

# CS 405G: Introduction to database systems

## Assignment 6

Assigned: April. 15 , 2011

Due: April. 22, 2011

Reading assignment: Chapter 17 and Chapter 18.1~18.3.

**Submission instructions:** Please try to prepare the assignment in word document (handwritten homework is also acceptable). Print and bring them to class on the due date. Put the course number and your name at the top left corner of the first page of your homework. Please follow the rules in the course syllabus regarding late homework and plagiarism.

Problem (1) A file has  $r=20000$  STUDENT records of **fixed-length**. Each record has the following fields: NAME (30 bytes), SSN (9 bytes), ADDRESS (40 bytes), PHONE (9 bytes), BIRTHDATE (8 bytes), SEX (1 byte), MAJORDEPTCODE (4 bytes), MINORDEPTCODE (4 bytes), CLASSCODE (4 bytes, integer), and DEGREEPROGRAM (3 bytes). An additional byte is used as a deletion marker. Assume the disk **seek time** is 5ms, **rotation delay** is 4.2ms, and **transfer time** is 0.5ms per block.

- (a) Calculate the record size  $R$  in bytes.
- (b) Calculate the **blocking factor**  $bfr$  and the number of file blocks  $b$  assuming an **unspanned** organization.
- (c) Calculate the average time it takes to find a record by doing a linear search on the file if
  - (i) the file blocks are stored contiguously and **double buffering** is used,
  - (ii) the file blocks are not stored contiguously.

Problem 2. Consider a disk with block size  $B=512$  bytes. A block pointer is  $P=6$  bytes long, and a record pointer is  $P R =7$  bytes long. A file has  $r=30,000$  EMPLOYEE records of fixed-length. Each record has the following fields: NAME (30 bytes), SSN (9 bytes), DEPARTMENTCODE (9 bytes), ADDRESS (40 bytes), PHONE (9 bytes), BIRTHDATE (8 bytes), SEX (1 byte), JOBCODE (4 bytes), SALARY (4 bytes, real number). An additional byte is used as a deletion marker.

- (a) Calculate the record size  $R$  in bytes.
- (b) Calculate the blocking factor  $bfr$  and the number of file blocks  $b$  assuming an unspanned organization.
- (c) Suppose the file is ordered by the key field SSN and we want to construct a **primary index** on SSN. Calculate
  - (i) the index blocking factor  $bfr_i$  (which is also the **index fan-out**);
  - (ii) the number of first-level index entries and the number of first-level index blocks;
  - (iii) the number of levels needed if we make it into a multi-level

index;

(iv) the total number of blocks required by the multi-level index; and

(v) the number of block accesses needed to search for and retrieve a record from the file--given its SSN value--using the primary index.

Problem (3) A PARTS file with Part# as key field includes records with the following Part# values: 23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20, 24, 28, 39, 43, 47, 50, 69, 75, 8, 49, 33, 38. Suppose the search field values are inserted in the given order in a **B + -tree** of order with maximum **fan-out** size as 4, show how the tree will expand and what the final tree looks like.