COMP 6231 - Distributed Systems Design
Concordia University
Fall 2023

● Course Description

This course emphasizes the principles of distributed systems in general and large-scale distributed data systems in particular, such as NoSQL systems. The emphasis will be on topics, such as processes and threads, concurrent programming, distributed interprocess communication, distributed process scheduling, virtualization, distributed file systems, the web and peer-to-peer systems, and security in distributed systems. Students will gain practical experience in designing, implementing, and debugging real distributed systems, such as Pregel, GraphLab, PowerGraph, Cassandra, Couchbase, MongoDB, or Elasticsearch. Students will also have practical experience with different cloud computing infrastructures, such as Amazon EC2, Google Compute Engine, or Microsoft Azure.

● Course Team

Instructor: Dr. Essam Mansour (essam.mansour@concordia.ca)
Instructor's Website: emansour.com, http://cods.ai

TAs:

Omij Mangukiya omij.mangukiya@mail.concordia.ca
Waleed Afandi waleed.afandi@mail.concordia.ca

DDDI - Monday 20:30 - 22:20 H 967 (Omij Mangukiya)
DDDJ - Monday 20:30 - 22:20 H 929 (Waleed Afandi)
DDDK - Monday 15:30 - 17:20 H 907 (Waleed Afandi)
DDDL - Thursday 15:30 - 17:20 H 811 (Omij Mangukiya)

Labs will be in person.

We will be available for help and support via Email or during our office hours. To receive help outside of office hours, it is preferred to schedule a meeting or send an email.

● Course Textbook

● *Distributed Systems, 3rd Edition, Andrew Tannenbaum and Maarten van Steen, 2018.*

All lectures will be prepared from this book. The book is available at moodle. plus some other material, we may need.

● Course Organization

The participation of students in the course will involve four forms of activities:

- Attending lectures and labs online.
- Solving assignments (involving writing and/or coding).
- A programming project of one phase.
- Participating in class discussions.

● Assessment

Each student will receive a numeric score with a corresponding letter grade, based on a weighted average of the following:

1. **Project:** The project will count for a total of **30% of your final score**. There will be **one phase** throughout the course in teams. Each team consists of 3 students. Each team will learn one of the following systems, Pregel, GraphLab, PowerGraph, Cassandra, Couchbase, MongoDB, or Elasticsearch. The team has to discuss the concepts of distributed systems that have been implemented in the chosen system. The team also should find a real dataset of at least one GB and demo the capabilities of the chosen system. The deliverables of the project are:

Demo	**15%** of the final grade
Presentation	**10%** of the final grade
Report	**05%** of the final grade

   The deliverables are due on **Sunday, Dec 17, 2023**. If you submit **one day late**, we will **deduct 25%** of the project score as a penalty. If you submit **two days late**, **50% will be deducted**. The project will not be graded (and you will receive a zero score on it) if you are more than two days late.

2. **Exams:** There will be two online midterm exams, which combined will account for **45% of your final score**, as the following:

Midterm one [**worth 20%**]	on Oct 23, 2023
Midterm two [**worth 25%**]	on Dec 04, 2023

3. **Assignments:** There will be 5 assignments that will test you on some problem analysis and solving skills. The 5 assignments will altogether contribute **25% towards your final score**. These assignments are:

ID	Topic	Weight	Release Date	Due Date
1	Socket	5%	Fri Sept 22	Sun Oct 08
2	*Multithreading & MPI*	*9%*	*Fri Oct 06*	*Sun Oct 29*
3	Docker	5%	Fri Oct 27	Sun Nov 12
4	SPARK Map-Reduce	6%	Fri Nov 10	Sun Nov 26

4. **Class/Recitation Participation:** Your attendance of both classes and labs as well as your participation in discussions during presentations, will help you to do well in your exams and assignments.

Table 1 below shows the breakdown of the main forms of activities that the course involves, alongside the quantity and the overall weight of each activity. Take into account that small differences in scores can make the difference between two letter grades. Letter grades will be determined by absolute standards. The total score will be plotted as a histogram. Cutoff points are determined by examining the quality of students' work on the borderlines. Individual cases, especially those near the cutoff points ***may be adjusted upward or downward based on factors such as attendance, class participation***, the improvement observed throughout the course, exam performance, and special circumstances.

Type	#	Weight
Project	1	30%
Exams	2	45%
Assignments	4	25%

Table 1: Breakdown of the main activities involved in the course.

● Policies

Working Alone on Assignments
Assignments that are assigned to students should be ***performed*** ***individually***.

Handing in Assignments/Projects
All assignments/projects are due at 11:59 PM (one minute before midnight) on the specified due date.

Appealing Grades
After each exam, assignment, and/or project is graded, you have **7** calendar days to appeal your grade. All your appeals should be provided in writing. If after appealing you are still not satisfied, please visit the instructors. If you have questions about an exam, an assignment or a project grade, please visit the instructors directly.

● Cheating

Each project or assignment must be the sole work of the student(s) turning it in. Projects and assignments will be closely monitored, and ***students may be asked to explain suspicious similarities with any write-up or piece of code available***. The following are guidelines on what cheating is and is not:

What is cheating?
1. Sharing code or other electronic files: either by copying, retyping, looking at, or supplying a copy of a file.
2. Sharing written assignments: either by re-writing, looking at, or supplying a copy of an assignment.

What is NOT cheating?
1. Clarifying ambiguities or vague points in-class handouts.
2. Helping others use computer systems, networks, compilers, debuggers, profilers, or other system facilities.
3. Helping others with high-level design issues.
4. Helping others debug their codes.

Consequently, be aware of what constitutes cheating (and what does not) when interacting with your colleague students. The same rules of cheating as above apply when collaborating with other students. In short, you cannot share written assignments, code, and/or other electronic files with other students. If you are unsure, ask the teaching staff.

Finally, be sure to store your work in protected directories. The penalty for cheating is severe, and might jeopardize your whole career as a student – cheating is not worth the trouble. By cheating in the course, you are cheating yourself; the worst outcome of cheating is missing an opportunity to learn. Besides, you will be removed from the course and assigned a failing grade. We also place a record of the incident in your permanent university profile.