

## Question: 1

Which statement is TRUE concerning optimizing the load performance?

- A. You can improve the performance by increasing the number of map tasks assigned to the load
- B. When loading large files the number of files that you load does not impact the performance of the LOAD HADOOP statement
- C. You can improve the performance by decreasing the number of map tasks that are assigned to the load and adjusting the heap size
- D. It is advantageous to run the LOAD HADOOP statement directly pointing to large files located in the host file system as opposed to copying the files to the DFS prior to load

**Answer: A**

Reference:

[https://www.ibm.com/support/knowledgecenter/en/SSCRJT\\_5.0.3/com.ibm.swg.im.bigsql.doc/doc/bigsql\\_loadperf.html](https://www.ibm.com/support/knowledgecenter/en/SSCRJT_5.0.3/com.ibm.swg.im.bigsql.doc/doc/bigsql_loadperf.html)

## Question: 2

Which of the following statements are TRUE regarding the use of Data Click to load data into BigInsights? (Choose two.)

- A. Big SQL cannot be used to access the data moved in by Data Click because the data is in Hive
- B. You must import metadata for all sources and targets that you want to make available for Data Click activities
- C. Connections from the relational database source to HDFS are discovered automatically from within Data Click
- D. Hive tables are automatically created every time you run an activity that moves data from a relational database into HDFS
- E. HBase tables are automatically created every time you ran an activity that moves data from a relational database into HDFS

**Answer: CE**

Reference:

[https://www.ibm.com/support/knowledgecenter/en/SSZJPZ\\_11.3.0/com.ibm.swg.im.iis.dataclick.doc/topics/hivetables.html](https://www.ibm.com/support/knowledgecenter/en/SSZJPZ_11.3.0/com.ibm.swg.im.iis.dataclick.doc/topics/hivetables.html)

## Question: 3

Which of the following statements regarding importing streaming data from InfoSphere Streams into Hadoop is TRUE?

- A. InfoSphere Streams can both read from and write data to HDFS
- B. The Streams Big Data toolkit operators that interface with HDFS uses Apache Flume to integrate with Hadoop
- C. Streams applications never need to be concerned with making the data schemas consistent with those on Hadoop
- D. Big SQL can be used to preprocess the data as it flows through InfoSphere Streams before the data lands in HDFS

**Answer: D**

### Question: 4

Which of the following is TRUE about storing an Apache Spark object in serialized form?

- A. It is advised to use Java serialization over Kryo serialization
- B. Storing the object in serialized form will lead to faster access times
- C. Storing the object in serialized form will lead to slower access times
- D. All of the above

**Answer: B**

Reference:

<https://spark.apache.org/docs/latest/rdd-programming-guide.html>

### Question: 5

Which ONE of the following statements regarding Sqoop is TRUE?

- A. HBase is not supported as an import target
- B. Data imported using Sqoop is always written to a single Hive partition
- C. Sqoop can be used to retrieve rows newer than some previously imported set of rows
- D. Sqoop can only append new rows to a database table when exporting back to a database

**Answer: C**

Reference:

<https://sqoop.apache.org/docs/1.4.1-incubating/SqoopUserGuide.html>

### Question: 6

Which one of the following statements is TRUE?

- A. Spark SQL does not support HiveQL
- B. Spark SQL does not support ANSI SQL

- C. To use Spark with Hive, HiveQL queries have to be rewritten in Scala
- D. Spark SQL allows relational queries expressed in SQL, HiveQL, or Scala

**Answer: D**

Reference:

<https://spark.apache.org/docs/1.2.1/sql-programming-guide.html> (overview)

### Question: 7

Which of the following statements regarding Big SQL is TRUE?

- A. Big SQL doesn't support stored procedures
- B. Big SQL can be deployed on a subset of data nodes in the BigInsights cluster
- C. Big SQL provides a SQL-on-Hadoop environment based on map reduce
- D. Only tables created or loaded via Big SQL can be accessed via Big SQL

**Answer: B**

Reference:

[https://books.google.com.pk/books?id=t13nCQAAQBAJ&pg=PA3&lpg=PA3&dq=Big+SQL+can+be+deployed+on+a+subset+of+data+nodes+in+the+BigInsights+cluster&source=bl&ots=RBbad0Xkel&sig=pMgmgDNLGUrkvOSXoVBj64xTMgk&hl=en&sa=X&redir\\_esc=y#v=onepage&q=Big%20SQL%20can%20be%20deployed%20on%20a%20subset%20of%20data%20nodes%20in%20the%20BigInsights%20cluster&f=false](https://books.google.com.pk/books?id=t13nCQAAQBAJ&pg=PA3&lpg=PA3&dq=Big+SQL+can+be+deployed+on+a+subset+of+data+nodes+in+the+BigInsights+cluster&source=bl&ots=RBbad0Xkel&sig=pMgmgDNLGUrkvOSXoVBj64xTMgk&hl=en&sa=X&redir_esc=y#v=onepage&q=Big%20SQL%20can%20be%20deployed%20on%20a%20subset%20of%20data%20nodes%20in%20the%20BigInsights%20cluster&f=false)

### Question: 8

The number of partitions created by DynamicPartitions in Hive can be controlled by which of the following?

- A. hive.exec.max.dynamic.partitions.pernode
- B. hive.exec.max.dynamic.partitions
- C. hive.exec.max.created.files
- D. All of the above

**Answer: A**

Reference:

<https://resources.zaloni.com/blog/partitioning-in-hive>

### Question: 9

Which of the following Jaq operators groups one or more arrays based on key values and applies an

aggregate expression?

- A. join
- B. group
- C. expand
- D. transform

**Answer: B**

Reference:

[https://books.google.com.pk/books?id=Qj-5BQAAQBAJ&pg=PA174&lpg=PA174&dq=Jaq+operators+groups+one+or+more+arrays+based+on+key+values+and+applies+an+aggregate+expression&source=bl&ots=zobr8AZzWy&sig=ZRCIH9ee4Un3Aam1hX8TzxfQI&hl=en&sa=X&redir\\_esc=y#v=onepage&q=Jaq%20operators%20groups%20one%20or%20more%20arrays%20based%20on%20key%20values%20and%20applies%20an%20aggregate%20expression&f=false](https://books.google.com.pk/books?id=Qj-5BQAAQBAJ&pg=PA174&lpg=PA174&dq=Jaq+operators+groups+one+or+more+arrays+based+on+key+values+and+applies+an+aggregate+expression&source=bl&ots=zobr8AZzWy&sig=ZRCIH9ee4Un3Aam1hX8TzxfQI&hl=en&sa=X&redir_esc=y#v=onepage&q=Jaq%20operators%20groups%20one%20or%20more%20arrays%20based%20on%20key%20values%20and%20applies%20an%20aggregate%20expression&f=false)

## Question: 10

Which of the following are CRUD operations available in HBase? (Choose two.)

- A. HTable.Put
- B. HTable.Read
- C. HTable.Delete
- D. HTable.Update
- E. HTable.Remove

**Answer: AC**

Reference:

[https://www.tutorialspoint.com/hbase/hbase\\_client\\_api.htm](https://www.tutorialspoint.com/hbase/hbase_client_api.htm)