Database Design
CS 6360.003: Homework #5

Due on Sunday November 20, 2016 at 11:59pm

Nurcan Yuruk

Hanlin He **(hxh160630)**
hanlin.he@utdallas.edu

Problem 1

Part (a)

The record size in bytes is:

$$R = 30 + 10 + 10 + 30 + 10 + 10 + 1 + 4 + 4 + 1 = 110 \text{ bytes.}$$

Part (b)

The blocking factor is:

$$bfr = \left\lfloor \frac{512}{110} \right\rfloor = 4 \text{ records per block.}$$

The number of blocks needed assuming an unspanned organization is:

$$b = \left\lceil \frac{3000}{4} \right\rceil = 750 \text{ file blocks.}$$

Part (c)

(i) Index size is $10 + 6 = 16$ bytes, so the index blocking factor is

$$bfr_i = \left\lfloor \frac{512}{16} \right\rfloor = 32 \text{ index records per block.}$$

(ii) Total 750 first level index entries (the same as file blocks), so the number of blocks needed for first level index is:

$$b_i = \left\lceil \frac{750}{32} \right\rceil = 24 \text{ first level index blocks.}$$

(iii) The second level need only 1 block to store 24 first level index blocks information. Thus, in total, 2 levels of indexes is needed.

(iv) Total number of blocks by multi-level index is $1 + 24 = 25$.

(v) 3 block accesses are needed.

Part (d)

(i) Index size is $10 + 6 = 16$ bytes, so the index blocking factor is

$$bfr_i = \left\lfloor \frac{512}{16} \right\rfloor = 32 \text{ index records per block.}$$

Problem 1 continued on next page. . .

(ii) Total 3000 first level index entries (the same as file records), so the number of blocks needed for first level index is:

$$b_i = \left\lceil \frac{3000}{32} \right\rceil = 94 \text{ first level index blocks.}$$

(iii) The second level need $\left\lceil \frac{94}{32} \right\rceil = 3$ blocks to store 94 first level index blocks.
The third level need only 1 block to store 3 second level index blocks information. Thus, in total, 3 levels of indexes is needed.

(iv) Total number of blocks by multi-level index is $1 + 3 + 94 = 98$.

(v) 4 block accesses are needed.

Part (e)

(i) Index size is $10 + 6 = 16$ bytes, so the index blocking factor is

$$bfr_i = \left\lfloor \frac{512}{16} \right\rfloor = 32 \text{ index records per block.}$$

(ii) Record pointer size = 7 bytes, thus a block can store at most $\left\lfloor \frac{512}{7} \right\rfloor = 73$ record pointers.
Given 3000 employees evenly distributed among 100 department, each department has 30 employees, which is smaller than 73, the capacity of one block to store record pointers, i.e. no cluster or linked list is needed.
Thus, 100 blocks representing 100 departments by the level of indirection are needed.

(iii) Total 100 first level index entries (the same as file blocks), so the number of blocks needed for first level index is:

$$b_i = \left\lceil \frac{100}{32} \right\rceil = 4 \text{ first level index blocks.}$$

(iv) The second level need only 1 block to store 7 first level index blocks information. Thus, in total, 2 levels of indexes is needed.

(v) Total number of blocks by multi-level index is $1 + 4 = 5$.
The blocks used in the extra level of indirection is 100.

(vi) $3 + 30 = 33$ block accesses are needed.

Part (f)

(i) Index size is $10 + 6 = 16$ bytes, so the index blocking factor is

$$bfr_i = \left\lfloor \frac{512}{16} \right\rfloor = 32 \text{ index records per block.}$$

Problem 1 continued on next page...

(ii) 100 first level index entries representing 100 departments are needed.
So the number of blocks needed for first level index is:

$$b_i = \left\lceil \frac{100}{32} \right\rceil = 4 \text{ first level index blocks.}$$

(iii) The second level need only 1 block to store 4 first level index blocks information. Thus, in total, 2 levels of indexes is needed.

(iv) Each block can store 4 records, each department has 30 employees. Thus each department needs $\left\lceil \frac{30}{4} \right\rceil = 8$ blocks to store all its employees records. In total, $8 \times 100 = 800$ blocks is needed.
Total number of blocks by multi-level index is $1 + 4 = 5$.

(v) $2 + 8 = 10$ block accesses are needed.

Part (g)

(i) For B+ tree:

$$\begin{aligned} & (p \times P) + ((p-1) \times V) \leq B \\ \Rightarrow \quad & (p \times 6) + ((p-1) \times 10) \leq 512 \\ \Rightarrow \quad & (15 \times p) \leq 522 \\ \Rightarrow \quad & \max\{p\} = 32 \end{aligned}$$

$$\begin{aligned} & (p_{leaf} \times (P_r + V)) + P \leq B \\ \Rightarrow \quad & (p_{leaf} \times (7 + 10)) + 6 \leq 512 \\ \Rightarrow \quad & (17 \times p_{leaf}) \leq 506 \\ \Rightarrow \quad & \max\{p_{leaf}\} = 29 \end{aligned}$$

(ii) On the average, each leaf node will hold $0.69 \times p_{leaf} = 0.69 \times 29$ or approximately 20 data record pointers.
Thus, leaf level blocks number is:

$$b_{leaf} = \left\lceil \frac{3000}{20} \right\rceil = 150 \text{ leaf level blocks.}$$

(iii) On the average, each internal node will have $0.69 \times p = 0.69 \times 34$ or approximately 22 pointers, and hence 21 values.
First internal level blocks number is:

$$b_1 = \left\lceil \frac{150}{21} \right\rceil = 8 \text{ first internal level blocks.}$$

Problem 1 continued on next page. . .

Second internal level (root level) block number is 1.
In total, there are 3 levels.

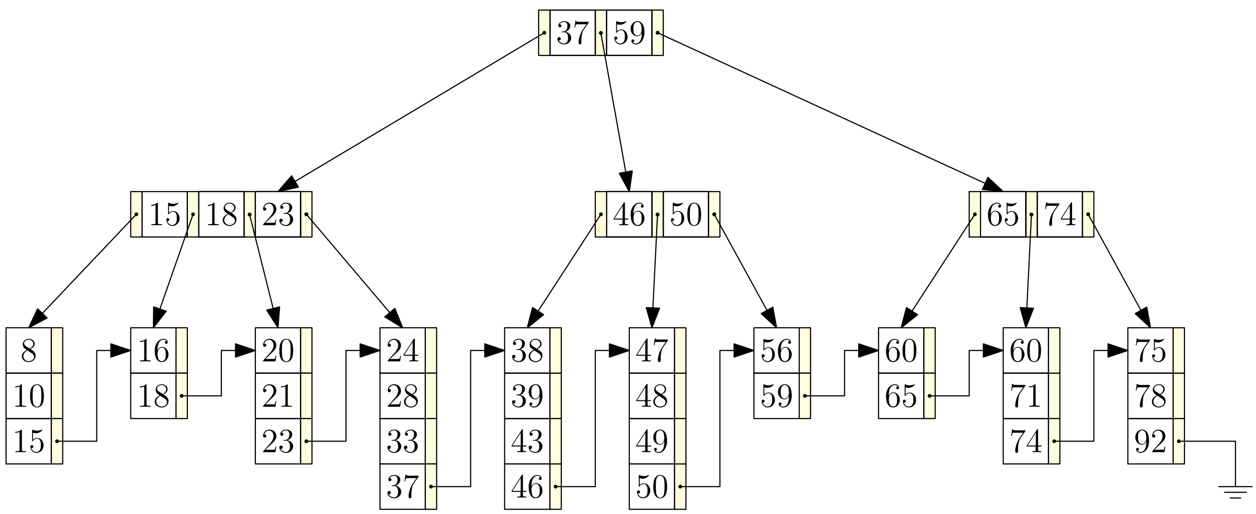
(iv) Total number of blocks required by B+ tree is:

$$B = 1 + 8 + 150 = 159 \text{ (blocks in total.)}$$

(v) $3 + 1 = 4$ block accesses are needed.

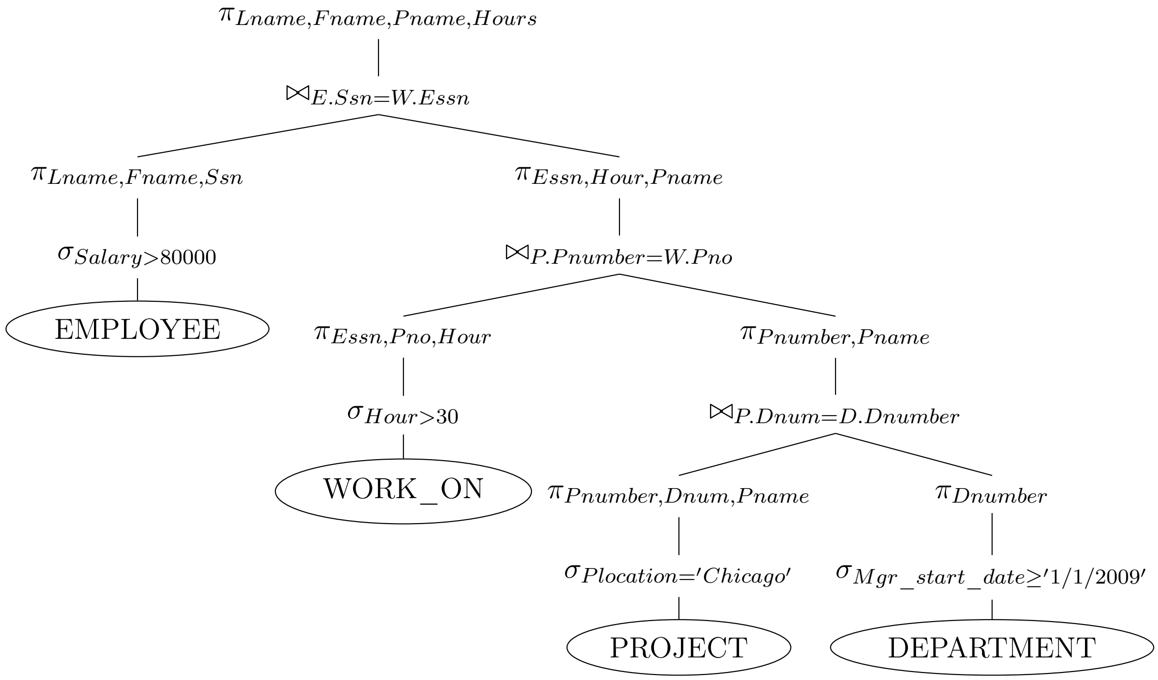
Problem 2

Figure 1: Final B+ tree



Problem 2 continued on next page...

Problem 3


$\pi_{Lname,Fname,Pname,Hours}$
$\bowtie_{E.Ssn=W.Essn}$
$\pi_{Lname,Fname,Ssn}$
$\sigma_{Salary>80000}$
EMPLOYEE
$\pi_{Essn,Hour,Pname}$
$\bowtie_{P.Pnumber=W.Pno}$
$\pi_{Essn,Pno,Hour}$
$\sigma_{Hour>30}$
WORK_ON
$\pi_{Pnumber,Pname}$
$\bowtie_{P.Dnum=D.Dnumber}$
$\pi_{Pnumber,Dnum,Pname}$
$\sigma_{Plocation='Chicago'}$
PROJECT
$\pi_{Dnumber}$
$\sigma_{Mgr\_start\_date\geq'1/1/2009'}$
DEPARTMENT