CS15-319: Cloud Computing

Lecture 1
Course Overview and Introduction
Prof. Majd F. Sakr
Our story…
Our Data Now…

Documents and Media

Personal Data

Emails, Calendars, Contacts, Location Information, etc…
We Live in a World of Data...
The World of Data

**Number of Emails Sent Every Second**: 2.9 million

**Data Consumed by Households Each Day**: 375 megabytes

**Video Uploaded to YouTube Every Minute**: 20 hours

**Data Per Day Processed by Google**: 24 petabytes

**Tweets Per Day**: 50 million

**Total Minutes Spent on Facebook Each Month**: 700 billion

**Data Sent and Received by Mobile Internet Users**: 1.3 exabytes

**Products Ordered on Amazon Per Second**: 72.9 items

*Sources: Cisco, comScore, Marketers, Radiant Group, PwC, YouTube*

*In the 21st Century, we live a large part of our lives online. Almost everything we do is reduced to bits and sent through cables around the world at light speed. But just how much data are we generating? This is a look at just some of the massive amounts of information that human beings create every single day.*
Big Data

- **Big data** is defined as large pools of data that can be captured, communicated, aggregated, stored, and analyzed.

- Data continues to grow:
  - In mid-2010, the information universe carried 1.2 zettabytes and 2020 predictions expect nearly 44 times more at 35 zettabytes coming our way.

- Applications are becoming *data-intensive*. 
What Do We Do With Data?

Store

Access

Encrypt

Share

Process

…. and more!

We want to do these seamlessly...
Using Diverse Interfaces & Devices

We also want to access, share and process our data from all of our devices, **anytime, anywhere!**
What About the Future?

How will you...

...work on documents?

...create, access, store and share media?

...get your news & info?

...navigate?

...communicate with friends and family?

...live in an intelligent home?

...
How Will We Manage Our Data?

Manage it ourselves?
- Personal, but time consuming.

How would you get access to your data wherever you are?
- Would you keep it on your devices?
- or would you keep it online?

What if it’s managed by someone else?
- and you can get this “service” for free or with a subscription?
Has this Happened Before?

Innovation  Product  Service
Think of it this Way …

+ Evolution of water Utility

Generate your own utility

Buy it as a product and manage it

Get a continuous supply of the utility through a dedicated connection
How About Electricity?

+ Transformation from a Product to a Service

- **Innovation**
  New Disruptive Technology

- **Product**
  Buy and Maintain the Technology

- **Service**
  Electric Grid, pay only for the electricity you use
...and Banking?

Evolution of Banking

- No Banks
  (Take care of your own money)

- Traditional Banking
  (Give your money to the bank)

- Banking Instruments
  (Cheques / Credit Cards)

- Internet Banking
  (...more services)
So What is Cloud Computing?
Can We Define Cloud Computing?

“Cloud Computing is the transformation of IT from a product to a service”
Cloud Computing

Transformation of IT from a Product to a Service

Innovation of IT
New Disruptive Technology

IT Products
Buy and Maintain the Technology

Cloud Computing
On-Demand IT services on a Pay-as-You-Go basis
So... how would you transform information technology into a Service?
Requirements to Transform IT to a Service

- Connectivity
  - For moving data around
- Interactivity
  - Seamless interfaces
- Reliability
  - Failure will affect many people, not just one
- Performance
  - It should not be slower or less efficient than what people already have
- Pay-as-you-Go
  - Should not pay an upfront fee for the service

- Ease of Programmability
  - Ease of development of complex services to users
- Manage Large Amounts of Data
  - Big Data
- Efficiency
  - Cost
  - Power
- Scalability & Elasticity
  - Flexible and rapid response to changing user needs
Requirements to Transform IT to a Service

- Connectivity
  + For moving data around
- Interactivity
  + Seamless interfaces
- Reliability
  + Fault-Tolerance
- Performance
  + Parallel/Distributed Programming
- Pay-as-you-Go
  + Utility Computing

- Ease of Programmability
  + Easy development of complex services to users
- Programming Model
- Manage Large Amounts of Data
  + Big Data
- Storage Technologies
- Efficiency
  + Cost
  + Power
- Scalability & Elasticity
  + Flexible and rapid response to changing user needs
- Virtualization Technologies

© Carnegie Mellon University in Qatar
Combine the Enabling Technologies…
Cloud Computing

+ Think of it as Internet Computing
+ Computation done over the Internet

Enabling Technologies
- High Bandwidth
- High Speed Internet
- Virtualization
- Utility Computing
- …
Cloud Computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices, as a metered service over a network.
Why Cloud Computing?

**Pay-as-You-Go** economic model
- Reduce capital expenditure
- No upfront cost
- Reduced Time to Market

**Simplified IT management**
- All you need is access to the internet.
- It’s the providers responsibility to manage the details.

**Scale quickly and effortlessly**
- Resources can be rented and released as required
- Software Controlled
- Instant scalability

**Flexible options**
- Configure software packages, instance types operating systems.
- Any software platform
- Access from any machine connected to the Internet

**Resource Utilization is improved**
- Reduce Idle resources by sharing and consolidation
- Better utilization of CPU / Storage and Bandwidth.

**Carbon Footprint decreased**
- Sharing of resources means less servers, less power and less emissions.

© Carnegie Mellon University in Qatar
Applications Enabled by Cloud Computing

- High Growth Applications
  - Startup Businesses
- Aperiodic Bursting Applications
  - Seasonal Businesses
- On-Off Applications
  - Research Computing
- Periodic Applications
  - Changing computational patterns over time
High Growth Applications

What do you do when your startup gains traction?

Can you grow quick enough?

friendster® vs. facebook

Could not keep up with the growth of their number of users.

Growing exponentially
High Growth Applications

Animoto’s Facebook Plugin doubled traffic to the site every 12 hours for 3 days. They could scale from 50 servers to 3500 and go back down using cloud computing services.

Users use it to produce video pieces from their photos, video clips and music.
Aperiodic Bursting Applications

- Website went down on 9/11/2001 due to traffic
- February 14th – Busiest Day of the Year
- US Holiday Season
- Website crashed within 10 minutes of the free trouser promotion during Superbowl 2010

Even if you design your website infrastructure to handle peak loads, won't it be idle during other times?
On-Off Applications

Researchers running large-scale scientific simulation using 1000s of computers.

Modern Drug Discovery

- Data-intensive simulation and tests to discover new compounds
- Large compute power required for simulation jobs
- Time to market is crucial

Why not rent computer time to run these simulations?
Periodic Applications

Sock Market Analysis

- Different computational requirements over time
- Mine market data during the day.
- Process and Analyze at night.

Dynamic and Flexible infrastructure can reduce costs and improve performance.
Technical Challenges

+ Programming is tricky but improving
+ Tools are continuously evolving
+ Moving large data is still expensive
+ Security
+ Quality of Service
+ Green computing
+ Internet Dependence
Non-Technical Challenges

+ Vendor Lock-In
+ Non-standardized
+ Security Risks
+ Privacy
+ Legal
+ Service Level Agreements
In 15-319

Find out what all the fuss is about!

Get a firm theoretical foundation behind cloud computing

Practical experience with real world tools and applications

A sought after skill in Industry and Research!
Course Objectives

1. Learn the core concepts of cloud computing
2. Hands-on experience in using cloud computing infrastructure
3. Work on a large research project in cloud computing
## Learning Outcomes

<table>
<thead>
<tr>
<th>LO1</th>
<th>Explain, apply and identify core concepts and emerging issues in the cloud computing paradigm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO1.1</td>
<td>Explain how and why this paradigm came about and the influence of several enabling technologies including programming models (e.g., MapReduce), virtualization (e.g., Xen and Vmware), distributed file systems and cloud storage (e.g., HDFS and HBase), and emerging cloud tools (e.g., Hive).</td>
</tr>
<tr>
<td>LO1.2</td>
<td>Examine the process of working on a large research project under the mentorship of a teaching staff member.</td>
</tr>
<tr>
<td>LO1.3</td>
<td>Identify some of the emerging cloud research challenges including cloud security, QoS, SLAs, and energy-efficient clouds.</td>
</tr>
</tbody>
</table>

- **Considered:** a reasonably critical and comprehensive perspective.
- **Thoughtful:** Fluent, flexible and efficient perspective.
- **Masterful:** a powerful and illuminating perspective.
Syllabus

+ Contract!
+ All you need to know about the course
  + Course Objectives
  + Learning Outcomes
  + Policies
  + Grading
  + Tentative Schedule
Target Audience

- CS Majors
- Juniors / Seniors
- Pre-requisites:
  - 15-213 – Introduction to Computer Systems
Instructors / Getting Help

Prof. Majd F. Sakr
msakr@qatar.cmu.edu
(2121)

Dr. Mohammad Hammoud
mhhammou@qatar.cmu.edu
(1013)

Dr. Fan Zhang
fanzhan1@qatar.cmu.edu
(1206)

Suhail Rehman
Feb 5th onwards
(2044)
Course Organization

Lectures

Projects

Project Status Reports and Presentations

In-Class Discussions
<table>
<thead>
<tr>
<th>Week</th>
<th>Session</th>
<th>Date</th>
<th>Topic</th>
<th>Teaching Method</th>
<th>Instructor</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>16 Jan</td>
<td>Administrivia and Course Motivation</td>
<td>Lecture</td>
<td>MFS</td>
<td>Project Start</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>18 Jan</td>
<td>Introduction to Cloud Computing and Cloud Software Stack</td>
<td>Lecture</td>
<td>MFS</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>23 Jan</td>
<td>Course Project and Amazon AWS</td>
<td>Lecture</td>
<td>MFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>25 Jan</td>
<td>Principles of Parallel Programming</td>
<td>Lecture</td>
<td>MFS</td>
<td>Project Update</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>30 Jan</td>
<td>Introduction to Programming Models</td>
<td>Lecture</td>
<td>MFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1 Feb</td>
<td>Introduction to MapReduce</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>6 Feb</td>
<td>MapReduce Algorithms</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8 Feb</td>
<td>Apache Mahout</td>
<td>Guest Lecture</td>
<td>Shannon</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>13 Feb</td>
<td>Pregel, Dryad, GraphLab</td>
<td>Lecture</td>
<td>MHH</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>15 Feb</td>
<td>Major Project Milestone Presentations</td>
<td>Project Presentation</td>
<td>MHH</td>
<td>Project Phase I End, Paper V1 Due, Project Phase II Start</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>20 Feb</td>
<td>Introduction to Cloud Storage and Distributed File Systems (DFSs)</td>
<td>Lecture</td>
<td>MFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>22 Feb</td>
<td>Cloud Storage and DFSs—HDFS, PVFS</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>27 Feb</td>
<td>Cloud Storage and DFSs - BigTable/HBASE</td>
<td>Lecture</td>
<td>MHH</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>29 Feb</td>
<td>Apache Pig, Hive, Zookeeper</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>12 Mar</td>
<td>Introduction to Virtualization</td>
<td>Lecture</td>
<td>MHH</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>14 Mar</td>
<td>Resource Virtualization (CPU/Memory/Disk/Network)</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>19 Mar</td>
<td>Virtualization Case Study: Xen</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>21 Mar</td>
<td>Virtualization Case Study: VMware</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>26 Mar</td>
<td>Benchmarking and Application Characterization</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>28 Mar</td>
<td>Major Project Milestone Presentations</td>
<td>Project Presentation</td>
<td>MHH</td>
<td>Project Phase II End, Project Phase III Start</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>2 Apr</td>
<td>Cloud Monitoring and Diagnostics</td>
<td>Lecture</td>
<td>MSR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4 Apr</td>
<td>Cloud Security</td>
<td>Lecture</td>
<td>Thierry</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>9 Apr</td>
<td>Cloud QoS and SLAs</td>
<td>Lecture</td>
<td>MFS</td>
<td>Poster Due</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>11 Apr</td>
<td>Green Cloud</td>
<td>Lecture</td>
<td>MFS</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>16 Apr</td>
<td>How to Write a Scientific Paper</td>
<td>Lecture</td>
<td>MFS</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>18 Apr</td>
<td>Guest Lecture</td>
<td>Lecture</td>
<td>MHH</td>
<td>Project Update</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>23 Apr</td>
<td>Final Project Presentations I</td>
<td>Project Presentation</td>
<td>MHH</td>
<td>Project End</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>28</td>
<td>Final Project Presentations II</td>
<td>Project Presentation</td>
<td>MHH</td>
<td>Project End</td>
</tr>
</tbody>
</table>

© Carnegie Mellon University in Qatar
Primary Textbook

Tom White, O'Reilly Media, 2010
Reference Books

• Virtual Machines: Versatile Platforms for Systems and Processes
  James E. Smith and Ravi Nair, MorganKauffman, 2005

• Programming Amazon EC2
  Jurg van Vilet and Flavia Paganelli, O'Reilly Media, 2011

• The Cloud at Your Service
  Jothy Rosenberg and Arthur Mateos, Manning Publications, 2010

• Mahout in Action
  Sean Owen, Robin Anil, Ted Dunning and Ellen Friedman, Manning Publications, 2011

• Hadoop in Action
  Chuck Lam, Manning Publications, 2011
Projects

+ This is a project-based Course
+ Select a domain and a mentor
  + NLP, Image Processing, DNA Sequencing
  + Dr. Hammoud, Mr. Rehman, Dr. Zhang
+ Three-phase project timeline.
  + Phase 1 – Introduction, Ramp-up, Timeline
  + Phase 2 – Application Implementation
  + Phase 3 – Characterization and Analysis
+ You will be using a Public Cloud Service
  + Amazon EC2
Overall Project Timeline

Project Start
- Amazon EC2 Introduction
- Run simple MapReduce Job
- Plan for Phases 2 & 3

Phase 1 Complete
Phase 2 Start
- Project Implementation
- Follow timeline set in project start
- Coding and Implementation of your project

Phase 2 Complete
Phase 3 Start
- Submit Application Implementation Code and Documentation
- Start Application Characterization and Analysis Phase

Project Completion
- Final Reports
- Final Presentation
- Paper / Poster Feasibility and Planning

In Addition:
Project Update Presentation every Wednesday.

© Carnegie Mellon University in Qatar
Special Note on Amazon EC2

+ Paid Cloud Service – you are billed by the hour.
+ Start a resource only when you need them.
+ Terminate resources as soon as you are done with them.
## Grading

<table>
<thead>
<tr>
<th>Course Elements</th>
<th>#</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Phases I, II &amp; III</td>
<td>3</td>
<td>75%</td>
</tr>
<tr>
<td>Project Update Presentations</td>
<td>6</td>
<td>20%</td>
</tr>
<tr>
<td>Class Participation and Attendance</td>
<td>28</td>
<td>5%</td>
</tr>
</tbody>
</table>
Questions?