

CS15-319 / 15-619

Cloud Computing

Recitation 1

Course Overview and Introduction

January 13 & 15 2015

<http://www.cs.cmu.edu/~msakr/15619-s15/>

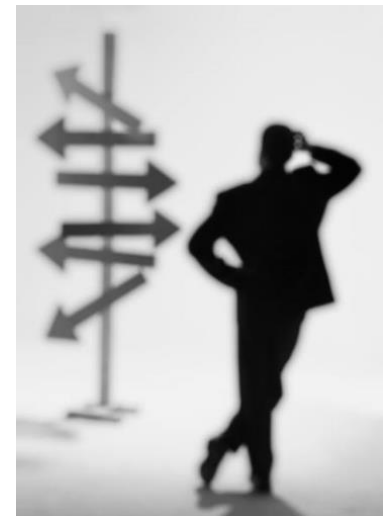
Outline

- **What is the course about?**
- What is an online course?
- Administrivia
- A couple of demos

**So What is Cloud
Computing?**

Data and Decision Making

- Analyzing data reflects reality
- Walmart: hurricane warning
 - Stock beer and strawberry pop-tarts
 - 7x increase in sales during large storms
- Government: resource allocation decisions
 - Data mining in Maryland → crime hotspots
 - Shuffle resource allocation, more to hotspots
 - violent crime down by 25%
 - \$20 million saved in the city of Baltimore



Data Science

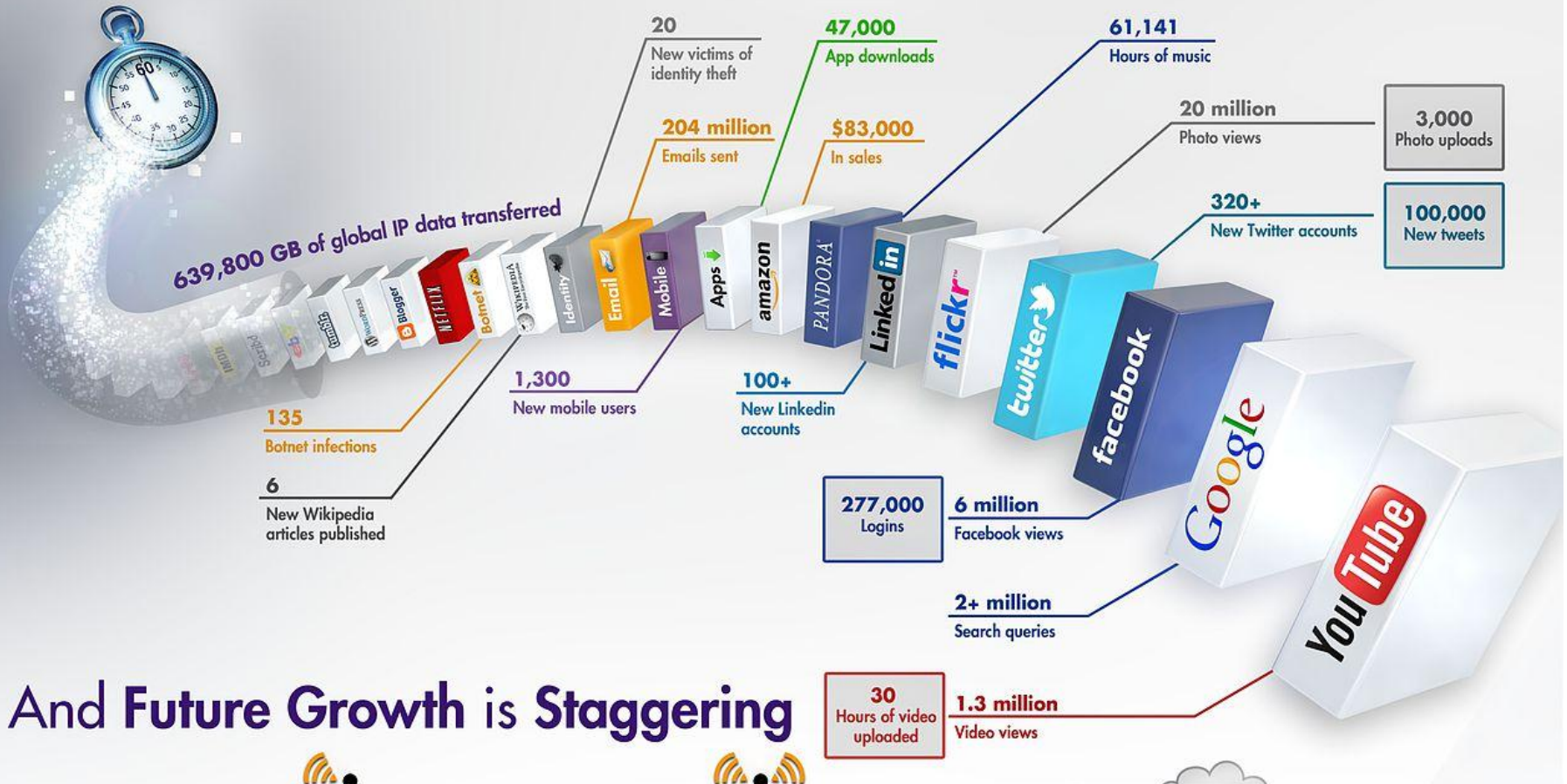
- Extraction of knowledge from data
- Employs statistical, machine learning and data mining techniques
 - Look for trends, patterns or anomalies in the data
- Affects research in many domains
 - Business, Economics and Finance
 - Biological Sciences and Bioinformatics
 - Social Sciences and Humanities
 - ...

An Increase in Data Capture

- Physical Sensors and Sensor Networks
 - Environmental, safety, transportation
- Social Media Interactions
 - Facebook, Twitter, Instagram
- Public Video and Image Capture
 - Surveillance, mobile phones, ...
- Customer Spending Habits
 - Loyalty programs and purchase data



What Happens in an Internet Minute?



And Future Growth is Staggering



Source: IntelFreePress

What is Big Data?

- Big Data
 - Volume, Velocity, Variety, Veracity
 - Data of next year >> data of this year
- Many Challenges
 - Store, share, analyze, search, transfer, visualize, and secure
- Traditional IT systems are insufficient

we need...

Large
Scale
Systems

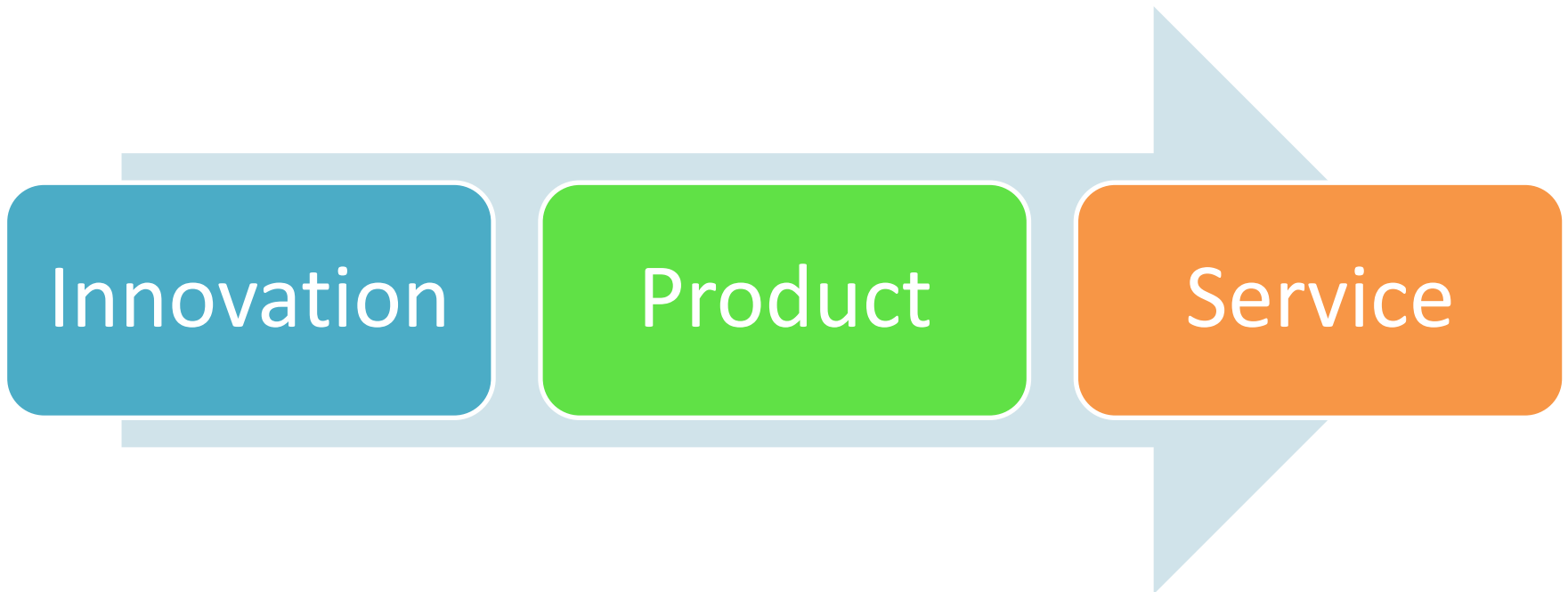
Large Scale System Challenges

- Lengthy procurement cycles
- Lengthy deployment effort
- Costly power and cooling
- Costly systems administration
- Low utilization
- Costly disaster recovery



Evolution of Computing

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



Evolution of Electricity



Innovation

New Disruptive
Technology



Product

Buy and Maintain
the Technology



Service

Electric Grid, pay
for what you use



A Cloud is ...

- Datacenter hardware and software that the vendors use to offer the computing resources and services



The Cloud



... for a more complete definition!



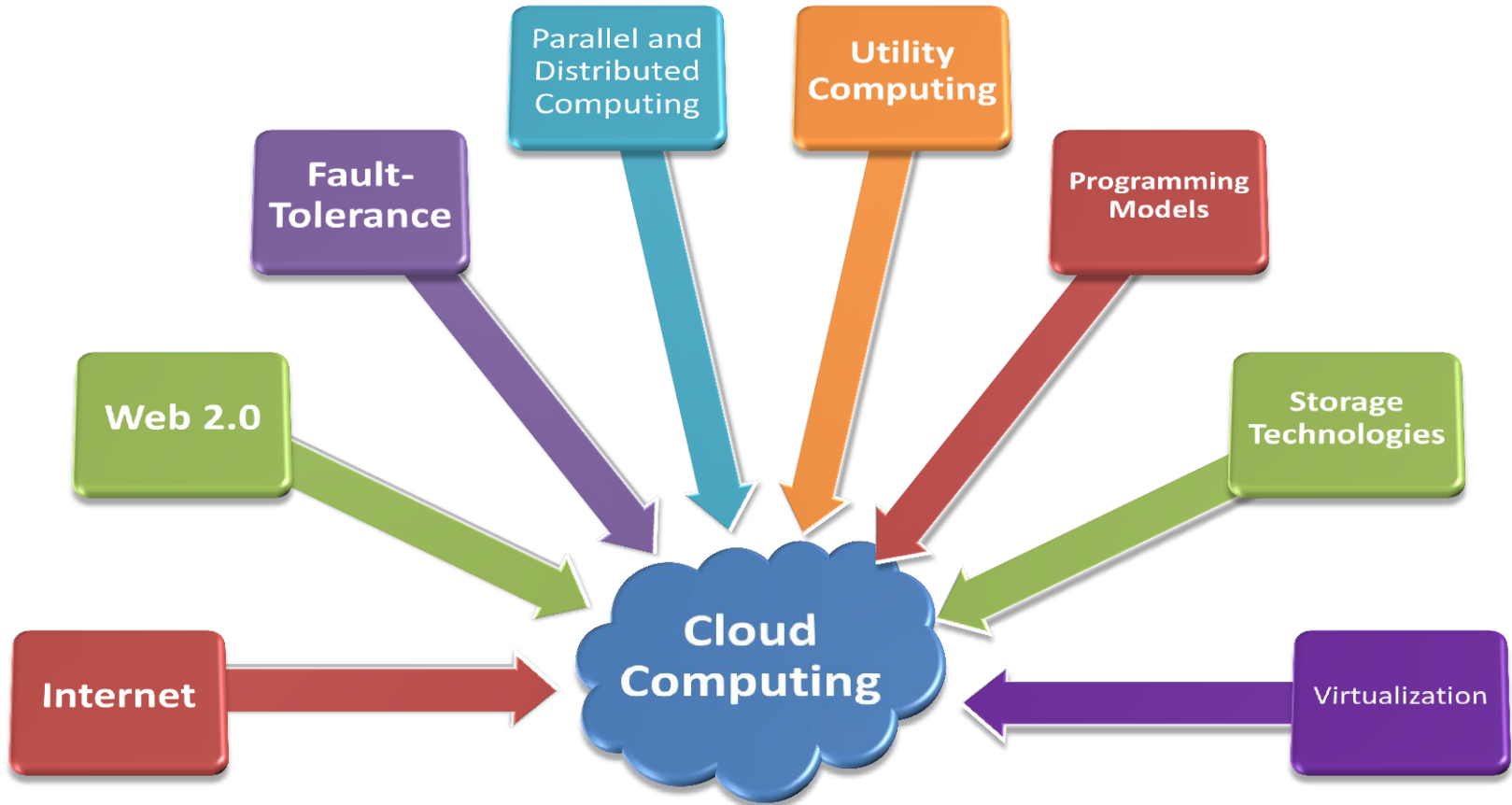
Cloud Computing is the delivery of computing as a **service** rather than a **product**,

whereby **shared resources, software, and information** are provided to computers and other devices,



as a **metered service** over a **network**.

Enabled by Maturing Technologies

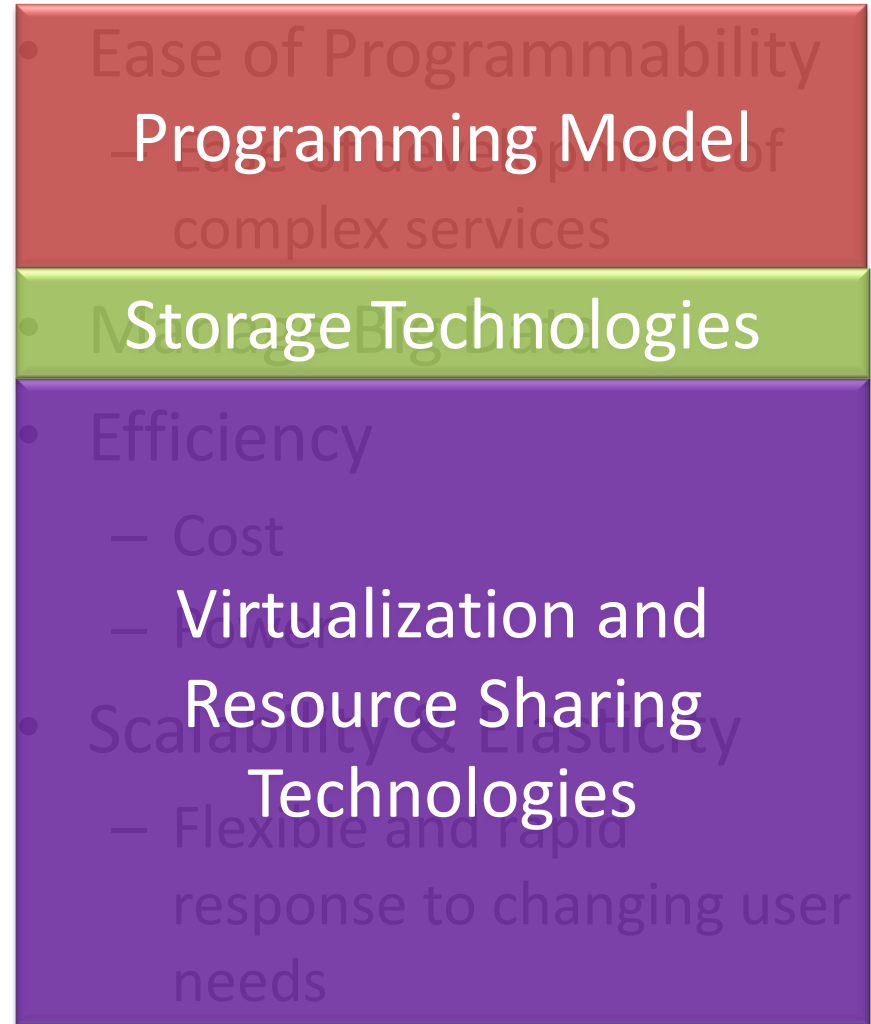
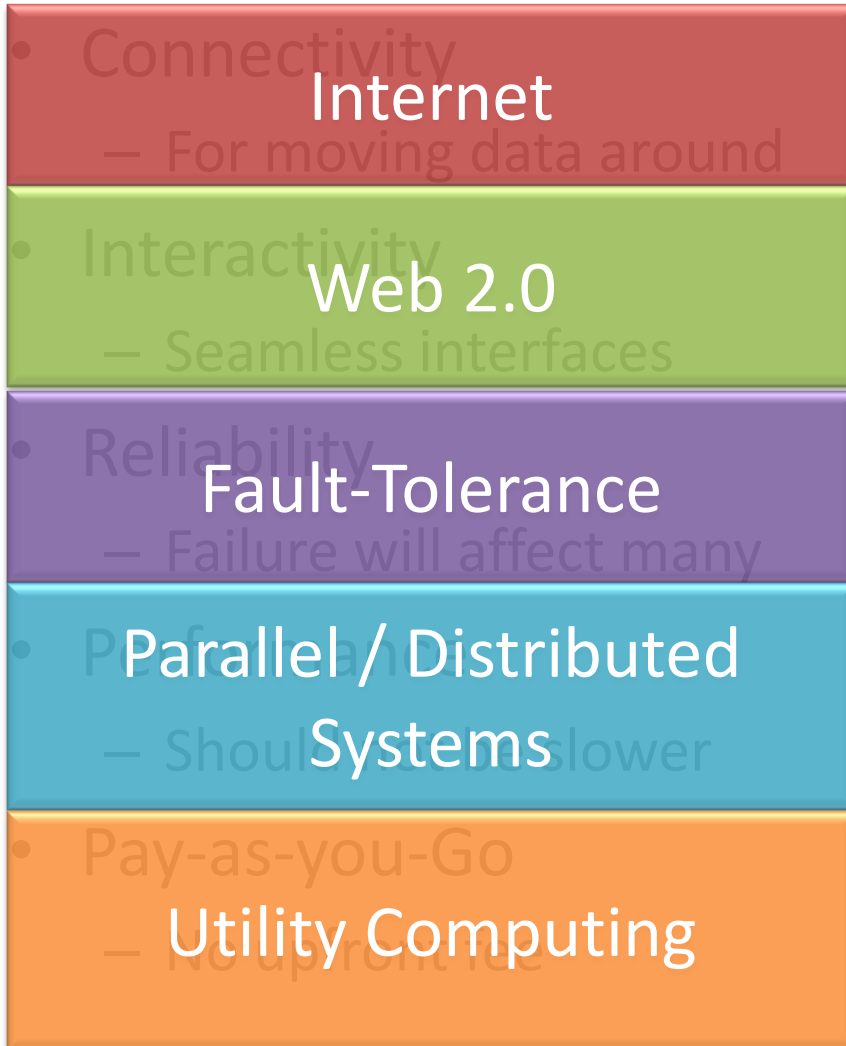


**So... how would you transform
information technology into a
Service?**

How to Transform IT to a Service?

- Connectivity
 - For moving data around
- Interactivity
 - Seamless interfaces
- Reliability
 - Failure will affect many
- Performance
 - Should not be slower
- Pay-as-you-Go
 - No upfront fee
- Ease of Programmability
 - Ease of development of complex services
- Manage Big Data
- Efficiency
 - Cost
 - Power
- Scalability & Elasticity
 - Flexible and rapid response to changing user needs

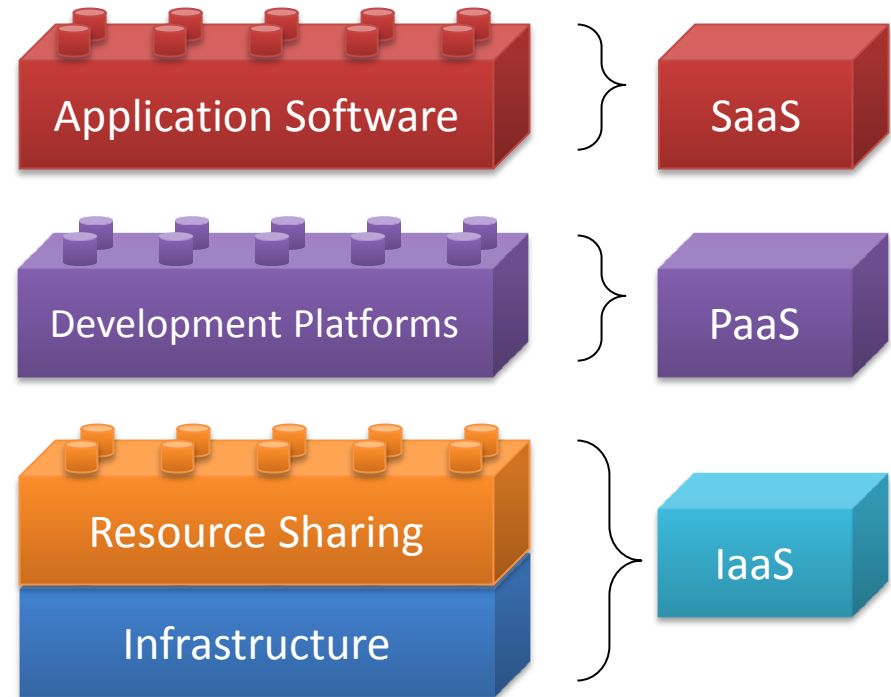
How to Transform IT to a Service?



Cloud Building Blocks

Cloud services are available in various forms, corresponding to the layer of abstraction desired by the user

- Software as a Service (**SaaS**)
- Platform as a Service (**PaaS**)
- Infrastructure as a Service (**IaaS**)



Software as a Service (SaaS)

- Software is delivered through the internet over a browser or mobile application
- Replace desktop software with cloud-based versions
- Webmail, Productivity Software, ERP, CRM etc.
- Centrally managed, globally available, automatically updated



Adobe® Creative Cloud



Platform as a Service (PaaS)

- Tools and APIs to develop and deploy cloud-based applications
- Create customized SaaS in the form of Web or mobile applications

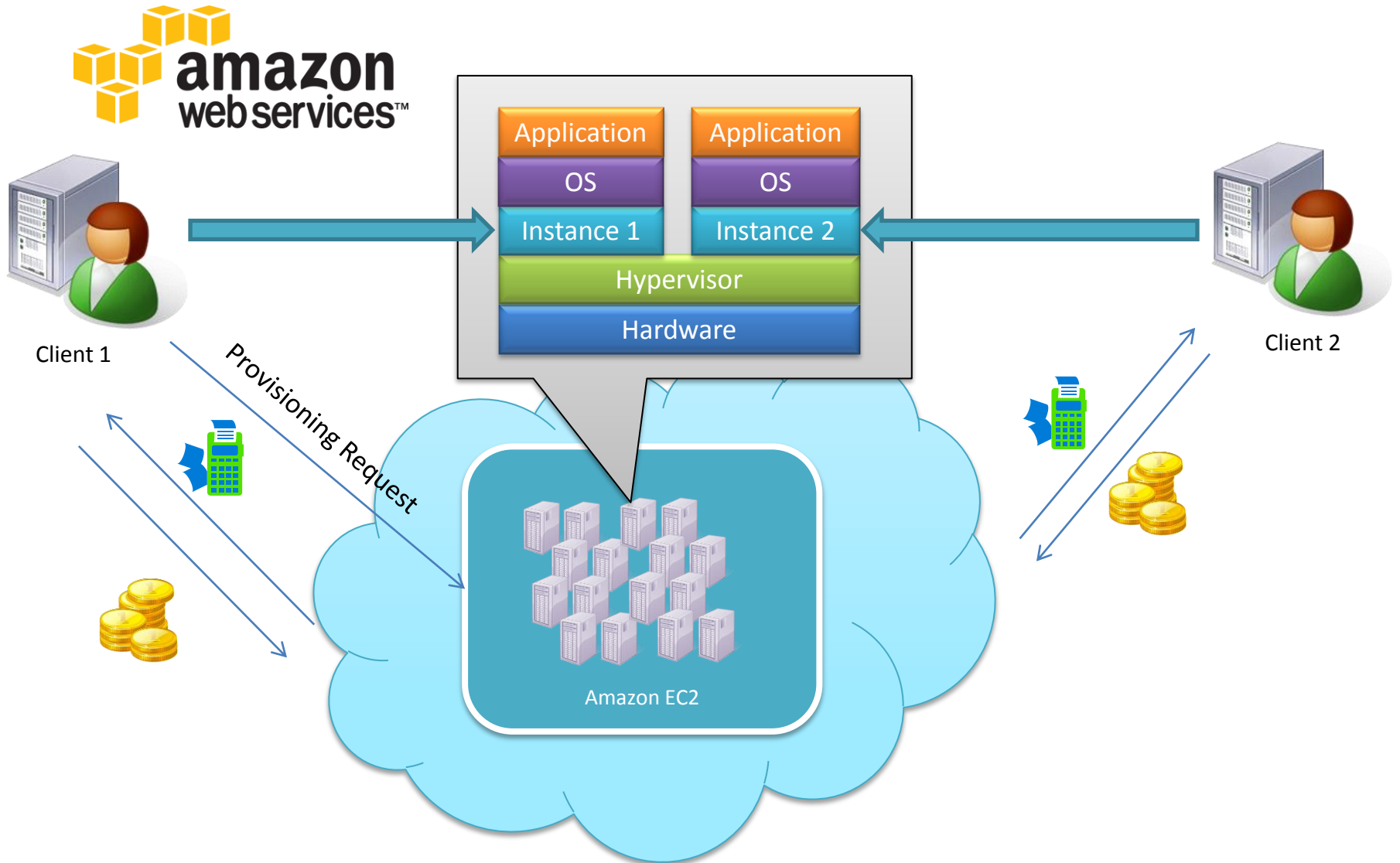


Infrastructure as a Service (IaaS)

- Compute, storage and network resources bundled in the form of virtual machines
- Fully flexible in terms of software and environment



Infrastructure as a Service



Benefits of Cloud Computing



Pay-as-You-Go
economic
model



Simplified IT
management



Elasticity
Scale quickly
& effortlessly



Customization
Flexible
options



Carbon
Footprint
decreased



Risks and Challenges



Migration



**Security &
Privacy**



**Vendor
Lock-In**



Legal



**Internet
Dependence**



Service Level Agreements and Objectives (SLA/SLO)

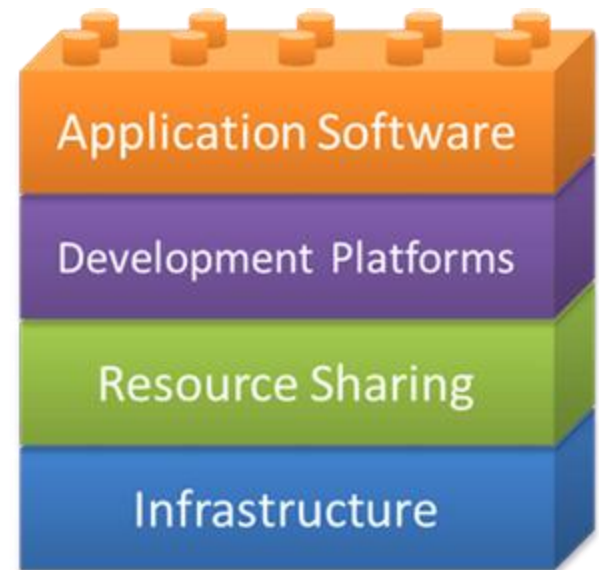
- SLA: Contract between cloud providers and users to define expected service
 - Service availability and delivery
 - Payment terms, bonuses and penalties for service
- SLO: Individual performance/service metrics regarding service delivery defined in the SLA
- Auditing: monitor resources to enforce SLOs and SLAs

Cloud Use Cases: Start-ups

- Infrastructure on demand
- Save money on data center real estate, servers, power and cooling
- Saving in capital expenditure which could be used to drive other areas of business growth
- Scale infrastructure as the business grows
- Levels the infrastructure playing field with established companies

Cloud Computing

- Applications
- Development Platforms
- Elasticity
 - APIs to enable automation, Alarms, protocols, triggers, etc...
- Sharing mechanisms
 - Virtualization, SDX, ...
- Distributed systems
 - Programming models
 - Storage
- Data centers

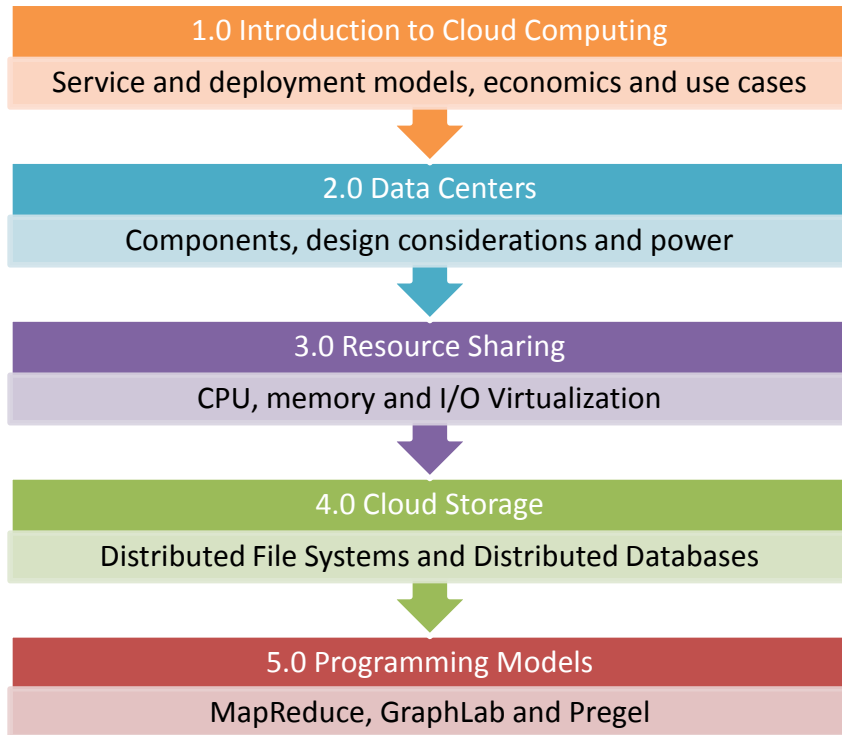


What is this course about?

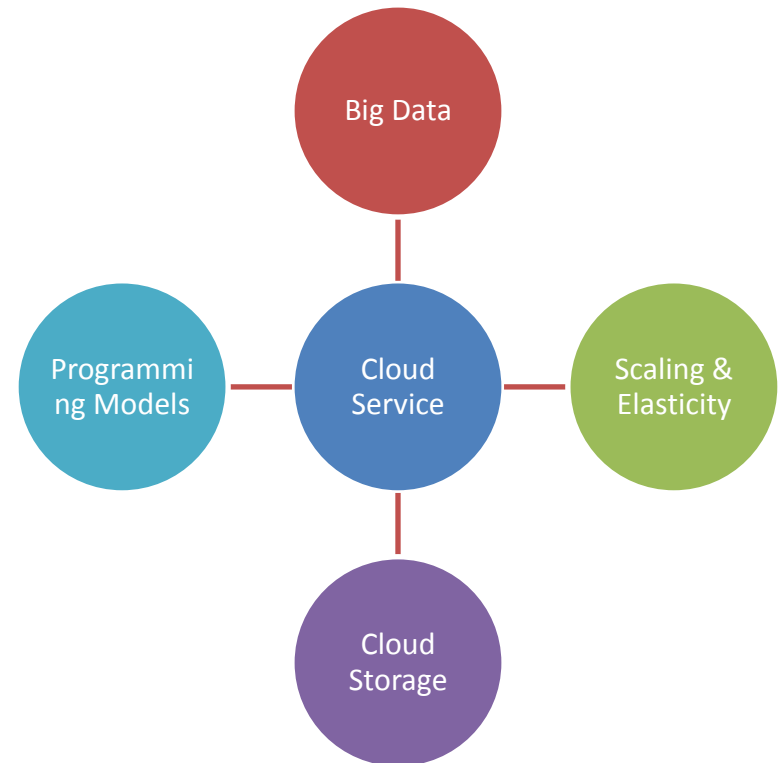
- Applied aspects of cloud computing
 - Between systems and services



Online content on OLI



Projects on AWS



Course Objectives

Students will learn:

- the fundamental ideas behind **Cloud Computing**;
- the basic ideas and principles in **data center** design and management;
- the resource sharing and **virtualization** techniques that serve in offering software, computation and storage services on the cloud;
- about **cloud storage** technologies and relevant distributed file systems;
- the variety of **programming models** and develop working experience in one of them.

Units

Unit #	Title	Modules and Description
1	Introduction	Introduction to Cloud Computing Building Blocks and Service Models in Cloud Computing Topics in Cloud Security Popular Cloud Stacks and Use Cases
2	Data centers	Historical Perspective Datacenter Components Design Considerations Power Calculations Software Defined Networks and Storage (SDN & SDS)
3	Virtualization	Resource Abstraction Resource Sharing (CPU, Memory, I/O) Sandboxing Case Study: Amazon EC2
4	Cloud Storage	Introduction to Storage Systems Cloud Storage Concepts Distributed File Systems Cloud Databases Case Study: Amazon Object Storage
6	Programming Models	Introduction to Programming Models Variety of Programming Models Case Studies: MapReduce, Spark, GraphLab

Projects

- Four Projects (all students):
 0. Primer (Complete by Sunday, **January 18, 2015**)
 1. Big Data
 2. Scalability and Elasticity
 3. Cloud Storage
 4. Programming Models
- 15-619 Project (extra 3-units)
 - One multi-week team project to build a complete web service

What this course is *not* about

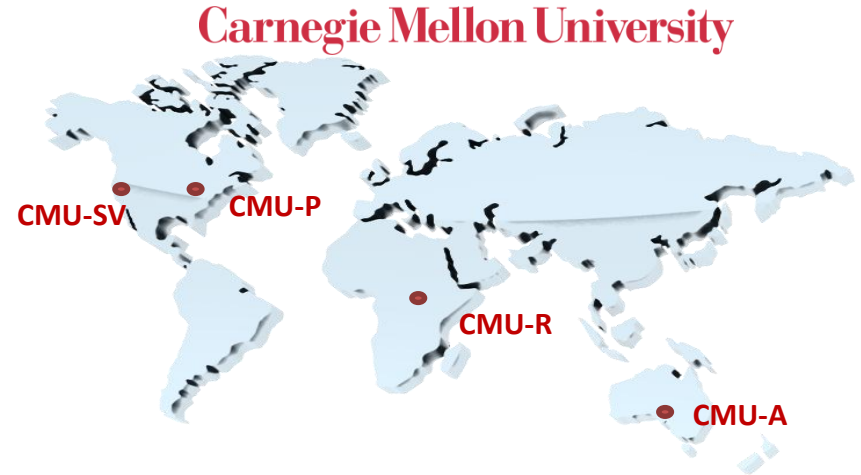
- Building Cloud Stack Modules
 - OpenStack
- Cloud Software Development
 - SaaS software engineering
- Distributed Systems
 - Synchronization, Consistency, ...
- Operating Systems
 - Developing a hypervisor
- Networks
 - Routing and switching protocols

Outline

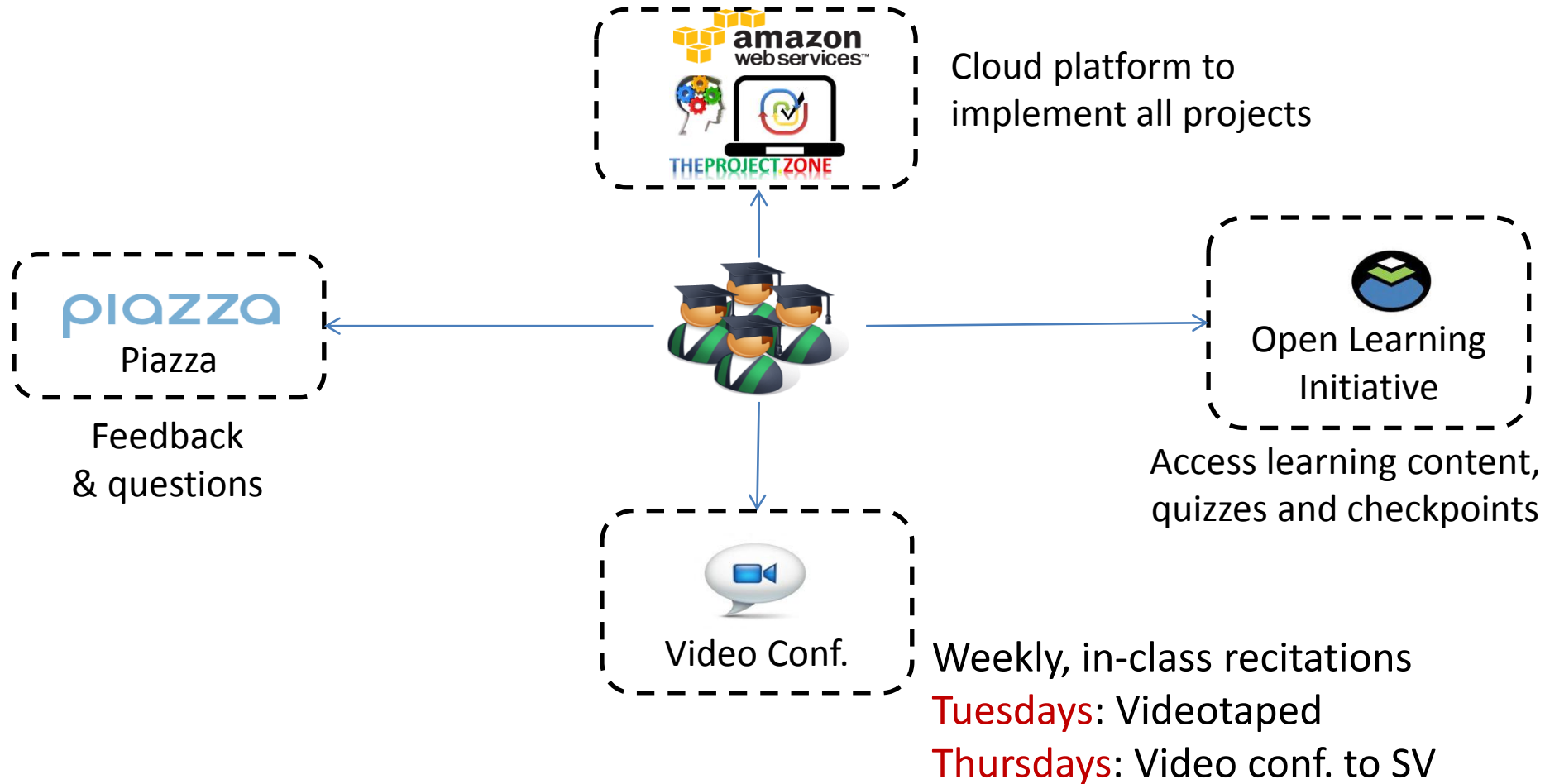
- What is the course about?
- **What is an online course?**
- Administrivia

Carnegie Mellon Global Course

Location	Students	Teaching Staff
CMU Pittsburgh	182	20
CMU Silicon Valley	16	1
CMU Rwanda	9	1
CMU Adelaide	18	1

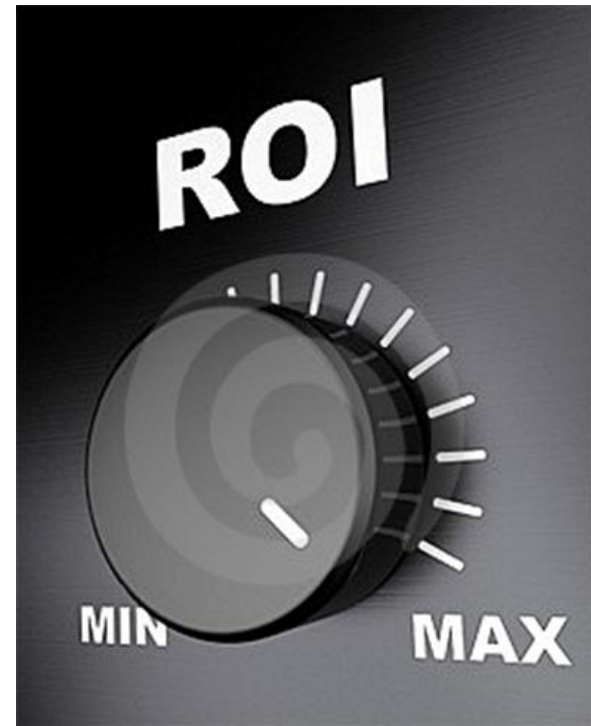


Online Course Engagement Model



Expectations

- Real world practical experience
 - Learn on your own
 - Languages, API, debugging
 - Overcome challenges
 - Deal with uncertainty
- Self paced learning
- Using experimental tools
 - Bleeding edge comes with risks



Outline

- What is the course about?
- What is an online course?
- **Administrivia**

Target Audience

- Technical Majors
- Undergraduate Juniors / Seniors
 - Pre-requisites:
 - 15213 – Introduction to Computer Systems
- Graduate Students
 - Experience:
 - Unix, scripting, python, & java

Course Organization



Course Units



Course Projects



Weekly Recitations



Office Hours



Discussions on Piazza

Getting Help

- TAs in Adelaide, Rwanda, Pittsburgh & Silicon Valley
- Piazza
 - Email does not scale
 - Discussion forum to support each other
- Recitations
 - Tuesdays (recorded)
 - At 8AM in GHC 4307
 - Thursdays (video conferenced to SV)
 - At 4:30PM in GHC 4307 (SV 212)
- Office Hours
 - Check Piazza for Office Hour schedule

Teaching Staff

Majd F. Sakr



msakr@cs.cmu.edu

GHC 7006

Office Hours:
Tuesdays, 3-4pm

Pittsburgh: Teaching Assistants

- Anshima Gupta



Pittsburgh: Teaching Assistants

- Chao Zhang



Pittsburgh: Teaching Assistants

- Debjani Biswas



Pittsburgh: Teaching Assistants

- Eryue Chen



Pittsburgh: Teaching Assistants

- Gongxun Liu



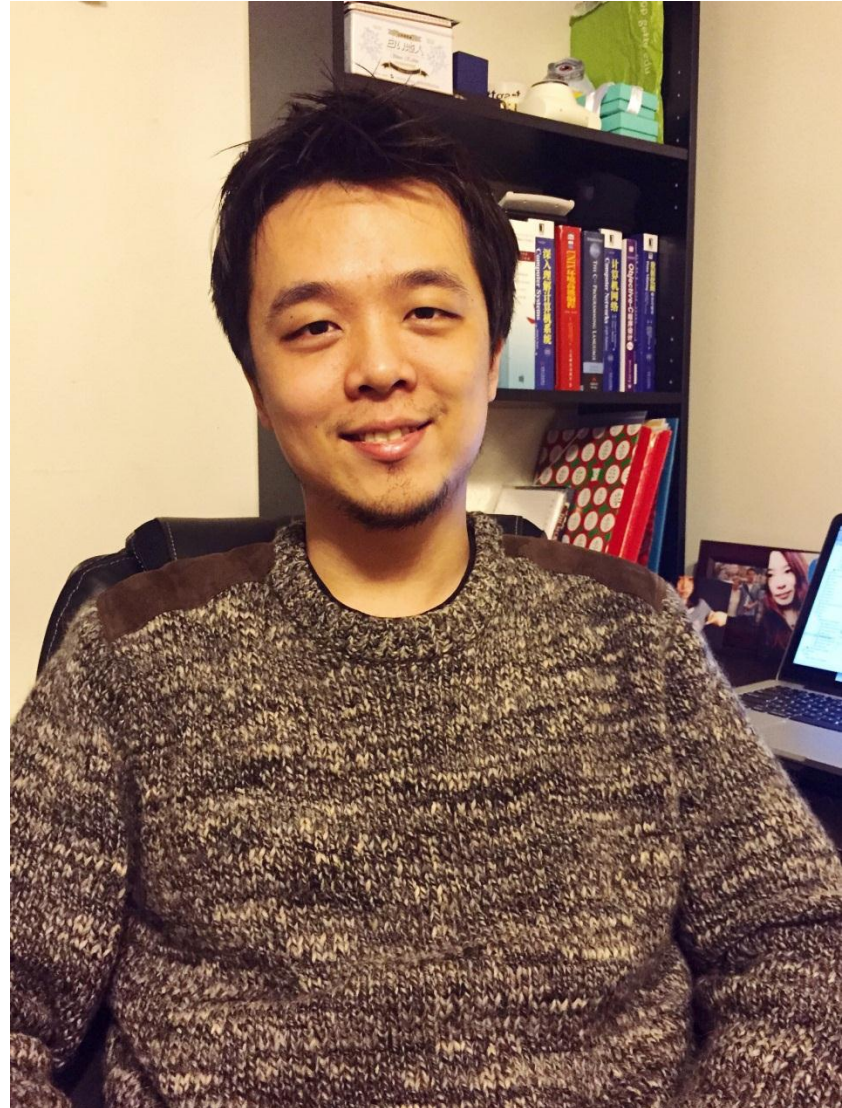
Pittsburgh: Teaching Assistants

- Haoliang Quan



Pittsburgh: Teaching Assistants

- Jiaduo He



Pittsburgh: Teaching Assistants

- Lei Wang



Pittsburgh: Teaching Assistants

- Lu Zeng



Pittsburgh: Teaching Assistants

- Mayank Singh Shishodia



Pittsburgh: Teaching Assistants

- Mengyu Yang
(Rainy)



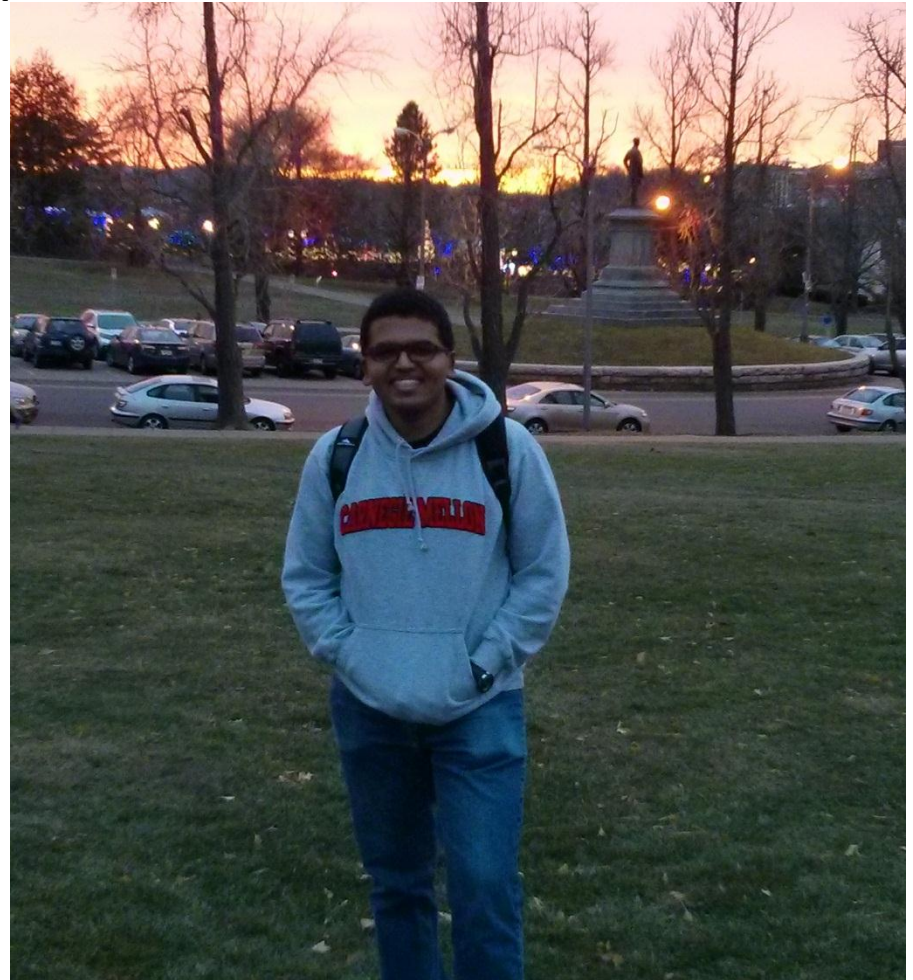
Pittsburgh: Teaching Assistants

- Mrigesh Kalvani



Pittsburgh: Teaching Assistants

- Prajwal Yadapadithaya



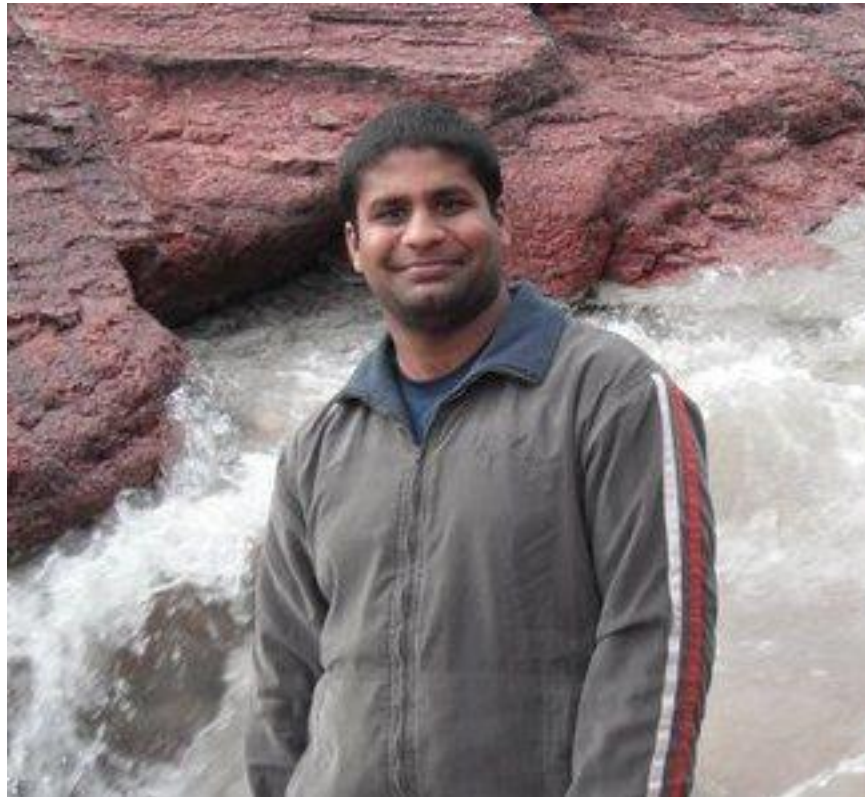
Pittsburgh: Teaching Assistants

- Pan Sun



Pittsburgh: Teaching Assistants

- Ravi Chandra Bandlamudi Venkata



Pittsburgh: Teaching Assistants

- Rohit Upadhyaya



Pittsburgh: Teaching Assistants

- Ru Jia



Pittsburgh: Teaching Assistants

- Satya Venkata Kamuju (Durga)



Pittsburgh: Teaching Assistants

- Siyuan Zhou



Pittsburgh: Teaching Assistants

- Suhail Rehman



Pittsburgh: Teaching Assistants

- Tianqi Wen



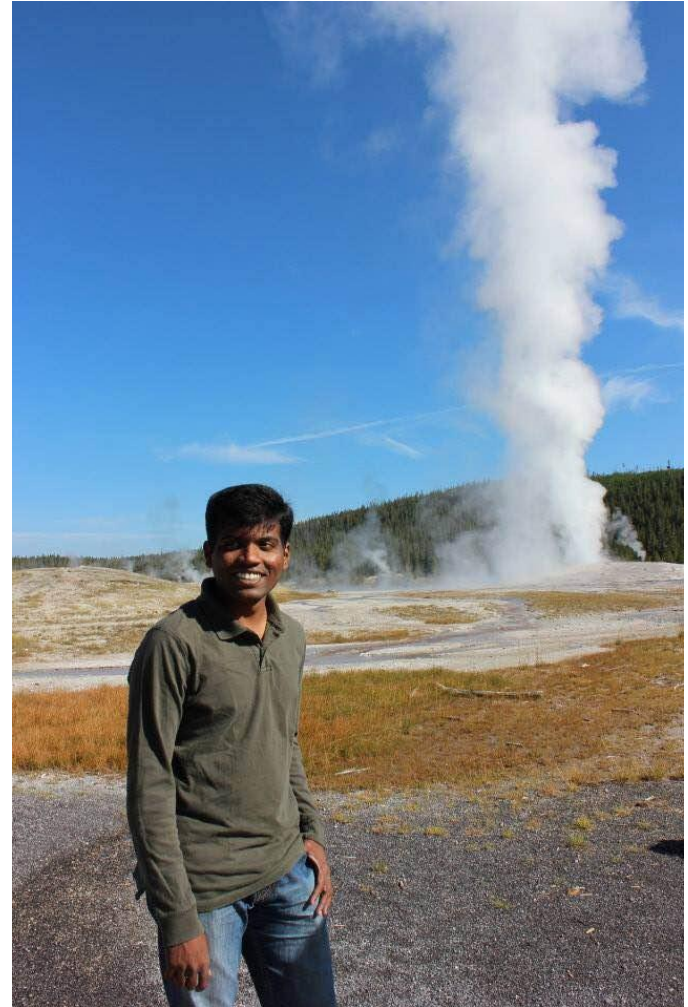
Pittsburgh: Teaching Assistants

- Zichang Feng



SV: Teaching Assistants

- Vinay Kumar Vavili



Rwanda: Teaching Assistant

- Cathy Bishop



Adelaide: Teaching Assistant

- Enrique Arango Lyons



Online Course Content - OLI

Course content is on the Open Learning Initiative:

- Students are automatically registered
- Access to OLI is through Blackboard
- Demo if time permits

The screenshot shows the OLI course interface. At the top, there is a red header with 'Carnegie Mellon University' and the 'Open Learning Initiative' logo, which includes the text 'Transforming higher education through the science of learning.' To the right, there are links for 'My Courses' and 'Help', and a user profile for 'Hello, Majd [sign out]'. Below the header, the course title 'Syllabus: S15-Cloud Computing (15319/15619): Jan - Jun 2015' is displayed, along with the instructor's name 'Instructor: Majd Sakr [msakr@ANDREW.CMU.EDU]'. A navigation bar contains links for 'Syllabus', 'Roster', 'Gradebook', and 'Unscored Activities'. A note states: 'Before you begin, [Test and Configure](#) your system for use with this course.' The main content is a table titled 'Cloud Computing' with columns for 'Assignment' and 'Status'. The table lists various modules and quizzes, including 'UNIT 1: Introduction to Cloud Computing', 'UNIT 2: Data Centers', and 'UNIT 3: Virtualizing Resources for the Cloud'. Each row provides a link to the module or quiz, a 'Checkpoint' status, and the availability date and time.

Assignment	Status
UNIT 1: Introduction to Cloud Computing	
Module 1: Introduction to Cloud Computing (Gradebook) (Learning Dashboard)	
Module 2: Building Blocks and Service Models (Gradebook) (Learning Dashboard)	Opens on 1/19/15 12:01 AM
Quiz 1: Introduction to Cloud Computing	Checkpoint Not yet available
UNIT 2: Data Centers	
Module 3: Data Center Trends (Gradebook) (Learning Dashboard)	Opens on 1/26/15 12:01 AM
Module 4: Data Center Components (Gradebook) (Learning Dashboard)	Opens on 1/26/15 12:01 AM
Module 5: Design Considerations (Gradebook) (Learning Dashboard)	Opens on 2/2/15 12:01 AM
Quiz 2: Data Centers	Checkpoint Not yet available
UNIT 3: Virtualizing Resources for the Cloud	
Module 6: Introduction and Motivation (Gradebook) (Learning Dashboard)	Opens on 2/9/15 12:01 AM
Module 7: Virtualization (Gradebook) (Learning Dashboard)	Opens on 2/9/15 12:01 AM
Module 8: Resource Virtualization - CPU (Gradebook) (Learning Dashboard)	Opens on 2/16/15 12:01 AM
Module 9: Resource Virtualization - Memory (Gradebook) (Learning Dashboard)	Opens on 2/23/15 12:01 AM
Module 10: Resource Virtualization - I/O (Gradebook) (Learning Dashboard)	Opens on 2/23/15 12:01 AM
Module 11: Case Study (Gradebook) (Learning Dashboard)	Opens on 2/23/15 12:01 AM
Quiz 3: Virtualizing Resources for the Cloud	Checkpoint Not yet available

Syllabus

- Updated on webpage
- Provides details on:
 - Course Objectives
 - Learning Outcomes
 - Policies
 - Grading
 - Tentative Schedule

15-319/15619: CLOUD COMPUTING

COURSE DESCRIPTION & SYLLABUS

CARNEGIE MELLON UNIVERSITY
SPRING 2015

1. OVERVIEW

Title: Cloud Computing

Units: 15-319 is 9 units and 15-619 is 12 units.

Pre-requisites for CMU Students: A "C" or better in 15-213.

Pre-requisites for Others: Knowledge of Computer Systems, Java programming.

OLI Course Link: <http://oli.cmu.edu>

The Project Zone: <http://TheProject.Zone>

Piazza Link: <http://piazza.com/cmu/spring2014/1531915619/home>

Recitation:

1. Tuesday, 8:00 AM – 8:50 AM, GHC 4307 (Videotaped)
2. Thursday, 4:30 PM – 5:20 PM, GHC 4307

Teaching Staff:

Prof. Majid F. Sakr

msakr@cs.cmu.edu

GHC 7006, +1-412-268-1161

Office hours: Tuesday, 3-4pm (Pittsburgh)

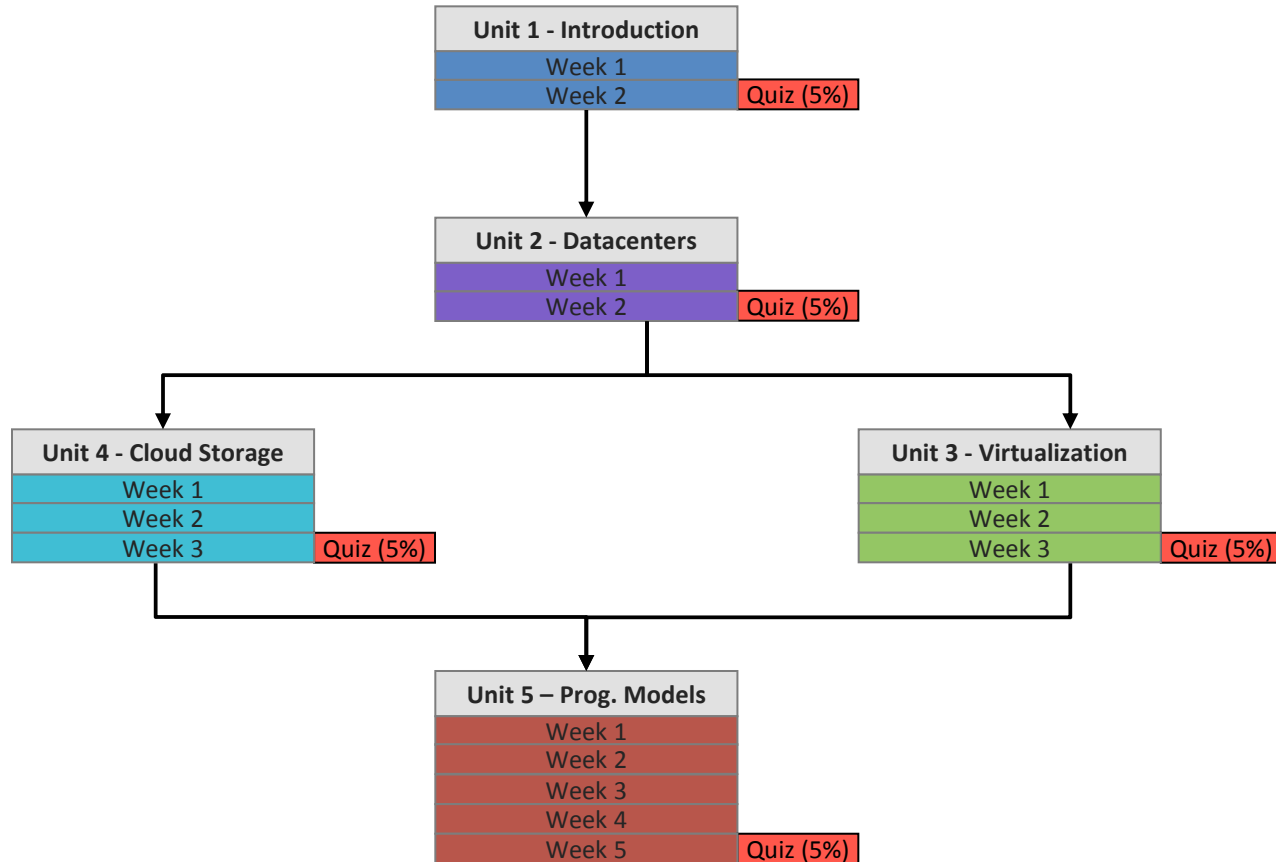
TAs in Pittsburgh typically hold office hours in GHC 5th Floor, office hours posted on Piazza:

- Anshima Gupta anshimag@andrew.cmu.edu
- Chao Zhang chaozhan@andrew.cmu.edu
- Cathy Bishop cbishop@rwanda.cmu.edu (Rwanda)
- Enrique Arango Lyons earangol@andrew.cmu.edu (Adelaide)
- Eryue Chen eryuec@andrew.cmu.edu
- Gongxun Liu gongxunli@andrew.cmu.edu
- Jiaduo He jiaduoh@andrew.cmu.edu
- Lei Wang leiwang@andrew.cmu.edu
- Lu Zeng luzeng@andrew.cmu.edu
- Mayank Singh Shishodia mshishod@andrew.cmu.edu
- Mengyu Yang mengyuy@andrew.cmu.edu
- Mrigesh Kalvani mkalvai@andrew.cmu.edu
- Prajwal Yadapadithaya pyadapad@andrew.cmu.edu
- Pan Sun pans@andrew.cmu.edu
- Ravi Chandra Bandlamudi Venkata rbandlam@andrew.cmu.edu
- Rohit Upadhyaya rupadhy@andrew.cmu.edu
- Ru Jia rjia@andrew.cmu.edu
- Satya Venkata Kamuju skamuju@andrew.cmu.edu
- Siyuan Zhou siyuanz@andrew.cmu.edu
- Suhail Rehman suhailr@andrew.cmu.edu
- Tianqi Wen tianqiw@andrew.cmu.edu
- Vinay Kumar Vavili vvavili@andrew.cmu.edu (Silicon Valley)
- Zichang Feng zfeng@andrew.cmu.edu

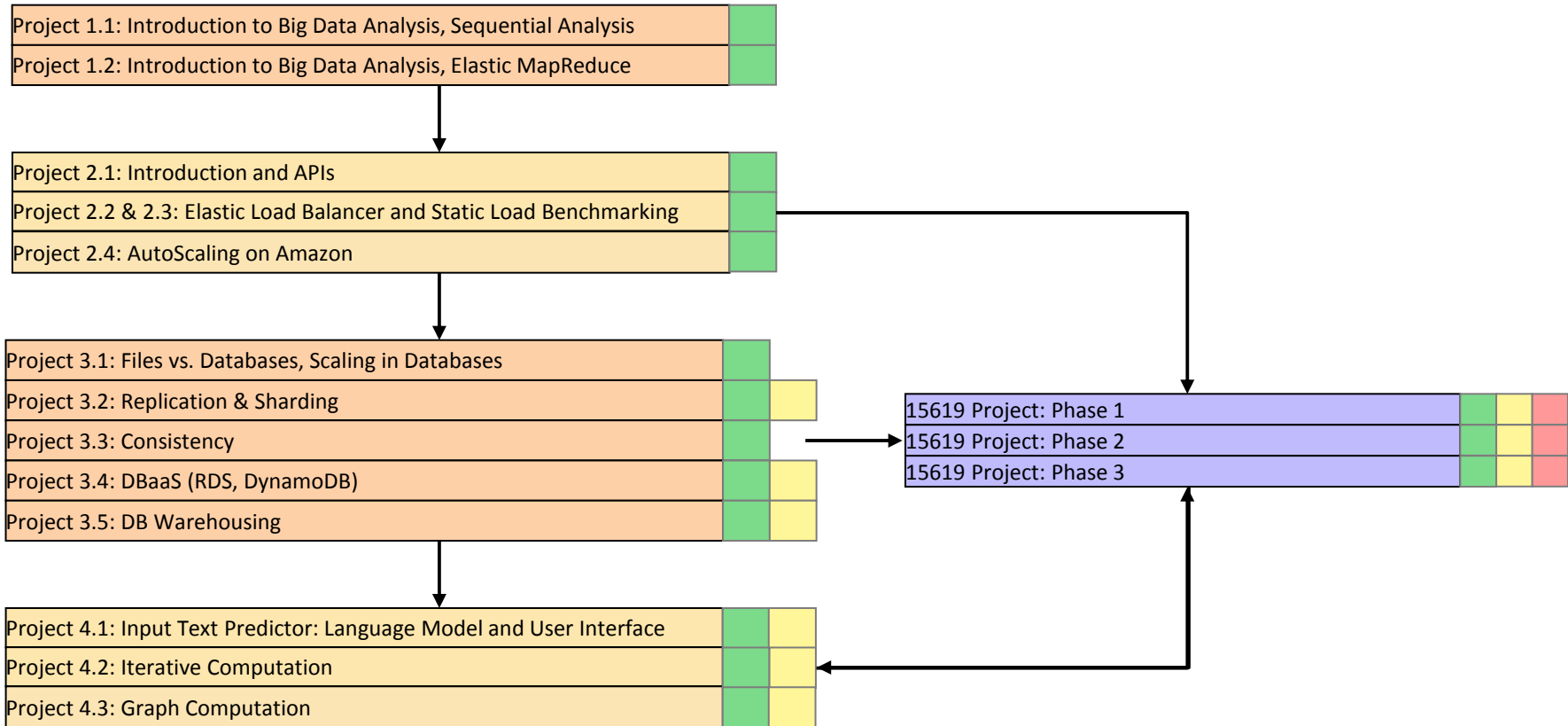
Tentative Schedule

Date	OLI Content	Quiz	Project	Extra Project	
12-Jan-15	Unit 1 – Introduction		Primer		
19-Jan-15		Unit 1 Checkpoint Quiz	Project 1		
26-Jan-15	Unit 2 –				
2-Feb-15	Datacenters	Unit 2 Checkpoint Quiz	Project 2		
9-Feb-15	Unit 3 – Virtualization				
16-Feb-15					
23-Feb-15		Unit 3 Checkpoint Quiz			
2-Mar-15	Unit 4 – Cloud Storage		Project 3		15-619 Extra Project
16-Mar-15					
23-Mar-15		Unit 4 Checkpoint Quiz			
30-Mar-15	Unit 5 – Programing Models		Project 4		
6-Mar-15					
13-Apr-15					
20-Apr-15					
27-Apr-15		Unit 5 Checkpoint Quiz			

Content: Timeline and Dependencies



Projects: Timeline and Dependencies



Grading

Course Elements	#	Weight
Projects	4 or 5	75%
OLI Unit Checkpoint Quizzes	5	25%

- All projects are equal weight
 - 18.75% for 15-319
 - 15% for 15-619
- All quizzes are equal weight
 - 5% for each quiz

Academic Integrity

It is the responsibility of each student to produce her/his own original academic work.

- Individual work:
 - Weekly Project Modules
 - Unit Checkpoint Quizzes
- Team work:
 - 15-619 Project

Read the [university policy on Academic Integrity](#).

The Penalties are Severe

- Cheating leads to several students being dismissed from the university every semester

LET IT NOT BE YOU!

What is Cheating

- Sharing code or other electronic files either by copying, retyping, looking at, or supplying a copy of any file.
- Copying answers to any checkpoint quiz from another individual, published or unpublished written sources, and electronic sources.
- Collaborating with another student or another individual on Unit Checkpoint Quizzes or Project Module Checkpoint Quizzes.
- Sharing written work, looking at, copying, or supplying work from another individual, published or unpublished written sources, and electronic sources.
- Collaboration in team projects is strictly limited to the members of the team.
- ...

Course Administration

- Students are automatically registered on OLI through blackboard.andrew.cmu.edu
- A *single* Piazza course page is created
 - We manually register students to Piazza
- Schedule of units and quizzes is on OLI
 - Content quizzes are due on Thursdays
- Schedule of weekly projects is on TheProject.Zone
 - Weekly project modules are due on Sundays

Special Note on Amazon EC2



- Paid Cloud Service – billed by the hour
- Start a resource only when you need it
- To explore, use a micro instance
 - You can keep one micro instance running 24x7
- Terminate all other resources as soon as you are done with them
- Students will be penalized for over usage
 - We have a fixed budget, do not abuse the resources!
 - Intentional or unintentional abuse → grade penalties
 - Resources need to be tagged, otherwise → penalties

This Week

- Become familiar with OLI
 - Content (Outline of Units 1-5)
- Projects (Primer)
- Check that you were enrolled on Piazza
- Create an account on AWS (**Deadline, Jan 18**)
 - Submit your AWS account number using the link provided
- Complete Project Primer by Sunday
- Start reading Unit 1, Module 1 on OLI

Questions?

