CSCI 381/780 Special Topics in Computer Science (Cloud Computing)

Welcome to the undergraduate/graduate course on cloud computing. This course broadly surveys key technical issues and solutions in cloud computing. We will cover industrial techniques and academic works on cloud computing, including the concepts and principles of cloud computing and the techniques of using cloud systems and developing cloud applications.

Topics: overview of cloud computing, data center, virtualization, cloud systems for big data storage & processing, cloud systems for machine learning, resource scheduling and resource management, consistency, and fault tolerance.

Course Information

- **Term**: Spring 2023
- **Instructor**: Jun Li
- **Email**: jun.li@qc.cuny.edu
- **Classroom**: SB-A101
- **Schedule**: Tu/Th 12:15 PM - 1:30 PM
- **Office**: SB-A342 (appointment required)
- **Online platform**: Microsoft Teams
- **Office Hours**: 10:30 AM -11:30 PM, every Tuesday (virtual office hour on Microsoft Teams)
- **Course Website**: https://phantom.cs.qc.cuny.edu/li/cc (all general information about the course will be posted here except the materials below)
    - Assignments & Grades: Blackboard
    - Announcements & Discussion: Microsoft Teams

Prerequisite

CSCI 340: Operating Systems Principles

Textbook

We don't have any textbook. Materials will be distributed on the course website.

Grading Rules

Your grade will be computed based on:

- Projects (50%)
- Mid-term Exam (20%)
    - the week of March 21
- Final Exam (30%)
    - May 18, SB-A101, 11:00 AM - 1:00 PM

Your final letters will generally be converted from your grade using the following rules:

- A+: >= 97%, A: [93%, 97%), A-: [90%, 93%)
- B+: [87%, 90%), B: [83%, 87%); , B-: [80%, 83%)
- C+: [75%, 80%), C: [70%, 75%)
- F: < 70%

Policies

Submission: Project code not submitted online before the deadline will be considered late and will be subject to penalty. The penalty for late submission without a pre-approved extension will be based on the lateness of the submission.
All submitted code should be reasonably documented. You are expected to submit your files on Blackboard. If there is something that you would like me to know while grading your assignment, please write it in a file called README and write your message there. Please do not mail your code to the instructor directly. You may discuss coursework with other students, but you must write up your coursework independently in your own words. You are not allowed to search the web for solutions.

Accessibility: If you need help due to a registered disability or you think you may have a disability, please talk to me and contact the Office of Special Services for Students with Disabilities as soon as possible. We will discuss how this course will accommodate your individual learning needs.

Academic Integrity and Honesty: Students are expected to adhere to the CUNY policy on academic integrity. Questions related to coursework and the academic honesty policy should be directed to the instructor. A sanction will certainly be imposed on the student committed to any academic fraud. It varies depending upon the instructor's evaluation of the nature and gravity of the offence. Possible sanctions include but are not limited to, the following: (1) Require the student to redo the assignment; (2) Require the student to complete another assignment; (3) Assign a grade of zero to the assignment; (4) Assign a final grade of zero for the whole course.

Recording of Remote Classes: Students who participate in this class with their camera on or use a profile image are agreeing to have their video or image recorded solely for the purpose of creating a record for students enrolled in the class to refer to, including those enrolled students who are unable to attend live. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate orally are agreeing to have their voices recorded. If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live.

(Tentative) Course Schedule

Note: topics and dates of future classes are tentative and subject to change.

Date	Details
January 26	Course Overview Introduction (1)
January 31	Class canceled
February 2	Introduction (2)
February 7	Data Center (1)
February 9	Data Center (2) Recommended reading: The Datacenter as a Computer
February 14	Data Center (3) Virtualization 1: Virtual Machine (1) **Project 0 released (due: 11:59 PM, February. 28)**
February 16	Virtualization 1: Virtual Machine (2) Virtualization 2: Container (1)
February 21	No class (following a Monday schedule)

Date	Details
February 23	Virtualization 2: Container (2) Recommended reading: Large-scale cluster management at Google with Borg, Borg: The Predecessor to Kubernetes
February 28	Hadoop (1) Recommended reading: Google File System
March 2	Hadoop (2) Recommended reading: MapReduce: Simplified Data Processing on Large Clusters **Project 1 released (due: 11:59 PM, Mar. 19)**
March 7	Hadoop (3) Recommended reading: Boosting Hadoop Performance and Cost Efficiency with Caching, Fast SSDs, and More Compute
March 9	Spark (1) Recommended reading: Spark: Cluster Computing with Working Sets
March 14	Spark (2) **Note: This is the end of the scope of the mid-term exam.**
March 16	Spark (3) Resource Scheduling (1)
March 21	Resource Scheduling (2)
March 23	Mid-term Review **Mid-term (take-home) exam released on Blackboard (due: 11:59PM, March 28)**
March 28	No class (for the mid-term exam)
March 30	Virtualization 3: Serverless Computing **Project 2 released (due 11:59PM, Apr. 23)**
April 4	Key-value Store (1) Recommended reading: Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications
April 6	No class (spring recess)
April 11	No class (spring recess)
April 13	No class (spring recess)
April 18	Key-value Store (2) Recommended reading: Dynamo: Amazon's Highly Available Key-value Store
April 20	Machine Learning (1)
April 25	Machine Learning (2)
April 27	Advanced Topic (1) **Project 3 released (due 11:59 PM, May 14)**
May 2	Advanced Topic (2)
May 4	Advanced Topic (3)

Date	Details
May 9	Advanced Topic (4)
May 11	Advanced Topic (5)
May 16	Final Review
May 18	**Final Exam** (11:00 AM - 1:00 PM, SB-A101)

Acknowledgments

Some materials are adapted from the following related courses/presentations:

- Mosharaf Chowdhury, EECS 598: Big Data Systems and Applications
- C. L. Wang, COMP7305A - Cluster and cloud computing
- Nick Mckeown's talk on SDN
- Michael Freedman and Kyle Jamieson, COS-418: Distributed Systems
- Ali Ghodsi and Ion Stoica, CS262a: Advanced Topics in Computer Systems
- Yiying Zhang, CSE 291 Virtualization
- Nael B. Abu-Ghazaleh, CS202: Advanced Operating Systems
- Vivienne Sze, Efficient Processing of Deep Neural Network: from Algorithms to Hardware Architectures

---

Last update: Tue Apr 4 13:39:42 EDT 2023