

# DBMS

Q.1 Consider a disk with the following characteristics:

There are 10 surfaces each with 100 tracks. Each track is divided into 18 sectors and a sector holds 512 bytes. Blocks consist of 2 sectors. 20% of the circumference of each track is occupied by gaps between sectors. The disk rotates at 6000 rotations per minute.

- a. What is the capacity of the disk?
- b. What is the transfer time for one block?

Q.2 Consider the following situation: Blocks are 1000 bytes long. There is no need for a block header. Records are 100 bytes long, of which 12 bytes are the key filed. Pointers take 8 bytes. A sequential file (sorted by the key filed) consists of 10000 records. Each block of the file contains as many records as possible.

- a. What is the minimum number of block required for a dense index on this file?
- b. What is the minimum number of blocks required for a sparse index on this file?

Q.3 Consider B+-tree of order 2. Give an example of a B+ tree with three levels whose set of keys could alternatively be represented in a B+ tree with two levels. Your example should consist of two trees, one with three levels and the equivalent one with two levels. Your trees should show all keys and pointers.

Q.4 Suppose we store relation R (a, b, c) in a partitioned hash table with 1024 buckets. That is, the hash function produces 10 bits. Queries about R all specify exactly one of the attributes a, b or c and each of the attributes is equally likely to be specified. The hash function produces 5 bits based only on the value of a, 3 bits based only on the value on b and 2 bits based only on c. What is the average number of buckets that must be searched to find matching tuples?

Q.5 Let R(A,B), S(B,C) and T(A,B,C) be three relations. Consider the following relational algebra expressions:

$$E1: \pi_{A,C}[\{\sigma_{B<10}(R) \bowtie S\} - T]$$

$$E2: \pi_{A,C}[\{R \bowtie \sigma_{B<10}(S) - \sigma_{B<10}(T)\}]$$

$$E3: \pi_{A,C}[\{\pi_A(R) \bowtie = \sigma_{B<10}(S)\} - T]$$

Are some of the expressions equivalents to each other?

Q.6 Let  $R(A,B)$ ,  $S(B,C)$  be relations with the following statistics:

$$T(R)=50$$

$$T(S)=100$$

$$V(R,A)=5$$

$$V(R,B)=5$$

$$V(S,B)=10$$

$$V(S,C)=20$$

Estimate the number of tuples of the following expression:  $\sigma_{A=1}(R) \bowtie S$

Q.7 We have three relations with the following statistics:

- $B(\text{Order})= 3000$  blocks
- $B(\text{Cust})= 1000$  blocks
- $B(\text{Book})=100$  blocks

Assume that the relations are stored contiguously, i.e, they are clustered relations. You should further assume that each operation uses memory efficiently. You can ignore final output I/O cost.

- a. We want to perform a selection of “Price<10” over block. We have 10 blocks of main memory. What is the required number of I/Os?
- b. We want to perform a one pass join of Order and Cust. How many main-memory-blocks do you need?
- c. We want to perform a one pass join of order and Cust and have sufficient main memory. What is the required number of I/Os?
- d. We want to perform a hash join (without the “hybrid optimization”) of order and Book. How main memory-blocks do you need?
- e. We want to perform a hash join (without the “hybrid optimization”) of order and Book and have sufficient main-memory. What is the required number of I/Os?