Module: Database Administration 381

Module name:	Database Administration 381
Code:	DBA381
NQF level:	7
Type:	Fundamental – Bachelor of Computing (Data Science stream)
Contact Time:	38 hours
Structured time:	6 hours
Self-Directed Time	46 hours
Notional hours:	90 hours
Credits:	9
Prerequisites:	DBD381

Purpose

This course introduces the student to the principles and practice of Administering Distributed database systems. Further understanding will be gained on how these distributed databases are logically interrelated and distributed over computer networks. In addition this module will discuss the fundamental principles of query processing, concurrency control and distributed reliability protocols.

Outcomes

Upon successful completion of this module, the student will be able to:

- Demonstrate integrated knowledge about relational database management systems, including an understanding of and the ability to apply and evaluate the key terms, concepts, facts, principles, rules and theories and how that knowledge relates to other fields, disciplines or practices.
- Demonstrate an understanding of a range of methods of enquiry in distributed transaction management, concurrency control and crash recovery components and their suitability to specific investigations within a practice.
- Identify, analyse, evaluate, and critically reflect on techniques related to the management of distributed storage structures addressing such complex problems, applying evidence-based solutions and theory-driven arguments.
- The ability to take decisions and act ethically and professionally when working with distributed databases and security concerns surrounding the information found within. Demonstrate the ability to justify those decisions and actions drawing on appropriate ethical values and approaches within a distributed database environment.
- Understand and manage processes in the ever growing world of distributed databases, recognising database performance is context and system bound, and does not occur in isolation.
- Demonstrate an understanding of knowledge surrounding concurrency control within distributed database systems and the ability to evaluate types of concurrency control mechanisms.

Assessment

Assessment is performed using a variety of instruments:

- Continuous evaluation of theoretical work through assignment, formative and a summative test.
- Continuous evaluation of project work, whereby the student must do capacity planning, installation, configuration, database design, migration, performance monitoring, security, troubleshooting, as well as backup and data recovery in a distributed Database environment.
- Final assessment through a written examination.
- The assignments or projects collectively will count 30% of your class mark.
- All tests will collectively account for 70% of your class mark.
- Your class mark contributes 30% towards your final mark for the subject, while the final assessment accounts for 70% of your final mark.

Teaching and Learning

Learning materials

Prescribed Material

- 🕮 Özsu, M.T., Valduriez, P. (2011), *Principles of Distributed Database Systems, Third Edition*, Springer-Verlag New York. [ISBN 978-1-4419-8833-1]
- 🕮 Elmasri, R., Navathe, S. (2010). *Fundamentals of Database Systems.* Addison Wesley Longman, Inc. [ISBN: 978-0136086208]
- 🕮 Mark L. Gillenson. (2004). *Fundamentals of Database Management Systems.* Wiley E -Books.

Learning activities

The teaching is a combination between presentation of theoretical concepts and exercises and discussions. It is practically oriented, with mandatory projects which must be completed during the course.

Notional learning hours

Activity	Units	Contact Time	Structured Time	Self-Directed Time
Lecture		27.0		14.0
Formative feedback		6.0		
Project	2	5.0		12.0
Assignment	1			3.0
Test	2		4.0	8.0
Exam	1		2.0	9.0
		38.0	**6.0**	**46.0**

Syllabus

- PART 1: QUERY OPTIMIZATION

- DDBMS –Relational Algebra for Query Optimization
  - Query Optimization Issues in DDBMS
  - Query Processing
  - Relational Algebra
  - Translating SQL Queries into Relational Algebra
  - Computation of Relational Algebra Operators
  - Computation of Selection
  - Computation of Joins
- DDBMS –Query Optimization in Centralized Systems
  - Query Parsing and Translation
  - Approaches to Query Optimization
  - DDBMS –Query Optimization in Distributed Systems
  - Distributed Query Processing Architecture
  - Mapping Global Queries into Local Queries
  - Distributed Query Optimization

- PART 2: CONCURRENCY CONTROL
  - DDBMS –Transaction Processing Systems
    - Transactions
    - Transaction Operations
    - Transaction States
    - Desirable Properties of Transactions
    - Schedules and Conflicts.
    - Serializability.
  - DDBMS – Controlling Concurrency
    - Locking Based Concurrency Control Protocols
    - Timestamp Concurrency Control Algorithms
    - Optimistic Concurrency Control Algorithm
    - Concurrency Control in Distributed Systems
  - DDBMS – Deadlock Handling
    - Deadlocks
    - Deadlock Handling in Centralized Systems
    - Deadlock Handling in Distributed Systems

- PART 3: FAILURE AND RECOVERY
  - DDBMS –Replication Control
    - Synchronous Replication Control
    - Asynchronous Replication Control
    - Replication Control Algorithms
  - DDBMS –Failure & Commit
    - Soft Failure
    - Hard Failure
    - Network Failure
    - Commit Protocols
    - Transaction Log
  - DDBMS–Database Recovery
    - Recovery from Power Failure
    - Recovery from Disk Failure