Nirma University

Institute of Technology

Supplementary Examination (SPE), March - 2023
M. Tech. in Computer Science and Engineering, Semester-I
6CS204 Advanced Database Systems

Roll/
Exam No

Supervisor's initial
with date

Time: 3 Hours

Max Marks: 100

Instructions :

1. **All questions are compulsory. (No optional questions)**
2. **Use section-wise separate answer books**
3. Figure to right indicate full marks
4. Draw neat sketches wherever necessary.

Section I

Q. 1   **Do as directed**   **18**

CLO2   A   Given the following SQL query: 12
BL5

Student (sid, name, age, address)
Book (bid, title, author)
Checkout (sid, bid, date)
SELECT S.name
FROM Student S, Book B, Checkout C
WHERE S.sid = C.sid
AND B.bid = C.bid   AND B.author = 'Olden Fames'
AND S.age > 12   AND S.age < 20

And assuming:

There are 10000 Student records stored on 1000 pages. There are 50000 Book records stored on 5000 pages. There are 300000 Checkout records stored on 15000 pages. There are 500 different authors. Student ages range from 7 to 24.

a. Show a physical query plan for this query, assuming there are no indexes and data is not sorted on any attribute.
b. Compute the cost of this query plan and the cardinality of the result.
c. Suggest two indexes and an alternate query plan for this query.
d. Compute the cost of your new plan.

CLO1   B   Demonstrate how pointer swizzling can improve data retrieval performance.   6
BL3

Q. 2   **Do as directed**   **16**

CLO3   A   Demonstrate the difference between legal and well-formed schedules with   6
BL4   suitable example.

CLO2   B   Relations $R(X, Y)$ and $S(Y, Z)$ covering 1000 and 500 blocks, respectively.   6
BL3   Assume ten tuples fit on one block, so $T(R) = 10,000$ and $T(S) = 5000$. Also, assume $V(S, Y) = 100$. Find out I/O cost for joining $(R \bowtie S)$ both the relations for the following cases,

1. R is clustered, and there is a clustering index on Y for S
2. R is clustered and there is a non-clustering index on Y for S
3. R and S both are clustered and there is no index on Y.
Neglect index retrieval cost wherever it is applicable.

CLO2   C   Consider these relations with the following properties: 4
BL3   R(A, B, C), 30,000 tuples, 25 tuples fit on 1 block
           S(C, D, E), 60,000 tuples, 30 tuples fit on 1 block
   Estimate the number of disk block accesses required for joining of R and S
           using a nested-loop join and block-nested loop join if R is used as the outer
           relation.

Q.3   Do as directed   **16**

CLO1   A   On the given hash index based on extendible hashing insert following values   8
BL3   and show the index status along with other parameters with every insertion.
           Keys to be inserted: 1011, 1111, 1000, 1010.

$$i = 1$$

1
0001

1
1001
1100
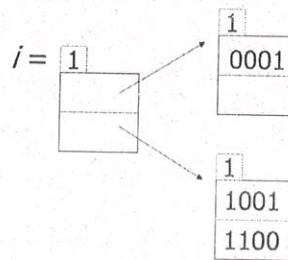
Output of h(k) is of 4 bits and maximum 2 keys/bucket.

CLO3   B   Demonstrate with suitable example the difference between Redo logging and   8
BL4   Undo logging applied for database recovery.

Section II

Q. 4   Do as directed   **18**

CLO3   A   Show the partial R-Tree nodes with the values in it as per the following   6
BL3   regions shown in the diagram below. Assume appropriate values for the
           boundaries of the various sub regions.
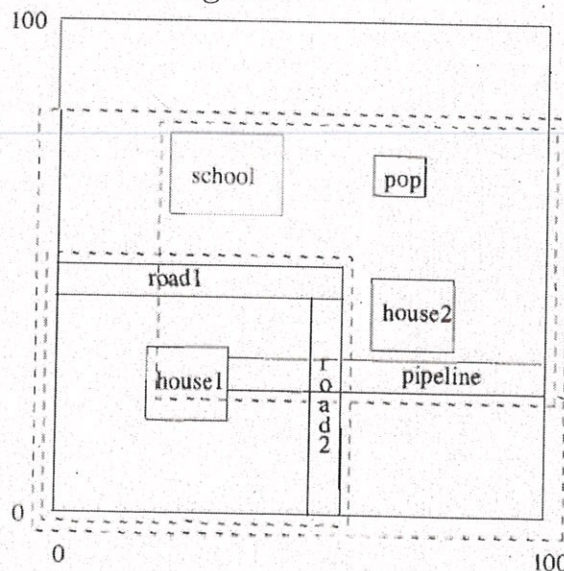
CLO2   B   Demonstrate working of any two page replacement policy with suitable   6

BL2 example.

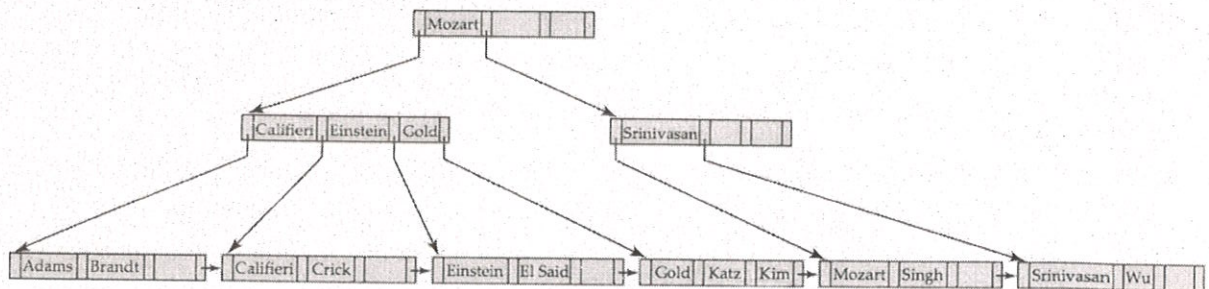CLO3 C Check for the serializability of following schedules: 6
BL4
    1. S: $r_2(A)$; $r_1(B)$; $w_2(A)$; $r_3(A)$; $w_1(B)$; $w_3(A)$; $r_2(B)$; $w_2(B)$;
    2. S_1: $r_2(A)$; $r_1(B)$; $w_2(A)$; $r_2(B)$; $r_3(A)$; $w_1(B)$; $w_3(A)$; $w_2(B)$;
    Justify your answer for each case.

Q. 5   Do as directed   16

CLO1 A For the given B+Tree, delete following values from it. Show the tree status 8
BL3 after every removal.

Values to be deleted in sequence: Srinivasan, Sing, Wu, Gold.

CLO1 B Construct required number of bitmap vectors for the following points. First 8
BL3 value in these points represent age and second value represent salary in thousand for employee database. Also demonstrate how the query to find the employees with age between 45-55 and salary the range of 100-200 will be answered.

1: (25, 60)   2: (45, 60)   3: (50, 75)   4: (50, 100)
5: (50, 120)   6: (70, 110)   7: (85, 140)   8: (30, 260)
9: (25, 400)   10: (45, 350)   11: (50, 275)   12: (60, 260)

Q. 6   Do as directed   16

CLO3 A Show how MongoDB handles replication requirements for the large-scale 4
BL4 applications.

CLO3 B Why there may be a need to physically optimize the space occupied by the 4
BL4 database table?

CLO2 C Given the following data file: EMPLOYEE (NAME, SSN, ADDRESS, JOB, 4
BL3 SAL, ...), record size R=150 bytes, block size B=512 bytes, total 30000 records, for an index on the SSN field, assume the field size VSSN=9 bytes, assume the record pointer size PR=7 bytes. Find out the following,

    1. Blocking factor for data blocks and total number of data blocks
    2. Size of an individual index entry
    3. Blocking factor for index blocks
    4. Total number of index blocks

CLO1 D Compare the performance of BTree and B+Tree in the context of various 4
BL4 operations on it.