332:494/02: TOPICS IN ELECTRICAL & COMPUTER ENGINEERING
Section 01: CLOUD COMPUTING

Fall 2016

Instructor

Maryam Mehri Dehnavi
maryam.mehri@rutgers.edu

Course Overview

This graduate-level course (also open to senior undergraduate students) provides an introduction to cloud computing. A combination of lectures, research presentations, and the course project will expose students to cloud computing concepts, programming models, and frameworks. Students will learn how to process large data sets on computer clusters built from commodity hardware.

Requirements:

The students should be comfortable programming in Python and Java. Familiarity with parallel & distributed computing and linear algebra is highly recommended.

Prerequisites:

14:332:331 – Computer Architecture and Assembly Language
14:332:434 – Introduction to Computer Systems

Class Evaluation:

The final grade is based on the course project, paper reviews and presentation, class quizzes, and programming tasks.

Presentation: You can either present research publications from the list of recommended papers or a topic of your choice if approved by the instructor. Each presentation should be around 15 minutes, after which the presenter will be in charge of leading a 10-minute open discussion.

Research paper summaries: The main propose of research presentations is to familiarize you with the process of reviewing conference papers. For sessions that you are not presenting, you should prepare a short summary of the paper being presented. Also prepare a list of questions that you would ask the presenter during his or her talk. Contributing in the open discussions of research presentations plays an important role in your final grade.

Project: An essential component of this course is a research project. You are expected to work on a cutting-edge research problem in one of the areas of cloud computing, mobile computing, or scientific computing and implement your ideas. You should form a team of at most three and prepare a one-page project proposal by early October. A midterm project report and a final report written in a conference or journal format are also required. Project ideas have to be discussed with the instructor.

Tentative Topics:

- Cloud computing infrastructure
- Parallel and distributed computing concepts
- Cloud computing frameworks: Hadoop, Spark, Yarn, Hive, Storm, etc.
- Cloud services: AWS, Microsoft Azure, Google Compute Engine

- The CAP theory and key value stores
- Virtualization and Openstack
- Cloud and mobile security
- Fault tolerance in cloud
- Data center outages
- Cloud applications

Tentative Schedule:

Week 1: Introduction to cloud computing
Week 2: Overview of parallel and distributed computing
Week 3: Applied parallel computing
Week 4: The mapreduce programming model and Hadoop
Week 5: Hadoop examples
Week 6: Distributed systems concepts
Week 7: SQL, NoSQL and the CAP theorem
Week 8: Overview of locality and introduction to Spark
Week 9: Advanced topics in Spark and applications
Week 10: Open discussion: data analytics on cloud
Week 11: The seven dwarfs of computing: Sparse and dense linear algebra
Week 12: The seven dwarfs of computing: Structured and unstructured grids
Week 13: The seven dwarfs of computing: Graphs
Week 14: Data center outages and fault tolerance in cloud
Week 15: Project presentations

Topics from paper discussions are not included in the above. Most sessions will also include paper reviews and discussions.

Outcome:

N = none S = Supportive H = highly related

Outcome	Level	Proficiency assessed by
(a) an ability to apply knowledge of Mathematics, science, and engineering	H	Homework, project, quizzes
(b) an ability to design and conduct experiments and interpret data	H	Homework, project
(c) an ability to design a system, component or process	S	
(d) an ability to function as part of a multi-disciplinary team	H	Project, homeworks
(e) an ability to identify, formulate, and solve ECE problems	S	Project
(f) an understanding of professional and ethical responsibility	S	Project
(g) an ability to communicate in written and oral form	H	Paper reviews, project
(h) the broad education necessary to understand the impact of electrical and computer engineeri	N	
(i) a recognition of the need for, and an ability to engage in life-long learning	S	Discussions during lectures
(j) a knowledge of contemporary issues	N	
(k) an ability to use the techniques, skills, and modern engineering tools necessary for electrical	H	Homework, project, quizzes
Basic disciplines in Electrical Engineering	S	
Depth in Electrical Engineering	S	
Basic disciplines in Computer Engineering	H	Homework, project, quizzes
Depth in Computer Engineering	H	
Laboratory equipment and software tools	H	Spark, Hadoop
Variety of instruction formats	H	Lecture, office hours, piazza forums

Recommended books:

- Distributed Systems: Concepts and Design, Coulouris, G., Dollimore, J., Kindberg, T., and Blair, G., Addison-Wesley, Fifth Edition, 2011
- Hadoop: The Definitive Guide, O'Reilly Media, Tom White, 4th Edition

- Advanced Analytics with Spark, Patterns for Learning from Data at Scale, Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills, O'Reilly Media, April 2015