CSC430/530 – Database Management Systems

Assignment #5 – Indexing Structures

Consider a disk with a block size **B = 1024 bytes** and a block pointer **P = 12 bytes**. Suppose file has **r = 69632** EMPLOYEE records of fixed length.

Each record has following attributes:

- First_name (*60 bytes*);
- SSN (*18 bytes*);
- Dnumber (*18 bytes*);
- Home_address (*80 bytes*);
- Contact_phone (*20 bytes*);
- BDate (*16 bytes*);
- Gender (*2 byte*);
- Job_code (*8 bytes*);
- Salary (*8 bytes*);
- Additional *2 bytes* are used as a deletion marker.

a. Assuming an unspanned organization, calculate following.
***To get full points, please, show all formulas and calculations.**

- Record size **R** (in bytes).

R = *60 + 18 + 18 + 80 + 20 + 16 + 2 + 8 + 8 + 2 = 232 bytes*

- Blocking factor **bfr.**

bfr = $\lfloor B/R \rfloor = \left\lfloor 1024\ bytes / 232\ bytes \right\rfloor$ **= 4 records/block**

- Number of file blocks **b**.

b = $\lceil r/bfr \rceil$ = $\lceil 69,632\ records / 4\ records/block \rceil$ = **17,408 blocks**

b. Assuming file is ordered by key attribute **Ssn** and primary index is constructed on this key attribute, calculate the following.

***To get full points, please show all formulas and calculations.**

- Index blocking factor $\mathbf{bfr_i}$.

R_i = SSN_Size + Pointer_Size = 18 bytes + 12 bytes = 30 bytes

bfr_i = $\lfloor B/R_i \rfloor$ = $\lfloor 1024/30 \rfloor$ = **34**

- Total number of indexes $\mathbf{r_i}$ and number of index blocks $\mathbf{b_i}$.

r_i = one for each data file block = b = 17,408 indexes

b_i = $\lceil r_i / bfr_i \rceil$ = $\lceil 17408/34 \rceil$ **R_i / bfr = 512 blocks**

- Total number of accesses needed to search for a record using primary index.

ceil($\log_2(b_i)$) + 1 = ceil($\log_2(512)$) + 1 = 10