NEW YORK CITY COLLEGE OF TECHNOLOGY
COMPUTER SYSTEMS TECHNOLOGY DEPARTMENT
COURSE OUTLINE
CST 3604 – Quality Database Implementation

COURSE DESCRIPTION

This course is the continuation of the course "Database Design" (CST3504). The course concentrates on: the physical design and implementation of databases; functionality of Database Management Systems (DBMS) in support of concurrent, secure, well-performing, reliable, scalable database solutions. Also discussed are: special database architectures and topics – data warehouses and business intelligence, and semi-structured stores (XML and Cloud). Database and data warehouse concepts are illustrated on Oracle RDBMS, XML and Cloud stores are illustrated on various open-source products.

4 Class Hours, 3 Credits

LEARNING OUTCOMES

After finishing the course, the students should be able to:

- Understand the goals of physical data design and implementation of a database. Design the physical model of a database in Oracle DBMS, including a distributed solution.
- Understand the basic features of DBMSs for support of query processing, transaction management, concurrent access, security, reliability
- Participate in resolving the goals of physical design of a database and understand the approaches to implementing user requirements on performance, security, consistency, and reliability with the help of a particular DBMS.
- Implement basic support of database security, performance improvement, reliability, and consistent transactions.
- Understand the purpose and architecture of special databases: data warehouses, semi-structured, Cloud
- Design and implement simple data warehouses and semi-structured stores

ASSESSMENT CRITERIA

Students will be assessed in exams, homework, case assignments, and through class participation. The major areas include:

- The basic understanding of the physical data model and features of a DBMS for its support.
- Implementing, supporting and maintaining a database (in Oracle).
- Understanding of the basic data warehouse design
- Understanding of the basics of semi-structured stores

General Education Outcomes

- **SKILLS/Inquiry/Analysis:** Students will employ scientific reasoning and logical thinking.
- **SKILLS/Communication:** Students will communicate in diverse settings and groups, using written (both reading and writing), oral (both speaking and listening), and visual means
- **VALUES, ETHICS, RELATIONSHIPS / Professional/Personal Development:** Students will work with teams, including those of diverse composition. Build consensus. Respect and use creativity.

PREREQUISITES

Completion of CST3504 with the grade C or higher.

REQUIRED TEXTBOOKS

1. *Essential Aspects of Physical Design and Implementation of Relational Databases*. T. Malyuta, A. Satyanarayana. Amazon, 2017.
2. *Database Systems: Introduction to Databases and Data Warehouses, N. Jukić, S Vrbsky, S. Nestorov,* Prospect Press, 2016, https://prospectpressvt.com/titles/jukic-database-systems/.
3. *A Guide to SQL, P. Pratt,* Course Technology, 2010.

Additional Resources

1. Oracle 11g Documentation, http://www.oracle.com/pls/db111/homepage. The books:
    a. Application Development – Concepts http://docs.oracle.com/cd/B28359_01/server.111/b28318/toc.htm
    b. Application Development – Security Guide http://docs.oracle.com/cd/B28359_01/network.111/b28531/toc.htm
    c. Database Administrator's Guide – Advance Security Administration Guide http://docs.oracle.com/cd/B28359_01/network.111/b28530/toc.htm
    d. Database Administrator's Guide – Performance Tuning Guide http://docs.oracle.com/cd/B28359_01/server.111/b28274/toc.htm
    e. Database Administrator's Guide – Distributed Database Concepts http://docs.oracle.com/cd/B28359_01/server.111/b28310/ds_concepts.htm
2. Oracle Database 11*g* Express Edition (free Oracle small footprint database) http://www.oracle.com/technetwork/database/database-technologies/express-edition/overview/index.html
3. Oracle Lite SQL (easy online access to Oracle database to test scripts and queries) https://livesql.oracle.com/apex/livesql/file/index.html

Academic Integrity Policy :

Students and all others who work with information, ideas, texts, images, music, inventions, and other intellectual property owe their audience and sources accuracy and honesty in using, crediting, and citing sources. As a community of intellectual and professional workers, the College recognizes its responsibility for providing instruction in information literacy and academic integrity, offering models of good practice, and responding vigilantly and appropriately to infractions of academic integrity. Accordingly, academic dishonesty is prohibited in The City University of New York and at New York City College of Technology and is punishable by penalties, including failing grades, suspension, and expulsion. The complete text of the College policy on Academic Integrity may be found in the catalog.

Tests:

Final Exam	30%
Midterm Exam	30%
Tests & Quizzes	15%
Case Assignments	15%
Homework	10%

Case Assignments

Case assignments include physical design of the centralized and distributed databases for a particular case (for this purpose the cases from Appendix 1 of [2] can be used). Design of the distributed database must include primary and derived horizontal fragmentation, vertical fragmentation, allocation and replication of fragments, and semantic control. The distribution of the database must be transparent to users.

On the local databases of the distributed solution and on the global distributed database the students must implement the basic security measures, define queries for the users' requests and suggest the measures for improving performance of these queries, and design a business transaction ensuring its consistency.

The assignments must be implemented in Oracle DBMS.

Grading Policy:

You cannot get a passing grade unless all case assignments are completed.
The professor preserves the right to ask you to defend any of your case assignments or tests.
Late submissions: next day after the deadline – 80%, before the next class after the deadline – 60%, after the discussion in class – 0%.

COURSE OUTLINE

Chapters of [1] are referenced directly; [2] is mentioned for its referenced chapters; for the chapters from the Oracle Documentation, the book of the documentation is referenced by the letter under which it is mentioned in the list of additional resources..

Week	*Subject*	*Source*
1.	The physical data model Goals of physical modeling Using features of the DBMS for physical data design Place of the physical design in the database life cycle Implementation of the database Review of the process of design of relational databases Operations of relational algebra	Chapter 1 Appendix 2
2.	The physical data model Tables. Columns' data types. Constraints. Indices Data storage organization Heap and organized data storage	Chapter 2 3a – Chapter 26 3a – Chapter 3
3.	The physical data model Distributed storage of data. Benefits and problems of design and implementation of distributed databases. Transparency of the physical model. 3-tier database architecture The physical data model in Oracle. Oracle distributed database *Practicing building physical data models in Oracle*	Chapter 2
4.	Distributed database design Primary and derived horizontal fragmentation Vertical fragmentation Allocation and replication	Chapter 3 3e
5.	Distributed database design Semantic control *Practicing distributed data design*	Chapter 3 3e
6.	Security Authentication and authorization. Users and schemas Views and stored procedures Security in distributed databases	Chapter 4 3a – Chapter 20 3b 3c
7.	Security in Oracle Security in centralized databases Security in distributed databases *Practicing implementing various security requirements in centralized and distributed databases*	Chapter 4 3a – Chapter 20 3c
8.	Query processing Overview of query processing Query decomposition Query optimization Factors that influence performance : data storage, indexes, clusters, database environment	Chapter 5
9.	Query processing in distributed databases Query decomposition and data localization Optimization of distributed queries Query processing and performance in Oracle. Types of optimization Statistics Hints *Practicing tuning performance in Oracle*	Chapter 5 3d
10.	Transaction management Basics of transaction management Concurrency control Serializability and recoverability Locking	Chapter 6 3a – Chapters 4, 13

	Timestamping Multiversionsing	
11.	Transaction management in distributed database Transaction management in Oracle Locking, multiversioning, different isolation levels, and additional tools *Practicing building transactions and analyzing results of concurrent execution of different simple transactions*	Chapter 6
12.	Transaction management in Oracle Examples of concurrent transactions for different isolation levels Recovery in centralized and distributed databases Failures and fault tolerance Transactions and Recovery	Chapters 6, 7 3a – Chapter 15
13.	Data Warehouses Differences between OLTP and OLAP databases Dimensional model *Practicing building a simple data warehouse*	[2] Chapters 7, 8. Handouts
14.	Semi-structured data stores Cloud store solutions, NoSQL databases Data on the Web, RDF	[2] Appendices I, J. Handouts
15.	Review Final test	