King Fahd University of Petroleum and Minerals

Information and Computer Science Department

ICS 424: Advanced Database Systems
Second Semester (062)

EXAM #2 (20%)

DO NOT OPEN UNTIL INSTRUCTED TO DO SO!!!!

Write clearly, precisely, and briefly!!

ID:	
Name:	

Grades		
Section	**Max**	**Scored**
A	18	
B	20	
C	26	
D	20	
E	18	
F	6	
TOTAL	**108**	

A. Questions from Chapter 15 (Query optimization) (18)

1. Assume that an EMPLOYEES table has 20000 records stored in 2000 disk blocks with blocking factor of 10 records/block and the following access paths:

   - A clustering index on SALARY, with 3 levels and the number of distinct SALARY values are 40. Assume that the distinct values are uniformly distributed.

   - A secondary index on the key attribute, SSN. Assume the index has 3 levels.

   - A secondary index on a nonkey attribute, DNO, with 2 levels. The first level index blocks are 4. Assume there are 125 distinct values of DNO and they uniformly distributed.

   Using the above assumptions calculate worst case costs (in terms of the number of blocks accessed) for each of the following queries.

   a. SELECT * FROM EMPLOYEE WHERE SSN = '1234';

   b. SELECT * FROM EMPLOYEE WHERE DNO = 5;

   c. SELECT * FROM EMPLOYEE WHERE SALARY = 1000;

B. Questions from Chapter 17 (Transaction processing)   (20)

1. Assume the following schedule S.

T_1	T_2
read(A)	
$A := A - 50$	
write(A)	
	read(A)
	$temp := A * 0.1$
	$A := A - temp$
	write(A)
read(B)	
$B := B + 50$	
write(B)	
	read(B)
	$B := B + temp$
	write(B)

a. What does conflict serializable mean?   (2)

b. Show whether S is conflict serializable or not.   (6)

    c. Show a conflict serialiazable schedule of T1 and T2 equivalent to a serial schedule T2 followed by T1. **(6)**

    d. Show a schedule of T1 and T2 which is not conflict serialiazable. Use precedence graph to prove that. **(6)**

C.

Questions from Chapter 18 (Concurrency Control)   (26)

1. Explain the relationship between two-phase locking and conflict serializable scheduling.   **(4)**

2. Show how you can run transactions T1 and T2 of the previous question using two-phase locking.   **(6)**

3. Assume 6 transactions (T1, T2, T3, T4, T5, and T6) are concurrently running and the RDBMS is using concurrency control based on time stamping. Assume T1 is older than T2, T2 is older than T3, T3 is older than T4, T4 is older than T5, and T5 is older than T6. Also assume that the youngest transaction who updated record B was T3 and the youngest transaction who read record B was T5.   **(12)**

    a. What happens if T4 tries to read record B?

    b. What happens if T4 tries to update record B?

c. What happens if T6 tries to read record B?

d. What happens if T6 tries to update record B?

**e.** What happens if T2 tries to read record B?

f. What happens if T2 tries to update record A?

4. What is the main benefit of multiversion concurrency control techniques? **(4)**

D. Questions from Chapter 19 (Recovery) (20)

1. The following figure shows a possible schedule of executing transactions just before a system crash. Assume the RDBMS uses write-ahead logging. What will the recovery manger do to recover from the crash
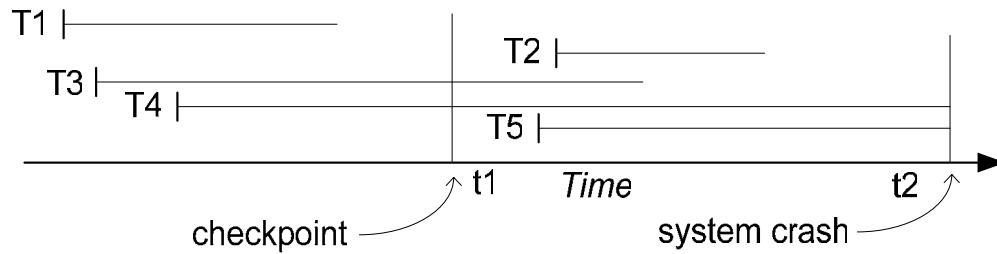
a. If the RDBMS was using **Steal/no-force** to update the data disk? **(10)**

b. If the RDBMS was using **No-steal/force** to update the data disk? **(10)**

E. **Questions from Chapter 23 (Security)**   **(18)**

1. In discretionary access control, show how the propagation of grants with GRANT OPTION can be controlled.   **(8)**

2. Using an example explain what inference control is   **(6)**

3. What is the use of Role-Based access control   **(4)**

F. **Questions from homework 1** **(6)**

1. Suppose we have a sequential (ordered) file of 200,000 records where each record is 240 bytes. Assume that B = 1200 bytes, s = 16 ms, rd = 8.3 ms, and btt = 0.8 ms. Suppose we want to make X independent random record read from the file. We could make X random block reads or we could perform one exhaustive read of the entire file looking for those X records. The question is to decide when it would be more efficient to perform one exhaustive read of the entire file than to perform X individual random reads. That is, what is the value of X when an exhaustive read of a file is more efficient than random X reads?