BRAINWARE UNIVERSITY

Term End Examination 2023
Programme – M.Tech.(CSE)-2021
Course Name – Advanced DBMS
Course Code - PCC-MCS102
(Semester I)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A

(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) How do integrity constraints play a role in maintaining data consistency within a distributed database?

a) They ensure that only centralized databases are used.
b) They prevent unauthorized access to distributed databases.
c) They enforce rules to ensure valid and consistent data across locations.
d) They facilitate data replication in distributed databases.

(ii) Explain the concept of global query optimization in distributed databases and why it is important.

a) Global query optimization ensures that all queries are executed globally.
b) Global query optimization involves optimizing queries for multiple distributed databases.
c) Global query optimization improves the performance of individual queries in a distributed environment.
d) Global query optimization refers to optimizing queries for centralized databases.

(iii) How does concurrency control based on timestamps work in a distributed database, and what benefits does it offer?

a) Concurrency control based on timestamps involves assigning timestamps to nodes in a network.
b) Concurrency control based on timestamps uses time-based ordering to manage concurrent access to data.
c) Concurrency control based on timestamps prevents all transactions from being executed simultaneously.
d) Concurrency control based on timestamps relies solely on transaction order.

(iv) Define the concept of reliability in a distributed database system and explain its significance.

a) Reliability refers to the speed of database operations in a distributed system.
b) Reliability is the ability of a database to handle only read-only operations.

c) Reliability ensures the availability and consistency of data in the face of failures or errors.

d) Reliability is the process of replicating data to multiple nodes.

(v) How do non-blocking commitment protocols contribute to ensuring reliability in distributed transactions?

a) Non-blocking commitment protocols prevent all transactions from being executed.

b) Non-blocking commitment protocols allow transactions to be executed without any restrictions.

c) Non-blocking commitment protocols ensure that transactions are not blocked by locking issues, enhancing reliability.

d) Non-blocking commitment protocols are only relevant for centralized databases.

(vi) Describe the purpose and function of checkpoints in the context of recovery mechanisms for distributed databases.

a) Checkpoints are points where transactions are terminated.

b) Checkpoints are used to save the current state of the database to support recovery in case of failures.

c) Checkpoints ensure that only authorized users can access the database.

d) Checkpoints are used to reverse transactions.

(vii) What is the role of distributed data dictionary management in a distributed database system?

a) Distributed data dictionary management is responsible for managing data in centralized databases.

b) Distributed data dictionary management handles data fragmentation and allocation.

c) Distributed data dictionary management ensures data consistency and integrity across distributed databases.

d) Distributed data dictionary management is only relevant for read-only applications.

(viii) Compare and contrast loosely coupled and tightly coupled architectures in the context of distributed databases.

a) Loosely coupled architecture involves strong interconnection between nodes, while tightly coupled architecture involves weak interconnection.

b) Loosely coupled architecture has minimal communication between nodes, while tightly coupled architecture involves extensive communication.

c) Loosely coupled architecture is only used in centralized databases, while tightly coupled architecture is used in distributed databases.

d) Loosely coupled and tightly coupled architectures are the same concept and can be used interchangeably.

(ix) What does distributed database administration entail, and what challenges does it pose compared to centralized databases?

a) Distributed database administration involves managing databases in multiple locations, and it poses challenges related to data consistency, security, and communication.

b) Distributed database administration focuses on optimizing query performance in a distributed environment.

c) Distributed database administration only deals with data replication strategies.

d) Distributed database administration is not relevant for modern databases.

(x) Define a heterogeneous distributed database system and explain its advantages in various real-world scenarios.

a) A heterogeneous distributed database system consists of identical nodes and offers limited flexibility.

b) A heterogeneous distributed database system involves nodes with different hardware and software, providing flexibility and integration across diverse environments.

c) A heterogeneous distributed database system can only store a single type of data.

d) A heterogeneous distributed database system is the same as a centralized database system.

(xi) Examine the challenges that might arise when translating a complex global query into fragmented queries for a distributed database system.

a) Complex queries require more network bandwidth.

b) Fragmented queries can result in increased query execution time.

c) Network latency is not relevant for query fragmentation.

d) Complex queries can only be executed on a centralized database.

(xii) Devise a protocol for deadlock detection and resolution in a distributed database system that involves multiple nodes, considering the potential interactions between transactions.

a) Abort all transactions involved in a deadlock.

b) Implement a timeout mechanism for transactions in a deadlock.

c) Use a graph-based algorithm to detect and break the deadlock.

d) Ignore deadlocks and let the system resolve them naturally.

(xiii) Evaluate the effectiveness of different recovery strategies, such as deferred updates and immediate updates, in a distributed database system.

a) Deferred updates ensure higher levels of data consistency.

b) Immediate updates are more suitable for read-only applications.

c) Deferred updates guarantee better fault tolerance.

d) Immediate updates reduce the need for recovery mechanisms.

(xiv) Develop a plan for recovering from a catastrophic node failure in a distributed database system, outlining the steps from identifying the failure to restoring normal operation.

a) Identify the failure, initiate a node restart, and resume transactions.

b) Trigger a system-wide restart to restore normal operation.

c) Identify the failure, replace the node, and recover data from backups.

d) Ignore the failure and let the system continue operating normally.

(xv) Analyze the benefits and challenges of using quorum-based systems for handling network partitions in distributed databases, considering scenarios with different numbers of nodes.

a) Quorum-based systems have no impact on data availability.

b) Smaller quorums are more resilient to network partitions.

c) Larger quorums can lead to lower system availability.

d) Quorum-based systems are only relevant for centralized databases.

Group-B

(Short Answer Type Questions) 3 x 5=15

2. Differentiate between distribution transparency levels for read-only applications and update applications. (3)
3. Define fragmentation in the context of distributed database design and mention the purpose of allocation in this process. (3)
4. Explain the concept of a Distributed Database Management System (DDBMS) and its role in managing data in a distributed environment. (3)
5. Illustrate the concept of distributed database administration. What are the key responsibilities of a distributed database administrator in managing and maintaining such systems? (3)
6. What is Distributed Data Dictionary Management, and why is it essential in a distributed database? How does it differ from a centralized data dictionary? (3)

OR

Differentiate between loosely coupled and tightly coupled heterogeneous distributed database systems. Provide examples of scenarios where each architecture is more suitable. (3)

Group-C
(Long Answer Type Questions) 5 x 6=30

7. Discuss the concept of integrity constraints in a distributed database. How do these constraints ensure data accuracy and consistency in a distributed environment? (5)
8. Elaborate on the concepts of fragmentation and allocation of fragmentation in distributed database design. How do these techniques contribute to optimizing data distribution and access efficiency? (5)
9. Explain the phases involved in the execution of a distributed query. How does the access plan impact the performance of a distributed query? (5)
10. Present an overview of the transaction management framework in a distributed database system. Highlight the roles of the transaction coordinator and participants, and explain how the two-phase commit protocol ensures transaction atomicity. (5)
11. Outline the concurrency control approach based on timestamps in a distributed database. Explain how this method utilizes timestamps to order transactions and ensure serializability. Discuss its advantages and limitations. (5)
12. Discuss the architectural aspects of handling node and link failures in a distributed database. How do distributed systems recover from these failures to maintain data availability and consistency? (5)

OR

Explain the challenges and strategies associated with recovering from node failures in a distributed environment. How does the availability of distributed backups influence recovery procedures? (5)

****