

National University of Computer and Emerging Sciences, Lahore Campus



Course:	Advance Database Concepts	Course Code:	CS4064
Program:	BS (Computer Science)	Semester:	Spring 2023
Out Date:	8-Mar-2023	Total Marks:	
Due Date:	Thu 16-Mar-2023 (Start of class)	Weight:	
Section		Page(s):	2
Assignment:	2 (File Structures & Hashing)		

Instructions:

- This assignment is an individual assignment.
- You are required to submit the hard copy of your assignment at the start of your class.
- Use any valid assumption where needed.
- For any query please contact your TA.

Q1. Consider a file system on a disk with block size $B=1024$ bytes. A file has $r=10,000,000$ STUDENT records of fixed length (un-spanned). Each record length is 100 bytes. Assume there are 200 departments and 50,000 students per department. Primary key of student file is RollNo.

Estimate the number of block fetches needed to compute the following queries:

- SELECT * FROM STUDENT WHERE rollno=1357; (Assume file is not ordered)
- SELECT * FROM STUDENT WHERE rollno=2468; (Assume file is ordered on rollno)
- SELECT * FROM STUDENT WHERE deptno= 11; (Assume file is not ordered)
- SELECT * FROM STUDENT WHERE deptno= 22; (Assume file is ordered on deptno)

Q2. A customer file has following customer ID values:

(2, 3, 5, 7, 11, 17, 19, 23, 29, 31)

- Consider these customer IDs as hash key values. The file uses 6 buckets named 0 to 5. One bucket cannot hold more than 3 records, means at max a bucket can hold 3 records. Load these records in file using hash function $h(k) = k \bmod 6$, in the given order.
- Calculate the average number of block accesses for random retrieval on customer ID.
- Load the given values in expandable hash files based on extendible hashing, show structure on each step, use hash function $h(k) = K \bmod 7$, max 3 records can be kept in one bucket.
- Load the given values in expandable hash files based on dynamic hashing, show structure on each step, use hash function $h(k) = K \bmod 7$, max 3 records can be kept in one bucket.
- Load the given values in expandable hash files based on linear hashing, show structure on each step, use hash function $h(k) = K \bmod 7$, max 3 records can be kept in one bucket.
- Show the structure after inserting an entry with customer ID value 13 into the final structure of part (c).
- Show the structure after inserting an entry with customer ID value 18 into the structure of part (f).

Q3. Consider a dynamic hash structure where buckets can hold up to three records. Initially the structure is empty. Then we insert the following records, in the order below, where we indicate the hashed key in brackets (in binary):

a	[010000]
b	[011010]
c	[111100]
d	[001110]
e	[010111]
f	[011010]
g	[101001]
h	[010111]
i	[000110]
j	[101001]

Show the linear hash structure after these records have been inserted. Assume that the threshold value is 2. (i.e., when the average number of keys per non-overflow bucket is greater than 2, we allocate another bucket).