Vendor: IBM

Exam Code: 000-M97

Exam Name: IBM InfoSphere Big Insights Techncial
Mastery

Version: DEMO

QUESTION NO: 1

To help firms analyze text-based data, BigInsights provides which of the following?

A. A runtime library of text annotators that can be invoked from Pig.

B. A runtime library of text annotators that can be invoked from HQL (Hive).

C. A runtime library of text annotators that can be invoked from Jaql.

D. A runtime library of text annotators that can be invoked from Flume.

Answer: C

QUESTION NO: 2

Hadoop is configured with FAIR scheduler. We have three queues defined : production, development, and QA. The production queue has a minimum share defined as 5. Other queues do not have minimum shares defined. QA has a weight defined as 2. Other queues do not have weight associated with them. Assuming that Production queue has demand of 100 tasks, Development of 30 tasks, and QA of 25 tasks with total of 30 tasks available in JobTracker for all queues, how many tasks per queue will be allocated?

A. Production: 5, Development: 10, QA: 15

B. Production: 10, Development: 10, QA: 10

C. Production: 8, Development: 8, QA: 14

D. Production: 6, Development: 6, QA: 18

Answer: C

QUESTION NO: 3

Which of the following statements is correct?

A. All nodes in the cluster should have the same hardware.

B. Master nodes need to have more reliable hardware than slave nodes.

C. Slave nodes need to have more reliable hardware than master nodes.

D. All nodes in the cluster must have reliable hardware except for disks.

Answer: B

QUESTION NO: 4

What mode is best suited for application developers?

A. Standalone (local) mode.

B. Pseudo-distributed mode.

C. Fully-distributed mode.

D. It does not matter, all are about the same.

Answer: B

QUESTION NO: 5

What will happen when a NameNode runs out of memory?

A. It will crash.

B. It will start swapping data to disk as long as swappiness is greater than 0.

C. It will move part of the workload to secondary name node.

D. It will start delaying some jobs to accommodate currently running jobs.

Answer: A

QUESTION NO: 6

Before it can be used, a new HDFS installation needs to be formatted. Which command will correctly format HDFS?

A. hadoop format

B. hadoop namenode ormathadoop namenode ?ormat

C. hadoop -format all

D. hadoop namenode ormathadoop namenode ?ormat

hadoop datanodes ormathadoop datanodes ?ormat

Answer: B

QUESTION NO: 7

If an administrator executes the following command: "hadoop balancer 15" what will happen to the cluster?

A. Hadoop will balance 15% of all of the nodes.

B. Hadoop will attempt to reserve 15% free space on each disk in the cluster.

C. Hadoop will attempt to rebalance the cluster so that all DataNodes are within 15% utilization of each other.

D. Hadoop will wait for 15 seconds before starting rebalancing.

Answer: C

QUESTION NO: 8

Which statement regarding the size of file blocks (splits) is correct?

A. They are determined by a configuration parameter.

B. They are always equal within a file.

C. They are 1GB by default.

D. They are 128MB by default.

Answer: A

QUESTION NO: 9

How can you can invoke Jaql query language?

A. Using Jaqlshell, Eclipse, and Embedding in Java and Python

B. From SQL

C. Embedding in machine languages

D. From XML query languages

Answer: A

QUESTION NO: 10

Your log dataset is an array of records. How do you select the record from sender 101 and project out only the message field? Example: data = [(from: 101, to: 102, msg: "Hello"), (from: 103, to: 104, msg: "World!"), (from: 105, to:106, msg: "Hello World")];

A. data -> filter $.from == 101

B. data->filter $.from == 101 -> expand (message: $.msg);

C. data->filter $.from == 101 -> transform (message: $.msg);

D. data->transform (message: $msg -> filter $.from == 101);

Answer: C

QUESTION NO: 11

What is good fit for HDFS?

A. Lots of small files.

B. Applications requiring low latency data access.

C. Multiple writers accessing the same file.

D. Applications requiring high throughput of data.

Answer: D

QUESTION NO: 12

What types of work are a good fit for Jaql?

A. Business Intelligence reporting.

B. Semi-structured data analysis.

C. Web service calls.

D. Information extraction for emails.

Answer: B

QUESTION NO: 13

Complete the following statement: "dfs.hosts" configuration property point to the file containing the list of machines that

A. Can access the cluster. If a machine is not on the list, it will be denied access to the cluster.

B. Are blocked from accessing the cluster. Anyone can access the cluster except for machines specified in the list.

C. Can directly access a DataNode. If the machine is not on the list, it will be denied access.

D. Belong to the cluster. Any machine not on the list is not considered part of the cluster and will not be used for any of the Hadoop operations.

Answer: A

QUESTION NO: 14

What compression format is best to use for large files without doing anything special in map jobs in order for map jobs to have data locality in HDFS?

A. GZIP

B. ZIP

C. LZO

D. BZIP2

Answer: D

QUESTION NO: 15

If a user enters following command:

what does the "3" indicate?

A. Expected replication factor for this file.

B. Actual number of times blocks in file are replicated.

C. This is third version of the file stored with this name.

D. The file was stored in 3 blocks.

Answer: A