



Free Questions for HDPCD by dumpssheet

Shared by Riddle on 06-06-2022

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

You want to Ingest log files Into HDFS, which tool would you use?

Options:

A- HCatalog

B- Flume

C- Sqoop

D- Ambari

Answer:

B

Question 2

Question Type: MultipleChoice

Which of the following tool was designed to import data from a relational database into HDFS?

Options:

A- HCatalog

B- Sqoop

C- Flume

D- Ambari

Answer:

B

Question 3

Question Type: MultipleChoice

Which HDFS command copies an HDFS file named foo to the local filesystem as localFoo?

Options:

- A- `hadoop fs -get foo LocalFoo`
- B- `hadoop -cp foo LocalFoo`
- C- `hadoop fs -ls foo`
- D- `hadoop fs -put foo LocalFoo`

Answer:

A

Question 4

Question Type: MultipleChoice

Which HDFS command displays the contents of the file x in the user's HDFS home directory?

Options:

- A- `hadoop fs -ls x`

B- hdfs fs -get x

C- hadoop fs -cat x

D- hadoop fs -cp x

Answer:

C

Question 5

Question Type: MultipleChoice

What is a SequenceFile?

Options:

A- A SequenceFile contains a binary encoding of an arbitrary number of homogeneous writable objects.

B- A SequenceFile contains a binary encoding of an arbitrary number of heterogeneous writable objects.

C- A SequenceFile contains a binary encoding of an arbitrary number of WritableComparable objects, in sorted order.

D- A SequenceFile contains a binary encoding of an arbitrary number key-value pairs. Each key must be the same type. Each value

must be same type.

Answer:

D

Explanation:

SequenceFile is a flat file consisting of binary key/value pairs.

There are 3 different SequenceFile formats:

Uncompressed key/value records.

Record compressed key/value records - only 'values' are compressed here.

Block compressed key/value records - both keys and values are collected in 'blocks' separately and compressed. The size of the 'block' is configurable.

Question 6

Question Type: MultipleChoice

Your cluster's HDFS block size is 64MB. You have a directory containing 100 plain text files, each of which is 100MB in size. The InputFormat for your job is TextInputFormat. Determine how many Mappers will run?

Options:

- A- 64
- B- 100
- C- 200
- D- 640

Answer:

C

Explanation:

Each file would be split into two as the block size (64 MB) is less than the file size (100 MB), so 200 mappers would be running.

Note:

If you're not compressing the files then Hadoop will process your large files (say 10G), with a number of mappers related to the block size of the file.

Say your block size is 64M, then you will have ~160 mappers processing this 10G file ($160 \times 64 \approx 10G$). Depending on how CPU intensive your mapper logic is, this might be an

acceptable blocks size, but if you find that your mappers are executing in sub minute times, then you might want to increase the work done by each mapper (by increasing the block size to 128, 256, 512m - the actual size depends on how you intend to process the data).

Question 7

Question Type: MultipleChoice

You want to run Hadoop jobs on your development workstation for testing before you submit them to your production cluster. Which mode of operation in Hadoop allows you to most closely simulate a production cluster while using a single machine?

Options:

- A-** Run all the nodes in your production cluster as virtual machines on your development workstation.
- B-** Run the hadoop command with the `--jt local` and the `--fs file:///options`.
- C-** Run the DataNode, TaskTracker, NameNode and JobTracker daemons on a single machine.
- D-** Run simldoop, the Apache open-source software for simulating Hadoop clusters.

Answer:

C

To Get Premium Files for HDPCD Visit

<https://www.p2pexams.com/products/hdpcd>

For More Free Questions Visit

<https://www.p2pexams.com/hortonworks/pdf/hdpcd>

