Exam. Code: 0919
Sub. Code: 33473

2076

B.E. (Computer Science and Engineering)
Seventh Semester
CS-702: Advance Database Systems

Time allowed: 3 Hours Max. Marks: 50

NOTE: Attempt five questions in all, including Question No. I which is compulsory and selecting two questions from each Unit.

x-x-x

I. Attempt the following:

a) Differentiate between schema and instance.

b) Define Data fragmentation with example.

c) Difference between OLTP and OLAP.

d) Explain the role of checkpoints in recovery techniques with example.

e) What is validation-based concurrency control? Explain with example. (5x2)

UNIT - I

II. a) A multinational logistic company stores the shipping details in a large database:
Shipment (SID, Date, OriginPort, DestinationPort, CustomerID, CustomerName, City, VesselID, VesselType, CargoID, CargoType, Weight)

i) Identify all the functional dependencies and multivalued dependencies.

ii) Describe the ***Shipment*** into BCNF and then 4NF and also explain the lossless join and dependency preservation.

b) A security analytics platform processes the multimedia data such as videos, audio logs, detected faces, events, objects and embeddings.

i) Propose an object-relational schema using object identifiers, user-defined types, inheritance and reference type.

ii) Explain various SQL3 features like table inheritance, ADTs, methods, polymorphic queries, collection types etc. and also explain how these features support scalable multimedia queries. (2x5)

P.T.O.

III. a) What is heuristic rule in query optimization? Explain transformation rules with example and relational tree algebra diagram.

b) Explain Two-phase locking with a detailed schedule describing the lock acquisition, lock release, and potential deadlock. Also compare the timestamp ordering and validation-based concurrency control with respect to conflicts and rollback behaviour. (2x5)

IV. a) A national bank wants to redesign the structure of the **Customer** table to include new KYC attributes, biometric data, and risk-scoring metrics. However, ATM services must continue to work without any application changes, Mobile banking should show additional fields in the API, and Core banking system uses compiled legacy code that should not break.

Explain how logical schema evolution and view-based abstraction can be used to support backward compatibility while extending the database. Provide a concrete strategy using views, mapping tables, and default values to achieve insulation from changes.

b) Explain the object-oriented data model with suitable examples of objects, classes, and inheritance. Also describe the major enhancements introduced in SQL3 which support object-relational features. (2x5)

UNIT - II

V. a) A financial institution has replicated data centres in Singapore, London and New York. It stores the following relations:

Accounts *(Acc_ID, CustomerID, Balance, Branch),*

Transactions *(Trans_ID, Acc_ID, Type, Amount, Timestamp),*

RiskProfile *(CID, Score, Country)*

Explain in detail the following questions:

i) Propose a hybrid distributed schema with horizontal, vertical, and derived fragmentation considering regulatory constraints and latency.

ii) Explain how the system ensures global serialization using distributed 2PL. Provide the timeline showing how anomalies are prevented.

b) Draw and explain the 3-tier Client-Server architecture of DDBMS with a suitable example. (2x5)

VI. a) Describe the advantages and disadvantages of log-based recovery techniques. How shadow paging is able to overcome various pitfalls of log-based techniques. Provide various disadvantages of shadow paging.

b) What are long-duration transactions? What are the adverse effects of the following concurrency protocols on long-duration transactions?

i) 2-phase locking

ii) Graph-based protocol (2x5)

VII. a) A telecom company wants to create a data warehouse for churn prediction and have the following data sources – Call Detail Records (CDR), Customer Profiles, Billing and Recharge Logs, and Network Quality Metrics. Explain in details the following questions:

i) Design a snowflake schema with appropriate fact and dimension tables. Also describe the complete ETL workflow including extraction from heterogenous sources, cleaning of missing logs and wrong tower IDs, transformation, loading and incremental refresh.

ii) Explain how OLAP operations like drill-down, slice, dice etc. helps in churn pattern discovery.

b) Explain how commercial DBMS like Oracle, IBM DB2, Microsoft SQL Server, MySQL and PostgreSQL support high availability, replication, in-memory processing, indexing mechanisms, and security models. Provide a real-world scenario where these features are essential. (2x5)

x-x-x