1. What is Hadoop MapReduce primarily used for?

a) Real-time data processing
b) Batch processing of large datasets
c) Stream processing
d) In-memory data processing

Answer: b) Batch processing of large datasets

Explanation: Hadoop MapReduce is designed for processing large datasets in batch mode, making it suitable for tasks like data warehousing, ETL (Extract, Transform, Load), and analytics where data can be processed in parallel across a distributed system.

2. Which of the following components are essential for creating Hadoop MapReduce jobs?

a) NameNode and DataNode
b) JobTracker and TaskTracker
c) ResourceManager and NodeManager
d) Mapper and Reducer

Answer: d) Mapper and Reducer

Explanation: The Mapper and Reducer are the key components of a Hadoop MapReduce job. The Mapper processes input data and emits intermediate key-value pairs, while the Reducer aggregates and processes these intermediate results to produce the final output.

3. What is the purpose of distributing data processing across server farms in Hadoop MapReduce?

a) To minimize network latency
b) To reduce the load on individual servers
c) To achieve fault tolerance and scalability
d) To increase power consumption

Answer: c) To achieve fault tolerance and scalability

Explanation: Distributing data processing across server farms in Hadoop MapReduce enables fault tolerance by replicating data and computations across multiple nodes. It also allows for scalability, as additional nodes can be added to handle increased workloads.

4. How can you monitor the progress of job flows in Hadoop MapReduce?

a) Using Hadoop Distributed File System (HDFS)
b) Through the ResourceManager web interface
c) By accessing the NameNode logs
d) Using the Hadoop shell commands

Answer: b) Through the ResourceManager web interface

Explanation: The ResourceManager web interface provides real-time information about the status and progress of Hadoop MapReduce jobs, including details such as job duration, resource usage, and task progress.

5. Which of the following are examples of Hadoop daemons?

a) NameNode and DataNode
b) Mapper and Reducer

c) ResourceManager and NodeManager
d) JobTracker and TaskTracker

Answer: a) NameNode and DataNode

Explanation: NameNode and DataNode are examples of Hadoop daemons responsible for managing the distributed file system (HDFS) and storing data blocks, respectively.

6. In Hadoop MapReduce, which execution mode is suitable for development and testing on a single machine?

a) Fully distributed mode
b) Pseudo-distributed mode
c) Local mode
d) Remote mode

Answer: c) Local mode

Explanation: Local mode in Hadoop MapReduce allows developers to test their code on a single machine without the need for a cluster setup. It simulates the Hadoop environment locally.

7. What is the primary function of the NameNode in Hadoop Distributed File System (HDFS)?

a) Storing data blocks
b) Managing metadata and namespace
c) Executing MapReduce tasks
d) Allocating system resources

Answer: b) Managing metadata and namespace

Explanation: The NameNode in HDFS is responsible for storing metadata and managing the namespace, including information about file locations, permissions, and replication.

8. Which execution mode in Hadoop MapReduce mimics a real distributed cluster but runs on a single machine?

a) Local mode
b) Pseudo-distributed mode
c) Fully distributed mode
d) Hybrid mode

Answer: b) Pseudo-distributed mode

Explanation: Pseudo-distributed mode in Hadoop MapReduce simulates a distributed cluster environment on a single machine, allowing developers to test their code in a setup similar to a real cluster.

9. What is the primary role of the ResourceManager in Hadoop YARN (Yet Another Resource Negotiator)?

a) Managing storage resources
b) Managing compute resources
c) Distributing data across nodes
d) Monitoring job progress

Answer: b) Managing compute resources

Explanation: The ResourceManager in Hadoop YARN is responsible for managing compute resources in a Hadoop cluster, including allocating resources to applications and tracking resource utilization.

10. Which Hadoop component is responsible for launching and monitoring MapReduce tasks on individual nodes in the cluster?

a) ResourceManager
b) NodeManager
c) JobTracker
d) TaskTracker

Answer: b) NodeManager

Explanation: NodeManager is responsible for launching and monitoring MapReduce tasks on individual nodes in a Hadoop cluster, managing resources and reporting status back to the ResourceManager.

11. What is the primary advantage of using Hadoop MapReduce for data processing?

a) Low latency
b) Real-time processing
c) Fault tolerance
d) Limited scalability

Answer: c) Fault tolerance

Explanation: Hadoop MapReduce provides fault tolerance by replicating data and

computations across multiple nodes in a distributed cluster, ensuring that processing can continue even if individual nodes fail.

12. Which execution mode in Hadoop MapReduce is suitable for production deployments across a cluster of machines?

a) Local mode
b) Pseudo-distributed mode
c) Fully distributed mode
d) Hybrid mode

Answer: c) Fully distributed mode

Explanation: Fully distributed mode in Hadoop MapReduce is suitable for production deployments across a cluster of machines, where data processing tasks are distributed across multiple nodes for parallel execution.

13. What is the role of the DataNode in Hadoop Distributed File System (HDFS)?

a) Managing metadata
b) Storing data blocks
c) Managing compute resources
d) Monitoring job progress

Answer: b) Storing data blocks

Explanation: DataNode in HDFS is responsible for storing actual data blocks on the disk and performing read and write operations as requested by clients or other Hadoop components.

14. Which Hadoop daemon is responsible for maintaining a global picture of the cluster and managing resources?

a) ResourceManager
b) NodeManager
c) NameNode
d) DataNode

Answer: a) ResourceManager

Explanation: ResourceManager in Hadoop YARN is responsible for maintaining a global view of the cluster and managing its resources, including allocating resources to applications and monitoring resource utilization.

15. What does the Reducer component do in a Hadoop MapReduce job?

a) Processes input data and emits intermediate key-value pairs
b) Aggregates and processes intermediate results to produce final output
c) Manages compute resources in a Hadoop cluster
d) Allocates storage resources in Hadoop Distributed File System

Answer: b) Aggregates and processes intermediate results to produce final output

Explanation: The Reducer component in Hadoop MapReduce aggregates and processes intermediate results generated by the Mapper, producing the final output of the job.

16. Which mode of execution in Hadoop MapReduce is recommended for production environments where data processing needs to be distributed across a large cluster?

a) Local mode
b) Pseudo-distributed mode
c) Fully distributed mode
d) Hybrid mode

Answer: c) Fully distributed mode

Explanation: Fully distributed mode in Hadoop MapReduce is recommended for production environments where data processing tasks need to be distributed across a large cluster of machines for parallel execution.

17. What is the primary function of the Job Tracker in Hadoop MapReduce?

a) Managing storage resources
b) Managing compute resources
c) Monitoring job progress
d) Distributing data across nodes

Answer: c) Monitoring job progress

Explanation: The JobTracker in Hadoop MapReduce is responsible for monitoring the progress of MapReduce jobs, scheduling tasks, and handling task failures in a Hadoop cluster.

18. In Hadoop MapReduce, what is the purpose of the Mapper component?

a) Aggregating intermediate results
b) Processing input data and emitting key-value pairs
c) Distributing data across nodes

d) Monitoring job progress

Answer: b) Processing input data and emitting key-value pairs

Explanation: The Mapper component in Hadoop MapReduce processes input data and emits intermediate key-value pairs based on the logic defined by the developer.

19. Which execution mode in Hadoop MapReduce simulates a real distributed cluster but runs on a limited number of machines?

a) Local mode
b) Pseudo-distributed mode
c) Fully distributed mode
d) Hybrid mode

Answer: b) Pseudo-distributed mode

Explanation: Pseudo-distributed mode in Hadoop MapReduce simulates a real distributed cluster environment but runs on a limited number of machines, making it suitable for testing and development purposes.

20. What is the primary benefit of distributing data processing across server farms in Hadoop MapReduce?

a) Minimizing network latency
b) Reducing the load on individual servers
c) Achieving fault tolerance and scalability
d) Increasing power consumption

Answer: c) Achieving fault tolerance and scalability

Explanation: Distributing data processing across server farms in Hadoop MapReduce enables fault tolerance and scalability by replicating data and computations across multiple nodes, allowing for parallel execution of tasks and handling of failures.

Related posts:

1. DESCRIPTIVE STATISTICS MCQs
2. INTRODUCTION TO BIG DATA MCQ
3. BIG DATA TECHNOLOGIES MCQs
4. PROCESSING BIG DATA 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. XML MCQs
191. System Security MCQs.
192. Linear Time- Invariant Systems mcqs
193. Control System MCQs: Basics, Feedback, and Analysis
194. OP-AMP applications MCQs
195. Radiation mcqs
196. NETWORKS mcqs
197. Satellite Services MCQs
198. NON-ELECTRICAL PARAMETER MEASUREMENTS mcqs
199. Practical Consideration and Technology in VLSI Design MCQs
200. Microwave Components and Circuits MCQs