Programme Code: TU856/TU857/TU858
Module Code: CMPU 4003
CRN: *22415, 30088, 31081*

TECHNOLOGICAL UNIVERSITY DUBLIN

City Campus Grangegorman

TU856 – BSc (Hons) in Computer Science
TU857 – BSc (Hons) in Computer Science (Infrastructure)
TU858 – BSc (Hons) in Computer Science (International)

Year 4

SEMESTER 1 EXAMINATIONS 2023/24

CMPU 4003 Advanced Databases

Internal Examiner(s):
Dr. Deirdre Lawless
Dr. Paul Doyle

External Examiner(s):
Ms. Sanita Tifentale
Ms. Caroline McEnroy

Instructions To Candidates: Answer any **THREE (3)** Questions
All questions carry equal marks (33 marks)
1 additional mark will given to each student.

Exam Duration: 2 hours

1. **(a)** In your opinion, can a relational database be replaced by a NoSQL database in all circumstances?

In your answer you should identify at least **TWO (2)** situations which support your opinion.

(5 marks)

(b) When migrating a relational database to a document NoSQL database the decision to create a single combined document collection or separate document collections depends on several considerations.

(i) Other than the particular use case, requirements, and the structure of the data, briefly discuss **FOUR (4)** of these considerations and how they could influence such a decision.

(4 x 2 marks)

(ii) Identify and briefly discuss **TWO (2)** challenges that need to be addressed in the Extract Transform and Load (ETL) process when migrating a relational database to a single document collection.

(2 x 2 marks)

Question 1 continues on the next page …

1. (c) Suppose that you are tasked with implementing a NoSQL document database for each of the **FOUR (4)** applications listed below.

For each application compare and contrast the suitability of both approaches, i.e. a single document collection or separate document collections, before making a recommendation of which approach to adopt.

In your answer you should include at least one advantage and one potential challenge for each approach for each application.

(4 x 4 mark)

a. A Content Management System (CMS) where users can create and manage various types of content, such as articles, blog posts, images, videos, and comments.

b. A chat application for students used during class time to interact with peers and lecturers. Messages need to be in order and those involved in a chat need to see all messages.

c. An Enterprise Resource Planning (ERP) system to manage and integrate various business processes and data across an organization, including finance, human resources, inventory, procurement, and sales.

d. A marketing application storing data from social networks and conducting sentiment analysis on this data to explore the impact of marketing campaigns involving TV advertising during prime time. Analysis is differentiated between weekday and weekend sentiment.

2. You are tasked with designing database replication for a large e-commerce website which operates in multiple geographic regions and experiences high traffic. Data consistency and availability are critical for the website.

(a) **(i)** Compare *master-slave* replication, *master-master* replication, and *masterless* replication in the context of the e-commerce website.

In your answer you should clearly explain each approach and briefly discuss **ONE (1)** advantage and **ONE (1)** challenge of the approach.

(3 x 4 marks)

(ii) Recommend one of the three approaches you discussed in part (a) (i) for use with the e-commerce website. Justify your answer.

(3 marks)

(b) Recommend **THREE (3)** mechanisms that can be employed to ensure that data needed for the e-commerce website is up to date and consistent across all replicas.

In your answer you should clearly explain each mechanism and justify your choice of these mechanisms for the e-commerce website.

(3 x 3 marks)

(c) Recommend **THREE (3)** mechanisms that can be employed to minimize *downtime* in case of a database server failure for the e-commerce website.

In your answer you should clearly explain each mechanism and justify your choice of these mechanism for the e-commerce website.

(3 x 3 marks)

3. (a) Suppose you trying to improve query performance.

(i) Identify and briefly discuss **THREE (3)** situations where you would consider using multiple indexes on a table/document. Justify your answer.

(3 x 3 marks)

(ii) Identify and briefly discuss **TWO (2)** situations where you would not consider using multiple indexes on a table/document. Justify your answer.

(2 x 4 marks)

(b) Suppose you are implementing the blog post feature of a Content Management System (CMS) and are designing a NoSQL database. Each post has an author name, title, content, and timestamp. Each post can be commented on by multiple users, multiple times.

(i) Consider a NoSQL document database with which you are familiar. Other than using indexes and partitioning, explain **TWO (2)** approaches which could be used to improve performance providing an example of how you would use each approach.

(2 x 4 marks)

(i) Consider a NoSQL wide column database with which you are familiar. Other than using indexes and partitioning, explain **TWO (2)** approaches which could be used to improve performance providing an example of how you would use each approach.

(2 x 4 marks)

4. (a) (i) Briefly discuss how a relational database, a NoSQL document database, and a wide column NoSQL database with which you are familiar addresses scalability.

(3 x 2 marks)

(ii) Briefly discuss how a relational database, a NoSQL document database, and a wide column NoSQL database with which you are familiar addresses consistency.

(3 x 2 marks)

(b) You are tasked with designing a social media messaging application to be used on web and mobile platforms. The application will provide a platform for users to post and share short messages, multimedia content, and engage with other users. In particular, users can create profiles, follow others, and stay updated with real-time content.

(i) For each database discussed in your answer to part (a), in terms of the approaches to scalability discussed, identify **TWO (2)** advantages if that database was adopted for the social media application.

(3 x 2 marks)

(ii) For each database discussed in your answer to part (a), in terms of the approaches to consistency discussed, identify **TWO (2)** challenges that would need to be overcome if that database was adopted for the social media application.

(3 x 2 marks)

Question four continues on the next page…

4. (c) You are tasked with designing and implementing a Content Management System (CMS) where users can create and manage various types of content, such as articles, blog posts, images, videos, and comments. Suppose the intention is for all data to be stored and managed by the organisation for which the CMS was built.

Briefly compare the suitability of a data warehouse, a data lake and a relational database to support the development of the CMS if content performance metrics such as page views, click-through rates are the main analysis required.

In your answer you should consider data structure, data consistency and access pattern requirements.

(3 x 3 marks)