FA24-CS643863 Cloud Computing

✎ Edit

CS 643 863 Fall 2024 - Syllabus

Instructor

- Cristian Borcea
- borcea@njit.edu (mailto:borcea@njit.edu)
- https://cs.njit.edu/~borcea (https://cs.njit.edu/~borcea)
- Online office hours: Mondays, 5pm-6pm
    - https://njit-edu.zoom.us/j/9788707366?pwd=R0s1aTBzTHNXVmZpZUR2R1BSYkw3QT09 (https://njit-edu.zoom.us/j/9788707366?pwd=R0s1aTBzTHNXVmZpZUR2R1BSYkw3QT09)

      Meeting ID: 978 870 7366

      Passcode: 133224

Short Description

The course presents a comprehensive view of cloud computing, from platforms and services to programming and infrastructure. The topics include: cloud computing platforms, with examples from Amazon Web Services (AWS), Google Cloud, and Microsoft Azure; cloud services for data analytics, machine learning, mobile computing, IoT, edge computing, security and privacy, and devops; programming frameworks for parallel computing in the cloud; distributed storage in the cloud; and virtualization and containerization. The course includes homework, programming assignments, and research paper presentations. The programming assignments will be done in AWS.

Learning Outcomes

Upon the successful completion of this course, the student should be able to:

- Analyze the trade-offs between deploying applications/services in the cloud and on premises

- Compare the advantages and disadvantages of different types of cloud platforms
- Deploy applications in public cloud platforms
- Program data intensive parallel applications in the cloud
- Analyze the performance, scalability, and availability of cloud systems and applications
- Identify security and privacy issues in the cloud
- Present state-of-the-art cloud research

Why Take This Course?

Cloud computing represents a major paradigm shift in computing from the era of personal computers to the era of computing as utility. Most major Internet services are already deployed in the "the cloud." We store most of our data in "the cloud" and execute most applications from "the cloud." This course is aimed at all graduate students (both M.S. and Ph.D. students) who want to learn how to design and program cloud services and applications as well as how to build and administer cloud systems. By studying real-world systems developed in industry, students will acquire cutting-edge knowledge that may be a major advantage when searching for a job.

Lectures and Readings

There is no book required for this class. Each module is based on research papers and/or online documentation covering a specific topic (i.e., readings). The readings should be used as reference material to clarify and add details to lectures. All the slides and voice-over video recordings will be posted by the first day of classes. You should allocate about one week for each class module.

Exams

There will two online exams: a midterm, and a final exam. Both exams are closed book (i.e., papers, notes, etc.), and students are not allowed any online access during the exams, except for the locked-down browser. The final exam will cover only the material taught after the midterm. In case of missing an exam, a make-up may be taken only after providing written documentation to the Dean of Students.

Homework

Homework will be assigned 6 times during the semester to prepare students with the type of questions encountered in exams. The solutions will be posted after the deadlines and, if needed, will be further discussed during office hours.

Programming Assignments

There will be two individual programming assignments. The first is to build an image recognition pipeline in Amazon AWS, using two EC2 instances, S3, SQS, and Rekognition. The assignment must be done in Java on Amazon Linux VMs. You will learn how to use the AWS cloud platform and how to develop an AWS application that uses existing cloud services. The second is to build a machine learning prediction model in Spark/MLlib over AWS. The model must be trained in parallel on multiple EC2 instances. The assignment must be implemented in Java, Scala, or Python on Ubuntu Linux. You will learn how to develop parallel machine learning applications in the AWS cloud platform. The anonymized best submitted solutions will be posted.

Research Paper Presentations

Students will present in groups of two one research paper. These papers cover state-of-the-art research in cloud computing. The slides and videos of the presentations must be uploaded in Canvas before the presentation deadline. Extra-credit is available for asking good questions about the presentations.

Grading

- Midterm exam - 25%
- Final exam - 25%
- Programming Assignment 1 - 10%
- Programming Assignment 2 - 20%
- Research presentation - 8%
- Homework - 12%

Schedule

- **Module 1: Course overview. Introduction to Cloud Computing.**
  - Form groups for research paper presentations by 8 September
- **Module 2: Cloud Platforms I: Infrastructure as a Service (IaaS), AWS.**
  - Homework 1 handed out on 11 September
  - Programming assignment 1 handed out on 15 September
- **Module 3: Cloud Platforms II: Serverless Computing; Function as a Service (FaaS).**
  - Homework 1 due on 18 September
- **Module 4: Cloud Services I: Data Analytics and Machine Learning.**

    - Homework 2 handed out on 25 September
- **Module 5: Cloud Services II: Mobile, IoT, and Edge Computing.**
    - Homework 2 due on 2 October
    - Homework 3 handed out on 2 October
    - Programming assignment 1 due on 6 October
- **Module 6: Cloud Services III: Security and Privacy, Devops.**
    - Homework 3 due on 9 October
- **Midterm on Saturday 12 October**
    - The exam must be completed in 90 minutes between 10AM and 3PM EDT
    - Discussion of midterm solutions during office hours on 14 October
- **Research Paper Presentations**
    - Presentations due on 19 October
- **Module 7: Parallel Programming in the Cloud I: Google's MapReduce, Apache's Hadoop, Yahoo's Pig Latin.**
    - Programming assignment 2 handed out on 23 October
- **Module 8: Parallel Programming in the Cloud II: Apache's Spark, Storm and Zookeper.**
    - Homework 4 handed out on 30 October
- **Module 9: Cloud Storage Systems I: Google's GFS and BigTable.**
    - Homework 4 due on 6 November
- **Module 10: Cloud Storage Systems II: Amazon's Dynamo and Other Cloud Databases.**
    - Homework 5 handed on 13 November
- **Module 11: Virtualization I: VMWare, XEN, Live VM Migration.**
    - Homework 5 due on 20 November
- **Module 12: Virtualization II: Containerization, Docker, Kubernetes.**
    - Homework 6 handed out on 4 December
- **Last Week of Classes**
    - Homework due on 11 December
    - Programming assignment 2 due on 11 December
- **Final Exam on Saturday 21 December**
    - The exam must be completed in 90 minutes between 10AM and 3PM

Academic Integrity

Academic Integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: **NJIT Academic Integrity Code (https://t.e2ma.net/click/fw4pfmb/3yu6novf/3qgyp0x)** .

Please note that it is my professional obligation and responsibility to report any academic misconduct to the Dean of Students Office. Any student found in violation of the code by cheating, plagiarizing or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university. If you have any questions about the code of Academic Integrity, please contact the Dean of Students Office at **dos@njit.edu (mailto:dos@njit.edu)** .

Modifications to Syllabus

The students will be consulted and must agree to any modifications or deviations from the syllabus throughout the course of the semester.