Alexandria University
Faculty of Engineering
Comp. & Comm. Engineering
CC471: Database Systems

جامعة الاسكندرية
كلية الهندسة
برنامج هندسة الحاسب والاتصالات
مادة قواعد البيانات
ربيع ٢٠٢١

Sheet4

DISC STORAGE, BASIC FILE STRUCTURES, HASHING

1) Can you think of techniques other than an unordered overflow file that can be used to make insertion in an ordered file more efficient?
2) Can you think of techniques other than chaining to handle bucket overflow in external hashing?
3) Consider a disk with the following characteristics (these are not parameters of any particular disk unit): block size B=512 bytes, interblock gap size G=128 bytes, number of blocks per track=20, number of tracks per surface=400. A disk pack consists of 15 double-sided disks.
   a) What is the total capacity of a track and what is its useful capacity (excluding interblock gaps)?
   b) How many cylinders are there?
   c) What is the total capacity and the useful capacity of a cylinder?
   d) What is the total capacity and the useful capacity of a disk pack?
   e) Suppose the disk drive rotates the disk pack at a speed of 2400 rpm (revolutions per minute); what is the transfer rate in bytes/msec and the block transfer time btt in msec? What is the average rotational delay rd in msec? What is the bulk transfer rate (see Appendix B)?
   f) Suppose the average seek time is 30 msec. How much time does it take (on the average) in msec to locate and transfer a single block given its block address?
   g) Calculate the average time it would take to transfer 20 random blocks and compare it with the time it would take to transfer 20 consecutive blocks using double buffering to save seek time and rotational delay.
4) A file has r=20,000 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. The file is stored on the disk whose parameters are given in the previous exercise.
   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, and (ii) the file blocks are not stored contiguously.
   d) Assume the file is ordered by SSN; calculate the time it takes to search for a record given its SSN value by doing a binary search.

How to submit the homework assignments?

- Solve the sheet individually without looking up the solution on the Internet. The sheet is to practice; it is a learning tool not an exam.
- Assignments are to be **handwritten**.
- Papers are to be scanned (I like camscanner app). Put all images in a pdf file (camscanner does that for you)
