

PRACTICE PROBLEM SET #7

CSC 261/461 (Database Systems), Spring 2018,
University of Rochester
03/28/2018

Sections to study

Chapter 16

- 16.1
- 16.2 (16.2.1, 16.2.2)
- 16.3
- 16.4
- 16.5
- 16.6
- 16.7
- 16.8 (Skip 16.8.3)

Chapter 17

- 17.1
- 17.2
- 17.3 (17.3.1, 17.3.2)

(Note: You need to study sections marked in red thoroughly. Other sections are important for comprehension. The sections those are not included can be skipped (but not recommended))

Problems

Chapter 16

Problem 1

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.

- What is the total capacity of a track and what is its useful capacity (excluding interblock gaps)?
- How many cylinders are there?
- What is the total capacity and the useful capacity of a cylinder?
- What is the total capacity and the useful capacity of a disk pack?
- Suppose the disk drive rotates the disk pack at a speed of 2400 rpm (revolutions per minute); what are the transfer rate in bytes/msec and the block transfer time btt in msec? What is the average rotational delay rd in msec?

- (f) Suppose the average seek time (time to go to the correct track mechanically) 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.

Solution:

- (a) Total track size = $20 * (512+128) = 12800$ bytes = 12.8 Kbytes
Useful capacity of a track = $20 * 512 = 10240$ bytes = 10.24 Kbytes
- (b) Number of cylinders = number of tracks = 400
- (c) Total cylinder capacity = $15 * 2 * 20 * (512+128) = 384000$ bytes = 384 Kbytes
Useful cylinder capacity = $15 * 2 * 20 * 512 = 307200$ bytes = 307.2 Kbytes
- (d) Total capacity of a disk pack = $15 * 2 * 400 * 20 * (512+128) = 153600000$ bytes = 153.6 Mbytes
Useful capacity of a disk pack = $15 * 2 * 400 * 20 * 512 = 122.88$ Mbytes
- (e) Transfer rate $tr = (\text{total track size in bytes}) / (\text{time for one disk revolution in msec})$
 $tr = (12800) / ((60 * 1000) / (2400)) = (12800) / (25) = 512$ bytes/msec
block transfer time $btt = B / tr = 512 / 512 = 1$ msec
average rotational delay $rd = (\text{time for one disk revolution in msec}) / 2 = 25 / 2 = 12.5$ msec
- (f) average time to locate and transfer a block = $s + rd + btt = 30 + 12.5 + 1 = 43.5$ msec
- (g) time to transfer 20 random blocks = $20 * (s + rd + btt) = 20 * 43.5 = 870$ msec
time to transfer 20 consecutive blocks using double buffering = $s + rd + 20 * btt = 30 + 12.5 + (20 * 1) = 62.5$ msec

Problem 2

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 Problem 1 .

- (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
- the file blocks are stored contiguously and double buffering is used, and
 - 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.

Solution:

(a) $R = (30 + 9 + 40 + 9 + 8 + 1 + 4 + 4 + 4 + 3) + 1 = 113$ bytes

(b) $bfr = \lfloor B/R \rfloor = \lfloor 512/113 \rfloor = 4$ records per block
 $b = \lceil (r/bfr) \rceil = \lceil 20000/4 \rceil = 5000$ blocks

(c) For linear search we search on average half the file blocks= $5000/2 = 2500$ blocks.

(i) If the blocks are stored consecutively, and double buffering is used, the time to read 2500 consecutive blocks

$$= s + rd + (2500 * (B/btr))$$

$$= 30 + 12.5 + (2500 * (512/409.6)) = 3167.5 \text{ msec} = 3.1675 \text{ sec}$$

(a less accurate estimate is $s + rd + (2500 * btt) = 30 + 12.5 + 2500 * 1 = 2542.5 \text{ msec}$)

(ii) If the blocks are scattered over the disk, a seek is needed for each block, so the time is:
 $2500 * (s + rd + btt) = 2500 * (30 + 12.5 + 1) = 108750 \text{ msec} = 108.75 \text{ sec}$

(d) For binary search, the time to search for a record is estimated as

$$\lceil \log_2 b \rceil * (s + rd + btt)$$

$$= \lceil \log_2 5000 \rceil * (30 + 12.5 + 1)$$

$$= 13 * 43.5 = 565.5 \text{ msec} = 0.5655 \text{ sec}$$

Chapter 17

Study all the examples (Examples 1-7 + Examples of Insertion into and Deletion from B+ trees from the textbook)

Solution: Please refer the textbook for solutions