Sl. No.

B2.3-R5 : ADVANCED DATABASED TECHNOLOGIES

NOTE :
1. **Answer question 1 and any FOUR questions from 2 to 7.**
2. **Parts of the same question should be answered together and in the same sequence.**

Time : 3 Hours   **Total Marks : 100**

1. (a) Explain the different problems of data access control in a distributed environment.
    (b) How is online analytical processing (OLAP) different from online transactional processing (OLTP) ?
    (c) What are conflict serializability and view serializability. Explain with an example.
    (d) Explain log-based mechanism and shadow paging in relation to database recovery. In which situations, is shadow paging useful over log-based recovery ?
    (e) What is a weak entity set ? Give an example. How are weak entity sets represented in a relational schema ?
    (f) Differentiate between Temporal and spatial database.
    (g) What is Data Warehousing ? Give the functions of Data warehouse tools and utilities.

                                                                                                                                              (7x4)

2. (a) What is two phase locking protocol ? How can cascading rollbacks be avoided using two-phase locking.
    (b) Explain how two-phase locking works in case of a multi versioning system.
    (c) What are the pitfalls of lock-based protocols ?

                                                                                                                                              (6+6+6)

3. (a) "The transformation of a high-level query (typically, in relational calculus) into an equivalent lower-level query in distributed query processing must achieve both correctness and efficiency". Explain.
    (b) Illustrate the process of distributed query processing in distributed databases.
    (c) How does the query optimization timing affect the execution time ? Explain based on the different timing strategies used.

                                                                                                                                              (4+10+4)

4. (a) Illustrate the top-down design of a distributed DBMS.
    (b) Explain the different fragmentation strategies used in distributed design.
    (c) What are the reasons for data fragmentation in distributed DBMS.

                                                                                                                                              (8+6+4)

5. (a) What is a data cube ? Explain the different tables used to implement a data cube.
    (b) Illustrate the differences between star schema and snowflake schema. Explain why the fact constellation schema is ideal for data warehouses and star or snowflake schema is used for data marts.
    (c) What are the different OLAP operations for querying and analyzing the multidimensional data.

                                                                                                                                              (4+6+8)

6. (a) Draw ER diagram for Hospital Management System (Use DOCTOR, PATIENT HOSPITAL and MEDICAL_RECORD Entity). Identify Primary Key and Foreign Key.

(b) Normalize database :

Employee(emp_id,emp_name,phone,skill,salary,deptno,dept_name,jobno,job_title) upto 3NF

(c) Write SQL query for the table EMP (empno, deptno, ename, salary, Designation, joiningdate, DOB, city)

(i) Display names of employees whose experience is more than 10 years

(ii) Display the maximum age of employees

(iii) Display average salary of all employee

(iv) Display name of employee who earned highest salary

(4+4+10)

7. (a) Discuss the shadow paging recovery scheme.

(b) What are the requirements of ODBC ? Describe the components required for implementation of ODBC.

(c) Consider the following SQL query that finds all applicants who want to major in CSE, live in Seattle, and go to a school ranked better than 10 (i.e., rank < 10). Relation Cardinality Number of pages Primary key Applicants (id, name, city, sid) 2,000 100 id Schools (sid, sname, srank) 100 10 sid Major (id, major) 3,000 200 (id,major) SELECT A.name FROM Applicants A, Schools S, Major M WHERE A.sid = S.sid AND A.id = M.id AND A.city = 'Seattle' AND S.rank < 10 AND M.major = 'CSE'

And assume that :

• Each school has a unique rank number (srank value) between 1 and 100.

• There are 20 different cities.

• Applicants.sid is a foreign key that references Schools.sid.

• Major.id is a foreign key that references Applicants.id.

• There is an unclustered, secondary B+ tree index on Major.id and all index pages are in memory.

What is the cost of the query plan below ? Count only the number of page I/Os

(4+4+10)

- o O o -