National University of Computer and Emerging Sciences
Lahore Campus

Advance Database Concepts (CS4064)

Date: Wed, 03 April 2024

Course Instructor(s)

Muhammad Ishaq Raza
Muhammad Naveed

Sessional-II Exam

Total Time (Hrs.): 1
Total Marks: 30
Total Questions: 3

_______________ _____________ _____________________________
Roll No Section Student Signature

Do not write below this line.

Attempt all the questions.

CLO # 2

Q. No 1: Suppose you have the following search-key values: 4, 15, 10, 16, 20, 12, 18, 5, 31, 17

a. Consider these search-key values. The file uses 6 buckets named 0 to 5 and each bucket cannot hold more than 2 records. Load these records in the file in the given order using hash function *h(k)= k mod 6*. [3]

Solution:

Buckets	0	1	2	3	4	5
Hashed Key	12,18	31	20	15,	4,10,16 overflow	5,17

b. Calculate the average number of block accesses for a random retrieval on a search-key. [2]

Solution: (1*(9/10)) + (2*(1/10)) =0.9+0.2= 1.1 block access

c. Load the records with the above search-key values into an expandable hash file based on extendible hashing. Show the structure of the directory at each step along with the global and local depths. Use the hash function *h(k)= k* and each bucket cannot hold more than 3 records. *Use higher bits (i.e., Left to Right) of hash value to determine a directory entry*. [5]

Solution:

H(Key)=key	4	15	10	16	20	12	18	5	31	17
Binary	00100	01111	01010	10000	10100	01100	10010	00101	11111	10001

d=0

4	15	10

d=1

0:	4	15	10	d’=1,	12 overflow
1:	16	20		d’=1	

d=2

00:	4	5		d’=2	
01:	15	10	12	d’=2	
10, 11:	16	20	18	d’=1	31 overflow

d=2

00:	4	5		d’=2	
01:	15	10	12	d’=2	
10:	16	20	18	d’=2	17 overflow
11:	31			d’=2	

d=3

000,001:	4	5		d’=2
010,011:	15	10	12	d’=2
100:	16	18	17	d’=3
101:	20			d’=3
111,110:	31			d’=2

CLO # 2

Q. No 2: Consider a disk with block size B=2048 bytes. A block pointer is P=6 bytes long. A file has r = 1,000,000 transaction records of fixed length. Record length R is 115 bytes long and customerID field is 10 bytes long. Suppose the file is ordered by the non-key field customerID and we want to construct a clustering index on customerID field that uses block anchors (i.e., every new value of customerID starts at the beginning of a new block). Assume there are 2500 distinct values of customerID, and that the transaction records are evenly distributed among these values. Calculate:

a. The index blocking factor (bfri). [2]
b. The number of first-level index entries (r1) and the number of first-level index blocks (b1). [2]
c. The number of levels needed (x) if we make it a multi-level index. [1]
d. The total number of blocks required by the multi-level index (bi). [1]
e. The number of block accesses needed to search for and retrieve all records in the file having a specific customerID value using the clustering index (assume that multiple blocks in a cluster are either contiguous or linked by pointers). [4]

Solution: B=2048 bytes, B.P=6 bytes, r=1000, 000, R=115 bytes,

customerID=10 bytes, distinct=2500

bfr=floor(2048/115)=17 records

a) $R_{i=}$ 10+6=16 bytes
Bfr_i=floor (B/R_i) =floor (2048/16) =128 records per block,

b) r1=2500 entries
b1=ceiling ($r1/bfr_i$) =ceiling (2500/128) =20 blocks

c) r2=20 entries
b2=ceiling ($r2/bfr_i$) =ceiling (20/128) =1 block
Hence: x=2 levels

d) Total blocks=b1+b2=20+1 = 21 blocks

e) The 400 records are clustered in ceiling (400/bfr) = ceiling (400/17) = 24 block.
Hence, total block accesses needed on average to retrieve all the records with a given customerID = x + 24 = 2 + 24 = 26 block accesses

CLO # 2

Q. No 3: Consider the following log at the point of system crash. Suppose that we use ARIES recovery algorithm to answer the following questions.

LSN	Last_LSN	Trans_ID	Type	Page_ID	Other_Info
1	0	T1	Update	A	…
2	1	T1	Update	B	…
3	2	T1	Commit		…
4	0	T2	Update	A	
5	0	T3	Update	C	
6	begin checkpoint				…
7	end checkpoint				…
8	5	T3	Commit		…
9	0	T4	Update	B	…
10	0	T5	Update	C	…
11	10	T5	Commit		…

a. What is the value of the LSN stored in the master log record (a special file)? [1]

b. Show the contents of transaction and dirty page table at the time of checkpoint. [2]

c. What is done during Analysis? Be precise about the points at which Analysis begins and ends and show the contents of transaction and dirty page table reconstructed in this phase. [3]

d. What is done during Redo? Be precise about the points at which Redo begins and ends. [2]

e. What is done during Undo? Be precise about the points at which Undo begins and ends. [2]

Solution:

a) LSN= 6

b)

Trans_ID	LSN	Status
T1	~~1,2~~,3	~~Inprogress~~,commit
T2	4	Inprogress
T3	5	Inprogress

Page_ID	LSN
A	1
B	2
C	5

c) Analysis: start=7 to till LSN=11

Transection Table

Trans_ID	LSN	Status
T1	~~1,2~~,3	~~Inprogress~~,commit
T2	4	Inprogress
T3	~~5~~,8	~~Inprogress~~,commit
T4	9	Inprogress
T5	~~10~~,11	~~Inprogress~~,commit

Dirty Page Table

Page_ID	LSN
A	1
B	2
C	5

d)

Begin_LSN (1) End_LSN(10)

LSNs(1,2,4,5,9,10) update the corresponding pages(A,B,A,C,B,C)

e)

Begin_LSN (9) End_LSN (4)

LSNs(9,4) corresponding pages (B,A) pages
Undo will perform on T2, T4