
UNIVERSITI SAINS MALAYSIA

First Semester Examination
2015/2016 Academic Session

December 2015/January 2016

CMT321 – Pengurusan & Kejuruteraan Pangkalan Data
[Management & Engineering of Databases]

Duration : 2 hours
[Masa : 2 jam]

INSTRUCTIONS TO CANDIDATE:

[ARAHAN KEPADA CALON:]

- Please ensure that this examination paper contains **FOUR** questions in **SEVEN** printed pages before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **EMPAT** soalan di dalam **TUJUH** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

- Answer **ALL** questions.

*[Jawab **SEMUA** soalan.]*

- You may answer the questions either in English or in Bahasa Malaysia.

[Anda dibenarkan menjawab soalan sama ada dalam bahasa Inggeris atau bahasa Malaysia.]

- In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi bahasa Inggeris hendaklah diguna pakai.]

1. (a) (i) Give the definition of a transaction.

(2/100)

- (ii) Given the following transaction and assume initial value for A and B is 100:

```
T1 : read(A);
      A := A - 50;
      write(A);
      read(B);
      B := B + 50;
      write(B);
```

Describe the **five (5)** possible states that the above transaction may go through in its execution (please use this transaction as an example to describe your answer).

(7/100)

- (b) Explain how does a deadlock occur in concurrent execution of a schedule and describe **two (2)** schemes to handle the prevention of this deadlock.

(7/100)

- (c) Determine whether the following schedule is view serializable to some of serial schedules. Explain how you derive this answer and if this schedule is view serializable then write the equivalent serial schedule.

Schedule A:

T10	T11
read (X);	read(Y);
write (X);	write(X);

(8/100)

2. (a) How do recovery using deferred update technique and recovery using immediate update technique are different in recovering a database? Explain.

(4/100)

- (b) Consider the following log entries and then use ARIES recovery algorithm to recover this schedule after the system crash:

<u>LSN</u>	<u>Log</u>
00	begin_checkpoint
10	end_checkpoint
20	update: T1 write B
30	update: T2 write A
40	T2 commit
50	update: T3 write A
	SYSTEM CRASH, RESTART

- (i) Identify the contents of transaction table and dirty page table after ANALYSIS phase has been completed.
- (ii) Identify the point in the log at which the REDO phase will start and identify the operations that need to be redone during this phase.
- (iii) Identify the operations that need to be undone during the UNDO phase.

(10/100)

- (c) Suppose a company selling computer hardware has enforced the following restriction on the access of its database using the following commands:

```
GRANT SELECT, UPDATE (model, speed, screen, price)
ON LAPTOP
TO Saleman1 WITH GRANT OPTION
```

Explain the function of these commands and identify the subject, object and privileges involved in these commands.

(6/100)

- (d) (i) State the restrictions that are imposed on database objects using the simple security property and the *- property in Bell-LaPadula model.

(3/100)

- (ii) Explain how these properties are used to enforce multilevel security for a database.

(4/100)

3. (a) Distributed database design has to consider three important issues relating to data i.e. Fragmentation, Replication and Allocation. Discuss the strategies for each of these issues in designing data for a distributed database.
- (9/100)
- (b) Suppose that you have been hired as a consultant to choose a database system for your client's application. For each of the following applications, state what type of database system (relational, OODB, object relational) you would suggest. Justify your suggestion.
- (i) A health information system that gathers a lot of digital signal data such as CAT scan and ultrasound imagery together with patient's personal information.
- (ii) The director of the university accomodation office requires a database to assist with the administration of the office.
- (iii) An educational software developer has asked you to develop a database of orchestral musical instruments. Orchestral instrument consists of four broad classes (strings, woodwinds, brass and percussion).
- (9/100)
- (c) Describe briefly how star, snowflake and starflake schema differ in data warehouse.
- (9/100)
4. (a) While working as a database analyst for a research institute, you are asked to be a part of its data warehouse project team. The group has assigned you to provide an OLAP overview before the group makes a commitment to OLAP. The group members are particularly concerned about the OLAP architecture requirements and how OLAP will fit into the existing environment. Your job is to convince them to use multidimensional OLAP (MOLAP) and to explain the main MOLAP components by using relevant architecture diagram and explanation.
- (10/100)
- (b) Explain briefly the basis of selecting data mining algorithms for data mining tools.
- (6/100)
- (c) Discuss how data mining can realize the value of a data warehouse.
- (6/100)

KERTAS SOALAN DALAM VERSI BAHASA MALAYSIA

[CMT321]

- 5 -

1. (a) (i) Beri definisi transaksi.

(2/100)

- (ii) Diberikan transaksi seperti berikut dan andaikan nilai awal A dan B adalah 100:

```
T1 : read(A);  
      A := A - 50;  
      write(A);  
      read(B);  
      B := B + 50;  
      write(B);
```

Huraikan **lima (5)** keadaan yang mungkin dilalui oleh transaksi di atas dalam pelaksanaannya (sila gunakan transaksi ini sebagai contoh untuk menghuraikan jawapan anda).

(7/100)

- (b) Terangkan bagaimana kebuntuan berlaku dalam pelaksanaan serentak suatu skedul dan jelaskan **dua (2)** skema untuk mengendali pencegahan kebuntuan ini.

(7/100)

- (c) Tentukan sama ada skedul berikut adalah berkebolehsirian pandangan kepada sesetengah skedul bersiri. Terangkan bagaimana anda memperoleh jawapan ini dan jika skedul ini adalah berkebolehsirian pandangan, sila tulis skedul bersiri yang setara.

Skedul A:

T10	T11
read (X); write (X);	read(Y); write(X);

(8/100)

2. (a) Bagaimana pemulihan yang menggunakan teknik pengemaskinian tertunda dan pemulihan yang menggunakan teknik pengemaskinian segera berbeza dalam memulihkan pangkalan data? Terangkan.

(4/100)

- (b) Pertimbangkan kemasukan log berikut dan seterusnya gunakan algoritma pemulihan ARIES untuk memulihkan skedul ini selepas kerosakan sistem berlaku:

<u>LSN</u>	<u>Log</u>
00	begin_checkpoint
10	end_checkpoint
20	update: T1 write B
30	update: T2 write A
40	T2 commit
50	update: T3 write A
	SYSTEM CRASH, RESTART

- (i) Kenal pasti kandungan jadual transaksi dan jadual halaman kotor selepas fasa ANALYSIS telah selesai.
- (ii) Kenal pasti titik dalam log di mana fasa REDO akan bermula dan kenal pasti operasi-operasi yang perlu dilakukan semula semasa fasa ini.
- (iii) Kenal pasti operasi-operasi yang perlu dibuat asal semasa fasa UNDO.

(10/100)

- (c) Andaikan sebuah syarikat menjual perkakasan komputer telah menguatkuasakan sekatan berikut ke atas akses kepada pangkalan datanya dengan menggunakan arahan berikut:

```
GRANT SELECT, UPDATE (model, speed, screen, price)
ON LAPTOP
TO Saleman1 WITH GRANT OPTION
```

Terangkan fungsi arahan-arahan ini dan kenal pasti subjek, objek dan privilej yang terlibat dalam arahan ini.

(6/100)

- (d) (i) Nyatakan sekatan yang dikenakan ke atas objek pangkalan data dengan menggunakan sifat keselamatan mudah dan sifat *- dalam model Bell-LaPadula.

(3/100)

- (ii) Terangkan bagaimana sifat-sifat ini digunakan untuk menguatkuasakan keselamatan berbilang aras untuk sebuah pangkalan data.

(4/100)

3. (a) Reka bentuk pangkalan data teragih perlu mengambil kira tiga isu penting yang berkaitan dengan data iaitu Penyerpihan, Replikasi dan Peruntukan. Bincangkan strategi-strategi untuk setiap isu ini dalam mereka bentuk data untuk pangkalan data teragih.
(9/100)
- (b) Andaikan anda telah dilantik sebagai perunding untuk memilih sistem pangkalan data bagi aplikasi pelanggan anda. Bagi setiap aplikasi berikut, nyatakan apakah jenis sistem pangkalan data (hubungan, OODB, objek hubungan) yang anda akan cadangkan. Berikan justifikasi di sebalik cadangan anda.
- (i) Satu sistem maklumat kesihatan yang mengumpulkan banyak data isyarat digital seperti imbasan CAT dan imej ultrasound bersama-sama dengan maklumat peribadi pesakit.
 - (ii) Pengarah pejabat penginapan universiti memerlukan pangkalan data untuk membantu dalam pentadbiran pejabat tersebut.
 - (iii) Pemaju perisian pendidikan telah meminta anda untuk membangunkan pangkalan data alat muzik orkestra. Alatan orkestra terdiri daripada empat kelas yang utama (bertali, tiup kayu, bras dan perkusi).
(9/100)
- (c) Huraikan secara ringkas bagaimana skema bintang, kepingan salji (*snowflake*) dan kepingan bintang (*starflake*) berbeza dalam penggudangan data.
(9/100)
4. (a) Semasa bekerja sebagai seorang penganalisis pangkalan data untuk sebuah institut penyelidikan, anda telah diminta untuk menjadi sebahagian daripada pasukan projek gudang data. Kumpulan ini telah memberi tugas kepada anda untuk memberikan gambaran keseluruhan tentang OLAP sebelum kumpulan ini membuat komitmen kepada OLAP. Ahli-ahli kumpulan amat mengambil berat tentang keperluan seni bina OLAP dan bagaimana OLAP akan disesuaikan ke dalam persekitaran sedia ada. Tugas anda adalah untuk meyakinkan mereka untuk menggunakan OLAP multi dimensi (MOLAP) dan untuk menerangkan komponen utama MOLAP dengan menggunakan gambar rajah seni bina yang berkaitan dan penjelasan.
(10/100)
- (b) Terangkan secara ringkas asas bagi pemilihan algoritma pelombongan data untuk alatan pelombongan data.
(6/100)
- (c) Bincangkan bagaimana perlombongan data boleh merealisasikan nilai sebuah gudang data.
(6/100)