



C2090-101^{Q&As}

IBM Big Data Engineer

Pass IBM C2090-101 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.passapply.com/c2090-101.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by IBM Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





QUESTION 1

To improve Big SQL query performance one should specify the INPUTFORMAT and OUTPUTFORMAT during which of the following?

- A. SELECT statements
- B. When the file is initially stored in HDFS
- C. CREATE and ALTER table statements
- D. While the files are accessed in your program

Correct Answer: C

The Hadoop environment within the IBM InfoSphere BigInsights Version 3.0 can read a large number of storage formats. This flexibility is partially because of the INPUTFORMAT and OUTPUTFORMAT classes that you can specify on the CREATE and ALTER table statements and because of the use of installed and customized SerDe classes. The file formats listed here are available either by using explicit SQL syntax, such as STORED AS PARQUETFILE, or by using installed interfaces, such as Avro.

Reference: <https://developer.ibm.com/hadoop/2014/09/19/big-sql-3-0-file-formats-usage-performance/>

QUESTION 2

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

- A. InfoSphere Streams utilizes Flume to interface to Hadoop
- B. The HDFSFileSink operator writes files in parallel to a Hadoop Distributed File System
- C. Buffering techniques are used to process incoming streams from InfoSphere Streams
- D. When you use the HDFS operators to access GPFS, you must install InfoSphere Streams on an InfoSphere Big Insights data node

Correct Answer: C

Reference: <https://books.google.com.pk/books?id=JWfRAGAAQBAJandpg=PA147andlpg=PA147anddq=IBM+Buffering+techniques+are+used+to+process+incoming+streams+from+InfoSphere+Streamsandsource=blandots=Z6XhA0Owkandsig=ACfU3U3T9ydrZHWTMvB31qQOyf6FtoDQgwandhl=enandsa=Xandved=2ahUKEwj6wsnvqvf0AhVUSxUIHWNDAYcQ6AEwA3oECBEQAQ#v=onepageandq=IBM%20Buffering%20techniques%20are%20used%20to%20process%20incoming%20streams%20from%20InfoSphere%20Streamsandf=false>

QUESTION 3

Which of the following is TRUE regarding a schema In Jaql?

- A. It specifies a particular type for each field
- B. It is used only when Jaql includes SQL statements



- C. It must be provided when the data format is not otherwise known
- D. It can be used both as a data constraint and to optimize processing and storage

Correct Answer: B

QUESTION 4

Which of the following statements regarding Sqoop is TRUE?

- A. Output files are always delimited text files in HDFS
- B. The import of a single database table results in multiple files in HDFS
- C. When exporting data back to a database, a target table will be created if it doesn't already exist
- D. Sqoop processes the entire database table as a single unit rather than row-by-row for enhanced performance

Correct Answer: A

Reference: <https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html>

QUESTION 5

Which of the following is the most effective method for improving query performance on large Hive tables?

- A. Indexing
- B. Bucketing
- C. Partitioning
- D. De-normalizing data

Correct Answer: B

Reference: <https://dzone.com/articles/how-to-improve-hive-query-performance-with-hadoop>

[C2090-101 PDF Dumps](#)

[C2090-101 Practice Test](#)

[C2090-101 Brainsdumps](#)