Branch : CSE & IT Batch : English

Database Management System

File Organization & Indexing

DPP 01

[MCQ]

1. Assume a relational database system that holds relation: C(colleges) with the following characteristics

- Records are stored as fixed length, fixed format records, length is 256 bytes.
- There are 16384 records.
- Records contains key attribute CollegeNumber (C.N), length 22 bytes and other fields.
- Unspanned organization is used to store the information or record.

Let's suppose we want to build a sparse primary index on C.N then how many numbers of 4096-byte blocks are needed to store the primary index when block pointer size is 10 bytes _____?

(a) 7 (b) 8

(c) 9 (d) 10

[NAT]

2. Assume a relational database system that holds relation: Product (P) with the following characteristics

- Records are stored as fixed length, fixed format records, with the length of 256 bytes.
- There are 262144 records.
- Records contain attribute P.I (The identifier of the product involved), with the length 24 bytes, and an attribute P.C (the cost of product), with the length 32 bytes and other fields.
- Unspanned organization is used to store the record.

Assume that we want to build a dense secondary index on P.C, then how many numbers of 4096-byte blocks needed to store the dense secondary index. When record pointer size is 32 bytes? _____.

[MCQ]

3. Consider a SQL statement SELECT P_1, P_2, P_3 from Q WHERE P_2 = 'Pavan' is frequently executed, which column(s) should be considered for indexing based only on the statement itself?

(a) P_1 only

(b) P_2 only

(c) P_3 only

(d) P_1, P_2 and P_3

[MCQ]

4. Consider the following specification of system-

Disk block size = 2048 bytes

Block pointer size = 16 bytes

Record pointer size = 20 bytes long

file contains 30,000 records.

Each record of the file has the following fields:

Fields	Size (in Bytes)
EmpName	5
EmpNum	10
DeptNum	9
Addr	20
PhNum	9
DOB	1
Sex	1
Job	3
Sal	5

An extra/additional byte is used per record to represent end of the record.

What is the block factor of the database file assuming unspanned file organization?

(a) 16 (b) 32

(c) 48 (d) 64

[MSQ]

5. Which one of the following statements is/are True regarding indexing?

(a) A database file can contain multiple clustered indexes.

(b) A database file can consist of only one clustered index with multiple secondary indexes.

(c) A database file can consist of multiple primary indexes.

(d) A database file can consist of both primary and clustered index.

[NAT]

6. Consider a database of fixed-length records stored as an ordered file. The database has 25,000 records with each records being 100 bytes, of which the non-key attribute on which clustering index is formed occupies 10 bytes. The data file is completely block aligned.

Suppose, block size, of the file system is 512 bytes and a pointer to the block occupy 5 bytes. You may assume that a binary search on an index file of b block may take $\lceil \log_2 b \rceil$ accesses in worst case.

Given that a cluster consumes 2 blocks, the number of block accesses required to identify the desired data in the worst case is _____.

[MCQ]

7. Consider the following statements-

S_1: If the records of a relation X are physically ordered over a non-key field P and an index is build over the key-field of relation X, then the index is necessarily a secondary index over key attribute.

S_2: More than one secondary indexes are possible.

Which of the given statement(s) is/are CORRECT?

(a) S_1 only

(b) S_2 only

(c) Both S_1 and S_2

(d) Neither S_1 nor S_2

Answer Key

1. **(b)**
2. **(4096)**
3. **(b)**
4. **(b)**
5. **(b)**
6. **(9)**
7. **(c)**

Hints & Solutions

1. (b)

In the primary index, number of entries in the index block equals to number of blocks of relation.

Number of database records in a single block B = 4096/256 = 16

Number of blocks of relation C = 16384/16 = 1024

Size of indexes = size of key field

+ size of block pointer

= 22 + 10 = 32 bytes

Number of indexes records present in single block = 4096/32 = 128

∴ Total number of blocks required to store primary index = 1024/128 = 8.

2. (4096)

In dense secondary index, number of entries in the index blocks equals to number of records of relation.

- Number of records in the relation P = 262144

  Size of index = size of key field

  + size of record pointer

  = 32 + 32 = 64 bytes

- Number of index entry in single block

  = 4096/64 = 64

  So, the total number of blocks required to store primary index = 262144/64 = 4096.

3. (b)

The column on which condition gets applied should be considered for indexing.

∴ P_2 is the answer.

4. (b)

Blocking factor (i.e number of records per block)

$= \dfrac{\text{Block size}}{\text{record size}}$

Record size of file = Sum of all field + additional bytes = 63 + 1 = 64

∴ Number of records per block = $\dfrac{2048}{64} = 32$

5. (b)

(a) **False:** A database file can contain one clustered index because the database is sorted on one field only.

(b) **True:** A database file can consist of one clustered index and multiple secondary index.

(c) **False:** The index on a unique field on which database is sorted is primary index and there can be only one primary index.

(d) **False:** A database file can consist of either a primary or clustered index but not both.

6. (9)

Block factor of database file = $\left\lfloor \dfrac{512}{100} \right\rfloor$

= 5 records/block

Number of blocks required to store 25,000 records

$= \left\lceil \dfrac{25000}{5} \right\rceil$

= 5000 blocks

Each cluster consumer 2 blocks

Number of entries in index file = $\dfrac{5000}{2} = 2500$

Block factor of index file = $\left\lfloor \dfrac{512}{15} \right\rfloor = 34$ entries/block

Number of blocks in index file = $\left\lceil \dfrac{2500}{34} \right\rceil = 74$

The number of block accesses in worst case

$= \lceil \log_2 74 \rceil + 1 + 1$

↓ for index file search; ↓ Accessing 1st block of cluster; ↓ Accessing 2nd block of cluster

= 7 + 1 + 1 = 9 blocks access required

7. (c)

S_1: Records are ordered over non-key field. It is unordered over key field.

Hence, secondary index if formed over unordered key-field. Hence, its CORRECT.

S_2: CORRECT. More than one secondary index is possible.