# 15-319: Introduction to Cloud Computing

Majd F. Sakr
Computer Science
Carnegie Mellon University

Spring 2010

## 1 Organization

**Instructors:**
Majd F. Sakr
msakr@qatar.cmu.edu
M1007, 454-8625
Tue, 3-5pm

**Teaching Assistant:**
Suhail Rehman, Office TBA, Phone TBA
suhailrehman@gmail.com

**Office Hours:**
To be decided in class and by appointment.

**Lecture:**
Tuesdays and Thursdays
10:30-11:50am
Room 2035

**Class Web Page:**
`http://www.qatar.cmu.edu/~msakr/15319-s10/`

**Class AFS space:**
`/afs/qatar.cmu.edu/course/15/319/`

## 2 Objectives

Our aim in CS 319 is to introduce you to the basics of the emerging cloud computing paradigm. It is important for you to learn how this paradigm came about and understand its enabling technologies. For this to happen we start by an overview of basic systems ideas as well as an introduction to parallel and distributed computing. It is important to cover parallel and distributed systems, their advantages and disadvantages. Understanding these systems is critical to understanding cloud computing systems. We want you to learn how the cloud is organized, provisioned and programmed. Further, we would like you to understand the computer systems constraints, tradeoffs and techniques in setting up and using the cloud to best serve the computing needs for different types of data and applications.

Why would a cloud computing system be preferable to existing computing systems? Cloud computing provides the flexibility to offer any platform that a user or company requires. Further,

it is highly scalable which means the resources can be expanded or shrunk based on current needs. This along with a pay-as-you-go economic model enables companies to have access to the platform of choice with the needed scale that is affordable to them. Further, the human resources required to maintain the cloud are reduced by the user or company utilizing these services. The data-center hosting the cloud will typically consolidate all requests on a large heterogeneous infrastructure to provide reliability, and in turn increase utilization as well as decrease its carbon footprint. This consolidation translates into efficient staffing needs for maintenance and backups. In summary, cloud computing offers the following advantages:

1. **Scalability** in terms of the resources. A company can start small and increase its hardware resources as it needs. This gives the illusion of having a large number of resources available on demand for the cloud user.
2. **Flexibility** in terms of the different software packages, multiple instance types, operating systems. As well as ease of system, application and data access from any networked computer.
3. **Pay-as-you-go** economic model borrowed from utility computing.
4. **Consolidation of system maintenance and management,** this overhead is shifted from the cloud users to its providers and is consolidated across many requests and systems.
5. **Reliability** where the system's fault tolerance is managed by the cloud providers and users no longer need to worry about it.
6. **High utilization & reduced carbon footprint** since typically a large number of custom servers is consolidated into a smaller number of shared servers.

Cloud services can be offered as three different flavors, platform as a service (PaaS) Software as a service (SaaS) or Infrastructure as a service (IaaS). An example of IaaS as cloud services is Amazon's Web Services (AWS).

By the end of the course you will understand the system perspective of the aforementioned advantages and will come to appreciate the benefits that cloud computing provides. You will also be able to design and implement parallel algorithms to efficiently distribute the intensive computation over the cloud machines and let them compute in parallel. You will get introduced to topics on parallel, distributed and large-scale data-intensive computing systems. You will have the foundation needed to match the future needs in this emerging programming paradigm.

# 3 Learning Outcomes

- To understand the emerging cloud computing paradigm, how it came about and how it relates to traditional models of computing.

- To understand the different technologies that enable cloud computing.

- To gain competence in Hadoop/MapReduce as a programming model for distributed processing of large datasets.

- To understand how different algorithms can be implemented and executed in the Hadoop framework.

- To gain competence in evaluating the performance and identifying bottlenecks when mapping applications to the cloud.

## 4 Textbook

The primary textbook for this course is:

- Tom White,
  *Hadoop: The Definitive Guide*, O'Reilly Media, 2009.

We have several reference books in the library covering most of the topics of the course. We will be mostly reading journal and conference publications on the subject.

In addition, it will be useful to have the following reference books:

- Tanenbaum and van Steen,
  *Distributed Systems: Principles and Paradigms*, Pearson, 2007.

- Jean Dollimore, Tim Kindberg, George Coulouris
  *Distributed Systems: Concepts and Design, Fourth Edition*,
  Addison Wesley, 2005.

- Randal E. Bryant and David R. O'Hallaron,
  *Computer Systems: A Programmer's Perspective*, Prentice Hall, 2003.

- Patterson and Hennessy,
  *Computer Organization and Design: The Hardware/Software Interface, Fourth Edition*,
  Morgan Kaufmann/Elsvier.

- Jason Venner,
  *Pro Hadoop*, Apress, 2009.

## 5 Course Organization

Your participation in the course will involve five forms of activity:

1. Attending and participating in the lectures

2. Assignments (including reading technical papers)

3. Projects

4. Exam

5. In-class Presentations and Discussions

Attendance will be taken at the beginning of each lectures, it will be worth 5% of your grade. You will be considered responsible for all material presented at the lectures.

# 6 Getting Help

For urgent communication with the teaching staff, it is best to send electronic mail (preferred) or to phone.

If you want to talk to a staff member in person, remember that our posted office hours are merely nominal times when we guarantee that we will be in our offices. You are always welcome to visit us outside of office hours if you need help or want to talk about the course. However, we ask that you follow a few simple guidelines:

- Prof. Sakr normally works with his office door open and welcomes visits from students whenever his doors are open. However, if his door is closed, he is busy with a meeting or a phone call and should not be disturbed.

We will use the Web as the central repository for all information about the class. The class home page is at

`http://www.qatar.cmu.edu/~msakr/15319-s10/`

We will use the following AFS folder for electronic data and assignment submission:

`/afs/qatar.cmu.edu/course/15/319/`

Using these locations, you can:

- Obtain copies of any handouts or assignments. This is especially useful if you miss class or you lose your copy.
- Find links to any electronic data you need for your assignments
- Read clarifications and changes made to any assignments, schedules, or policies.

Post messages to make queries about the course, specific labs, or exams by emailing the following dlist:

`15319-s10@lists.qatar.cmu.edu`

# 7 Policies

### Working Alone on Assignments/ Projects

You will work on all assignments/projects by yourself unless instructed by the teaching staff.

### Handing in Assignments/ Projects

All assignments/projects are due at 11:59pm (one minute before midnight) on the specified due date. All handins are electronic using the AFS file system:

`/afs/qatar.cmu.edu/course/15/319/handin/userid/`

## Making up Exams, Assignments and Projects

Missed exam more than 2 days late can be made up on a case by case basis, but only if you make prior arrangements with Prof. Sakr. However you should have a good reason for doing so. It is your responsibility to get your projects done on time. Be sure to work far enough in advance to avoid unexpected problems, such as illness, unreliable or overloaded computer systems, etc.

## Appealing Grades

After each exam and project is graded, you have seven calendar days to appeal your grade. Your appeal should be in writing.

If you are still not satisfied, please come and visit Prof. Sakr. If you have questions about an exam grade, please visit Prof. Sakr directly.

## Final Grade Assignment

Each student will receive a numeric score for the course, based on a weighted average of the following:

- **Participation:** Your attendance and participation in the different discussions held in class are going to account for 10% of your final grade.

- **Exam:** There will be one in-class exam, counting 15%.

- **Assignments:** The Assignments will count a combined total of 15% of your score. The assignments will involve readings, writing and in class discussions/presentations about different cloud-computing related papers and topics.

- **Projects:** The Projects will count a combined total of 60% of your score. There are 4 projects throughout the course. The first and second projects are worth 10% each of your final grade. The third project is worth 15%. The final project is worth 25% as it will involve a presentation and a paper as well as the project code. Since small differences in scores can make the difference between two letter grades, you will want to make a serious effort on each project.

| Grade Distribution | |
|---|---|
| Participation | 10% |
| Exam | 15% |
| Assignments | 15% |
| Projects | 60% |
| Total | 100% |

Grades for the course will be determined by absolute standards. The total score will be plotted as a histogram. Cutoff points are determined by examining the quality of work by students on the borderlines. Individual cases, especially those near the cutoff points may be adjusted upward or downward based on factors such as attendance, class participation, improvement throughout the course, exam performance, and special circumstances.

### Cheating

Each project must be the sole work of the student turning it in. Projects will be closely monitored by automatic cheat checkers, and students may be asked to explain any suspicious similarities with any piece of code available. The following are guidelines on what collaboration is authorized and what is not:

**What is Cheating?**

- *Sharing code or other electronic files*: either by copying, retyping, looking at, or supplying a copy of a file.
- *Sharing written assignments*: Looking at, copying, or supplying an assignment.

**What is NOT Cheating?**

- Clarifying ambiguities or vague points in class handouts.
- Helping others use the computer systems, networks, compilers, debuggers, profilers, or other system facilities.
- Helping others with high-level design issues.
- Helping others debug their code.

Be sure to store your work in protected directories.

The usual penalty for cheating is to be removed from the course with a failing grade. We also place a record of the incident in the student's permanent record.

## 8 Facilities: Qatar Cloud Computer System: Qloud

We have 14 Linux-based 64-bit servers specifically for 15-319, that we will use for all projects. The class Web page has details.

# 9 Class Schedule

Figure 1 shows the tentative schedule for the class. The reading assignments are all from journal, conference publications and books. The schedule also indicates the project activities. Any changes will be announced on the class distribution list (`15319-s10@lists.qatar.cmu.edu`). An updated schedule will be maintained on the class Web page.

| Class | Date | Day | Topic | Assignments | Projects | Lecturer |
|---|---|---|---|---|---|---|
| 1 | 12-Jan | Tue | Course Overview/ Introduction | Assignment A Out | | MFS |
| 2 | 14-Jan | Thu | Introduction to Cloud Computing | | Project 1 Out | MFS |
| 3 | 19-Jan | Tue | Cloud and Hadoop Demo | | | MFS |
| 4 | 21-Jan | Thu | Systems Review | Assignment A Due, Assignment B Out | | MFS |
| 5 | 26-Jan | Tue | Assignment Presentations and Discussion | | | MFS |
| 6 | 28-Jan | Thu | Parallel Processing I | | Project 1 Due, Project 2 Out | MFS |
| 7 | 2-Feb | Tue | Parallel Processing II | | | MFS |
| 8 | 4-Feb | Thu | Distributed Systems | | | MFS |
| 9 | 9-Feb | Tue | Paper Presentation and Discussion | Assignment B Due, Assignment C Out | | MFS |
| 10 | 11-Feb | Thu | Virtualization | | | MFS |
| 11 | 16-Feb | Tue | Distributed Storage | | | MFS |
| 12 | 18-Feb | Thu | MapReduce | | Project 2 Due, Project 3 Out | MFS |
| 13 | 23-Feb | Tue | Hadoop | | | MFS |
| 14 | 25-Feb | Thu | Paper Presentation and Discussion | Assignment C Due, Assignment D Out | | MFS |
| | 2-Mar | Tue | Spring Break | | | n/a |
| | 4-Mar | Thu | Spring Break | | | n/a |
| 15 | 9-Mar | Tue | Algorithm: Searching | | | MFS |
| 16 | 11-Mar | Thu | Algorithm: Sorting | | | MFS |
| 17 | 16-Mar | Tue | Exam | | | MFS |
| 18 | 18-Mar | Thu | Cloud Business Model I | | | Guest |
| 19 | 23-Mar | Tue | Cloud Business Model II | | Project 3 Due, Project 4 Out | Guest |
| 20 | 25-Mar | Thu | Guest Speaker | | | Guest |
| 21 | 30-Mar | Tue | Case Study Presentations | Assignment D Due, Assignment E Out | | MFS |
| 22 | 1-Apr | Thu | Guest Speaker | | | Guest |
| 23 | 6-Apr | Tue | Project 4 Status Report Presentations | | | MFS |
| 24 | 8-Apr | Thu | Guest Speaker | | | Guest |
| 25 | 13-Apr | Tue | Paper Presentation and Discussion | Assignment E Due | | MFS |
| 26 | 15-Apr | Thu | Guest Speaker | | | Guest |
| 27 | 20-Apr | Tue | Student Project Presentations | | | MFS |
| 28 | 22-Apr | Thu | Student Project Presentations | | Project 4 Due | MFS |

Figure 1: CS 319 Class Schedule