

**Vendor:** IBM

**Exam Code:** P2090-032

**Exam Name:** IBM Big Data Fundamentals Technical  
Mastery Test v1

**Version:** Demo

**Question No : 1**

Which of the following options is CORRECT?

- A.** InfoSphere Data Explorer provides powerful navigation capabilities across all the important information stored exclusively into Hadoop Distributed File System in a single view. No other file systems are supported.
- B.** InfoSphere Data Explorer is not able to mirror pre-existing security frameworks, therefore it doesn't make use of industry-standard authentication and authorization processes already in place.
- C.** InfoSphere Data Explorer can find, extract and deliver content regardless of format or where it resides.
- D.** InfoSphere Data Explorer uses a vector-based index for unique search and indexing flexibility.

**Answer: C**

**Question No : 2**

Which of the following InfoSphere BigInsights features provides a vast library of extractors enabling actionable insights from large amounts of native textual data?

- A.** Text Analytics.
- B.** Adaptive MapReduce.
- C.** General Parallel File System.
- D.** BigSheets.

**Answer: A**

**Question No : 3**

Which of the following options contain security enhancements available in InfoSphere BigInsights (Choose two) ?

- A.** LDAP authentication
- B.** Secure file transfers through SFTP protocol.
- C.** Trusted Context.
- D.** Kerberos authentication protocol.

---

**Answer: A,B**

**Question No : 4**

In regards of InfoSphere Streams, Which of the following options is CORRECT?

- A.** InfoSphere Streams is a powerful analytic computing platform capable of gathering large quantities of data, manipulating the data, and storing it on disk.
- B.** InfoSphere Streams is a powerful analytic computing platform capable of analyzing data in real time with micro-latency.
- C.** InfoSphere Streams is an extract, transform, and load (ETL) platform that is capable of integrating small volumes of data across a wide variety of data sources and target applications.
- D.** InfoSphere Streams is web administration graphical user interface (GUI) capable of setting up a secure communication channel to stream post-processed data from a Hadoop cluster into a relational database, such as IBM DB2.

**Answer: B**

**Question No : 5**

The following types of indexes are available in the InfoSphere BigInsights? Large Scale indexing feature, EXCEPT:

- A.** MapReduce index
- B.** Parallel index
- C.** Real-time index
- D.** Partitioned index

**Answer: A**

**Question No : 6**

How do enterprises leverage big data platforms?

- A.** By storing all of the data in its native business object format, so the enterprise can get

---

value out of it through massive parallelism using readily available components.

**B.** By modifying the data being streamed using pre-existing ETL transformations, and storing the final formatted data into a data warehouse for further enterprise analysis.

**C.** By sitting on top of a large data warehouse solution acting as transparent abstract conversion layer allowing enterprises to query unstructured data.

**D.** By isolating workloads into a single node, and further processing all the data in sequence.

**Answer: A**

#### **Question No : 7**

What is the difference between Hadoop's MapReduce and IBM's Adaptive MapReduce feature available in InfoSphere BigInsights?

**A.** Hadoop's MapReduce is optimized for operating on small files or splits, while IBM's AdaptiveMapReduce is optimized for operating on large partitioned files.

**B.** Hadoop's MapReduce is optimized for operating on large files, while IBM's Adaptive MapReduce is configurable to operate optimized on large or small files or splits.

**C.** Hadoop's MapReduce is optimized for operating on small files or splits, while IBM's Adaptive MapReduce is optimized for operating on large files stored in individual blocks.

**D.** Hadoop's MapReduce is optimized for operating on small partitioned tables stored in the HBase component, while IBM's Adaptive MapReduce is optimized for operating on large partitioned files.

**Answer: B**

#### **Question No : 8**

Which of the following options are CORRECT (Choose two)?

**A.** The Stream Processing Language provides a language that works with the Streams runtime framework to support streaming applications.

**B.** Users can develop Streams applications using Data Studio Web Console, an Eclipse-

---

based Integrated Development Environment (IDE).

**C.** InfoSphere Streams perform complex analytics on data at rest.

**D.** Users can deploy existing data mining scoring models in Streams applications for real time insights as opposed to running those models on persistent, or stored data.

**Answer: A,D**

**Question No : 9**

How is data stored in a Hadoop cluster?

**A.** The data is divided into blocks, and copies of these blocks are replicated across multiple servers in the Hadoop cluster.

**B.** The data converted into a single block, and the block is stored in just one of the servers in the Hadoop cluster.

**C.** The data is divided into blocks, each block is stored in a different server in the Hadoop cluster, but the blocks are not replicated.

**D.** The data is converted into a single block, and copies of this block are replicated across multiple servers in the Hadoop cluster.

**Answer: A**

**Question No : 10**

What is good fit for Hadoop Distributed File System (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**