Number: P2090-032
Passing Score: 800
Time Limit: 120 min
File Version: 4.3

CertKiller

000-N32

IBM Big Data Fundamentals Technical Mastery Test v1

✔ Don't panic, just study from this dump and definite your success.

✔ valid test preparation - take the exam today with 10 new questions

✔ It's still valid and there are a lot of fresh and updated questions.But if you know the answers of that dump you will succeed easily because the questions are   similar.

✔ Exactly same questions! great work

✔ 100% Valid in US, UK, Australia, India and Emirates. All my friends in group have these same questions.

Exam A

QUESTION 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.

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

QUESTION 2
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.

Correct Answer: AB
Section: (none)
Explanation

Explanation/Reference:

QUESTION 3
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.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
well defined.

QUESTION 4
Which of the following toolkits is NOT provided by InfoSphere Streams?

A. Intranet toolkit.
B. Finance toolkit
C. Standard toolkit.
D. Database toolkit.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
proper answer.

QUESTION 5
Which of the following components is NOT included in the BigInsights Basic Edition distribution?

A. Hadoop Distributed File System.
B. Hive.
C. Pig.
D. BigSheets.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:

QUESTION 6
Which of the following statements is NOT CORRECT?

A. InfoSphere Streams provides support for reuse of existing Java or C++ code, as well as Predictive Model Markup Language (PMML) models.
B. InfoSphere Streams supports communications to Internet Protocol version 6 (IPv6) networks.
C. InfoSphere Streams jobs must be coded using either HiveQL or Jaql languages.
D. InfoSphere Streams supports both command line and graphical interfaces to administer the Streams runtime and maintain optimal performance and availability of applications.

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:
appropriate answer.

QUESTION 7
How do big data solutions interact with the existing enterprise infrastructure?

A. Big data solutions must substitute for the existing enterprise infrastructure; therefore there is no interaction between them.
B. Big data solutions are placed on top of the existing enterprise infrastructure, acting as a transparent layer converting unstructured raw data into structured, readable data, and storing the final results in a traditional data warehouse.
C. Big data solutions must be isolated into a separate virtualized environment optimized for sequential workloads, so that it doesn?t interact with existing infrastructure.
D. Big data solutions works in parallel with the existing enterprise infrastructure leveraging all the unstructured raw data that cannot be processed and stored in a traditional data warehouse solutions.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:

QUESTION 8
Which of the following options is CORRECT?

A. InfoSphere Streams submits queries to structured static data.
B. InfoSphere Streams submits queries to structured dynamic data.

C. InfoSphere Streams submits queries to unstructured dynamic data.
D. InfoSphere Streams submits dynamic data to pre-existing queries.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
true answer.

QUESTION 9
InfoSphere Data Explorer accommodates data variety through (choose three):

A. Broad connectivity to a wide range of data management systems and applications.
B. Sophisticated security mapping, including cross-domain and field-level security.
C. Support for new ?virtual multi-dimensional node? technology capable of aggregating documents created from multiple sources or tables.
D. Federated connectivity in the cloud and on-premise.

Correct Answer: ABC
Section: (none)
Explanation

Explanation/Reference:

QUESTION 10
Which of the following options is NOT CORRECT?

A. Big data solutions are ideal for analyzing not only raw structured data, but semi- structured and unstructured data from a wide variety of sources.
B. Big data solutions are ideal when all, or most, of the data needs to be analyzed versus a sample of the data; or a sampling of data isn?t nearly as effective as a larger set of data from which to derive analysis.
C. Big data solutions are ideal for Online Transaction Analytical Process (OLTP) environments.
D. Big data solutions are ideal for iterative and exploratory analysis when business measures on data are not predetermined.

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

rightful.

QUESTION 11
Which of the following options best describes the proper usage of MapReduce jobs in Hadoop environments?

A. MapReduce jobs are used to process vast amounts of data in-parallel on large clusters of commodity hardware in a reliable, fault-tolerant manner.
B. MapReduce jobs are used to process small amounts of data in-parallel on expensive hardware, without fault-tolerance.
C. MapReduce jobs are used to process structured data in sequence, with fault-tolerance.
D. MapReduce jobs are used to execute sequential search outside the Hadoop environment using a built-in UDF to access information stored in non-relational databases.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
right answer.

QUESTION 12
Which of the following components is a feature from InfoSphere Data Explorer?s Discovery module?

A. Auto-commit
B. Auto-correction
C. Auto-classification
D. Auto-save

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:
rectified answer.

QUESTION 13
What is the purpose of IBM InfoSphere Streams Studio?

A. IBM InfoSphere Streams Studio is an Eclipse-ready tool that enables you to create, edit, visualize, test, debug, and run MapReduce jobs for virtualized applications.
B. IBM InfoSphere Streams Studio is an Eclipse-ready tool that enables you to create, edit, visualize, test, debug, and run Streams Processing Language (SPL) and SPL mixed-mode applications.

C. IBM InfoSphere Streams Studio provides an integrated, modular environment for database development and administration of DB2 databases, Informix, and other non-IBM databases.

D. IBM InfoSphere Streams Studio improves performance and cuts costs by providing expert advice on writing high quality queries and improving database design.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
valid answer.

QUESTION 14
Which of the following options contains the main components of Hadoop? (Choose three)

A. Text Analytics.
B. MapReduce framework.
C. Hadoop Distributed File System (HDFS).
D. Apache Commons.

Correct Answer: BCD
Section: (none)
Explanation

Explanation/Reference:

QUESTION 15
Which of the following options best describes the differences between a traditional data warehouse environment and a Hadoop environment?

A. Traditional data warehousing environments are mostly ideal for analyzing structured data from various systems, while a Hadoop environment is well suited to deal with semi-structured and unstructured data, as well as when a data discovery process is needed.

B. Hadoop environments are mostly ideal for analyzing structured and semi-structured data from a single system, while traditional data warehousing environment is well suited to deal with unstructured data, as well as when a data discovery process is needed.

C. Typically, data stored in Hadoop environments is cleaned up before storing in the distributed file- system.

D. Typically, data stored in data warehousing environments is rarely filtered and pre-processed. On the other hand, data injected into Hadoop environments is always pre-processed and filtered.

Correct Answer: A
Section: (none)
Explanation

QUESTION 16
Which of the following compression algorithms is used by InfoSphere BigInsights to provide an additional compression option over the ones that come with the base Hadoop distribution?

A. gzip.
B. brzip.
C. lza.
D. lzo.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:

QUESTION 17
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

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:

QUESTION 18
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.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:

QUESTION 19
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 Adaptive MapReduce 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.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:

QUESTION 20
Which of the following options are CORRECT (Choose two)?

A. The Stream Processing Language provides a language that works with the Streams run-time 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.

Correct Answer: AD
Section: (none)
Explanation

Explanation/Reference:

QUESTION 21
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.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
fresh answer.

QUESTION 22
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.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:

QUESTION 23
What does ?Big Data? represent?

A. A Hadoop feature capable of processing vast amounts of data in-parallel on large clusters of commodity hardware in a reliable, fault-tolerant

manner.

B. Large amounts of unstructured, semi-structured, and structured raw data that cannot be stored, processed or analyzed using traditional relational data warehouses.
C. A database feature capable of converting pre-existing structured data into unstructured raw data.
D. Only data stored in the BIGDATA table in any relational database.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
updated.

QUESTION 24
Which of the following options is CORRECT?

A. InfoSphere BigInsights is based on a branched Hadoop distribution, and therefore backwards compatibility is not guaranteed.
B. InfoSphere BigInsights is a distributed file system used as base for Hadoop distributions.
C. InfoSphere BigInsights is based on the nonforked core Hadoop distribution, but backwards compatibility with the Apache Hadoop project is not guaranteed, therefore applications written for Hadoop might not run on BigInsights.
D. InfoSphere BigInsights is based on the nonforked core Hadoop distribution, and backwards compatibility with the Apache Hadoop project will always be maintained. Therefore, all applications written for Hadoop will run on BigInsights.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
explained answer.

QUESTION 25
Streams jobs can be monitored using the following tools (choose three):

A. Streams Studio.
B. Streams universal web management interface.
C. Streams console.
D. Streamtool command.

Correct Answer: ACD
Section: (none)

Explanation

Explanation/Reference:
actual answer.

QUESTION 26
What is HADOOP?

A. Hadoop is a single-node file system used as a base for storing traditional formatted data.
B. Hadoop is a framework that allows for the distributed processing of large data sets across clusters of computers using a simple programming model.
C. Hadoop is a universal Big Data programming language used to query large datasets.
D. Hadoop is framework capable of transforming raw, unstructured data into plain, regular data readable by traditional data warehouses.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:

QUESTION 27
Hadoop environments are optimized for:

A. Processing transactions (random access).
B. Low latency data access.
C. Batch processing on large files.
D. Intensive calculation with little data.

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

QUESTION 28
Which of the following options is CORRECT?

A. InfoSphere Streams optimizes its workload by aggregating an entire job into a single node.

B. InfoSphere Streams is only able to process traditional structured data from a variety of sources.

C. InfoSphere Streams does not allow you to dynamically add hosts and jobs

D. InfoSphere Streams high availability feature allows for processing elements (PEs) on failing nodes to be moved and automatically restarted, with communications re-routed, to a healthy node.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
real answer.

QUESTION 29
In a traditional Hadoop stack, which of the following components provides data warehouse infrastructure and allows SQL developers and business analysts to leverage their existing SQL skills?

A. Avro.

B. Hive.

C. Zookeeper.

D. Text analytics.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:

QUESTION 30
Which of the following tools can be used to configure the InfoSphere Data Explorer environment (choose two) ?

A. Data Studio Web Console

B. InfoSphere Data Explorer?s web-based interface

C. REST/SOAP APIs

D. Data Explorer Virtual Desktop

Correct Answer: BC
Section: (none)
Explanation

Explanation/Reference:
efficient answer.

QUESTION 31
Which of the following connectivity modules is provided by InfoSphere Data Explorer?

A. Federation Module
B. Navigation Module
C. Discovery Module
D. Language Module

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
true answer.

QUESTION 32
What are the ?4 Vs? that characterize IBM?s Big Data initiative?

A. Variety, Versions, Velocity, Volatility
B. Velocity, Volatility, Variety, Veracity
C. Veracity, Variety, Volume, Velocity
D. Volume, Volatility, Velocity, Variety

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

QUESTION 33
Which of the following options is CORRECT regarding InfoSphere Data Explorer?s annotators?

A. InfoSphere Data Explorer?s annotators allow users to create groups of search results.
B. InfoSphere Data Explorer?s annotators is an add-on feature capable of handling of a variety of data formats and types, including structured, semi-structured and unstructured, as well as the special demands of rich media and transactional data.
C. InfoSphere Data Explorer?s annotators allow users to interact with search results by providing feedback about the result's value, and by

adding useful information and communication with other users.

D. InfoSphere Data Explorer?s annotators allow users to save results in a private/public folder for later review or sharing.

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

QUESTION 34
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.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
exact answer.