COSC 404 / IGS 509K - Database System Implementation
2015 Winter Term 2

Instructor: **Dr. Ramon Lawrence**
Class Schedule: 11:00 a.m. – 12:30 p.m. Tuesday/Thursday
Location: ARTS 104
Lab times/locations: **L01:** 8:30 a.m. – 10:30 a.m. Wednesdays at ASC 165
L02: 12:30 p.m. – 2:30 p.m. Wednesdays at FIP 129
Office Hours: 9:00-10:00 a.m. Tuesdays/Thursdays, in labs, or by appointment
Office Location: ASC 349
Phone: 807-9390
E-mail: ramon.lawrence@ubc.ca (preferred contact method)
Course URL: http://people.ok.ubc.ca/rlawrenc/teaching/404/

Course Description

Official Calendar: Fundamental concepts in constructing database systems including file organizations, storage management, system architectures, query processing/optimization, transaction management, recovery, and concurrency control. Additional topics may include distributed databases, mobile databases, and integration. [3-2-0]

Prerequisite

- COSC 304 – Introduction to Database Systems

Evaluation Criteria and Grading

Programming Assignments	**20 %**	(weekly assignments)
Written Assignments	**15 %**	(weekly assignments)
Clickers	**5 %**	
Midterm Exam	**20 %**	
Final Exam	**40 %**	(cumulative, three hours)

IGS Graduate Student Evaluation:

Programming Assignments	**10 %**	(weekly assignments)
Written Assignments	**5 %**	(weekly assignments)
Clickers	**5 %**	
Project	**20 %**	
Midterm Exam	**20 %**	
Final Exam	**40 %**	(cumulative, three hours)

- **Graduate students are responsible for a substantial research and development project. Further, optional or bonus assignment questions for undergraduates will often be required for grad students.**
- **A student must receive a combined grade of at least 50% on the exams (midterm and final) to pass the course. Otherwise, the student will be assigned a maximum overall grade of 45.**

Textbook and Reference Material:

- All notes and reference material will be on the web site. Clickers are required.
- Optional textbook: Hector Garcia-Molina, Jeffrey Ullman and Jennifer Widom, *Database Systems: The Complete Book (2nd edition)*, Prentice Hall, ISBN 0-131-87325-3, 2008.

Expectations

- Attend **all** classes and prepare before attending class.
- Read the lecture notes **before** the lecture.
- Learn the material in the course and undertake sufficient effort to produce all the programming assignments and quality projects.
- Enjoy attending class and feel free to participate according to your personality. Feel free to ask questions by raising your hand or speaking out at appropriate times.
- Please actively participate in class discussions, questions, and problem solving exercises.
- **I want all students to pass the course, receive a good grade, and feel the course was beneficial.**

Homework Expectation

For this course, it is expected that you will spend ***at least six hours per week in out-of-class preparation***.

Grievances and Complaints Procedures

A student who has a complaint related to this course should follow the procedures summarized below.

• The student should attempt to resolve the matter with the instructor first. Students may talk first to someone other than the instructor if they do not feel, for whatever reason, that they can directly approach the instructor.

• If the complaint is not resolved to the student's satisfaction, the student should go to the departmental chair John Braun at SCI 388, 807-8032.

Academic Integrity

The academic enterprise is founded on honesty, civility, and integrity. As members of this enterprise, all students are expected to know, understand, and follow the codes of conduct regarding academic integrity. At the most basic level, this means submitting only original work done by you and acknowledging all sources of information or ideas and attributing them to others as required. This also means you should not cheat, copy, or mislead others about what is your work. Violations of academic integrity (i.e., misconduct) lead to the breakdown of the academic enterprise, and therefore serious consequences arise and harsh sanctions are imposed. For example, incidences of plagiarism or cheating may result in a mark of zero on the assignment or exam and more serious consequences may apply if the matter is referred to the President's Advisory Committee on Student Discipline. Careful records are kept in order to monitor and prevent recurrences. A more detailed description of academic integrity, including the University's policies and procedures, may be found in the Academic Calendar at: http://okanagan.students.ubc.ca/calendar/index.cfm?tree=3,54,111,0.

Disability Assistance

If you require disability-related accommodations to meet the course objectives, please contact the Diversity Advisor of Disability Resources located in the University Centre, Room 227. For more information about Disability Resources or academic accommodations, please visit the website at: http://www.ubc.ca/okanagan/students/drc/.

Equity, Human Rights, Discrimination and Harassment

UBC Okanagan is a place where every student, staff and faculty member should be able to study and work in an environment that is free from human rights based discrimination and harassment. If you require assistance related to an issue of equity, discrimination or harassment, please contact the Equity Office, your administrative head of unit, and/or your unit's equity representative. **UBC Okanagan Equity Advisor: ph. 250-807-9291; email** equity.ubco@ubc.ca
Web: www.ubc.ca/okanagan/equity

Missing an Exam

Only students who miss the final exam for a reason that corresponds to the University of British Columbia Okanagan's policy on excused absences from examinations will be permitted to take the final exam at a later time. A make-up exam may have a question format different from the regular exam. **There will be no make-up midterm exams.** If the reason for absence is satisfactory, the student's final exam will be worth more of the final grade. Further information on Academic Concession can be found under Policies and Regulation in the Okanagan Academic Calendar http://www.calendar.ubc.ca/okanagan/index.cfm?tree=3,48,0,0.

Objectives and Learning Outcomes

- Experience using and developing on many different SQL and NoSQL databases.
- Manipulate data in memory and storage and use index structures for improved performance.
- Understand query processing including parsing, translation, optimization, and execution.
- Apply principles of transactions, concurrency, recovery, and distribution for databases.
- Use knowledge of database techniques to be better users with the ability to use different database systems, compare their properties, and adapt database techniques when developing software.

Course Outline

The course has a substantial amount of material to be covered in a short time. This requires the student make a strong effort to prepare before class so that the material can be practiced in class. Below is an outline of the topics. The professor is not bound to these topics and timelines as they only serve as a general reference.

Date	Topics Covered and Description
January 5 (T)	**First day of classes. Introduction to course. Database architecture.**
January 7 (TH)	Storage I: Accessing Data on Devices: Memory, Hard drives, SSDs, RAID
January 12 (T)	Storage II: Representing Data Using Records and Blocks
January 14 (TH)	Indexing I: Index Types, Primary Indexes, Multi-level Indexes, Secondary Indexes
January 19 (T)	Indexing II: B-Trees (insertion, deletion), B+-Trees
January 21 (TH)	Indexing III: B+-Trees, R-Trees
January 26 (T)	Indexing IV: Hash Indexes, SQL Indexing in Practice
January 28 (TH)	Query processing I: SQL/RA Review, Types of Operators, Iterators, One-pass Algorithms
February 2 (T)	Query processing II: Nested-Loop Joins, External Sorting, Two-Pass Sorting Algorithms, Sort-Join, Sort-Merge-Join
February 4 (TH)	Query processing III: Hash Partitioning, Two-Pass Hash Algorithms, Hybrid Hash Join
February 9 (T)	**No classes during Midterm Break.**
February 11 (TH)	**No classes during Midterm Break.**
February 16 (T)	Query optimization I: Query Parsing/Translation, Relational Algebra Laws
February 18 (TH)	**Midterm Exam**
February 23 (T)	Query optimization II: Heuristic Optimization, Physical Query Plans
February 25 (TH)	Query optimization III: Cost-based Query Optimization
March 1 (T)	Transaction processing I: ACID Properties, Schedules, Conflict Serializablity
March 3 (TH)	Transaction processing II: View Serializablity, Schedule Properties
March 8 (T)	Concurrency control I: Two-Phase Locking (2PL), Multiple Granularity Locking, Deadlock Handling, Wait-for Graphs
March 10 (TH)	Concurrency control II: Timestamp Protocols, Validation Protocols, Multi-versioning, Snapshot isolation
March 15 (T)	Concurrency control III: SQL Isolation Levels, Phantom Phenomenon, CC in systems
March 17 (TH)	Recovery I: Types of Failures, Log-Based Recovery
March 22 (T)	Recovery II: Undo/Redo Logging
March 24 (TH)	Distribution I: Architectures, Semi-joins, Two-Phase Commit
March 29 (T)	Distribution II: Fragmentation, Partitioning, Sharding
March 31 (TH)	Distribution III: Replication: Master-Master and Master-Slave, CAP Theorem
April 5 (T)	Architecture I: Comparison of database architectures: Relational, Key-Value, In-Memory
April 7 (TH)	**Review for final exam.**

Laboratory times: The laboratory time will be spent on programming assignments.

Week	Dates	Topics Covered and Description
1	January 6	**No Lab First Week of Class**
2	January 13	Lab 1: MySQL vs. PostgreSQL – Creating and Querying Data
3	January 20	Lab 2: MySQL vs. PostgreSQL – Indexing for Performance
4	January 27	Lab 3: Implementing a Text Database and JDBC Driver
5	February 3	Lab 4: Query Processing with Iterators
6	February 10	**No Lab During Midterm Break**
7	February 17	Lab 5: Query Parsing with JavaCC
8	February 24	Lab 6: Storing JSON Documents: MongoDB and PostgreSQL
9	March 2	Lab 7: Map-Reduce
10	March 9	Lab 8: Transactions with Microsoft SQL Server
11	March 16	Lab 9: VoltDB – In-memory database
12	March 23	Lab 10: Recovering from a Database Failure
13	March 30	Lab 11: Scaling MySQL: Master-Slave Replication and Partitioning
14	April 6	**No Lab Last Week of Class**

Written assignments: Written assignments practice fundamental skills. They will be done at various times including in class, in lab, and before and after class time.

Week	Dates	Topics Covered and Description
1	January 6	**No Written Assignment First Week of Class**
2	January 13	Assign 1: Storage Performance Calculations
3	January 20	Assign 2: Index Performance Calculations
4	January 27	Assign 3: B-Trees/B+-Trees and Linear Hashing
5	February 3	Assign 4: Query Processing
6	February 10	**No Assignment During Midterm Break**
7	February 17	No assignment.
8	February 24	Assign 5: Query Optimization
9	March 2	Assign 6: Transactions
10	March 9	Assign 7: Concurrency Control Protocols
11	March 16	Assign 8: Deadlock Handling
12	March 23	Assign 9: Recovery
13	March 30	Assign 10: Distribution and Partitioning
14	April 6	**No Assignment Last Week of Class**