Branch : CSE & IT **Batch : Hinglish**

Database Management System

File Org & Indexing

DPP 02

[NAT]

1. Consider the following specification of system with disk block size 2048 bytes, block pointer size 14 bytes, record pointer size 18 bytes long and file size 60,000 records. Each record of file is 256 bytes long and record of the size is sorted on the key field. If the primary index (sparse) is built on the key field (ESN) which is 18 bytes long. What is the Index blocking factors (That is number of indexes per block)
Assuming unspanned file organization _________.

[NAT]

2. Common data for next two Questions:
Consider a disk blocking size B = 1024 bytes. A block pointer (BP) = 12 bytes long and a record pointer (RP) = 7 bytes long. A file has r = 60,000 patient records of fixed length. The size of record is 230 bytes. Suppose the file is not ordered by the key field PSN (18 bytes) and we want to construct a secondary index on key attributes (PSN).

The number of first level index entries are x and number of second level index entries are y then find the value of x + y?

[MCQ]

3. The number of first level index blocks are x and number of second level index blocks are y then x +y______?

(a) 1600 (b) 1500
(c) 45 (d) 1545

[NAT]

4. Consider a file of r = 40,000 records, each record is R = 100 bytes long and its key field is of size v = 20 bytes. The file is ordered on a key field, and the file organization is unspanned. The file is stored in a file system with block size B = 2000 bytes, and size of block pointer is 20 bytes. If the primary index is built on the key field of the file and multilevel index scheme is used to store the primary index, then the total number of blocks required by the multilevel index is______.

[MCQ]

5. Assume that we have an ordered file with r = 60,000 records stored on a disk with block size B = 2048 bytes. File record are of fixed size & are unspanned with record length R = 200 bytes. Now assume that the ordering key field of the file is V= 18 bytes long, a block pointer P = 12 bytes long, and we have construed a primary index for the file. Let p and q be the number of blocks required to access a record in case of without index and with primary index using binary search respectively, Then the values of p + q is ________.

(a) 18 (b) 19
(c) 20 (d) 21

[MCQ]

6. Consider an unordered file of 10^6 records with records size of 200 bytes stored on blocks of 8KB with a spanned records organization. We will assume that no system related information is stored within a block, then how many blocks would it be need to store this file?

(a) 24400 (b) 24405
(c) 24410 (d) 24415

[MCQ]

7. Consider the following statements:
S_1: For any given data file, it is possible to create two different sparse first level indexes on various keys.
S_2: For any given data file, it is possible to create two different denes first level indexes on various keys.
Select the correct statements.

(a) Only S_1 correct
(b) Only S_2 correct
(c) Both S_1 and S_2 is correct
(d) Neither is S_1 nor S_2 is correct.

[MCQ]

8. Which of the following is NOT a benefit of using Indexes in a database?

(a) Improved query performance
(b) reduced disk I/O
(c) Increased storage space
(d) Faster data retrieval

[MCQ]

9. Which of the following best describes an index in a database.

(a) A column that stores unique identifiers for each row in a table.

(b) A data structure that allows for fast searching and retrieval of data, based on certain criteria.

(c) A set of constraints that enforce rules for data integrity

(d) None of the above.

Answer Key

1. (64)
2. (61500)
3. (d)
4. (41)
5. (21)
6. (d)
7. (b)
8. (c)
9. (b)

Hints & Solutions

1. (64)

Number of indexes per blocks = $\left\lfloor \frac{2048}{14+18} \right\rfloor$ = 64

2. (61500)

Number of first level index entries r_1 = Number of files records r = 60,000.

So, x will be 60000.

Now, block factor of the first level index = $\left\lfloor \frac{1024}{18+7} \right\rfloor = 40$ index entries per block.

$\therefore$ Number of index block at first level (that is entries for second level = y) = $\left\lceil \frac{60000}{40} \right\rceil = 1500$ blocks

$\therefore$ x + y = 60000 + 1500 = 61500

3. (d)

Now, block factor of the first level index = $\left\lfloor \frac{1024}{18+7} \right\rfloor = 40$ index entries per block.

$\therefore$ Number of index block at first level (that is entries for second level = x) = $\left\lceil \frac{60000}{40} \right\rceil = 1500$ blocks

Now, block factor of the second level index = $\left\lfloor \frac{1024}{18+12} \right\rfloor = 34$ index entries per block.

$\therefore$ Number of index block at second level (that is entries for second level = y) = $\left\lceil \frac{1500}{34} \right\rceil = 45$ blocks

$\therefore$ x + y = 1500 + 45 = 1545

4. (41)

File blocking factor bfr = $\lfloor (B/R) \rfloor$

$= \lfloor 2000/100 \rfloor$

= 20 records per block

Number of blocks needed for file = $\lceil r/bfr \rceil$

= $\left\lceil \frac{40,000}{20} \right\rceil$ = 2000 database blocks

Index records size R_i = (V + P) = (20 + 20) = 40 bytes

Index blocking factors bfr_i = $\lfloor B/R_i \rfloor$ =

$\left\lfloor \frac{2000}{40} \right\rfloor$ = 50 index records per block

Number for first level index entries (r_i) will be equal to number of file blocks b = 2000 entries

Number of 1st Level index blocks $b_1 = \lceil r_1 / bfr_i \rceil$

= $\left\lceil \frac{2,000}{50} \right\rceil$ = 40

Number of 2nd level index entries r_2 = number of 1st level block b_1 = 40 entries.

Number of 2nd level index $b_2 = \lceil r_2 / bfr_i \rceil = \lceil 40/50 \rceil = 1$

Since the 2nd level has only 1 block, it is the top index level. Hence, the index has 2 levels.

The total Number of blocks required by the multilevel index (b_i) = $b_1 + b_2$ = 40 + 1 = 41 blocks.

5. (d)

The blocking factor for the file will be bfr = $\lfloor (B/R) \rfloor$

= $\lfloor 2048/200 \rfloor$ = 10 records per block.

The number of blocks needed for the file is b = $\lceil (r/bfr) \rceil = \lceil (60,000/10) \rceil$ = 6000 blocks.

A binary search on the data file would need approximately $\lceil \log_2 b \rceil$ = $\lceil \log_2 6000 \rceil$ = 13 block accesses.

The size of each index entry is R_i = (18 + 12) = 30 bytes.

So, the blocking factors for the index is bfr_i = $\lfloor (B/R_i) \rfloor$

= $\left\lfloor \left(\frac{2048}{30} \right) \right\rfloor$ = 68 entries per block

The total number of index entries r_i is equal to number of blocks in the data file, which is 6000 blocks.

Hence, the number of index blocks is b_i = $\lceil (r_i / bfr_i) \rceil$

= $\left\lceil \frac{6000}{68} \right\rceil$ = 89 blocks.

To perform a binary search on the index file it would it need $\lceil (\log_2 / b_i) \rceil = \lceil \log_2 (89) \rceil$ = 7 blocks access.

To search for a record using the index, we need one additional block access to data file which makes total of(q) 7 + 1 = 8 = blocks access, an improvement over binary search on data file which required 13 block access.

$\therefore$ p + q = 13 + 8 = 21

6. (d)

Blocks size = 8 KB
Records size = 200 bytes

The number of records in a block = $\frac{8192}{200}$ = 40.96 records (Spanned Organization).
As it is spanned hence it takes whole as 40.96
1 block contains 40.96 records.

$\therefore$ The number of file blocks $= \left\lceil \frac{10^6}{40.96} \right\rceil$ = 24415 blocks.

7. (b)

S_1: (false): It is not possible because the requirement of sparse indexing is that the database must be stored and as we know that database is sorted only on one column.

S_2:(True): Any number of dense indexes is possible to construct. In the dense indexing we have index entries for each file records.

8. (c)

The correct answer is 'c'.
'Increase storage space'. This is because adding indexes to a database can actually increase the amount of storage space required as the index data structure needs to be stored alongside the data. The other answer choices are all benefits of using indexes.

9. (b)

Indexing in a database is a data structure that allows for fast searching and retrieval of data based on certain criteria.