CCD-410^{Q&As}

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.geekcert.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

In a MapReduce job, the reducer receives all values associated with same key. Which statement best describes the ordering of these values?

A. The values are in sorted order.

B. The values are arbitrarily ordered, and the ordering may vary from run to run of the same MapReduce job.

C. The values are arbitrary ordered, but multiple runs of the same MapReduce job will always have the same ordering.

D. Since the values come from mapper outputs, the reducers will receive contiguous sections of sorted values.

Correct Answer: B

Note:

*

 Input to the Reducer is the sorted output of the mappers.

*

 The framework calls the application\\\'s Reduce function once for each unique key in the sorted order.

*

 Example:

For the given sample input the first map emits:

The second map emits:

QUESTION 2

What is the disadvantage of using multiple reducers with the default HashPartitioner and distributing your workload across you cluster?

A. You will not be able to compress the intermediate data.

B. You will longer be able to take advantage of a Combiner.

C. By using multiple reducers with the default HashPartitioner, output files may not be in globally sorted order.

D. There are no concerns with this approach. It is always advisable to use multiple reduces.

Correct Answer: C

Multiple reducers and total ordering

If your sort job runs with multiple reducers (either because mapreduce.job.reduces in mapred- site.xml has been set to a number larger than 1, or because you\\'ve used the -r option to specify the number of reducers on the command-line), then by default Hadoop will use the HashPartitioner to distribute records across the reducers. Use of the HashPartitioner means that you can\\'t concatenate your output files to create a single sorted output file. To do this you\\'ll need total ordering,

Reference: Sorting text files with MapReduce

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 totals, spread across five file splits. The first four file splits each have two control characters and the last split has four control characters.

Indentify 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 4

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. mXn (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

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.

QUESTION 5

You have written a Mapper which invokes the following five calls to the OutputColletor.collect method:

output.collect (new Text ("Apple"), new Text ("Red")) ;

output.collect (new Text ("Banana"), new Text ("Yellow")) ; output.collect (new Text ("Apple"), new Text

("Yellow")) ; output.collect (new Text ("Cherry"), new Text ("Red")) ;

output.collect (new Text ("Apple"), new Text ("Green")) ;

How many times will the Reducer\\'s reduce method be invoked?

A. 6

B. 3

C. 1

D. 0

E. 5

Correct Answer: B

reduce() gets called once for each [key, (list of values)] pair. To explain, let\\'s say you called:

out.collect(new Text("Car"),new Text("Subaru");

out.collect(new Text("Car"),new Text("Honda");

out.collect(new Text("Car"),new Text("Ford");

out.collect(new Text("Truck"),new Text("Dodge");

out.collect(new Text("Truck"),new Text("Chevy");

Then reduce() would be called twice with the pairs

reduce(Car,)

reduce(Truck,)

Reference: Mapper output.collect()?

[CCD-410 PDF Dumps](https://www.geekcert.com/ccd-410.html)   [CCD-410 VCE Dumps](https://www.geekcert.com/ccd-410.html)   [CCD-410 Study Guide](https://www.geekcert.com/ccd-410.html)