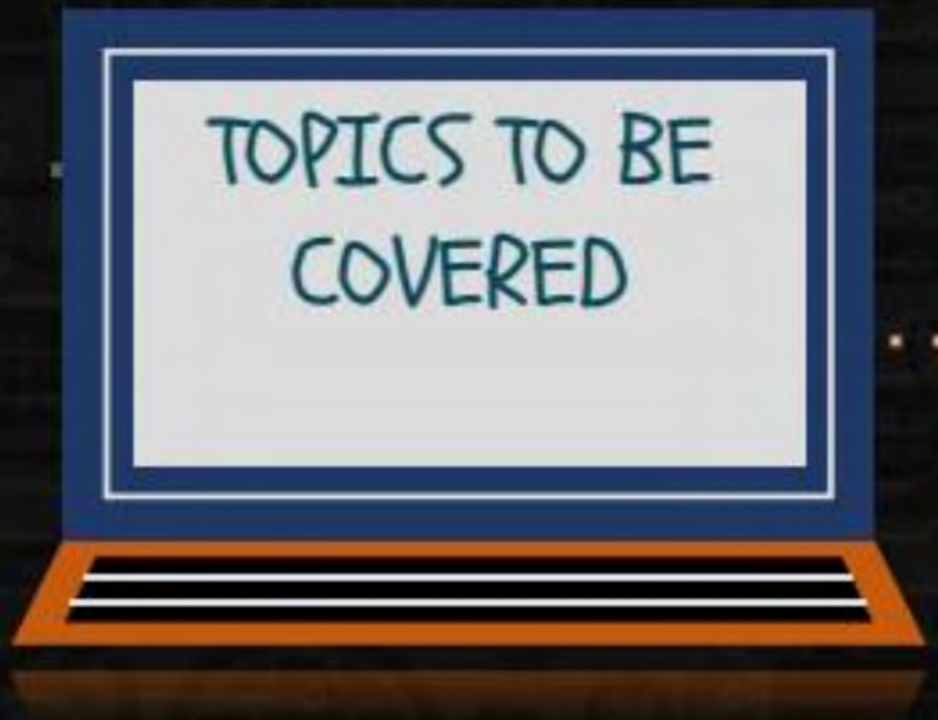
CS & IT ENGINEERING

Database Management System

File Org & Indexing

DPP – 01 Discussion Notes

By- Vijay Agarwal sir



01 Question

02 Discussion

Q.1 [MCQ]

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

Ans (B)

B. 8

C. 9

D. 10

Record size = 256 Byte

Records = 16,384 Records

Fixed Length Unspanned & SPARSE Primary Index

Key = 22 Byte & Bp = 10 Byte

Block Size = 4096 Byte

Soln

$$\text{Block factor of DB File } [BF_{DB}] \text{ [unspanned]} = \left\lfloor \frac{\text{Block Size}}{\text{Record Size}} \right\rfloor \Rightarrow \frac{4096 \text{ Byte}}{256 \text{ Byte}} \Rightarrow \frac{2^{12}}{2^8} = 2^4$$

$\Rightarrow$ 16 Record Per Blocks

$$\text{Total \# of DB Blocks} = \left\lceil \frac{16,384}{16} \right\rceil \Rightarrow \frac{2^{14}}{2^4} = 2^{10} \Rightarrow$$ 1024 DB Block.

One Index Record size = Size of Key + Size of Block Pointer $\Rightarrow 22 + 10 = 32 \text{ Byte} \Rightarrow 2^5$ Byte

Block Size = 4096 Byte.

$$\text{Block factor of Index file } [BF_I] = \frac{4096 \text{ B}}{32 \text{ B}} \Rightarrow \frac{2^{12}}{2^5} = 2^7 \Rightarrow$$ 128 Index Entries Per Block.

Primary Index ⇒ SPARSE

BF_I = 128 Index Entries Per Block

Sparse Index ⇒ Total # Index entries = # of DB Blocks = 1024

Total # Index Blocks = $\frac{1024}{128} \Rightarrow \frac{2^{10}}{2^7} = 2^3$

⇒ 8 Index Block. Ans

Q.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?

4096.

Ans (4096)

[NAT]

Record Size = 256 Byte

Total # Records = 262144 Records $[2^{18}]$, P.I = 24 Byte

Product Cost = 32 Byte
[P.I] ← key

Block Size = 4096 B

Bp = 32 Byte

Secondary ⇒ Dense

Soln

One Index Record Size = Size of key + Bp = 32 + 32 = 64 Byte

Block Size = 4096 Byte

Block factor of Index file = $\frac{4096}{64} \Rightarrow \frac{2^{12}}{2^6} = 2^6 \Rightarrow$ 64 Index entries per Block
$[B_{F_I}]$

Secondary Index

Dense: No. of Index Entries = Number of DB Records = 262144 Records $[2^{18}]$

Total # Index Blocks $= \left\lceil \frac{262144}{64} \right\rceil$ $\Rightarrow \frac{2^{18}}{2^6} = 2^{12}$

$\Rightarrow$ 4096 Index Blocks

Ans

Q.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?

[MCQ]

A. P_1 only

B. P_2 only

C. P_3 only

D. P_1, P_2 and P_3

SELECT P_1, P_2, P_3
FROM Q
WHERE P_2 = 'PAVAN'

Ans (B)

Q.4

[MCQ]

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:
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?

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

63 + 1 Byte

A. 16

B. 32

Ans (B)

C. 48

D. 64

Record Size = 64 Byte

Block Size = 2048 Byte

Block factor of DB file (unspanned) $B_{f_{DB}} = \left\lfloor \frac{\text{Block Size}}{\text{Record Size}} \right\rfloor$

$= \left\lfloor \frac{2048 \text{ Byte}}{64 \text{ Byte}} \right\rfloor = \frac{2^{11}}{2^6} = 2^5$

$\Rightarrow$ 32 Record per Block

Q.5 Which one of the following statements is/are True regarding indexing? [MSQ]

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.

× Primary & clustered Both Not possible

One Primary Index possible

P.I + S.I ✓

C.I + S.I ✓

Ans (B)

Q.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

9 Ans

Ans (9)

[NAT]

Records = 25,000 Record Size = 100 Byte Key = 10B Bp = 5 Byte

Block Size = 512 Byte Each cluster Consume 2 Blocks

Block factor of DB file = $\left\lfloor \frac{\text{Block Size}}{\text{Record Size}} \right\rfloor = \left\lfloor \frac{512B}{100B} \right\rfloor = 5$ Record Per Block

Total # DB Block = $\left\lceil \frac{25000}{5} \right\rceil = 5000$ DATA Block

Each cluster Consume = 2 Blocks

Total Number of Index Entries = $\frac{5000}{2} = 2500$ Entries

One Index Record Size = 10 + 5 = 15 Byte

Blouc factor of Index file $= \left\lfloor \frac{512B}{15B} \right\rfloor = 34$ Index Entries Per Block

Total # Index Blocks $=$

$$\left\lceil \frac{2500}{34} \right\rceil = 74 \text{ Index Blocks}$$

$$\# \text{ Blocks Access Required} = \lceil \log_2 74 \rceil + 1 + 1$$

(1^{st} $\rightarrow$ 1, 2^{nd} $\rightarrow$ 1)

$$\Rightarrow 7 + 1 + 1$$

$$= (9) \text{ Ans}$$

Q.7 [MCQ]

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?

S1 ORDERED ⇓ Non key Attribute

Ordering Done on Non key →

Index ⇓ key + Unordered

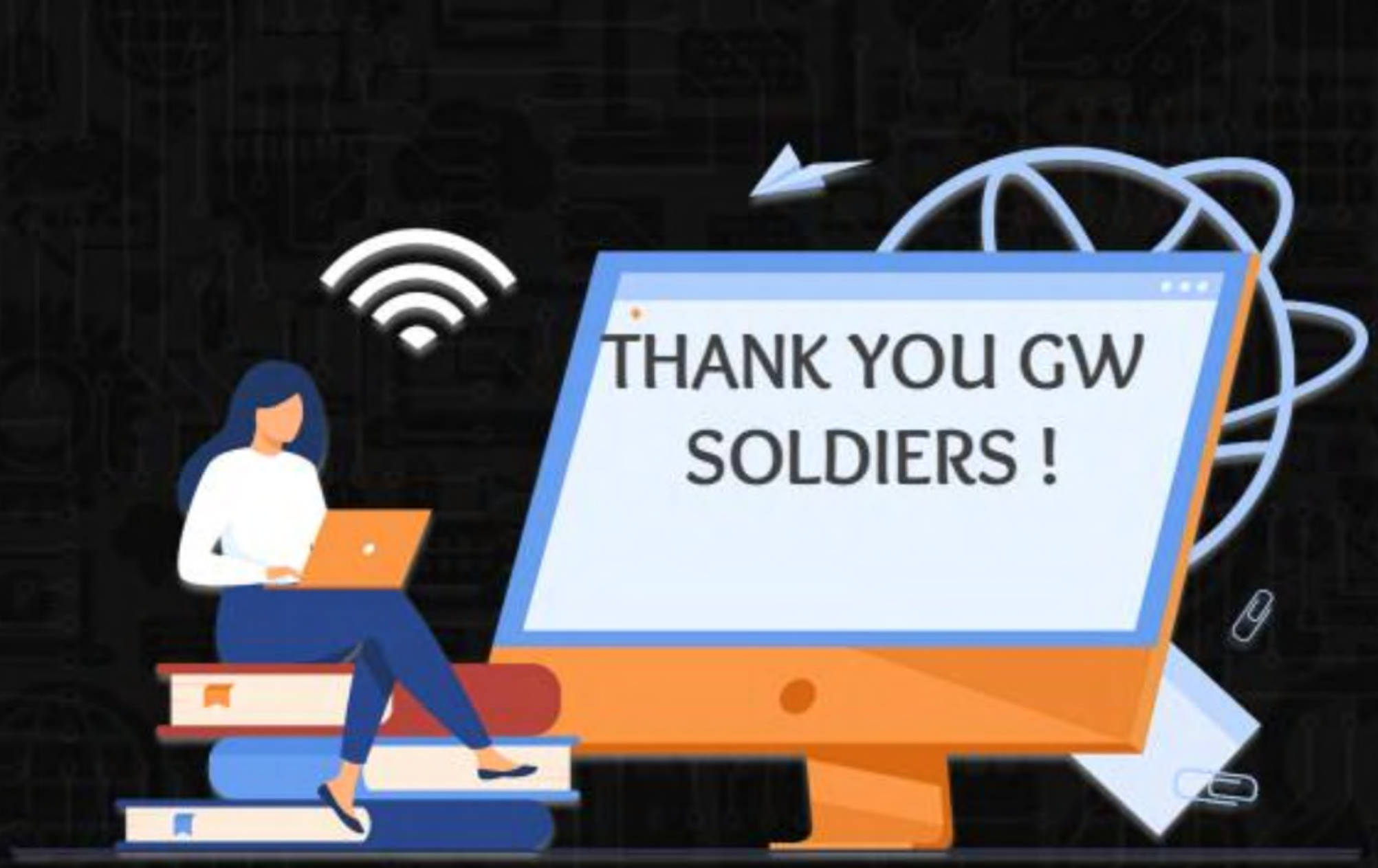
A. S_1 only

B. S_2 only

C. Both S_1 and S_2 only

D. Neither S_1 nor S_2

Ans (C)


THANK YOU GW
SOLDIERS !