writable
- E. InputFormat
- F. Combiner

Correct Answer: F

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.

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

---

**QUESTION 3**

In the reducer, the MapReduce API provides you with an iterator over Writable values. What does calling the next () method return?

- A. It returns a reference to a different Writable object time.
- B. It returns a reference to a Writable object from an object pool.
- C. It returns a reference to the same Writable object each time, but populated with different data.
- D. It returns a reference to a Writable object. The API leaves unspecified whether this is a reused object or a new object.
- E. It returns a reference to the same Writable object if the next value is the same as the previous value, or a new Writable object otherwise.

Correct Answer: C

Calling Iterator.next() will always return the SAME EXACT instance of IntWritable, with the contents of that instance replaced with the next value.

Reference: manipulating iterator in mapreduce

---

**QUESTION 4**

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

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 5

Determine which best describes when the reduce method is first called in a MapReduce job?

- A. Reducers start copying intermediate key-value pairs from each Mapper as soon as it has completed. The programmer can configure in the job what percentage of the intermediate data should arrive before the reduce method begins.
- B. Reducers start copying intermediate key-value pairs from each Mapper as soon as it has completed. The reduce method is called only after all intermediate data has been copied and sorted.
- C. Reduce methods and map methods all start at the beginning of a job, in order to provide optimal performance for map-only or reduce-only jobs.
- D. Reducers start copying intermediate key-value pairs from each Mapper as soon as it has completed. The reduce method is called as soon as the intermediate key-value pairs start to arrive.

Correct Answer: B

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers , When is the reducers are started in a MapReduce job?

[CCD-410 PDF Dumps](#)

[CCD-410 Study Guide](#)

[CCD-410 Exam Questions](#)