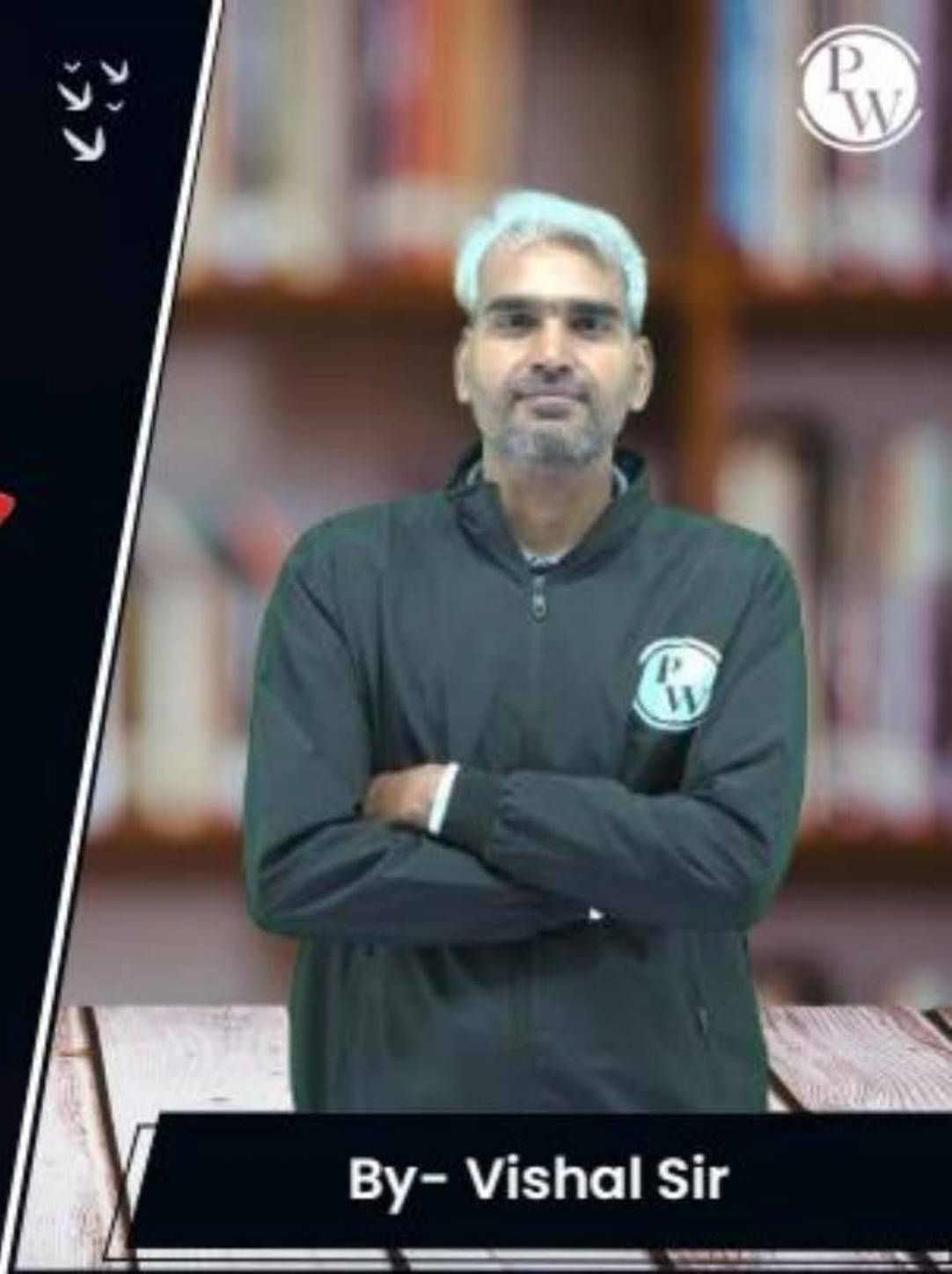
DS & AI

Database Management System

Super 1500+

Lecture No. 07



By- Vishal Sir

Recap of Previous Lecture

Topic Structure Query Language (SQL)

Slide

Topics to be Covered

Topic	File organization and Indexing

Slide

#Q.36 Consider a file of 10,00 records, and disk blocks of 1000 bytes. There are two alternate options to create a single level index file:

- Option 1: Index file is created on an unordered key field, where key field is 12 bytes long.
- Option 2: Index file is created on an ordered nonkey field, where non-key field is 20 bytes long.

Block pointer size is 10 bytes long and unspanned organization is used.

Let 'X' is the blocking factor of index block using option 1 and 'Y' is the blocking factor of index block using option 2 then |X – Y| is 12.

Secondary index

Clustering index

$X = \left\lfloor \frac{\text{Block size}}{(12+10)} \right\rfloor = \left\lfloor \frac{1000}{22} \right\rfloor = 45$

$Y = \left\lfloor \frac{\text{Block size}}{(20+10)} \right\rfloor = \left\lfloor \frac{1000}{30} \right\rfloor = 33$

Ans

#Q.37 Consider an unordered file of 10,0000 records with a record size of 100 bytes stored on blocks of 1KB with an unspanned record organization. We will assume that no system related information is stored within a block. How many blocks would be needed to store this file? _____

1024 Bytes

$$\text{No. of blocks required} = \frac{\text{Total no. of records}}{(\text{No. of records per block})} = \frac{10,0000}{\left\lfloor \left(\frac{B.S.}{R.S.}\right) \right\rfloor = \left\lfloor \frac{1024}{100} \right\rfloor}$$

$$\left\lfloor \frac{10,0000 * 100}{1024} \right\rfloor$$

$$= \frac{10,000\cancel{0}}{1\cancel{0}}$$

$$= 10,000$$

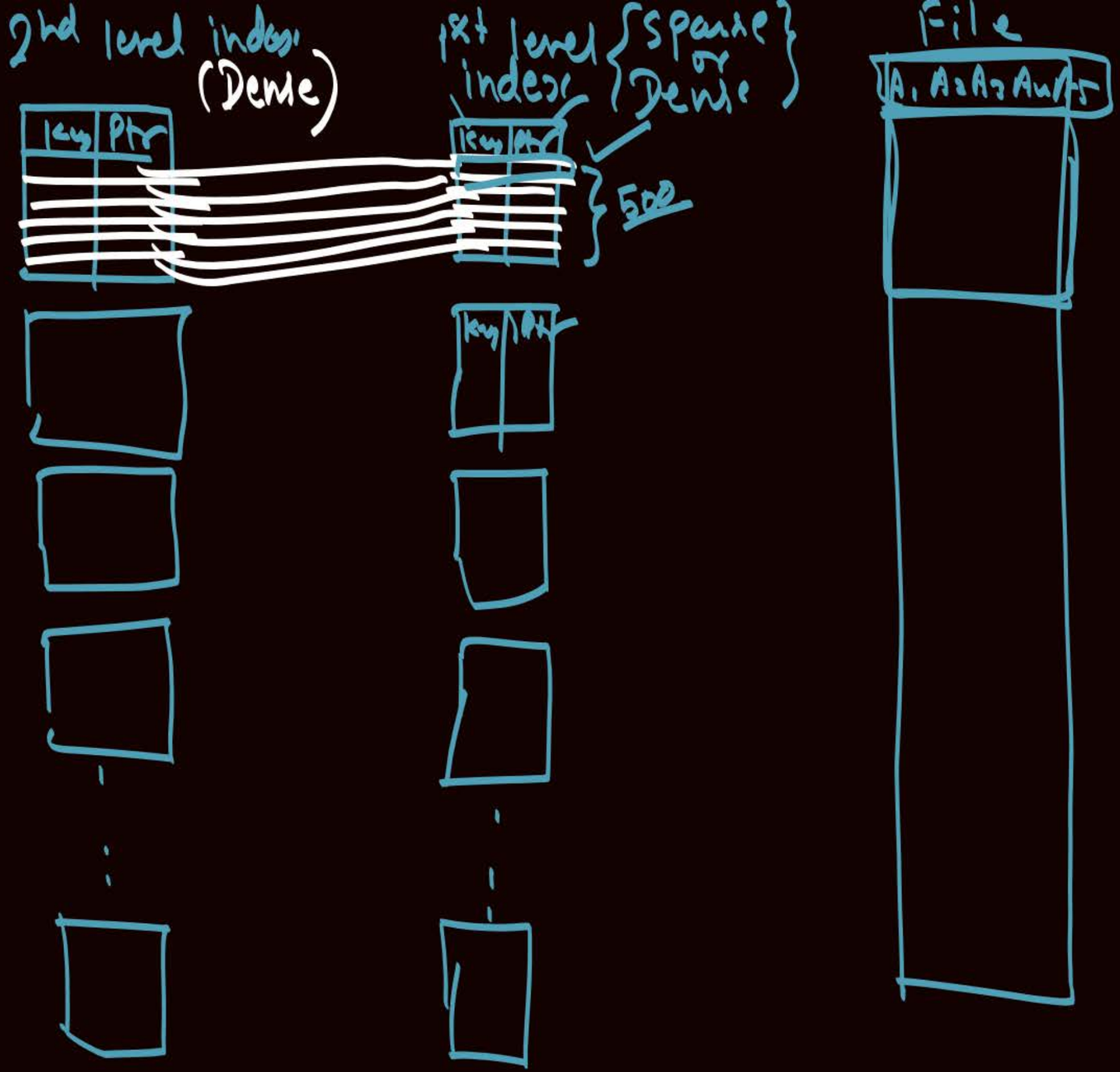
#Q.38 Which of the following is/are true for a multi-level index on a database file?

A. It is possible to construct more than one sparse first level indexes on different keys

B. It is possible to construct more than one dense first level indexes on different keys

C. First level index can be sparse or dense.

D. There is no benefit of using a dense second level index.


2nd level index
(Dense)
1st level index
{Sparse or Dense}
Key
Ptr
500
File

#Q.39 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 = Total no. of records EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

The record size of the file is_____bytes

$= (30 + 9 + 9 + 40 + 9 + 8 + 1 + 4 + 4) + 1$

$= (115)$

Record size = 115 bytes

#Q.40 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

The number of disk blocks required to store the file assuming an unspanned organization is used_____

$$= \frac{\text{No. of records}}{\text{No. of records per block}} = \frac{30,000}{\lfloor 512/115 \rfloor} = \frac{30,000}{4} = 7500$$

Blocks to store file : 7500 ; R.S = 115 ; B.F of database block = 4

#Q.41 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

The blocking factor of index block is_____

$$\left\lfloor \frac{512}{9+6} \right\rfloor = 34$$

#Q.42 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

Total number of entries in first level index is = No. of blocks to store DB file = 7500

Primary index is always sparse

In general 1 entry per block of database file

#Q.43 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

Total number of disk blocks required to store first level index is_____

$$\#\text{ Entries needs to be stored} = 30000$$

$$\text{No. of entries per block} = \left\lfloor \frac{512}{34} \right\rfloor = 15$$

$$\left\lceil \frac{30000}{15} \right\rceil = \lceil 2000 \rceil = 2000$$

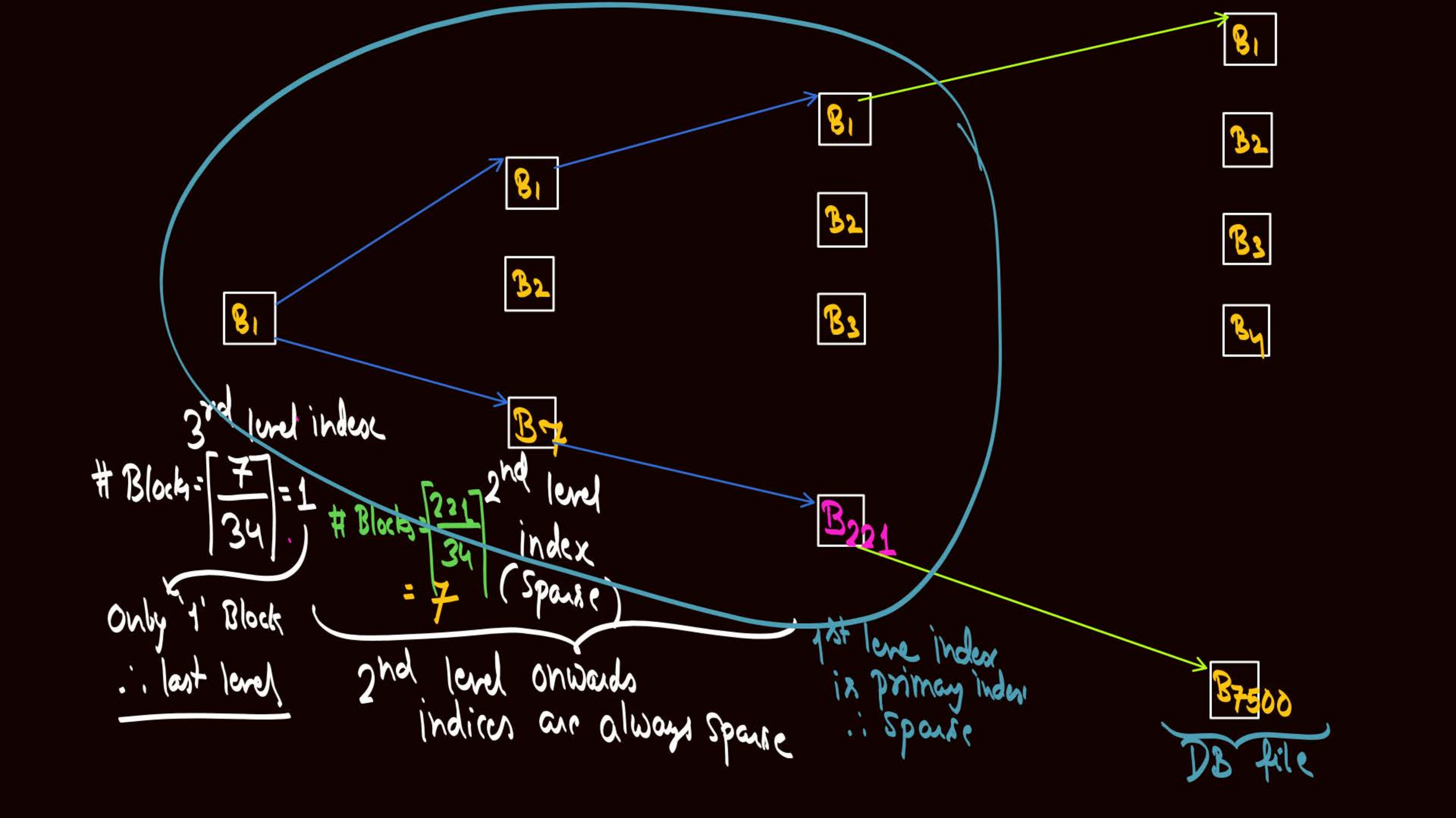
#Q.44 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

Total number of disk blocks required to store first level index is_____

#Q.45 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

- If we decide to create a multi-level index, then total number of levels needed in multi-level index is "3"

IO Cost using Multi-level index = (No. of levels + 1) = 3 + 1 = 4

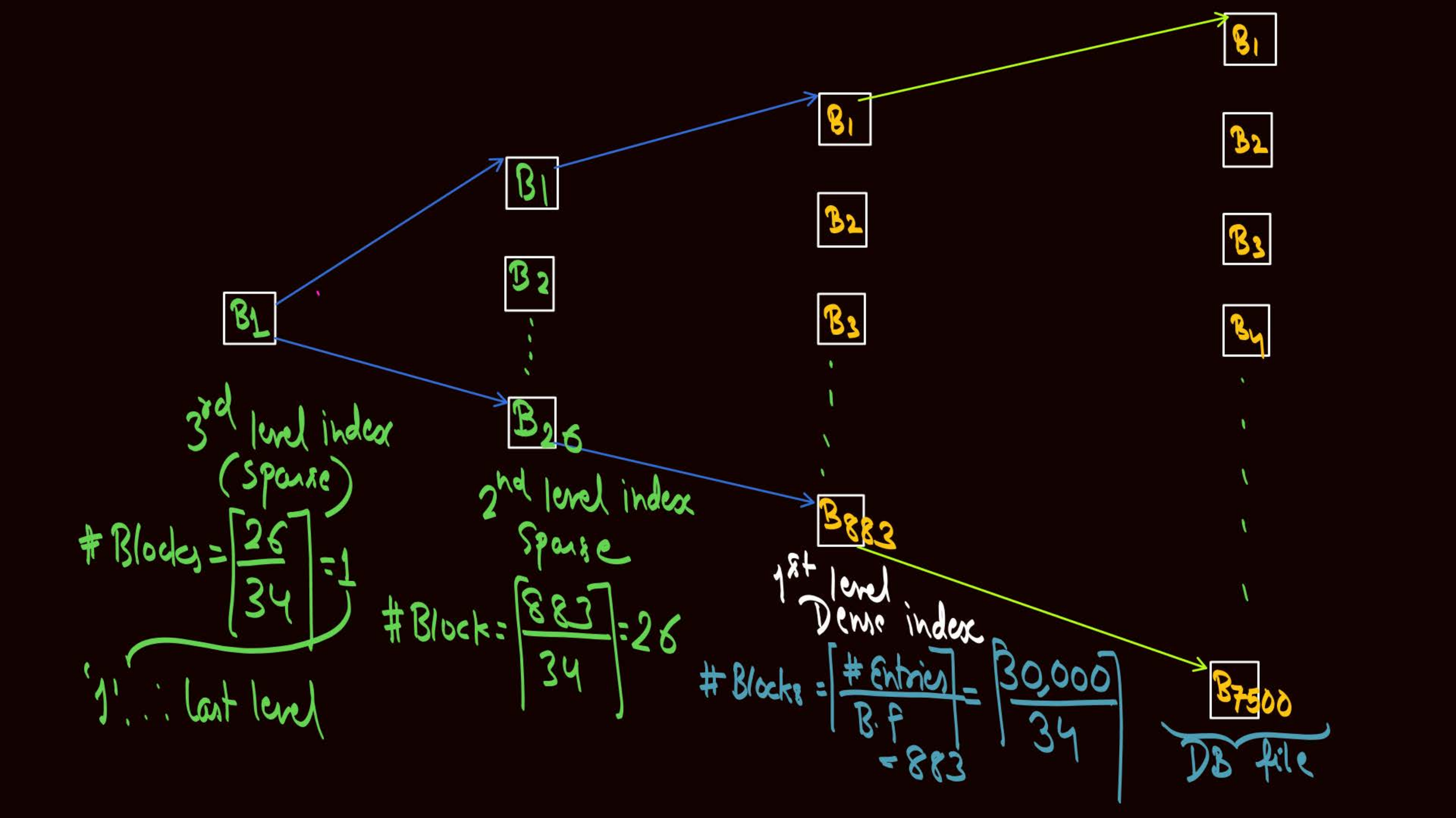

B1
B1
B2
B1
B2
B3
B1
B2
B3
B4
B7
B221
B7500
3rd level index
Blocks = ⌈7/34⌉ = 1
Only '1' Block
∴ last level
Blocks = ⌈221/34⌉ = 7
2nd level index (Sparse)
2nd level onwards indices are always Sparse
1st level index is primary index ∴ Sparse
DB file

#Q.45 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

- If we decide to create a multi-level index, then total number of levels needed in multi-level index is 3

(given first level index is dense index)

IO Cost using multi-level index = (No. of levels + 1) = 3 + 1 = 4


B1
B1
B2
B3
B4
B1
B1
B2
B3
B2
B26
B883
B7500
3rd level index (sparse)
Blocks = ⌈26/34⌉ = 1
'1' ... last level
2nd level index Sparse
Block = ⌈883/34⌉ = 26
1st level Dense index
Blocks = ⌈# Entries / B.f⌉ = ⌈30,000/34⌉
= 883
DB file

#Q.46 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

If we decide to create a multi-level index, then total number disk blocks required to store multi-level index is_____

{when first level index is primary index} = 221+7+1 = 229

#Q.46 Consider a disk with block size B = 512 bytes. A block pointer is PB = 6 bytes long, and a record pointer is PR = 7 bytes long. A file has 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Ename (30 bytes), Eid (9 bytes), Did(9 bytes), Location (40 bytes), Contact (9 bytes), DOB(8 bytes), Gender (1 byte), Role(4 bytes), and Salary (4 bytes) Other than the specified fields, an additional byte is used as a deletion marker in each record. Suppose the file is ordered by the key field Eid and we want to construct a primary index on Eid.

If we decide to create a multi-level index, then total number disk blocks required to store multi-level index is_____

{when first level index is secondary index} = 883+26+1 = 910

∴

#Q.47 Consider a B tree in which the maximum numbers of keys in a node is '9',
The minimum numbers of keys a non-root may have is ____ Ans = 4

If order of a node of B tree is 'p'
then Maximum no. of keys a node can have = $p-1 = 9$
$\Rightarrow p = 10$
End

Minimum No. of keys a non-root node must have is $\lceil \frac{p}{2} \rceil - 1$
$= \lceil \frac{10}{2} \rceil - 1 = 5-1 = 4$

#Q.48 Let 'X' denotes the order of an internal node of B+tree, and 'Y' denotes order of a leaf node of B+tree.

If disk block size is 1024 bytes, search key is 15 bytes long, block pointer is 10 bytes long and record pointer is 10 byte long then maximum value of X+Y is (41+40) = 81

= Ans

$(X * Bp\ size + (X-1) * key\ size) \leq Block\ size$

$(X*10 + (X-1)*15) \leq 1024$

$25X \leq 1024+15$

$X \leq 41.56 \Rightarrow X_{max} = 41$

$(Y * (key\ size + Rp\ size) + 1 * Bp\ size) \leq Block\ size$

$Y*(15+10)+10 \leq 1024$

$25Y \leq 1014 \Rightarrow Y \leq 40.56$ $\quad Y_{max} = 40$

#Q.49 Let 'X' denotes the order of an internal node of B tree, and 'Y' denotes order of a leaf node of B tree.

If disk block size is 1024 bytes, search key is 15 bytes long, block pointer is 10 bytes long and record pointer is 10 byte long then maximum value of X+Y is $29+29 = 58$ Ans

* For a B tree, order of internal node is same as order of leaf node i.e. $x = y$

$$(x * Bp\,size) + (x-1)\left(\frac{key}{size} + \frac{R.p}{size}\right) \leq \frac{Block}{size}$$

$$x * 10 + (x-1)(15+10) \leq 1024$$

$$35x \leq 1024 + 25$$

$$35x \leq 1049$$

$$x \leq 29.97$$

$x_{max} = 29$, $y_{max} = 29$

2 mins Summary

Topic File organization and Indexing

Slide

THANK - YOU

Slide