

CSCI E-90 Cloud Computing

Lecturer: Zoran B. Djordjević, PhD, Enterprise Architect, NTT Data, Inc.

Lectures: Fridays beginning Sept. 4th, 5:30-7:30PM, Rm 306, 1 Story St, Cambridge

Optional labs: Saturdays beginning Sept. 12th , 10:00-11:30AM online only

Course Description: Cloud Computing has transformed the IT industry by opening the possibility for infinite or at least highly elastic scalability in the delivery of enterprise applications and software as a service (SaaS). Amazon Elastic Cloud, Microsoft's Azure, Google App Engine, and many other Cloud offerings give mature software vendors and new start-ups the option to deploy their applications to systems of infinite computational power with practically no initial capital investment and with modest operating costs proportional to the actual use. The course examines the most important APIs used in the Amazon and Microsoft Cloud, including the techniques for building, deploying, and maintaining machine images and applications. We will learn how to use Cloud as the infrastructure for existing and new services. We will use open source implementations of highly available clustering computational environments, as well as RESTful Web services, to build very powerful and efficient applications. We also learn how to deal with not trivial issues in the Cloud, such as load balancing, caching, distributed transactions, and identity and authorization management. In the process we will also become very familiar with Linux operating system.

Prerequisites: Familiarity with Java, Python or C#.

Course Materials: Detailed handouts with references to material on the Web will be handed every week. There is no required text book.

Grading: Practically every class will be followed by a homework assignment. Grades on the solutions for class assignments constitute approximately 85% of the final grade. 15% of the grade will be earned through the final project. Final projects will be assigned/selected three weeks before the end of the class. You will produce a paper (10+ pages of MS Word text, 10+ PowerPoint Slides, a working demo, 15 minute YouTube Video of your presentation and a brief 2 minute YouTube video that might be presented to the class on the day of final presentations. Several students will be invited to present their final projects live to the entire class. 95% or higher gives you an A, 90-95% gives you an A-, 85-90% a B+, etc.

Communications: zdjordj@fas.harvard.edu, cscie90@fas.harvard.edu

Tentative List of Class Topics

Class	Date	Topic
1	09/04/2015	Cloud computing: General Benefits and Architecture, Business Drivers, Main players in the Field, Overview of Security Issues, XaaS Cloud Based Service Offerings
2	09/11/2015	Key Amazon offerings: EC2, SimpleDB, S3, Simple Queue, Simple Relational Database, Elastic MapReduce, Virtual Amazon Cloud. S3 Command Line tool

3	09/18/2015	Bundling Amazon instances: We will learn how to create and manipulate Amazon instances with command line tools, transfer application software to instances and bundle them into new AMI-s that could be offered to the public.
4	09/25/2015	Elastic bloc storage (EBS) provides persistence storage in the cloud. We will learn how to move application code and data from non-EBS instance into EBS volumes, and create our own EBS based AMI-s
5	10/02/2015	AWS Identity Management and Security in the Cloud
6	10/09/2015	Virtua Private Cloud (VPC) and Directory Service
7	10/16/2015	Java AWS DK, S3 API, Relational Database Service, SimleDB Service , NoSQL Databases
8	10/23/2015	Messaging in the Cloud. We will review details of AWS Simple Notification and Simple Queuing Service
9	10/30/2015	RESTFul WebServices. AWS APIs are sufficiently rich to allow you easy interaction with AWS service. However, in order to establish connectivity between your own modules in the Cloud you should use RESTFul Web Services
10	11/06/2015	Elastic Load Balancing and Auto Scaling allow automation of resource manipulation
11	11/13/2015	Introduction to Microsoft Cloud. Microsoft offers a set of resources and features that are of great utility to those who are restricted to programming in .Net Environment
12	11/20/2015	MapReduce performs large distributed computation as a set of distributed operations on data sets composed of key-value pairs producing a reduced set of key-value pairs. We will learn the basics of Hadoop, an open source implementation of MapReduce, and its Java API. Hadoop Distributed File System (HDFS) features
	11/27/2015	Thanksgiving Break
13	12/04/2015	HIVE is a convent SQL like language that allow us to write scripts that run atop of HDFS
14	12/11/2015	OPSWorks is an application management service that provides an integrated experience for overseeing the complete application lifecycle. With AWS OpsWorks, we can provision AWS resources, supervise their configuration, deploy applications to those resources, and monitor their health.
15	12/18/2015	Presentations of selected final projects