VIT UNIVERSITY

SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING

CAT -1 Master of Computer Applications FALL Semester 2022-2023

Course Name: Database Technologies | Maximum Marks: 50

Course Code: ITA 5008 | Slot: B2

Common for : Prof. Bimal kumar Roy Prof Shashikiran V(Class id: VL2022230105105) Prof Vinay M

Answer any **FIVE** Questions. All questions carry EQUAL marks

[1] Construct Initial Canonical Query Tree and Equivalent optimal query tree for the SQL Statement given below.

SELECT S.sname, P.pname FROM Suppliers S, Parts P, Supply Y WHERE S.sid = Y.sid AND Y.pid = P.pid AND S.city = 'Madison' AND P.price ≤ 1,000

(or)

How does a query tree represent a relational algebra expression? What is meant by an execution of a query tree? Discuss the rules for transformation of query trees, and identify when each rule should be applied during optimization. Give Examples.

[2] Suppose that a disk unit has block size B = 1024 bytes. An EMPLOYEE file has the following fields: Ssn, 9 bytes; Last_name, 20 bytes; First_name, 20 bytes; Middle_init, 1 byte; Birth_date, 10 bytes; Address, 35 bytes; Phone, 12 bytes; Supervisor_ssn, 9 bytes; Department, 4 bytes; Job_code, 4 bytes; Salary, 4 bytes; Gender, 1 byte. The EMPLOYEE file has r = 140,000 records. Fixed-length format, and un-spanned blocking. Write appropriate formulas and calculate the following values for the above EMPLOYEE file:

a. Calculate the record size R (including the deletion marker), the blocking factor , and the number of disk blocks b. [2]

b. Calculate the wasted space in each disk block because of the un-spanned organization. [2]

c. Calculate the average number of block accesses needed to search for an arbitrary record in the Data file, using linear search and binary search. [2]

d. Create a primary index file on ordering key field (say First_name). Assume the block pointer size is 9bytes. Calculate Blocking factor and number of disk blocks needed for the primary index file. Calculate number of block accesses needed for arbitrary record search using Linear and Binary search [4]

(or)

(a) Describe the two ways of implementing clustering index. Distinguish between dense index and sparse index. [2]

(b) Which of the three basic file organizations would you choose for a file where the most frequent operations are as follows?

(b.1) Search for records based on a range of field values. (b.2) Perform inserts and scans where the order of records does not matter.

(b.3) Search for a record based on a particular field value [5]

[3] A PARTS file with Part# as the key field includes records with the following Part# values: 23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 33, 38. Suppose that the search field values are inserted in the given order in a B+-tree of order p = 4 and p_{leaf} = 3; show how the tree will expand and what the final tree will look like.

(or)

(a) Distinguish between homogeneous and heterogeneous DDBMS with scheme architecture diagrams [6]

(b) Highlight the access path to local internal schema from local applications and global applications defined to run on Homogenous DDBMS [2]

(c) How does global system catalog (GSC) keep track of fragments? Does Local users exist in homogeneous DDBMS, if so or not why? [2]

[4] Consider four different applications are defined run at four different sites by accessing the following relation by executing stated queries below.

Relation: Employee (SSN ENAME SAL LOC)

Q1: SELECT SAL FROM Employee WHERE SSN=Value ; Q2: SELECT ENAME, SAL FROM Employee;

Q3: SELECT ENAME FROM Employee WHERE LOC=Value; Q4: SELECT SUM(SAL) FROM Employee WHERE Loc=Value

Assuming ref(q1), ref(q2), ref(q3), ref(q4) from all the sites as 1. The application access frequency is stated below.

Acc1(q1) = 15 Acc2(q1) = 20 Acc3(q1) = 10 ; Acc1(q2) = 5 Acc2(q2) = 15 Acc3(q2) = 15

Acc1(q3) = 25 Acc2(q3) = 15 Acc3(q3) = 15 ; Acc1(q4) = 3 Acc2(q4) = 15 Acc3(q4) = 15

(a) Construct Attribute usage matrix and Attribute affinity matrix. [5]

(b) Construct clustered affinity matrix and vertically fragment the relation. [5]

(or)

[a] Consider the following relations.

BOOKS(Book#, Primary_author, Topic, Total_stock, Sprice) ; BOOKSTORE(Store#, City, State, Zip, Inventory_value)

STOCK(Store#, Book#, Qty) ;

1. Give an example of two simple predicates (or conditions) that would be meaningful for the BOOKSTORE relation for horizontal partitioning. [2]

2.How would a derived horizontal partitioning of STOCK be defined based on the partitioning of BOOKSTORE? [2]

3. Show predicates by which BOOKS may be horizontally partitioned by topic. [2]

4. Show how the STOCK may be further partitioned from the partitions in (2) by adding the predicates. [2]

[b] How is a vertical partitioning of a relation specified? How can a relation be put back together from a complete vertical partitioning? [2]

[5] Consider the three transactions T1, T2, and T3, and the Histories S1 and S2 given below. Draw the serializability (precedence) graphs for S1 and S2, and state whether each History (or schedule) is serializable or not. If a schedule is serializable, write down the equivalent serial schedule(s).

T1: r1 (X); r1 (Z); w1 (X); T2: r2 (Z); r2 (Y); w2 (Z); w2 (Y); T3: r3 (X); r3 (Y); w3 (Y);

S1: r1 (X); r2 (Z); r1 (Z); r3 (X); r3 (Y); w1 (X); w3 (Y); r2 (Y); w2 (Z); w2 (Y);

S2: r1 (X); r2 (Z); r3 (X); r1 (Z); r2 (Y); r3 (Y); w1 (X); w2 (Z); w3 (Y); w2 (Y);

(or)

How do Locking protocols prevent deadlocks in a distributed environment? Discuss the significance of deadlock detection methods in centralized, hierarchical and distributed scenarios.

[6] Use Odd-even Transposition sort on one dimensional mesh processor array model to sort the sequences

2 13 9 8 53 96 33 14

(or)

Use Bitonic Merge for sorting the sequence below.

1 3 4 5 7 8 10 11 12 16 15 14 13 9 6 2