Exam Questions C2090-102

IBM Big Data Architect

https://www.2passeasy.com/dumps/C2090-102/

NEW QUESTION 1
In a typical Hadoop HA cluster, two separate machines are configured as which of the following?

A. Data Nodes
B. Edge Nodes
C. Name Nodes
D. None of the Above

Answer: A

Explanation:
Reference:
http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/HDFSHighAvailabilityWithQJM.html

NEW QUESTION 2
The Yarn Resource Managers (RM) have an option to embed an ActiveStandbyElector to decide which RM should be the Active. Upon which of the following is it based?

A. Job Tracker
B. Zookeeper
C. Task Tracker
D. Name Node

Answer: C

NEW QUESTION 3
A smart meter project is being undertaken by anElectric utility headquartered in San Francisco. They are looking for technology for monitoring thousands of meters, multiple times within an hour, andperforming corresponding downstream advisory action to influence usage behavior. Which of the following would you recommend to function at the heart of this process?

A. Hadoop
B. SPSS
C. Reporting engine and a portal
D. Infosphere Streams

Answer: B

NEW QUESTION 4
The analysis layer reads the data digested by the layer massaging and store layer. In some cases, the analysis layer accesses the data directly from the data source. Designing the analysis layer requires careful forethought and planning. Decisions must be made with regard to how to manage the tasks to do which of the following?

A. Reduce the volume of data
B. Locate the data sources that can provide data for the required entities
C. Identify multi-dimension data sources
D. Identify data import and export tools to adopt and deliver an ETL strategy

Answer: B

Explanation:
References:
http://www.ibm.com/developerworks/library/bd-archpatterns3/

NEW QUESTION 5
It's helpful to look at the characteristics of big data along certain lines – for example, how the data is collected, analyzed and processed. There are many characteristics to consider. Which one of the following is NOT a characteristic that should be considered?

A. Data frequency and size
B. Software
C. Data source
D. Processing methodology

Answer: B

Explanation:
References:
http://www.ibm.com/developerworks/library/bd-archpatterns1/

NEW QUESTION 6
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 7
Which of the following is NOT a typical Service Level Agreement (SLA) metric?

A. Abandonment rate
B. Average speed to answer
C. Time service factor
D. Average response time

Answer: D

NEW QUESTION 8
A media company wants to measure the effectiveness of their advertising campaign. Before they release a movie they prepare and run a campaign for promotion. Based on the response on Twitter and Facebook they want to decidewhetheror not they should continue a particular campaign. Which of the following should be selected to meet these requirements?

A. Hadoop
B. Streams
C. Unica
D. Pure Data for Analytics

Answer: C

NEW QUESTION 9
Big data is often defined as the ability to derive new insights from data that has scaled up along three axes known as the three v's. Which of the following is the fourth v? (Hint: It has something to do with the uncertainty.)

A. volume
B. variety
C. velocity
D. veracity

Answer: C

NEW QUESTION 10
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 10
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 11
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 16
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 19
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, onlyin 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 23
For company B, 85% of their analytics queries only involve about 25% of their data; another 10% of the queries will touch 35% of the rest of the data, and only 5% of the queries will touch the remaining 40% of the data. The estimated volume is 50TB growing at 1 TB per year. Which of the following would provide the best value (business benefit) and lowest TCO?

A. Place the entire set of data in a data warehouse with proper partitioning and indexing
B. Place the entire set of data in a hadoop environment – using commodity HW
C. Place the top 25% of data (used by 85% of the query) in a hadoop environment, and the rest in a data warehouse
D. Place the top 25% of data (used by 85% of the query) in a data warehouse, and the rest in a hadoop environment

Answer: C

NEW QUESTION 26
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 27
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 30
Data in motion is data that is constantly appended with new information. This type of data can be imported into the distributed file system by using which of the following?

A. Scoop
B. Flume
C. Big SQL LOAD
D. Hadoop shell commands

Answer: B

Explanation:
References: http://www.ibm.com/support/knowledgecenter/SSPT3X_2.1.2/com.ibm.swg.im.infosphere.biginsights.import.doc/doc/data_motion.html

NEW QUESTION 35
A company has to design a new data system. They will need to support several OLTP applications. Every three days a batch job will run to load specific data into a set of 10 large tables (with historical data) where OLAP analytics will be performed. Performance for both OLTP and OLAP queries is important. Which of the

following designs would you suggest to the company?

A. Use a NoSQL data store such as MongoDB or Cloudant on the cloud to provide needed scalability
B. Use DB2 Data Partition Feature (DPF), partitioning all tables into different partitions
C. Use DB2 Data Partition Feature (DPF), only partitioning the 10 tables where Analytics will be run
D. Use DB2 with BLU Acceleration, use columnar store for the 10 tables where Analytics will be run

Answer: D

NEW QUESTION 39
Which of the following is NOT an example of unstructured data?

A. HBase table
B. Tweet
C. Netezza table
D. Internet Protocol Detail Record

Answer: C

Explanation:
‒References: https://www.ibm.com/support/knowledgecenter/SSULQD_7.2.0/com.ibm.nz.gsg.d oc/c_gets trt_nzsql_help.html

NEW QUESTION 42
Which of the following is NOT a valid Big Data platform integration?

A. Platform plugins
B. Intraplatform integration
C. Enterprise integration with other repositories
D. Network integration

Answer: D

NEW QUESTION 47
A component of IBM Industry Model forms the basis of the "Logical Data Warehouse Model" that spans across the traditional RDBMS and Hadoop technology. It defines all of the data structures that would be expected to be defined in the Detailed System of Record. What is the name of this component?

A. Business data model
B. Atomic warehouse model
C. Dimensional warehouse models
D. Metadata management

Answer: A

Explanation:
‒References:
http://www.ibm.com/support/knowledgecenter/SS9NBR_9.1.0/com.ibm.ima.using/comp/bdm/intro.dita

NEW QUESTION 48
Which of the following approaches can an organization take to solve the impact on the performance and capacity originated by the variety of data types?

A. Define a data catalog in a traditional data warehouse
B. Create different solutions to handle every kind of data
C. Store a wide range of data formats on the same platform
D. Define a comprehensive taxonomy and constantly review

Answer: D

Explanation:
Reference:
http://www.redbooks.ibm.com/redpapers/pdfs/redp5070.pdf Page: 15

NEW QUESTION 50
Company A is searching for a browser-based visualization tool to perform analysis
on vast amounts of data in any structure. They want to execute operations such as pivot, slice and dice, among others. Which of the following would meet these requirements?

A. Streams
B. BigSheets
C. Aginity Workbench
D. Watson Explorer

Answer: B

Explanation:
References:
http://www.dotgroup.co.uk/wp-content/uploads/2014/11/Harness-the-Power-of-Big-Data- The-IBM-Big-DataPlatform.pdf Page: 132

NEW QUESTION 51
Which of the following statements is TRUE regarding Cloud deployment models?

A. Performance and scalability requirements are a critical factor for deciding between Platform as a Serviceand Infrastructure as a Service deployment models
B. In a platform as a Service offering, the customer has root access to the servers
C. Applications with extremely high transactions volumes are good candidates for Platform as a Service
D. In an infrastructure as a service deployment, the cloud provider provides security patching, monitoring andfail over capabilities

Answer: A

NEW QUESTION 53
You are designing storage for a new Hadoop cluster. Which of the following statements is TRUE regarding the usage of SAN or NAS?

A. SAN or NAS should not be used to set up HDFS
B. SAN or NAS must be used, if available, to provide backup capabilities
C. SAN or NAS can be used to support retention policies
D. SAN or NAS cannot be used if your Hadoop cluster spans 2 sites

Answer: A

Explanation:
References:
http:// www-01.ibm.com/software/data/infosphere/hadoop/hdfs/

NEW QUESTION 57
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 62
......

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual C2090-102 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the C2090-102 Product From:

https://www.2passeasy.com/dumps/C2090-102/

Money Back Guarantee

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 FirstTry

* C2090-102 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year