DUBLIN INSTITUTE OF TECHNOLOGY

DT211C/4 BSc. (Honours) Degree in Computer Science (Infrastructure)

DT228/4 BSc. (Honours) Degree in Computer Science

DT282/4 BSc. (Honours) Degree in Computer Science (International)

DT8900/1 International Pre Masters for MSc in Computing

WINTER EXAMINATIONS 2017/2018

ADVANCED DATABASES [CMPU4003]

DR. PIERPAOLO DONDIO
DR. DEIRDRE LILLIS
MR. ALAN FAHEY – DT211C
DR. MARTIN CRANE – DT228/DT282

WEDNESDAY 17TH JANUARY 2.00 P.M. – 4.00 P.M.

TWO HOURS

INSTRUCTIONS TO CANDIDATES:

ANSWER **FOUR** QUESTIONS OUT OF **FIVE**.

ALL QUESTIONS CARRY EQUAL MARKS.

Question 1. Raid and Indexes [25 marks in total]

a. Suppose you have 10 disks, each of them with a capacity of 100GB. 8 are fast, with a transfer rate of 1000 Mbits/sec and 2 are slow, with a transfer rate of 500Mbits/sec. Your DB has a stable size of 750 GB.

(i) Show how to implement different levels of RAID (0,1,3,5,10,0+1) and for the ones that can be implemented compute:

[4]

(ii) The storage efficiency (disk space actually used for data / total disk space used)

[4]

(iii)The average read access time in an ideal situation (i.e. how many Mbits per second?)

[4]

You are not requested to use all the disks, but only the one you need to accommodate your database of 750GB

[4 marks each, 12 marks]

b. Describe the concept of primary, secondary and multi-level index providing an example for each of them

[6 marks]

c. Describe the concept of dynamic hashing, providing an example. Discuss the advantages with respect to a static hashing scheme.

[7 marks]

Question 2. **[25 marks in total]**

a.

(i) What is a Bitmap Index in Oracle? When is it better to use it? When is it not?

[3 marks]

(ii) Suppose you have a million records and a field "*weekday*" that can take only seven values (one for each day of the week). How big will the bitmap index be?

[4 marks]

b. Explain the CAP theorem for distributed databases, providing examples of one system that drops Consistency, a system that drops Availability and a system that drops Partition-tolerance

[7 marks]

c. Insert the following index values in a (2,3) b-tree:

28, 61, 69, 30, 33, 55, 11, 78, 89, 99

[11 marks]

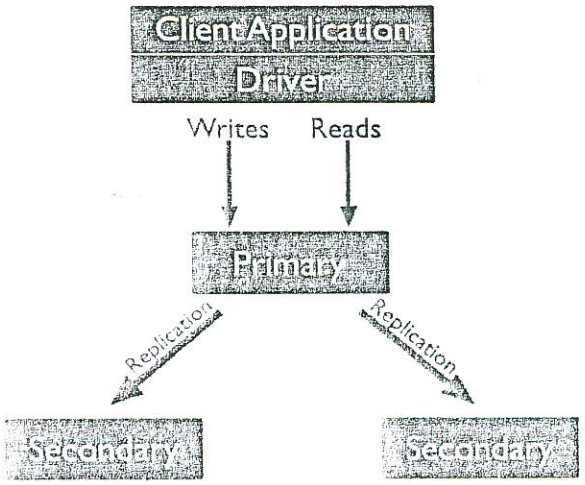
Question 3. **[25 marks in total]**

a. Describe the strengths of a Dimensional Model, including the advantages over an ER model

[9 marks]

b. Referring to the following figure, explain the concept of *replica set* in MongoDB. Discuss the functioning of a replica set under normal conditions and during a failover.
What is the difference between *sharding* and *replication* in MongoDB?

[7 marks in total]



c. Describe the ACID properties of a relational database. [3]
Describe the BASE properties of a NOSQL database. [3]
Discuss the difference between the two approaches and for which application the BASE approach is more suitable [3]

[9 marks in total, 3 for each question]

Question 4. **[25 marks in total]**

a. The European Union requires designing of a data warehouse to record university students. The following ER diagram is available (primary keys are underlined, FK = foreign keys). For each University, the database contains information about the courses offered by each Department and the lecturers delivering each course. The database also contains the list of students and the marks they achieved in each course. Course date has the format DD/MM/YYYY.

LECTURER (lCode, LectureName, Salary, YearsofService, Nationality)
COURSES (cCode, CourseDescription, CourseDate)
STUDENT (sCode, Name, Address, Phone, Gender, Age)
MARKS (sCode (FK), cCode (FK), Date, Mark)
DEPARTMENT (DepartmentID, DName, Address, UCode (FK))
REGION (RegionID, RegionName, Country)
UNIVERSITY (UCode, UniversityName)

You are require to:

(i) Produce a star schema for the above ER diagram. The diagram should support queries about students' marks in each course, at different periods of time (down to weekly reports) at each university/department, including information about lecturers and geographical region. Justify your choices (grain, facts and dimensions)

[15 marks]

(ii) Write an SQL query to get the average marks for each department in each University in 2009

[5 marks]

(iii) Write an SQL query to get, for each course, the total number of times an exam course was attempted and failed for each year and country.

[5 marks]

Question 5. **[25 marks in total]**

a. Describe the concept of write concern in MongoDB. Discuss the difference between the following 4 write concern levels:
   (i) Unacknowledged
   (ii) Acknowledged
   (iii) Journaled
   (iv) Replica Acknowledged

[8 marks]

b. Explain the following dimensional model concepts:
   (i) Surrogate Key
   (ii) Business Key
   (iii) Staging Area
   (iv) Grain

[4 marks]

c. A database contains information about bank transactions between accounts. The following not normalized table displays the available information. A Customer might have more than one account, each account belongs to exactly one bank and an account has exactly one owner.

FIELD	DESCRIPTION
TransactionID	Unique ID of the transaction
Tr_Date	Date of the transaction
Amount	Amount of money transferred
Acc_Num_From	Account number of the sender
Acc_Num_From_Type	Type of the account of the sender (basic, premium or saving)
Cust_ID_From	Unique ID of the customer sending money
Cust_Name_From	Name of the customer sending money
Cust_Age_From	Age of the customer sending money
Cust_From_BankID	Unique ID of the bank sending money
Cust_From_BankName	Name of the bank sending money
Cust_From_BankAddress	Address of the bank r sending money
Acc_Num_TO	Account number of the receiver
Acc_Num_TO_Type	Type of the account of the receiver (basic, premium or saving)
Cust_ID_TO	Unique ID of the customer receiving money
Cust_Name_TO	Name of the customer receiving money
Cust_Age_TO	Age of the customer receiving money
Cust_TO_BankID	Unique ID of the bank receiving money
Cust_TO_BankName	Name of the bank receiving money
Cust_TO_BankAddress	Address of the bank receiving money
Status	Status of the transaction (completed or not)

(i) Produce a normalized relational model to store the above information. For each table, clearly show its fields, primary key and foreign key(s).

[7 marks]

(ii) Show how the above information can be stored in a Graph Database. Show the nodes, relationships and their attributes.

[6 marks]