# IBM

## Exam C2090-102

## IBM Big Data Architect

**Version: 7.0**

**[ Total Questions:   110 ]**

**Question No : 1**

What are the two levels documented in the Operational Model? (Choose two.)

**A.** Logical
**B.** Rational
**C.** Theoretical
**D.** Physical
**E.** Middleware

**Answer: A,C**

**Question No : 2**

The inputs to the Architectural Overview document do NOT include which of the following?

**A.** Architectural Goals
**B.** Key Concepts
**C.** Architectural Overview Diagram
**D.** Component Model

**Answer: D**

**Question No : 3**

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

**Question No : 4**

"The programming model for client developers will hide the complexity of interfacing to legacy systems" is an example of which of the following?

**A.** A use case
**B.** An architectural decision
**C.** A client imperative
**D.** An empathy statement

**Answer: B**

## Question No : 5

Which of the following is the artifact that assists in ensuring that the project is on the right path toward success?

**A.** Component Model
**B.** Empathy Map
**C.** Viability Assessment
**D.** Opportunity Plan

**Answer: C**

## Question No : 6

An upstream Oil and Gas Producer needs to optimize the performance of its assets. It needs to calculate Key Performance Indicators for flow rate sensors deployed to monitor the output flow and temperature and pressure for multiple pipelines in an oil field with hundreds of wells. Which of the following would you recommend to meet these requirements?

**A.** Datastage ETL jobs should be created
**B.** Infosphere Streams should be used
**C.** Cognos reports would suffice here
**D.** Hadoop based datastorage engine should be used

**Answer: B**

## Question No : 7

Which of the following statements is TRUE?

**A.** A good use of the BigInsights is as a query-ready archival system for your data warehouse to quickly accessdata
**B.** BigInsights and Hadoop based systems in general are best used in high concurrency transactional systems
**C.** Hadoop map reduce based computing engine is ideal for use as a real-time or near real-time processingextension to your existing business intelligence reporting
**D.** The system ML engine is the preferable option to add unstructured text sentiment analysis to the customerservice reporting system

**Answer: A**

**Explanation:**

References:
http://www.ibm.com/support/knowledgecenter/SSPT3X_2.1.2/
com.ibm.swg.im.infosphere.biginsights.product.doc/doc/bi_enterprise_architecture.html

## Question No : 8

In designing a new Hadoop system for a customer, the option of using SAN versus DAS was brought up. Which of the following would justify choosing SAN storage?

**A.** SAN storage provides better performance than DAS
**B.** SAN storage reduces and removes a lot of the HDFS complexity and management issues
**C.** SAN storage removes the Single Point of Failure for the NameNode
**D.** SAN storage supports replication, reducing the need for 3-way replication

**Answer: D**

## Question No : 9

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**

Reference:

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

## Question No : 10

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**

## Question No : 11

Which of the following can NOT be performed by the Open Data Platform (ODP)?

**A.** Provide development and governance of upstream projects
**B.** Contribute to the Apache Software Foundation projects
**C.** Produce a set of tools and methods that enable members to create and test differentiated offerings based on the ODP core
**D.** Accelerate the delivery of Big Data solutions by providing a well-defined core consisting of all the components of the Apache Software Foundation projects

**Answer: C**

## Question No : 12

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.** JobTracker
**B.** Zookeeper
**C.** Task Tracker
**D.** Name Node

**Answer: C**

---

**Question No : 13**

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**

---

**Question No : 14**

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**
Reference:

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

## Question No : 15

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**

## Question No : 16

You are building a pricing application combining the output from your internal customer database along with an online feed that contains up to the minute competitor pricing adjustments. Which of the following is the best choice for data level real time output?

**A.** Big R
**B.** Big SQL
**C.** Hadoop native hdfs map reduce based queries
**D.** Infosphere streams

**Answer: D**
**Explanation:**

References: http://www-03.ibm.com/software/products/en/ibm-streams

## Question No : 17

Which of the following statements regarding Big R is TRUE?

**A.** Unless specified otherwise, Big R automatically assumes all data to be integers
**B.** Big R's 'bigr.frame' is equivalent to R's 'data.frames'
**C.** When you execute Big R "apply" function, Big R transparently extracts data out of HDFS into the Big R engine

**D.** A data analyst using Big R employs MapReduce programming principles

**Answer: A**

Reference:

https://developer.ibm.com/hadoop/docs/biginsights-value-add/big-r/bigr-tutorial/

## Question No : 18

SPSS and BigInsights integration is enabled by which of the following?

**A.** SPSS Analytic Server
**B.** InfoSphere DataStage
**C.** SPSS Modeler
**D.** SPSS Collaboration & Deployment Services

**Answer: A**

**Explanation:**

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

## Question No : 19

Which of the following is true about Cloudant?

**A.** It is a byte-value store
**B.** It is available as a database-as-a-service (DBaaS)
**C.** It runs remotely in the cloud and not on a local system
**D.** It offers SQL access to data

**Answer: B**

**Explanation:**

References: http://www-01.ibm.com/software/data/cloudant/

## Question No : 20