Sri Lanka Institute of Information Technology

B.Sc. Special Honours Degree/Diploma
in
Information Technology

Final Examination
Year 2, Semester 1 (2017)
Regular Intake

IT202 - Database Management Systems II

Duration: 3 Hours

Instructions to Candidates:

- This paper has 5 questions. Answer all questions.
- Write answers in the booklet given.
- Total marks 100.
- This paper contains 7 pages including the cover page.
- Electronic devices capable of storing and retrieving text, including calculators and mobile phones are not allowed.

Question 1 (20 Marks)

Consider the following relations in a database created for a taxi service

Vehicle (*vehicleID* : char(4), *type* : varchar(10), *fuelType*: varchar(6), *mileage*: real)

Driver (*driverID* : char(4), *name* : varchar(50), *phone* : char(10))

Customer (*customerID* : char(4), *name*: varchar(5), *phone* :char(10))

Booking (bookingID: *int, vehicleID*: char(4), *driverID:* char(4), *customerID*: char(4), *date*: datetime, *from* : varchar(15), *to*: varchar(15), *distance*: real)

Payment (bookingID: *int*, fee: *real*)

The database stores information on vehicles owned by the service in the **Vehicle table**. The table stores a unique vehicle ID (*vehicleID*), *type* ('Hatchback', 'Sedan' or 'Van'), *fuel type* ('Petrol' or 'Diesel') and the total distance the vehicle have travelled so far (*mileage*) for each vehicle owned by the service. Information of drivers working at the taxi service is stored in the **Driver table**. The table stores a unique ID (*driverID*), *name* and the *phone* number of each driver working at the service. For each customer, the service stores a customer ID (*customerID)* in the **Customer table** in the database along with the *name* of the customer and his/her *phone* number. When a customer books a taxi, the service assigns a unique ID (*bookingID*) for the booking, and stores the booking ID along with the ID of the driver assigned (driverID), ID of vehicle assigned (*vehicleID*), customer ID of the customer (*customerID*), booked date (*date*), from where the customer is taking the taxi from (*from*) and the destination the customer want to reach (*to*) in the **Booking table**. Once the customer has completed the journey, the distance of the journey is updated in kilometers in the Booking table. **Payment table** stores the booking ID and the payment to be made for each booking.

Use SQL queries to answer following questions.

i. Find the names of the customers who has booked the vehicle number 'V001' to reach 'Maharagama' (2 marks)

ii. For each driver who had driven a 'Sedan' for a total distance exceeding 5000km display the name of the driver and the total distance he had driven. (4 marks)

iii. Find the ID and type of all vehicles which have been only booked for travelling more than 100 kilometers between 1st January 2016 and 31st December 2016. (3 marks)

b) The charging scheme for the taxis booked with the taxi service is as follows: Initial charge for the taxi is Rs. 300 and the customer can travel for 5kms without an additional payment. For each km travelled after the first 5km, a payment of Rs.50 per kilometer would be charged. Create a stored procedure to calculate the payment due for a booking, given the distance the customer has travelled and the booking ID, and store the payment amount to be made along with the booking ID in the Payment table. (5 marks)

c) Create a trigger to update the mileage of the vehicle assigned and add a record to the payment table (use the stored procedure created in part b)), assuming that the Booking table is only updated when the *distance* is assigned at the completion of a booking. (6 marks)

Question 2 (20 marks)

a) Briefly explain an approach that could be used to *minimize seek time and rotational delay* in accessing a data file stored in a disk. (3 marks)

b) "*Data striping is a techniques used in RAID to improve the reliability*" Do you agree with this statement? Justify your answer. (2 marks)

c) Consider a disk with a block size of 512 bytes, 400 tracks per surface, 20 sectors per track and 15 double-sided platters. Seek time is 30 msec. Suppose that the sector size is 256 bytes is chosen. Suppose the disk drive rotates the disk head at a speed of 2400rpm (revolutions per minute). The Disk Space Manager is requested to read blocks P, Q and R respectively.

Block P is in platter 1 – track 100
Block Q is in platter 1 – track 200
Block R is in platter 1 – track 300

i. Assume ½ revolution for average rotational delay, how much of time does it take (on average) in seconds to read blocks P, Q and R? (6 marks)
ii. How much time does it take if the blocks (P, Q and R) are placed next to each other on the same track? (3 marks)

d) Explain with an aid of a diagram, the alternatives available for storing variable length fields in a record? (2 marks)

e) Consider the following buffer pool with 5 frames. Each frame has a frame number (Frame No), page id (PageID) of page, pin count (PinCount) and a dirty bit (Dirty).

Frame No: 0	Frame No: 1	Frame No: 2	Frame No: 3	Frame No: 4
PageID: 52	PageID: 15	PageID: 20	PageID: 35	PageID: 19
PinCount: 1	PinCount: 1	PinCount: 0	PinCount: 1	PinCount: 0
Dirty: 1	Dirty: 1	Dirty: 1	Dirty: 0	Dirty: 1

Assume that the following times states the last time the particular frame was accessed (on the same day):

Frame number 0 at 8:15 am
Frame number 1 at 8:17 am
Frame number 2 at 7:25 am
Frame number 3 at 7:12 am
Frame number 4 at 7:05 am

Describe the steps followed by the buffer manager when a request for the page with PageID 32 occurs if MRU (Most-recently-used) replacement policy is used. (4 marks)

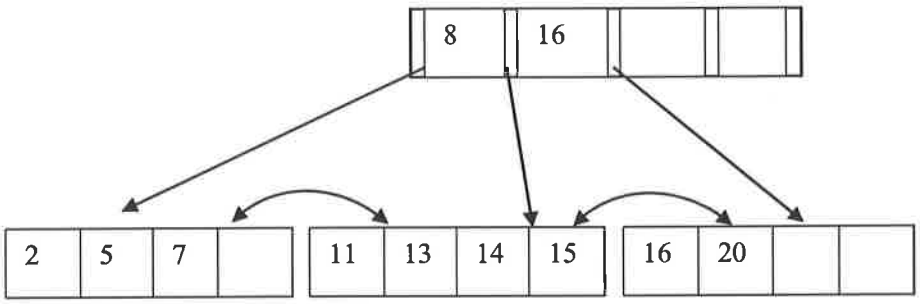
Question 3 (20 marks)

a) Briefly explain the three main file organizations (3 marks)

b) "Equality search in Hash file organization is always easy". Accept or refute the above statement providing reasons. (3 marks)

c) Briefly explain the two alternatives used in DBMS for managing the fixed length records. Provide one limitation of each alternative. (4 marks)

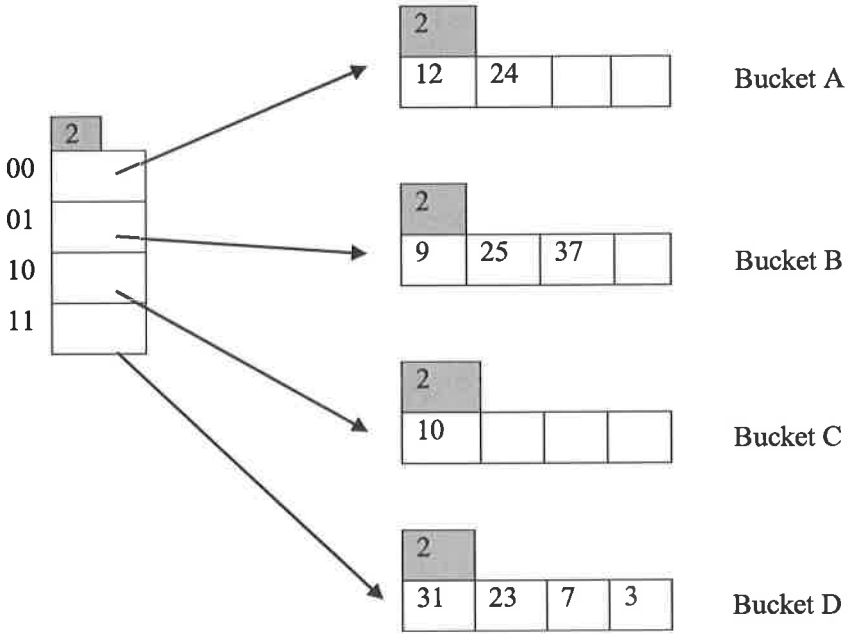
d) Consider the following B+ tree of order 2.



i. Illustrate B+ tree after inserting 10 to the tree above. (3 marks)

ii. Illustrate the B+ tree after deleting 5 and 7 from the **original tree**. (3 marks)

e) Consider the following extendible hashing index. The hashing function considers the last 2 digits of the binary representation.



Show the extendible hashed file after inserting 11 and 14. (4 marks)

Question 4 (15 marks)

a) Briefly explain how the query evaluation engine performs query optimization. (3 marks)

b) Briefly explain why it is often advantageous to do selections before joins in a query plan. (2 marks)

c) Consider the following relations

Student (sid, sname, birthdate, gpa)
Grades (sid, cid, year, semester, grade)

Student table contains 200 pages, each page containing 100 tuples. Grades table contains 1000 pages, each page containing 200 tuples. Both *Student* and *Grades* tables have clustered B+ tree indexes on sid. Assume that the indexes on *Student* and Grades have 100 and 1000 leaf pages respectively and the heights of their B+ trees are 2 and 3 respectively. Each student follows 40 courses during his enrollment period. There are 20 buffer frames available to run the query. Assume that the student name is unique.

Following query is written on the relations above

Query:
```
Select g.grade
From Student s, Grades g
Where s.sid=g.sid and s.name='Amila'
```

i. Draw query tree for the best query plan? (2 marks)
ii. Estimate the cost of the selection s.sname = 'Amila' in page I/Os. (3 marks)
iii. Estimate the cost of the join s.sid=g.sid for index nested loop join. (4 marks)
iv. Estimate the total cost for the above query. (1 mark)

Question 5 (25 marks)

a) Briefly explain the properties of a transaction (4 marks)

b) Why do DBMSs interleave actions of multiple transactions? Why is interleaving actions in the schedules desirable in serializable fashion? (3 marks)

c) Briefly explain the rules in Strict 2 Phase Locking Protocol. (3 marks)

d) Explain with an example how 'WR conflict' could be avoided using the Strict 2PL Protocol. (3 marks)

e) Briefly explain the term 'Deadlock' associates with schedules in DBMS. (2 marks)

f) Briefly explain the wound-wait algorithm used for deadlock prevention. (2 marks)

g) Consider the following part of the schedule.

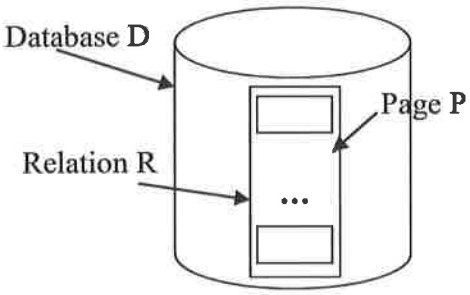
T_1	T_2	T_3	T_4
S(A)			
R(A)			
	X(B)		
	W(B)		
S(B)			
		S(C)	
		R(C)	
	X(C)		
			S(D)
			R(D)
		X(D)	
			X(A)

Assume that Transaction T_i is higher priority than transaction T_{i+1}(i.e. transaction T_1 has higher priority than T_2; T_2 has higher priority than T_3; and T_3 has higher priority than T_4).

i. Draw a Wait-For-Graph for the schedule given above. (2 marks)

ii. Draw the schedule again considering deadlock prevention using Wound-wait approach (3 marks)

h) Consider the following scenario:



Database *D* contains a relation *R*. Relation *R* contains 1000 pages. Assume that multiple granularity locking scheme is used.

Describe the locks acquired when all rows in page P are modified. (3 marks)