

# Implement of DBMS 2023/2024 Question

## Question 1. Mark = 4

Consider a disk with the following characteristics:

There are 50 surfaces with 20,000 tracks each

Tracks hold on average 1000 sectors of 512 bytes each

The disk rotates at 3,000 rpm

The time it takes the head to move n tracks is  $1 + 0.001n$  ms

30% of each track is used for gaps

A block has 32 sectors

a) What is the capacity of the disk?

b) What is the transfer time of a block?

## Question 2. Mark 4

Suppose we use a RAID level 4 scheme with four data disks and one redundant disk.

Assume that the blocks are a single byte. Give the block of the redundant disk if the corresponding blocks of the data disks are:

01010110, 11000000, 00111011, and 11111011

Recover the block while the redundant disk holds 11001011.

What changes to the

corresponding blocks on the other disks must be made?

## Question 3. Mark 6

### Prof. given tree and we need to insert value and delete value:

Insert the keys 20, 40, 10, 30, 15, 35, 7, 26, 18, 22, 36, and 42 in this order into an initially empty B+-tree of order 2.

Delete from the B+-tree of order 2 you produced in task 1 the keys 7 in this order.

## Question 4. Mark 6

Let R and S be relations, p a predicate with only R attributes, q a predicate with only S attributes and m a predicate with attributes from R and S. Show that the following rule holds. Use in your proof only the given rules in the box. Indicate in each step which rule you are using.  $\text{mpq}(R \ A \ S) = m([p(R)] \ A \ [q(S)])$

## Question 5. Mark 6

Below are some statistics for four relations W, X, Y and Z.

W(a, b) X(b, c) Y(c, d) Z(d, e)

$T(W) = 100$   $T(X) = 200$   $T(Y) = 300$   $T(Z) = 400$

$V(W, a) = 20$   $V(X, b) = 50$   $V(Y, c) = 50$   $V(Z, d) = 40$

$V(W, b) = 60$   $V(X, c) = 100$   $V(Y, d) = 50$   $V(Z, e) = 100$

- a) Consider a query optimizer that uses statistical data. In particular, the following information is known about an attribute A of relation R. Attribute A is of type integer. Consider the query  $A=7(R)$ . How many tuples are expected in the answer, assuming values are uniformly distributed over possible  $V(R, A)$  values?
- b) Consider the query  $A=17(R)$ . How many tuples are expected in the answer, assuming values are uniformly distributed over possible domain values?

**Question 6. Mark 8**

Suppose that blocks can hold either ten records or 99 keys and 100 pointers. Also assume that each B+-tree node is 60% full, i.e., it will have 59 keys and 60 pointers in case of interior nodes and 60 keys and 60 pointers to records in case of leaf nodes. We have a data file that is a sequential file, and the B+- tree is a sparse index, but each primary block of the data file has one overflow block. The primary blocks are full, and the overflow blocks are half full.

However, records are in no particular order within primary block and its overflow block.

- a) Calculate the total number of blocks needed for a 3,240,000-record file and the index.
- b) Calculate the average number of disk I/O's needed to retrieve a record given its search key. You may assume that nothing is in memory initially, and that the search key is the primary key for the records.

**Question 7. Mark 7**

Consider a clustered relation  $R(A, B, C, D)$  that has a clustering index on  $A$  and a non-clustering

index on each of the other attributes. The relevant parameters are:  $B(R) = 1000$ ,  $T(R) = 5000$ ,  $V(R, A)$

$= 20$ ,  $V(R, B) = 1000$ ,  $V(R, C) = 5000$ , and  $V(R, D) = 500$ . Give the best query plan for the following

selection and the corresponding number of disk I/O's. You can ignore the cost for accessing the index