Fall 2020

CS 643-101: Cloud Computing

Cristian Borcea

CS 643 101 Fall 2020 - Syllabus

Instructor

- Cristian Borcea
- borcea@njit.edu
- https://cs.njit.edu/~borcea (Links to an external site.)

- Synchronous online class: Wednesdays, 6pm-8:50pm, Webex: https://njit.webex.com/meet/borcea (Links to an external site.)
- Live office hours: Mondays, 6pm-7pm, Webex: https://njit.webex.com/meet/borcea (Links to an external site.)
    - You can also join by phone: 1-650-479-3207 *Call-in toll number (US/Canada).* Access code: 928 115 913

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 in the cloud and over local infrastructure
- Compare the advantages and disadvantages of different types of cloud platforms
- Deploy applications over commercial cloud platforms
- Program data intensive parallel applications in the cloud
- Analyze the performance, scalability, and availability of the cloud systems and applications
- Identify security and privacy issues in the cloud
- Collaborate to 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 lecture is based on research papers and/or online documentation covering a specific topic (i.e., readings). The slides for each lecture will be posted before each class. Each lecture will be delivered during our synchronous classes. After each class, the voice-over video recordings of each sub-module of the lecture will be posted as well. The readings should be used as reference material to clarify and add details to lectures.

Exams

There will be two exams: a midterm, and a final exam. Both exams are closed book (i.e., papers, notes). The final exam will cover only the material taught after the midterm. Both exams will take place online using ProctorU and Canvas. 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 5 times during the semester to prepare students with the type of questions encountered in exams. The solutions will be discussed in during our synchronous class.

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.

Research Paper Presentations

Students will present, in groups of three, one research paper during the semester. These papers cover state-of-the-art research in cloud computing. The video recordings of voice over slides will be uploaded in Canvas for the whole class to watch. The presentations must be uploaded in the week to which the papers are assigned. 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 - 10%
- Homework - 10%

Schedule

- **Week 1: Course overview. Introduction to Cloud Computing.**
  - Form groups for research paper presentations by 9 September
- **Week 2: Cloud Platforms I: Infrastructure as a Service (IaaS), AWS.**
  - Homework 1 handed out on 9 September
  - Research papers assigned to groups by 11 September
- **Week 3: Cloud Platforms II: Platform as a Service (PaaS), Google App Engine, Windows Azure.**
  - Homework 1 due on 16 September
  - Programming assignment 1 handed out on 16 September
- **Week 4: Cloud Platforms III: Serverless Computing; Function as a Service (FaaS).**
  - Homework 2 handed out on 23 September
  - Research paper presentations 1 and 2 due on 23 September
- **Week 5: Cloud Services I: Data Analytics and Machine Learning.**
  - Homework 2 due on 30 September
  - Research paper presentations 3 and 4 due on 30 June
- **Week 6: Cloud Services II: Mobile, IoT, and Edge Computing.**
  - Homework 3 handed out on 7 October
  - Research paper presentations 5 and 6 due on 7 October
- **Week 7: Cloud Services III: Security and Privacy, Devops.**
  - Programming assignment 1 due on 14 October
  - Homework 3 due on 14 October
- **Week 8: Midterm on 21 October (Online using ProctorU and Canvas)**
  - Discussion of midterm solutions during office hours on Monday, 26 October
- **Week 9: Parallel Programming in the Cloud I: Google's MapReduce, Apache's Hadoop, Yahoo's Pig Latin.**
  - Research paper presentation 7 due on 28 October
  - Programming assignment 2 handed out on 28 October
- **Week 10: Parallel Programming in the Cloud II: Apache's Spark, Storm and Zookeper.**
  - Research paper presentations 8 and 9 due on 4 November
- **Week 11: Cloud Storage Systems I: Google's GFS and BigTable.**
  - Research paper presentations 10 and 11 due on 11 November
  - Homework 4 handed out on 11 November
- **Week 12: Cloud Storage Systems II: Amazon's Dynamo and Other Cloud Databases.**
  - Homework 4 due on 18 November
  - Research paper presentations 12 and 13 due on 18 November
- **Week 13: Virtualization I: VMWare, XEN, Live VM Migration.**
  - Homework 5 handed out on 2 December

- **Week 14: Virtualization II: Containerization, Docker, Kubernetes.**
  - Homework 5 due on 9 December
  - Programming assignment 2 due on 9 December
- **Week 15: Final Exam is on 16 December (Online using ProctorU and Canvas)**

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 this link: University Policy on Academic Integrity (Links to an external site.).

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.

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.