

# CCD-410 - Cloudera Certified Developer for Apache Hadoop (CCDH)

<http://www.certleader.com/CCD-410-dumps.html>



1. When is the earliest point at which the reduce method of a given Reducer can be called?

- A. As soon as at least one mapper has finished processing its input split.
- B. As soon as a mapper has emitted at least one record.
- C. Not until all mappers have finished processing all records.
- D. It depends on the InputFormat used for the job.

**Answer: C**

2. 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.

**Answer: A**

3. You are developing a combiner that takes as input Text keys, IntWritable values, and emits Text keys, IntWritable values. Which interface should your class implement?

- A. Combiner <Text, IntWritable, Text, IntWritable>
- B. Mapper <Text, IntWritable, Text, IntWritable>
- C. Reducer <Text, Text, IntWritable, IntWritable>
- D. Reducer <Text, IntWritable, Text, IntWritable>
- E. Combiner <Text, Text, IntWritable, IntWritable>

**Answer: D**

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

**Answer: D**

5. How are keys and values presented and passed to the reducers during a standard sort and shuffle phase of MapReduce?

- A. Keys are presented to reducer in sorted order; values for a given key are not sorted.
- B. Keys are presented to reducer in sorted order; values for a given key are sorted in ascending order.
- C. Keys are presented to a reducer in random order; values for a given key are not sorted.
- D. Keys are presented to a reducer in random order; values for a given key are sorted in ascending order.

**Answer: A**

6. Assuming default settings, which best describes the order of data provided to a reducer's reduce method:

- A. The keys given to a reducer aren't in a predictable order, but the values associated with those keys always are.
- B. Both the keys and values passed to a reducer always appear in sorted order.
- C. Neither keys nor values are in any predictable order.
- D. The keys given to a reducer are in sorted order but the values associated with each key are in no predictable order

**Answer: D**

7. 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 totals, 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

**Answer: E**

8. You want to populate an associative array in order to perform a map-side join. You've decided to put this information in a text file, place that file into the DistributedCache and read it in your Mapper before any records are processed.

Identify which method in the Mapper you should use to implement code for reading the file and populating the associative array?

- A. combine
- B. map
- C. init
- D. configure

**Answer: D**

9. You've written a MapReduce job that will process 500 million input records and generated 500 million key-value 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

**Answer: F**

10. 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.

**Answer: A**

11. You have just executed a MapReduce job. Where is intermediate data written to after being emitted from the Mapper's map method?

- A. Intermediate data is streamed across the network from Mapper to the Reducer and is never written to disk.
- B. Into in-memory buffers on the TaskTracker node running the Mapper that spill over and are written into HDFS.
- C. Into in-memory buffers that spill over to the local file system of the TaskTracker node running the Mapper.
- D. Into in-memory buffers that spill over to the local file system (outside HDFS) of the TaskTracker node running the Reducer
- E. Into in-memory buffers on the TaskTracker node running the Reducer that spill over and are written into HDFS.

**Answer: C**

12. You want to understand more about how users browse your public website, such as which pages they visit prior to placing an order. You have a farm of 200 web servers hosting your website. How will you gather this data for your analysis?

- A. Ingest the server web logs into HDFS using Flume.
- B. Write a MapReduce job, with the web servers for mappers, and the Hadoop cluster nodes for reducers.
- C. Import all users' clicks from your OLTP databases into Hadoop, using Sqoop.
- D. Channel these clickstreams into Hadoop using Hadoop Streaming.

E. Sample the weblogs from the web servers, copying them into Hadoop using curl.

**Answer: A**

13. 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.

**Answer: BD**

14. You need to run the same job many times with minor variations. Rather than hardcoding all job configuration options in your driver code, you've decided to have your Driver subclass `org.apache.hadoop.conf.Configured` and implement the `org.apache.hadoop.util.Tool` interface.

Identify which invocation correctly passes `mapred.job.name` with a value of `Example` to Hadoop?

- A. `hadoop "mapred.job.name=Example" MyDriver input output`
- B. `hadoop MyDriver mapred.job.name=Example input output`
- C. `hadoop MyDriver -D mapred.job.name=Example input output`
- D. `hadoop setproperty mapred.job.name=Example MyDriver input output`
- E. `hadoop setproperty ("mapred.job.name=Example") MyDriver input output`

**Answer: C**

15. You are developing a MapReduce job for sales reporting. The mapper will process input keys representing the year (`IntWritable`) and input values representing product identifiers (`Text`).

Identify what determines the data types used by the Mapper for a given job.

- A. The key and value types specified in the `JobConf.setMapInputKeyClass` and `JobConf.setMapInputValuesClass` methods

- B. The data types specified in HADOOP\_MAP\_DATATYPES environment variable
- C. The mapper-specification.xml file submitted with the job determine the mapper's input key and value types.
- D. The InputFormat used by the job determines the mapper's input key and value types.

**Answer: D**

## Thank You for Trying Our Product

\* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

\* One year free update

You can enjoy free update one year. 24x7 online support.

\* Trusted by Millions

We currently serve more than 30,000,000 customers.

\* Shop Securely

All transactions are protected by VeriSign!

**100% Pass Your CCD-410 Exam with Our Prep Materials Via below:**

<http://www.certleader.com/CCD-410-dumps.html>