

1. What is the primary purpose of integrating disparate data stores in big data processing?

- a) To increase the complexity of data analysis
- b) To simplify data retrieval processes
- c) To ensure data consistency and accuracy
- d) To minimize the need for data transformation

Answer: c) To ensure data consistency and accuracy

Explanation: Integrating disparate data stores helps in bringing together various sources of data into a unified format, ensuring consistency and accuracy in the data being processed.

2. In big data processing, what does mapping data to the programming framework involve?

- a) Assigning geographical coordinates to data points
- b) Linking data sets to a specific programming language
- c) Defining how data will be processed within the chosen programming model
- d) Sorting data into alphabetical order

Answer: c) Defining how data will be processed within the chosen programming model

Explanation: Mapping data to the programming framework involves defining how data will be processed within the selected programming model, ensuring efficient utilization of computational resources.

3. Which stage of big data processing involves connecting and extracting data from storage?

- a) Data transformation

- b) Data integration
- c) Data retrieval
- d) Data mapping

Answer: c) Data retrieval

Explanation: Data retrieval involves connecting to data storage systems and extracting the required data for processing.

4. What is the purpose of transforming data in big data processing?

- a) To increase data complexity
- b) To reduce storage requirements
- c) To prepare data for analysis
- d) To remove redundant data

Answer: c) To prepare data for analysis

Explanation: Transforming data involves converting raw data into a format suitable for analysis, including cleaning, filtering, and structuring it appropriately.

5. Which technique is commonly used for subdividing data in preparation for Hadoop MapReduce?

- a) Data sharding
- b) Data serialization
- c) Data replication
- d) Data compression

Answer: a) Data sharding

Explanation: Data sharding involves dividing large datasets into smaller, manageable chunks for parallel processing in frameworks like Hadoop MapReduce.

6. What does the acronym "HDFS" stand for in the context of big data processing?

- a) Hadoop Distributed File System
- b) High Data Flow System
- c) Hierarchical Data Formatting System
- d) Heterogeneous Data Fusion System

Answer: a) Hadoop Distributed File System

Explanation: HDFS (Hadoop Distributed File System) is a distributed file system designed to store large volumes of data across multiple machines in a Hadoop cluster.

7. Which of the following is a key advantage of using Hadoop MapReduce for big data processing?

- a) Real-time data processing
- b) High throughput for small datasets
- c) Fault tolerance and scalability
- d) Simple programming model

Answer: c) Fault tolerance and scalability

Explanation: Hadoop MapReduce provides fault tolerance and scalability by distributing

computation across a cluster of commodity hardware, enabling efficient processing of large datasets.

8. What role does Apache Spark play in big data processing?

- a) Data storage
- b) Data visualization
- c) Data processing and analytics
- d) Data indexing

Answer: c) Data processing and analytics

Explanation: Apache Spark is a fast and general-purpose distributed computing system used for big data processing, including data processing and analytics tasks.

9. Which component of big data processing frameworks is responsible for job scheduling and resource management?

- a) YARN (Yet Another Resource Negotiator)
- b) HBase
- c) Kafka
- d) Spark SQL

Answer: a) YARN (Yet Another Resource Negotiator)

Explanation: YARN (Yet Another Resource Negotiator) is a key component of Hadoop that manages resources and schedules jobs in a distributed environment, facilitating efficient resource utilization.

10. What is the purpose of data serialization in big data processing?

- a) To optimize data storage
- b) To convert data into a readable format
- c) To transfer data over a network
- d) To enable efficient data processing

Answer: c) To transfer data over a network

Explanation: Data serialization involves converting complex data structures into a format suitable for transmission over a network, facilitating efficient data transfer between different components of a big data processing system.

Related posts:

1. DESCRIPTIVE STATISTICS MCQs
2. INTRODUCTION TO BIG DATA MCQ
3. BIG DATA TECHNOLOGIES MCQs
4. HADOOP MAPREDUCE MCQs
5. BIG DATA TOOLS AND TECHNIQUES MCQs
6. Introduction to Energy Science MCQ
7. Ecosystems MCQ
8. Biodiversity and its conservation MCQ
9. Environmental Pollution mcq
10. Social Issues and the Environment MCQ
11. Field work mcq
12. Discrete Structure MCQ
13. Set Theory, Relation, and Function MCQ

14. Propositional Logic and Finite State Machines MCQ
15. Graph Theory and Combinatorics MCQ
16. Relational algebra, Functions and graph theory MCQ
17. Data Structure MCQ
18. Stacks MCQ
19. TREE MCQ
20. Graphs MCQ
21. Sorting MCQ
22. Digital Systems MCQ
23. Combinational Logic MCQ
24. Sequential logic MCQ
25. Analog/Digital Conversion, Logic Gates, Multivibrators, and IC 555 MCQ
26. Introduction to Digital Communication MCQ
27. Introduction to Object Oriented Thinking & Object Oriented Programming MCQ
28. Encapsulation and Data Abstraction MCQ
29. MCQ
30. Relationships - Inheritance MCQ
31. Polymorphism MCQ
32. Library Management System MCQ
33. Numerical Methods MCQ
34. Transform Calculus MCQ
35. Concept of Probability MCQ
36. Algorithms, Designing MCQ
37. Study of Greedy strategy MCQ
38. Concept of dynamic programming MCQ
39. Algorithmic Problem MCQ
40. Trees, Graphs, and NP-Completeness MCQ

41. The Software Product and Software Process MCQ
42. Software Design MCQ
43. Software Analysis and Testing MCQ
44. Software Maintenance & Software Project Measurement MCQ
45. Computer Architecture, Design, and Memory Technologies MCQ
46. Basic Structure of Computer MCQ
47. Computer Arithmetic MCQ
48. I/O Organization MCQ
49. Memory Organization MCQ
50. Multiprocessors MCQ
51. Introduction to Operating Systems MCQ
52. File Systems MCQ
53. CPU Scheduling MCQ
54. Memory Management MCQ
55. Input / Output MCQ
56. Operating Systems and Concurrency
57. Software Development and Architecture MCQ
58. Software architecture models MCQ
59. Software architecture implementation technologies MCQ
60. Software Architecture analysis and design MCQ
61. Software Architecture documentation MCQ
62. Introduction to Computational Intelligence MCQ
63. Fuzzy Systems MCQ
64. Genetic Algorithms MCQ
65. Rough Set Theory MCQ
66. Introduction to Swarm Intelligence, Swarm Intelligence Techniques MCQ
67. Neural Network History and Architectures MCQ

68. Autoencoder MCQ
69. Deep Learning MCQs
70. RL & Bandit Algorithms MCQs
71. RL Techniques MCQs
72. Review of traditional networks MCQ
73. Study of traditional routing and transport MCQ
74. Wireless LAN MCQ
75. Mobile transport layer MCQ
76. Big Data MCQ
77. Hadoop and Related Concepts MCQ
78. Hive, Pig, and ETL Processing MCQ
79. NoSQL MCQs Concepts, Variations, and MongoDB
80. Mining social Network Graphs MCQ
81. Mathematical Background for Cryptography MCQ
82. Cryptography MCQ
83. Cryptographic MCQs
84. Information Security MCQ
85. Cryptography and Information Security Tools MCQ
86. Data Warehousing MCQ
87. OLAP Systems MCQ
88. Introduction to Data & Data Mining MCQ
89. Supervised Learning MCQ
90. Clustering & Association Rule mining MCQ
91. Fundamentals of Agile Process MCQ
92. Agile Projects MCQs
93. Introduction to Scrum MCQs
94. Introduction to Extreme Programming (XP) MCQs



95. Agile Software Design and Development MCQs
96. Machine Learning Fundamentals MCQs
97. Neural Network MCQs
98. CNNs MCQ
99. Reinforcement Learning and Sequential Models MCQs
100. Machine Learning in ImageNet Competition mcq
101. Computer Network MCQ
102. Data Link Layer MCQ
103. MAC Sub layer MCQ
104. Network Layer MCQ
105. Transport Layer MCQ
106. Raster Scan Displays MCQs
107. 3-D Transformations MCQs
108. Visualization MCQ
109. Multimedia MCQs
110. Introduction to compiling & Lexical Analysis MCQs
111. Syntax Analysis & Syntax Directed Translation MCQs
112. Type Checking & Run Time Environment MCQs
113. Code Generation MCQs
114. Code Optimization MCQs
115. INTRODUCTION Knowledge Management MCQs
116. Organization and Knowledge Management MCQs
117. Telecommunications and Networks in Knowledge Management MCQs
118. Components of a Knowledge Strategy MCQs
119. Advanced topics and case studies in knowledge management MCQs
120. Conventional Software Management MCQs
121. Software Management Process MCQs

122. Software Management Disciplines MCQs
123. Rural Management MCQs
124. Human Resource Management for rural India MCQs
125. Management of Rural Financing MCQs
126. Research Methodology MCQs
127. Research Methodology MCQs
128. IoT MCQs
129. Sensors and Actuators MCQs
130. IoT MCQs: Basics, Components, Protocols, and Applications
131. MCQs on IoT Protocols
132. IoT MCQs
133. INTRODUCTION Block Chain Technologies MCQs
134. Understanding Block chain with Crypto currency MCQs
135. Understanding Block chain for Enterprises MCQs
136. Enterprise application of Block chain MCQs
137. Block chain application development MCQs
138. MCQs on Service Oriented Architecture, Web Services, and Cloud Computing
139. Utility Computing, Elastic Computing, Ajax MCQs
140. Data in the cloud MCQs
141. Cloud Security MCQs
142. Issues in cloud computing MCQs
143. Introduction to modern processors MCQs
144. Data access optimizations MCQs
145. Parallel Computing MCQs
146. Efficient Open MP Programming MCQs
147. Distributed Memory parallel programming with MPI MCQs
148. Review of Object Oriented Concepts and Principles MCQs.

149. Introduction to RUP MCQs.
150. UML and OO Analysis MCQs
151. Object Oriented Design MCQs
152. Object Oriented Testing MCQs
153. CVIP Basics MCQs
154. Image Representation and Description MCQs
155. Region Analysis MCQs
156. Facet Model Recognition MCQs
157. Knowledge Based Vision MCQs
158. Game Design and Semiotics MCQs
159. Systems and Interactivity Understanding Choices and Dynamics MCQs
160. Game Rules Overview Concepts and Case Studies MCQs
161. IoT Essentials MCQs
162. Sensor and Actuator MCQs
163. IoT Networking & Technologies MCQs
164. MQTT, CoAP, XMPP, AMQP MCQs
165. IoT MCQs: Platforms, Security, and Case Studies
166. MCQs on Innovation and Entrepreneurship
167. Innovation Management MCQs
168. Stage Gate Method & Open Innovation MCQs
169. Innovation in Business: MCQs
170. Automata Theory MCQs
171. Finite Automata MCQs
172. Grammars MCQs
173. Push down Automata MCQs
174. Turing Machine MCQs
175. Database Management System (DBMS) MCQs

176. Relational Data models MCQs
177. Data Base Design MCQs
178. Transaction Processing Concepts MCQs
179. Control Techniques MCQs
180. DBMS Concepts & SQL Essentials MCQs
181. Pattern Recognition MCQs
182. Classification Algorithms MCQs
183. Pattern Recognition and Clustering MCQs
184. Feature Extraction & Selection Concepts and Algorithms MCQs
185. Pattern Recognition MCQs
186. Understanding Cybercrime Types and Challenges MCQs
187. Cybercrime MCQs
188. Cyber Crime and Criminal justice MCQs
189. Electronic Evidence MCQs
190. Data Science MCQs
191. DBMS Normalization MCQs
192. Advanced Computer Architecture MCQ
193. Introduction to Information Security MCQ
194. HTML MCQs
195. Basic concepts of OOP MCQs
196. File System MCQs
197. Social Issues and the Environment mcqs
198. FM Modulation & Transmission MCQs
199. Introduction to ICs and Op-Amps MCQs
200. Efficient Computation of the DFT mcqs