Assignment 1

Table 1: Disk Parameters

disk pack configuration		2 double-sided disks
read/write heads		4
blocks per track		400
tracks per surface		20,000
block size	B	512 bytes
rotational velocity	p	7200 rpm
average seek time	s	6 msec
block pointer size	P	6 bytes
record pointer size	P_R	7 bytes

Table 2: EMPLOYEE Relation

number of records	500,000		
storage	unspanned, contiguous, ordered on PK (EMPID)		
schema	**attribute**	**domain**	**size (bytes)**
	FNAME	CHAR(15)	20
	LNAME	CHAR(15)	20
	EMPID	CHAR(9)	9
	BDATE	DATE	4
	ADDRESS	CHAR(30)	30
	SEX	CHAR	1
	SALARY	DECIMAL(10,2)	8
	SUPERSSN	CHAR(9)	9
	DNO	INT	4
	TITLE	CHAR(15)	15

Record Storage: Answer the following for the EMPLOYEE relation:

(a) Compute the record size (R):

Assignment 1

(b) Compute the blocking factor (bfr):

(c) How many blocks are required to store the relation?

Primary Indexes: Compute the following for a *primary index* on the EMPLOYEE relation.

(a) What is the index record size?

(b) What is the index blocking factor (bfr_i)?

(c) How many disk blocks are required to store a *single-level* primary index?

(d) How many disk blocks are required to store a *multi-level* primary index? (Assume the top level is one block.) (e) How many block reads are required to execute the following parameterized query: SELECT * FROM EMPLOYEE WHERE EMPID = [parameter];

With no index?
With a single-level primary index?
With a multiple-level primary index?

(f) Estimate the *time* required to execute the query in (e), for each of the three conditions.

Clustering Indexes: Suppose the EMPLOYEE relation is ordered on field DNO (department number), rather than the primary key. There are 100 departments and employees are evenly distributed among the departments.

(a) What is the index blocking factor (bfr_i)?

(b) How many blocks are required to store a clustering index on DNO?

(c) Would a multi-level index on DNO be effective?

Secondary Indexes:

(a) Assume the company is 60% male and 40% female. Would a secondary index on the field SEX be effective?

(b) Assume there are 2000 departments and the employees are evenly distributed among those departments. Assume that the primary key EMPID orders the EMPLOYEE relation, and that a multilevel primary index exists.
Design a secondary index on DNO.

(c) How many blocks are required to store your secondary index?

(d) How many block accesses (on average) are required to execute the following parameterized query, using your secondary index?
SELECT * FROM EMPLOYEE WHERE DNO =

[parameter]; (e) Design a secondary index on employee

names.

Assignment 1

For the following disk and B+-tree, estimate:

1) height of the tree
2) blocks required to store the tree

Assume a fill factor of 70% for all nodes.

Disk Parameters

block size	B	1024 bytes
block pointer size	P	6 bytes

Data File and Tree Parameters

number of data records	r	1,800,000 records
record size	R	128 bytes
size of index value for B+-tree	V	19 bytes