

# 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.pass2lead.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

For each input key-value pair, mappers can emit:

- A. As many intermediate key-value pairs as designed. There are no restrictions on the types of those key-value pairs (i.e., they can be heterogeneous).
- B. As many intermediate key-value pairs as designed, but they cannot be of the same type as the input key-value pair.
- C. One intermediate key-value pair, of a different type.
- D. One intermediate key-value pair, but of the same type.
- E. As many intermediate key-value pairs as designed, as long as all the keys have the same types and all the values have the same type.

Correct Answer: E

Mapper maps input key/value pairs to a set of intermediate key/value pairs.

Maps are the individual tasks that transform input records into intermediate records. The transformed intermediate records do not need to be of the same type as the input records. A given input pair may map to zero or many output pairs.

Reference: Hadoop Map-Reduce Tutorial

---

### QUESTION 2

Which best describes what the map method accepts and emits?

- A. It accepts a single key-value pair as input and emits a single key and list of corresponding values as output.
- B. It accepts a single key-value pairs as input and can emit only one key-value pair as output.
- C. It accepts a list key-value pairs as input and can emit only one key-value pair as output.
- D. It accepts a single key-value pairs as input and can emit any number of key-value pair as output, including zero.

Correct Answer: D

```
public class Mapper
```

```
extends Object
```

Maps input key/value pairs to a set of intermediate key/value pairs.

Maps are the individual tasks which transform input records into a intermediate records. The transformed intermediate records need not be of the same type as the input records. A given input pair may map to zero or many output pairs.

Reference: org.apache.hadoop.mapreduce

```
Class Mapper
```

---

**QUESTION 3**

Can you use MapReduce to perform a relational join on two large tables sharing a key? Assume that the two tables are formatted as comma-separated files in HDFS.

- A. Yes.
- B. Yes, but only if one of the tables fits into memory
- C. Yes, so long as both tables fit into memory.
- D. No, MapReduce cannot perform relational operations.
- E. No, but it can be done with either Pig or Hive.

Correct Answer: A

Note:

\*

Join Algorithms in MapReduce A) Reduce-side join B) Map-side join C) In-memory join / Striped Striped variant variant / Memcached variant

\*

Which join to use? / In-memory join > map-side join > reduce-side join / Limitations of each? In-memory join: memory  
Map-side join: sort order and partitioning

Reduce-side join: general purpose

---

**QUESTION 4**

Identify the utility that allows you to create and run MapReduce jobs with any executable or script as the mapper and/or the reducer?

- A. Oozie
- B. Sqoop
- C. Flume
- D. Hadoop Streaming
- E. mapred

Correct Answer: D

Hadoop streaming is a utility that comes with the Hadoop distribution. The utility allows you to create and run Map/Reduce jobs with any executable or script as the mapper and/or the reducer.

Reference: <http://hadoop.apache.org/common/docs/r0.20.1/streaming.html> (Hadoop Streaming, second sentence)

---

**QUESTION 5**

Which describes how a client reads a file from HDFS?

- A. The client queries the NameNode for the block location(s). The NameNode returns the block location (s) to the client. The client reads the data directory off the DataNode(s).
- B. The client queries all DataNodes in parallel. The DataNode that contains the requested data responds directly to the client. The client reads the data directly off the DataNode.
- C. The client contacts the NameNode for the block location(s). The NameNode then queries the DataNodes for block locations. The DataNodes respond to the NameNode, and the NameNode redirects the client to the DataNode that holds the requested data block(s). The client then reads the data directly off the DataNode.
- D. The client contacts the NameNode for the block location(s). The NameNode contacts the DataNode that holds the requested data block. Data is transferred from the DataNode to the NameNode, and then from the NameNode to the client.

Correct Answer: A

Reference: 24 Interview Questions and Answers for Hadoop MapReduce developers, How the Client communicates with HDFS?

[CCD-410 Practice Test](#)

[CCD-410 Exam Questions](#)

[CCD-410 Braindumps](#)