



IBM

Exam Questions C2090-102

IBM Big Data Architect

NEW QUESTION 1

The downside of cloud computing, relative to SLAs, is the difficulty in determining which of the following?

- A. Root cause for service interruptions
- B. Turn-Around-Time (TAT)
- C. Mean Time To Recover (MTTR)
- D. First Call Resolution (FCR)

Answer: A

Explanation:

References:

https://en.wikipedia.org/wiki/Service-level_agreement

NEW QUESTION 2

Service Level Agreements (SLAs) are defined at different levels. Which of the following is NOT a valid level?

- A. Customer-Based SLA
- B. Service-Based SLA
- C. Multilevel SLA
- D. Third-Party SLA

Answer: C

NEW QUESTION 3

Financial company A needs to ingest, analyze and store 8 billion structured transactions per day (around 2TB to 4TB of data) which translates full table scans, 24/7 access without any downtime, trend analysis, and the search for evidence of fraudulent activity. Which of the following technologies would you recommend to the company to meet these requirements?

- A. IBM Cognos and Unica
- B. Pure Data System for Analytics
- C. Infosphere Streams
- D. The Advanced Text Analytic Toolkit

Answer: B

NEW QUESTION 4

Which one of the following statements is TRUE?

- A. Big SQL uses Hadoop MR framework to process query tasks in parallel
- B. Big SQL executes queries locally on Big SQL server single node on a multi node cluster
- C. Big SQL can process queries in parallel and executes queries locally
- D. Big SQL only works with HDFS

Answer: A

Explanation:

Reference:

https://www-01.ibm.com/support/knowledgecenter/SSPT3X_3.0.0/com.ibm.swg.im.infosphere.biginsights.analyze.doc/doc/bigsql_performance.html

NEW QUESTION 5

A telecommunication company is experiencing high churn rate. They have more than five million customers. They have more than 400TB of call detail records per year. The majority of their customers are prepaid, so they can change the telecom provider any time. In order to remedy the situation, they want to understand the customer behavior. For that reason they want to build the individual profile of each customer. They also want to enrich the profile by incorporating social media data. Given this set of requirements, which of the following would you recommend?

- A. Hadoop
- B. RDBMS
- C. Storm
- D. IBM Cloudant

Answer: C

Explanation:

References:

<http://www.ibm.com/developerworks/library/os-twitterstorm/>

NEW QUESTION 6

Data stored in Netezza can be used for model building, scoring, and model refresh. How is SPSS Modeler connectivity to Netezza enabled?

- A. Open Database Connectivity (ODBC) driver provided by Netezza
- B. DataDirect Connect for ODBC drivers
- C. OLE DB drivers provided by Netezza
- D. Java Database Connectivity (JDBC) driver provided by Netezza

Answer: A

Explanation:

Reference: http://www.ibm.com/developerworks/data/library/cognos/infrastructure/cognos_sp ecific/pag e569.html

NEW QUESTION 7

Considering a service level requirement (SLR) of less than 3 milliseconds, what task must be performed to meet the SLR?

- A. Collect and analyze SNMP MIB data
- B. Survey a representative amount of end users
- C. Measure switch failure frequency
- D. Inquire when prime network hours occur

Answer: C

NEW QUESTION 8

Which of the following is the section of the Component Model that details how the solution integrates?

- A. Component Relationship Diagram
- B. Component Interface Diagram
- C. Component Interaction Diagram
- D. Component Reaction Diagram

Answer: A

NEW QUESTION 9

Which of the following statements regarding SPSS is TRUE?

- A. SPSS software provides a security framework
- B. SPSS analytics are primarily accessed through a scripting language
- C. SPSS models can only be applied to data at rest
- D. SPSS can directly use BigR syntax such as “bigr.list”

Answer: A

Explanation:

Reference:

<http://www.ibmbigdatahub.com/blog/7-really-good-reasons-partner-spss-analytics-r>

NEW QUESTION 10

A reputable market research firm wants to explore more business opportunities. They have great in house skill in python and machine learning. Their business model is simple, they build the solutions for customers using python and machine learning algorithms and give these solutions to the customer's engineering team for implementation. Given this scenario, which of the following would you recommend?

- A. Netezza
- B. Spark
- C. Cloudant
- D. Hadoop

Answer: D

NEW QUESTION 10

Which of the following is a requirement for data retention and archival?

- A. A format and storage repository for archived data
- B. Public cloud
- C. Hosting location
- D. Solid-state technology

Answer: A

NEW QUESTION 12

The AQL query language is the easiest and most flexible tool to pull structured output from which of the following?

- A. Hive data structures
- B. Unstructured text
- C. Hbase schemas
- D. JDBC connected relational data marts

Answer: A

Explanation:

Reference:

<http://www.ibm.com/developerworks/library/bd-sqltohadoo2/>

NEW QUESTION 13

Which language is used to define secondary views in a Cloudant database?

- A. Ruby
- B. Java, C++, or Perl
- C. Javascript
- D. Python

Answer: C

NEW QUESTION 17

As you explore the data for a BigSheets workbook, you must run the workbook against the full data set to get the most current results for analysis. Which statement is TRUE regarding running and visualizing data in a workbook?

- A. You can create graphs for more than one sheet within the same workbook
- B. By default, the first sheet in your workbook is named the Results sheet
- C. When you save and run the workbook, the data in a Child Workbook is the output for that workbook
- D. When you add sheets to workbooks, saving the sheets runs the individual data for the sheet but not for the full workbook

Answer: C

Explanation:

Reference:

https://www-01.ibm.com/support/knowledgecenter/SSPT3X_4.1.0/com.ibm.swg.im.infosphere.biginsights.analyze.doc/doc/bigsheets_con_workbooks.html

NEW QUESTION 22

A media company wants to enhance their subscription services to their end customers by incorporating Twitter, Wikis, Blogs, and Analysts data. The company likes to provide a real-time dashboard view for news articles, press releases, and analysts perspective to end customers based on their subscriptions. What solution blueprint would offer the best fit for this critical business requirement?

- A. Watson Analytics
- B. Big Insights, Spark, and Big SQL
- C. Pure Data for Analytics PDA, and SPSS
- D. InfoSphere Streams, BigInsights, and Watson Explorer

Answer: D

NEW QUESTION 26

Data Architects assess data according to several common characteristics. One of these characteristics is the type of data. Which of the following are data types that fall into this ontology? (Choose two.)

- A. Master data
- B. Meta data
- C. Boolean
- D. Composite
- E. Integer

Answer: AD

NEW QUESTION 29

Which of the following statements is TRUE regarding cloud based solutions?

- A. In a Platform as a Service Cloud deployment, the customer chooses the operating system they want to use
- B. Automated recovery from hardware or network failures is not possible in a public cloud implementation, only in a private clouds
- C. There are benefits to use the cloud even for small-scale applications
- D. Using firewalls to create network boundaries is sufficient for ensuring cloud security

Answer: C

Explanation:

References:

<http://www.ibm.com/developerworks/cloud/library/cl-cloudappdevelop/>

NEW QUESTION 31

If the recovery point objective (RPO) is low, which of the following techniques would be the most appropriate?

- A. Clustering
- B. Database Shadowing
- C. Remote Journaling
- D. Tape Backup

Answer: A

Explanation:

References:

<http://whatis.techtarget.com/definition/recovery-point-objective-RPO>

NEW QUESTION 32

Defining your need to enrich existing customer data you realize that you need to process large quantities of Geospatial data and output to your data warehouse in a standard GeoJSON format. Which of the following would provide a business analyst with the desired output?

- A. Big SQL
- B. BigSheets
- C. Hive queries
- D. Text Analytics

Answer: C

NEW QUESTION 35

Company A has decided to implement a new data system to support their rapidly growing business. They have an existing 20 TB worth of raw data, with an expected weekly incoming rate of 50 GB of new raw data. The data is mostly text based and unstructured. A typical query can involve pulling in 10 GB of data. Historically, performance has been an issue and currently needs to be addressed. Which of the following would you suggest to support these requirements?

- A. Set up a Hadoop system with commodity HW for scalability
- B. Utilize de-duplication and compression technology
- C. Use a mixture of different disk-types to provide hot/cold storage
- D. Create range partitions for the data

Answer: A

NEW QUESTION 39

What term applies to the data elements in Infosphere Streams?

- A. Tuples
- B. Operators
- C. Sink adapters
- D. Composite operators

Answer: B

Explanation:

Reference:

http://www-01.ibm.com/support/knowledgecenter/SSCRJU_3.2.1/com.ibm.swg.im.infosphere.streams.glossary.doc/doc/glossary_Streams.html?lang=en

NEW QUESTION 42

Which of the following statements is TRUE regarding cloud applications?

- A. Migrating a legacy application to the cloud is a simple solution to drive down cost
- B. Architecting and deploying a scalable cloud application requires a private cloud implementation
- C. To be truly elastic, a cloud application must be tightly coupled to the physical cloud environment
- D. Leveraging a private v
- E. public cloud may result in sacrificing some of the core advantages of cloud computing

Answer: B

NEW QUESTION 45

The NameNode determines the rack id to which each DataNode belongs via the process outlined in which of the following?

- A. HDFS APIs
- B. Hadoop Rack Awareness
- C. Job Tracker
- D. NameNode High Availability

Answer: D

NEW QUESTION 46

Which of the following statements is TRUE regarding cloud computing solutions?

- A. Cloud security is planned, developed, and layered on top of an application after the application development process is complete
- B. Stateless applications are better candidates for cloud services than applications that maintain state
- C. Cloud solutions rely on scaling up (vertical) scaling v
- D. scale out (horizontal) scaling
- E. Server virtualization is a requirement in a cloud implementation

Answer: D

NEW QUESTION 51

Which architecture document is used to help organize projects, manage the complexity of the solution, and ensure that all architecture requirements have been addressed?

- A. Operational Model
- B. Component Model

- C. Connection Model
- D. API Model

Answer: B

NEW QUESTION 56

A large application vendor wants to port their existing distributed applications to run on Hadoop. In order to be competitive they need to provide monitoring and keep the size of the monitored applications consistent with the configuration. This implies the ability to deploy a replacement, for example, for any failed components. Which of the following would be a workable solution?

- A. Nagios with YARN
- B. Slider with YARN
- C. Oozie with Lucene
- D. OPTIM Performance Manager

Answer: B

NEW QUESTION 59

On startup, the NameNode enters a special state called Safemode. Replication of data blocks does not occur when the NameNode is in the Safemode state. The NameNode receives Heartbeat and Blockreport messages from the DataNodes. Which of the following contains the list of data blocks that a DataNode is hosting?

- A. Rack Awareness
- B. Snapshot
- C. Replica
- D. Blockreport

Answer: D

Explanation:

References:

https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html

NEW QUESTION 62

Which of the following statements regarding Big R is TRUE?

- A. Missing data values must be handled by ETL processes prior to analyzing data with Big R
- B. A bigr.frame loads data in memory for optimal performance
- C. A Big R user is responsible for parallelizing the execution of the R functions being used in the R program
- D. Performing a mathematical operation on a Big R vector variable will automatically loop through each item in the vector

Answer: D

Explanation:

Reference:

<http://www.computerworld.com/article/2497319/business-intelligence-beginner-s-guide-to-r-syntax-quirks-you-llwant-to-know.html>

NEW QUESTION 66

A news organization wants to analyze all the news stories coming in real time and make them available to their users based on the user's interest. Given this requirement, which of the following would you recommend?

- A. Spark
- B. Hadoop
- C. Netezza
- D. Cloudant

Answer: A

NEW QUESTION 68

Company K is designing their Big Data system. In their enterprise, they anticipate every 9 months there will be a big spike of new data on the order of multiple TB. Their company policy also dictates that data older than one year will be archived with a major clean up every 5 years. Cost is also a big issue. Which of the following provides the best design for these requirements?

- A. Estimate the peak volume over a 5 year period and set up a Hadoop system with commodity HW and storage to accommodate that volume
- B. Estimate the peak volume over a 3 year period and set up a Hadoop system with NAS to accommodate the expected volume
- C. Use Cloud elasticity capabilities to handle the peak and valley data volume
- D. Use SAN storage with compression to handle the peak and valley data volume

Answer: A

NEW QUESTION 73

Which of the following statements is TRUE about BigSheets?

- A. It is a tool that enables line-of-business users to harness the power of Hadoop using a familiar spreadsheet interface
- B. It is a browser-based visualization and analysis tool designed to help programmers work with Big Data
- C. It is a component of Pure Data for Analytics
- D. It is a browser-based visualization and analysis tool included in Streams

Answer: C

Explanation:

Reference:

<http://www-01.ibm.com/software/data/puredata/analytics/>

NEW QUESTION 78

Which one of the following statements about Big SQL is NOT TRUE?

- A. BIG SQL supports querying data but doesn't support updates and deletes
- B. BIG SQL is ANSI 2011 Compliant
- C. BIG SQL is supported to work with Cloudera
- D. BIG SQL supports Fluid Query

Answer: A

Explanation:

Reference:

<http://www.ibm.com/developerworks/library/bd-bigsq/>

NEW QUESTION 82

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

C2090-102 Practice Exam Features:

- * C2090-102 Questions and Answers Updated Frequently
- * C2090-102 Practice Questions Verified by Expert Senior Certified Staff
- * C2090-102 Most Realistic Questions that Guarantee you a Pass on Your First Try
- * C2090-102 Practice Test Questions in Multiple Choice Formats and Updates for 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The C2090-102 Practice Test Here](#)