

BCS THE CHARTERED INSTITUTE FOR IT

BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 6 Professional Graduate Diploma in IT

ADVANCED DATABASE MANAGEMENT SYSTEMS

Monday 4th October 2021 – Afternoon

Answer **any** THREE questions out of FIVE. All questions carry equal marks.

Time: THREE hours

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

The marks given in brackets are **indicative** of the weight given to each part of the question.

Calculators are NOT allowed in this examination.

Section A
Answer Section A questions in Answer Book A

A1.

- a) Database security usually comes under scrutiny when things go wrong, with many high profile attacks reported in the daily press.
- i) Describe **physical security**, highlighting what typically happens in a breach along with measures that can mitigate against the breach. **(5 marks)**
 - ii) Name **TWO** common aspects that **logical security** addresses. **(2 marks)**
- b) Explain the following concepts in the context of database security:
- i) Data confidentiality; **(3 marks)**
 - ii) Data integrity. **(3 marks)**
- c) Discuss the following security risks and describe the security controls that can be put in place in a database context to reduce the risk of them occurring:
- i) SQL injection; **(4 marks)**
 - ii) Data entry mistakes. **(4 marks)**
- d) An SME (Small or Medium Enterprise) hosts its database server and web applications server on the same physical machine. Briefly discuss the advantages and disadvantages of this approach when compared to having the servers on separate machines. **(4 marks)**

A2.

- a) Database design generally aims at high levels of normalisation (such as 3rd Normal Form or Boyce Codd Normal Form). However, industrial practice sees **denormalised databases** used as well. Discuss the advantages and disadvantages of using denormalised databases.

(5 marks)

- b) Consider the following tables:

```
CD (CDID, title, year)
Artist (AID, name)
Recorded (CDID, AID)
```

And the following query:

```
SELECT CD.title
FROM CD
      JOIN Recorded
      ON CD.CDID = Recorded.CDID
      JOIN Artist
      ON Recorded.AID = Artist.AID
WHERE Artist.name = 'Queen'
      AND CD.year = 1975;
```

- i) Draw a query tree that corresponds to the most efficient way of processing this query.

(12 marks)

- ii) Assume there is a B-Tree index on the column "title" of the table "CD". For each of the following queries, explain how this index could be used when executing each query:

X. SELECT * FROM CD WHERE title = 'Let It Be'

(2 marks)

Y. SELECT COUNT(title) FROM CD

(3 marks)

Z. SELECT * FROM CD ORDER BY title

(3 marks)

[Turn Over]

A3.

a) Database systems need to operate reliably both in the context of failure and when simultaneously accessed by various users. In this context consider the following questions:

i) Using your own example and simple diagrams explain the concept of guaranteed **data consistency** via **ACID** principles. **(5 marks)**

ii) Using your own example, explain the concepts of **ROLLBACK** and **COMMIT**. **(5 marks)**

iii) Using a diagram, explain the mechanism of the 2PC (2 Phase Commit Protocol). **(5 marks)**

b) Your database uses REDO/UNDO with Checkpointing.

i) Briefly describe what this means for your database in operation. **(4 marks)**

ii) Consider the following log and describe what needs to be done for each of the transactions T1, T2 and T3 in the case of a server crash just after LogID = 11 is created. **(6 marks)**

LogID	TxID	Time	Operation	Object
1	T1	8:15	START	
2	T1	8:16	UPDATE	Account 122
3	T2	8:16	START	
4	T2	8:19	INSERT	Account 145
5	T2	8:20	DELETE	Account 153
6	T2	8:21	UPDATE	Customer 89
7	T3	8:21	START	
8	T1	8:21	COMMIT	
9		8:23	CHECKPOINT	T2, T3
10	T2	8:24	COMMIT	
11	T3	8:25	INSERT	Customer 46

Section B
Answer Section B questions in Answer Book B

B4.

- a) Fundamentally, a true distributed database system consists of a collection of physically distributed databases connected via a network. The result should be that, to a user, the system behaves exactly as if it were one single database.

It should have the following properties:

- i) Location Transparency.
- ii) Distributed Query Processing (Each query is able to efficiently access data from multiple sites).
- iii) Distributed Transaction Management.

Describe **EACH** property and explain the practical problems that would arise if they were enabled and implemented.

(12 marks)

- b) Given the problems that arise from part a), outline the pros and cons of the following alternative approaches to a true distributed database approach:

Approach 1: strongly consistent data replicas of tables across all sites
Approach 2: fragmentation of tables of data across sites

(6 marks)

- c) Consider the following scenario:

Scenario:

XYZ is a business that has a large portfolio of products that it sells over the internet. Most customers order products without any problem but about 25% of customers return a product and get a refund because it was unsuitable. To alleviate the high cost of handling returns and to improve customer experience 20 large shops (or retail outlets) have been opened in different cities/regions allowing customers to purchase products directly from a shop rather than over the internet. Currently a centralised database holds the following tables:

```
Customers (CustomerID, Name, Address, Phone, ...)  
Products (ProductID, Productname, Description, UnitPrice)  
Orders (OrderID, date)  
OrderLine (OrderID, ProductID, QuantityOrdered)  
CustomerOrders (CustomerID, OrderID, ShopID)
```

Describe a Fragmentation strategy to data distribution that could be applied to meet the needs of XYZ.

(7 marks)

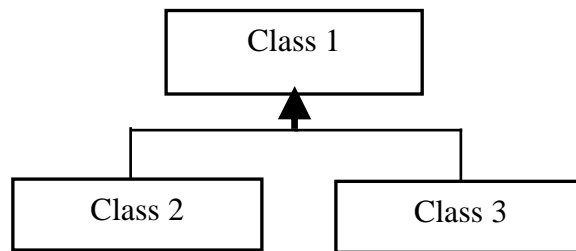
[Turn Over]

B5.

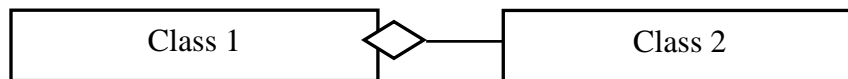
- a) UML Class models are often used to extend the Entity Relationship model to encompass Object Oriented (OO) modelling concepts and constructs.

Using example Class names, describe the type of OO modelling construct that is represented by each of the following models.

Model 1



Model 2



Model 3



(9 marks)

- b) Explain with reference to the following headings what is meant by the term “Impedance mismatch”, when describing the disparity between OO concepts and Relational model concepts.

- i) Identity (of objects);
- ii) Associations (between objects);
- iii) Data access and navigation.

(9 marks)

- c) Object Relational Mapping (ORM) is a software component, part of an N-tier client server architecture that is designed to reconcile the impedance mismatch between Relational databases and object oriented programming languages such as Java.

Explain briefly how ORM addresses the problem of impedance mismatch.

(7 marks)

End of Examination