

CONFIDENTIAL



UNIVERSITI TUN HUSSEIN ONN MALAYSIA

**FINAL EXAMINATION
SEMESTER I
SESSION 2017/2018**

COURSE NAME : DISTRIBUTED DATABASE
COURSE CODE : BIW 33103
PROGRAMME CODE : BIW
EXAMINATION DATE : DECEMBER 2017 / JANUARY 2018
DURATION : 3 HOURS
INSTRUCTION : ANSWER ALL QUESTIONS

TERBUKA

THIS QUESTION PAPER CONSISTS OF SIX ⁷~~6~~ PAGES

CONFIDENTIAL

SECTION A

Q1 State either **TRUE** or **FALSE** for each of the following questions.

- (a) A transaction can have multiple database access operations. (1 mark)
- (b) In distributed database, all database access operations of the transaction have been executed successfully and their effect on the database has been recorded in log. (1 mark)
- (c) Query decomposition consists of the following steps in sequence, which are normalization, analysis, redundancy elimination, and rewriting. (1 mark)
- (d) Fragmentation may cause performance degradation in terms of database access. (1 mark)
- (e) Centralized Two-Phase Locking (2PL) consists of two phases, which are, Join Phase and Disjoin Phase. (1 mark)
- (f) Distributed Two-Phase Locking (2PL) is the best concurrency control option for distributed database. (1 mark)
- (g) Site catalog is not something important in distributed database system and we can always refer to the global catalog at all time. (1 mark)
- (h) Ignoring Cartesian product in query processing will always improve the performance of the processing at both centralized and distributed environment. (1 mark)

TERBUKA

- (i) Partial replication is different with fragmentation. (1 mark)
- (j) Vertical partitioning requires *minterm predicates*. (1 mark)

SECTION B

- Q2** (a) Give **TWO (2)** differentiation of centralized database and distributed database. (6 marks)
- (b) Outline **THREE (3)** requirements to decide the allocation of databases at different sites. (3 marks)
- (c) Compare primary horizontal fragmentation and derived horizontal fragmentation using visualisation or examples. (4 marks)

Q3 Consider the following schema:

```
Suppliers(sid: integer, sname: string, address: string)
Parts(pid: integer, pname: string, color: string)
Catalogue(sid: integer, pid: integer, cost: real)
```

The key fields are underlined, and the domain of each field is listed after the field name. Therefore sid is the key for Suppliers, pid is the key for Parts, and sid and pid together form the key for Catalogue. The Catalogue relation lists the prices charged for parts by Suppliers.

- (a) State what the following queries compute.
- (i) $\pi_{sname}(\pi_{sid}((\sigma_{color='red'} Parts) \bowtie (\sigma_{cost < 100} Catalogue)) \bowtie Suppliers)$ (2 marks)

TERBUKA

(ii)

$$(\pi_{sname}((\sigma_{color='red'}Parts) \bowtie (\sigma_{cost < 100}Catalog) \bowtie Suppliers)) \cap (\pi_{sname}((\sigma_{color='green'}Parts) \bowtie (\sigma_{cost < 100}Catalog) \bowtie Suppliers))$$

(2 marks)

- (b) Write the following queries in relational algebra:
- (i) Find the name of suppliers who supply some red part. (2 marks)
- (ii) Find the `sid` of suppliers who supply some red or green part. (2 marks)

Q4 A Distributed Database Management System (DDBMS) must ensure that no two sites create a database object with the same name. One solution to this problem is to create a central name server.

- (a) State **TWO (2)** disadvantages with this approach (2 marks)
- (b) Propose an alternative approach that overcomes these disadvantages. (3 marks)

Q5 (a) Describe **TWO (2)** levels of transactions defined in Distributed Relational Database Architecture (DRDA) based on International Business Machines (IBM). (4 marks)

- (b) Give example for each of the level described in **Q5(a)**. (4 marks)

TERBUKA

Q6 Consider the database:

```
PRODUCTION (SerialNumber, PartType, Model, Quantity,
            Machine)
PICKUP (SerialNumber, Lot, Client, SalesPerson, Amount)
CLIENT (Name, City, Address)
SALESPERSON (Name, City, Address)
```

Assume four production centres located in Dublin, San José, Zurich and Taiwan and three sales points, located in San José, Zurich and Taiwan. Each production center is responsible for one type of part; the parts are Central Processing Unit (CPU), Keyboard, Screen and Cable. Suppose also three sales points, located in San José, Zurich and Taiwan. Suppose that the sales are distributed by geographic location; thus, Zurich clients are served only by salespeople in Zurich (assume that the sales point in Zurich also serves Dublin). Assume that each geographic area has its own database (that is, databases are available in Dublin, San José, Zurich and Taiwan).

- (a) Design the horizontal fragmentation of the tables PRODUCTION, PICKUP, CLIENT and SALESPERSON. (8 marks)
- (b) Write the following queries on transparency levels of fragmentation, allocation and language:
- (i) Determine the available quantity of the product 77y6878. (5 marks)
 - (ii) Determine the clients who have bought a lot from the retailer Wong, who has an office in Taiwan. (5 marks)
 - (iii) Determine the machines used for the production of the parts type Keyboard sold to the client Brown. (5 marks)
 - (iv) Modify the address of the retailer Brown, who is moving from '27 Church St.' in Dublin to '43 Park Hoi St.' in Taiwan. (5 marks)

TERBUKA

- Q7** (a) Justify **TWO (2)** the importance of concurrency control. (4 marks)
- (b) Illustrate distributed Two-Phase (2P) Commit Protocols using a diagram. (5 marks)
- Q8** Chantek.com is one of international online e-commerce system which mainly runs from the United States (the headquarters). However, it also has several distributed sites to support the system at different locations. Using the skills that you have learned in the tutorials,
- (a) Construct your own estimations of the system. One of the points that need to be estimated is the possible tables that should exist in the database. (4 marks)
- (b) Construct **TWO (2)** possible tables that might exist in the database which you already mentioned in **Q8(a)**. (4 marks)
- (c) Propose at least **ONE (1)** vertical fragmentation and **ONE (1)** horizontal fragmentation from the table in **Q8(b)**, and then justify using your own words how the fragments will benefit Chantek.com. You need to show proper steps for the horizontal fragmentation. (7 marks)
- (d) Let the following diagram in **Figure Q8(d)** be a network of your distributed database system where r_n is the fragmented relations. Propose **TWO (2)** execution strategies for query q , where

TERBUKA

$$q = r_2 \otimes r_3 \otimes r_6 \text{ at node 1}$$

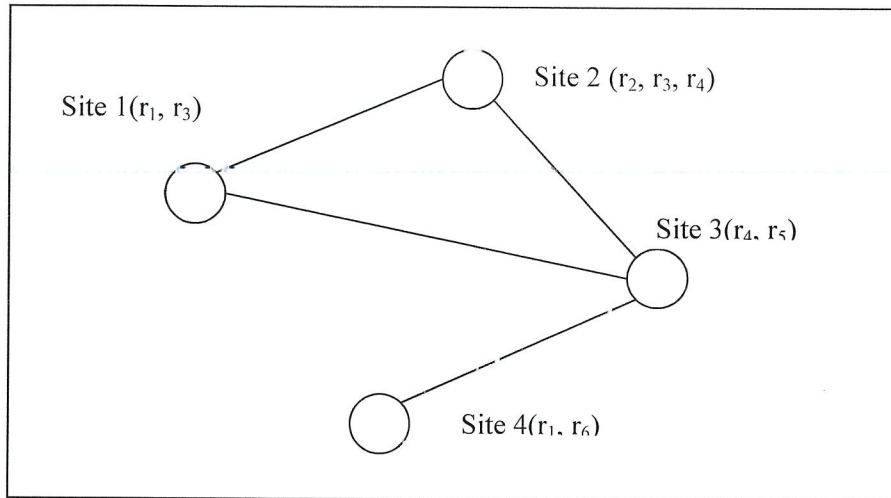


Figure Q8(d)

(4 marks)

– END OF QUESTION –

TERBUKA

