CSE 4331/5331 DBMS Models and Implementation Techniques
Department of Computer Science and Engineering
The University of Texas at Arlington

Offering: Fall 2014
Time: Tuesday/Thursday 12:30 Pm to 1:50 Pm
Place: Room LS 424 (Liberal Science Hall)
Instructor: Sharma Chakravarthy, 632 ERB
Phone: 272-2082, email: sharma@cse.uta.edu,
Course URL: https://www.wweb.uta.edu/faculty/sharmac/
Research URL: http://itlab.uta.edu/sharma
Office Hours: Tu/Th 2pm to 3:20pm and by appointment
TA:
TA Office Hours:

Prerequisites: CSE 3330/CE 5330 or consent of instructor

Catalog Description: DBMS system implementation techniques include query optimization, transaction processing, concurrency control, buffer management, and recovery. Object-oriented, object-relational, and XML databases. Introduction to advanced database models, such as active, distributed, temporal, spatial, and data warehousing. Prerequisite: CSE 3330/CSE 5330, or consent of instructor.

Objective: The objective of this course is to understand the theoretical underpinnings as well as design and implementation of various components of a relational database management system. Through a series of projects, students will understand and appreciate the implementation techniques used for various modules, such as query processor/optimizer, B+ tree index, buffer manager, transaction manager and recovery subsystems. We will introduce newer paradigms of data processing such as map/reduce and NoSQL to contrast it with relational DBMSs.

Course Outline: This course will cover various components of a Database Management system mainly from the systems/implementation viewpoint (as opposed to the users’ viewpoint as in Database I). Topics will be covered under four modules: i) storage and indexing (File types, B+ trees, Hash indexes, access analysis), ii) Transaction Management and concurrency control (ACID properties and their need, two-phase commit, ARIES recover, relationship between this and rest of the system), iii) Cloud computing and Big Data (need for new paradigm, map/reduce details, big data applications)m and iv) query optimization (cost model, system R query optimization).

Textbook:

Project: Since the emphasis of this course is on implementation techniques, there will be a number of hands-on implementation projects in Java/C++ as part of this course. I plan on using Minibase for implementing Heap files, join algorithms, components of a transaction manager, buffer pool manager, and/or access methods. Students may be asked to present the details of their implementation in the class. Evaluation of the project will have an optional discussion session with the students to discuss the approach taken.

Grading: There will be 3 hands-on projects that will constitute approximately 45% of the grade. There will be 3 in-class tests that will constitute approximately 50% of the grade. This is an initial proposal. Class attendance and presentation will carry 5% of the grade. The instructor reserves the right to re-distribute the percentages if deemed necessary. The students may be asked to make an in-class presentation on the project experiences. Attendance and class participation is important for doing well on the course. Grading for undergraduates will be done separately from graduates in a combined offering. Based on past observations, typically, class average corresponds to a B grade. One standard deviation above the class average is likely to earn an A; Passing grade is 50%. Note that you have to do consistently well on all projects and exams to earn an A grade.

Academic Honesty: I strictly adhere to The University of Texas at Arlington rules and guidelines for handling violations of academic dishonesty. Please refer to the pamphlet “CHEATING: Definitions and Consequences” for additional information. You are required to sign and return two statements about academic dishonesty. If anyone is caught for cheating, plagiarism or collusion either on the project or on the quiz/exam, the grade for the course will be an automatic Fail grade (F) and will be reported. For projects, all members of the team will receive the same penalty; so make sure you are aware of what your partner is doing! Projects can be checked for cheating/plagiarism at any time during the semester and brought to trial.

How to Do Well in This Course: Students who get the most out of this course will be the ones who follow the material taught and ask a lot of questions. If you want to do well, attend all the lectures, read the assigned sections of the book/papers, and start early on your projects. Working out the assigned sample questions and questions from book chapters will immensely help in doing well on quizzes/exams. If you are having difficulty, you owe it to yourself to get help. We will be more than happy to help you. Don't be afraid to come and see us. Please make use of the office hours which are meant to help you. If you can't make it to office hours but really need help, contact one of us for an appointment. I sincerely want all of you to do well. It is your responsibility to check the course web site and the black board at least twice a week.

NOTE 1: The class schedule, exam, and project due dates are tentative. Project deadlines and exam dates may be changed (with sufficient notice) based on the progress made in the class. No makeup quizzes or exams will be given unless there is a justifiable documented reason.

NOTE 2: Once the grade of a quiz/exam/project is distributed, you will have 5 business days to dispute it and get it re-evaluated. No re-evaluation will be entertained after the 5 day period. For projects, as part of the document, what has been designed and implemented by each partner (if it is done as a team) should be clearly stated and documented. All team members will get the same grade on the project.