



IBM C2090-102

IBM BIG DATA ARCHITECT CERTIFICATION QUESTIONS & ANSWERS

Exam Summary – Syllabus – Questions

C2090-102

IBM Certified Big Data Architect

55 Questions Exam – 60% Cut Score – Duration of 90 minutes

www.AnalyticsExam.Com

Table of Contents

Know Your C2090-102 Certification Well:	2
C2090-102 IBM Big Data Architect Certification Details:	2
C2090-102 Syllabus:	3
IBM C2090-102 Sample Questions:	4
Study Guide to Crack IBM Big Data Architect C2090-102 Exam:	7

Know Your C2090-102 Certification Well:

The C2090-102 is best suitable for candidates who want to gain knowledge in the IBM Data and AI - Platform Analytics. Before you start your C2090-102 preparation you may struggle to get all the crucial IBM Big Data Architect materials like C2090-102 syllabus, sample questions, study guide.

But don't worry the C2090-102 PDF is here to help you prepare in a stress free manner.

The PDF is a combination of all your queries like-

- What is in the C2090-102 syllabus?
- How many questions are there in the C2090-102 exam?
- Which Practice test would help me to pass the C2090-102 exam at the first attempt?

Passing the C2090-102 exam makes you IBM Certified Big Data Architect. Having the IBM Big Data Architect certification opens multiple opportunities for you. You can grab a new job, get a higher salary or simply get recognition within your current organization.

C2090-102 IBM Big Data Architect Certification Details:

Exam Name	IBM Certified Data Architect - Big Data
Exam Code	C2090-102
Exam Duration	90 minutes
Exam Questions	55
Passing Score	60%
Exam Price	\$200 (USD)
Training	BigInsights Analytics for Programmers
Exam Registration	PearsonVUE
Sample Questions	IBM Big Data Architect Certification Sample Question
Practice Exam	IBM Big Data Architect Certification Practice Exam

C2090-102 Syllabus:

Objective	Details	Weight
Requirements	<ul style="list-style-type: none"> - Define the input data structure - Define the outputs - Define the security requirements - Define the requirements for replacing and/or merging with existing business solutions - Define the solution to meet the customer's SLA - Define the network requirements based on the customer's requirements 	16%
Use Cases	<ul style="list-style-type: none"> - Determine when a cloud based solution is more appropriate vs. in-house (and migration plans from one to the other) - Demonstrate why Cloudant would be an applicable technology for a particular use case - Demonstrate why SQL or NoSQL would be an applicable technology for a particular use case - Demonstrate why Open Data Platform would be an applicable technology for a particular use case - Demonstrate why BigInsights would be an applicable technology for a particular use case - Demonstrate why BigSQL would be an applicable technology for a particular use case - Demonstrate why Hadoop would be an applicable technology for a particular use case - Demonstrate why BigR and SPSS would be an applicable technology for a particular use case - Demonstrate why BigSheets would be an applicable technology for a particular use case - Demonstrate why Streams would be an applicable technology for a particular use case - Demonstrate why Netezza would be an applicable technology for a particular use case - Demonstrate why DB2 BLU would be an applicable technology for a particular use case - Demonstrate why GPFS/HPFS would be an applicable technology for a particular use case - Demonstrate why Spark would be an applicable technology for a particular use case - Demonstrate why YARN would be an applicable technology for a particular use case 	46%
Applying Technologies	<ul style="list-style-type: none"> - Define the necessary technology to ensure horizontal and vertical scalability - Determine data storage requirements based on data volumes - Design a data model and data flow model that will meet the business requirements - Define the appropriate Big Data technology for a given customer requirement (e.g. Hive/HBase or Cloudant) - Define appropriate storage format and compression for given customer requirement 	16%
Recoverability	<ul style="list-style-type: none"> - Define the potential need for high availability 	11%

Objective	Details	Weight
	<ul style="list-style-type: none"> - Define the potential disaster recovery requirements - Define the technical requirements for data retention - Define the technical requirements for data replication - Define the technical requirements for preventing data loss 	
Infrastructure	<ul style="list-style-type: none"> - Define the hardware and software infrastructure requirements - Design the integration of the required hardware and software components - Design the connectors / interfaces / API's between the Big Data solution and the existing systems 	11%

IBM C2090-102 Sample Questions:

Question: 1

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: 2

Faced with a wide area network implementation, you have a need for asynchronous remote updates. Which one of the following would best address this use case?

- a) GPFS Active File Management allows data access and modifications even when remote storage cluster is unavailable
- b) HDFS Cluster rebalancing is compatible with data rebalancing schemes. A scheme might automatically move data from one DataNode to another if the free space on a DataNode falls below a certain threshold
- c) GPFS File clones can be created from a regular file or a file in a snapshot using the mmclone command
- d) HDFS NameNode The NameNode keeps an image of the entire file system namespace and file Blockmap in memory. This key metadata item is designed to be compact, such that a NameNode with 4 GB of RAM is plenty to support a huge number of files and directories

Answer: c

Question: 3

A telecommunication company needs a Big Data solution that could store and analyze multiple years worth of call detail records (CDRs, aprox. 17 billion events per day) containing switch, billing, and network event data for its millions of subscribers.

Which of the following would you recommend for these requirements?

- a) Infosphere DataStage
- b) DB2
- c) Pure Data System for Analytics
- d) SPSS

Answer: c

Question: 4

A large Retailer (online and “brick & mortar”) processes data for analyzing marketing campaigns for their loyalty club members. The current process takes weeks for processing only 10% of social data.

What is the most cost effective platform for processing and analyzing campaign results from social data on a daily basis using 100% dataset?

- a) Enterprise Data Warehouse
- b) BigInsights Open Data Platform
- c) High Speed Mainframe Processing
- d) In Memory Computing

Answer: b

Question: 5

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

Question: 6

A bank wants to build a system that tracks all ATM and online transactions in realtime. They want to build a personalized model of their customer's financial activity by incorporating enterprise data as well as social media data.

The system must be able to learn and adapt over a period of time. These personalized models will be used for real time promotions as well as for any fraud or crime detections.

Given these requirements, which of the following would recommend?

- a) Spark
- b) Hadoop
- c) Cloudand
- d) Netezza

Answer: d

Question: 7

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 vs. scale out (horizontal) scaling
- d) Server virtualization is a requirement in a cloud implementation

Answer: d

Question: 8

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

Answer: b

Question: 9

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

Question: 10

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 decide whether or 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

Study Guide to Crack IBM Big Data Architect C2090-102 Exam:

- Getting details of the C2090-102 syllabus, is the first step of a study plan. This pdf is going to be of ultimate help. Completion of the syllabus is must to pass the C2090-102 exam.
- Making a schedule is vital. A structured method of preparation leads to success. A candidate must plan his schedule and follow it rigorously to attain success.
- Joining the IBM provided training for C2090-102 exam could be of much help. If there is specific training for the exam, you can discover it from the link above.
- Read from the C2090-102 sample questions to gain your idea about the actual exam questions. In this PDF useful sample questions are provided to make your exam preparation easy.
- Practicing on C2090-102 practice tests is must. Continuous practice will make you an expert in all syllabus areas.

Reliable Online Practice Test for C2090-102 Certification

Make AnalyticsExam.Com your best friend during your IBM Certified Data Architect - Big Data exam preparation. We provide authentic practice tests for the C2090-102 exam. Experts design these online practice tests, so we can offer you an exclusive experience of taking the actual C2090-102 exam. We guarantee you 100% success in your first exam attempt if you continue practicing regularly. Don't bother if you don't get 100% marks in initial practice exam attempts. Just utilize the result section to know your strengths and weaknesses and prepare according to that until you get 100% with our practice tests. Our evaluation makes you confident, and you can score high in the C2090-102 exam.

Start Online Practice of C2090-102 Exam by Visiting URL

<https://www.analyticsexam.com/ibm-certification/c2090-102-ibm-certified-data-architect-big-data>