15-319 / 15-619
Cloud Computing

Recitation 5

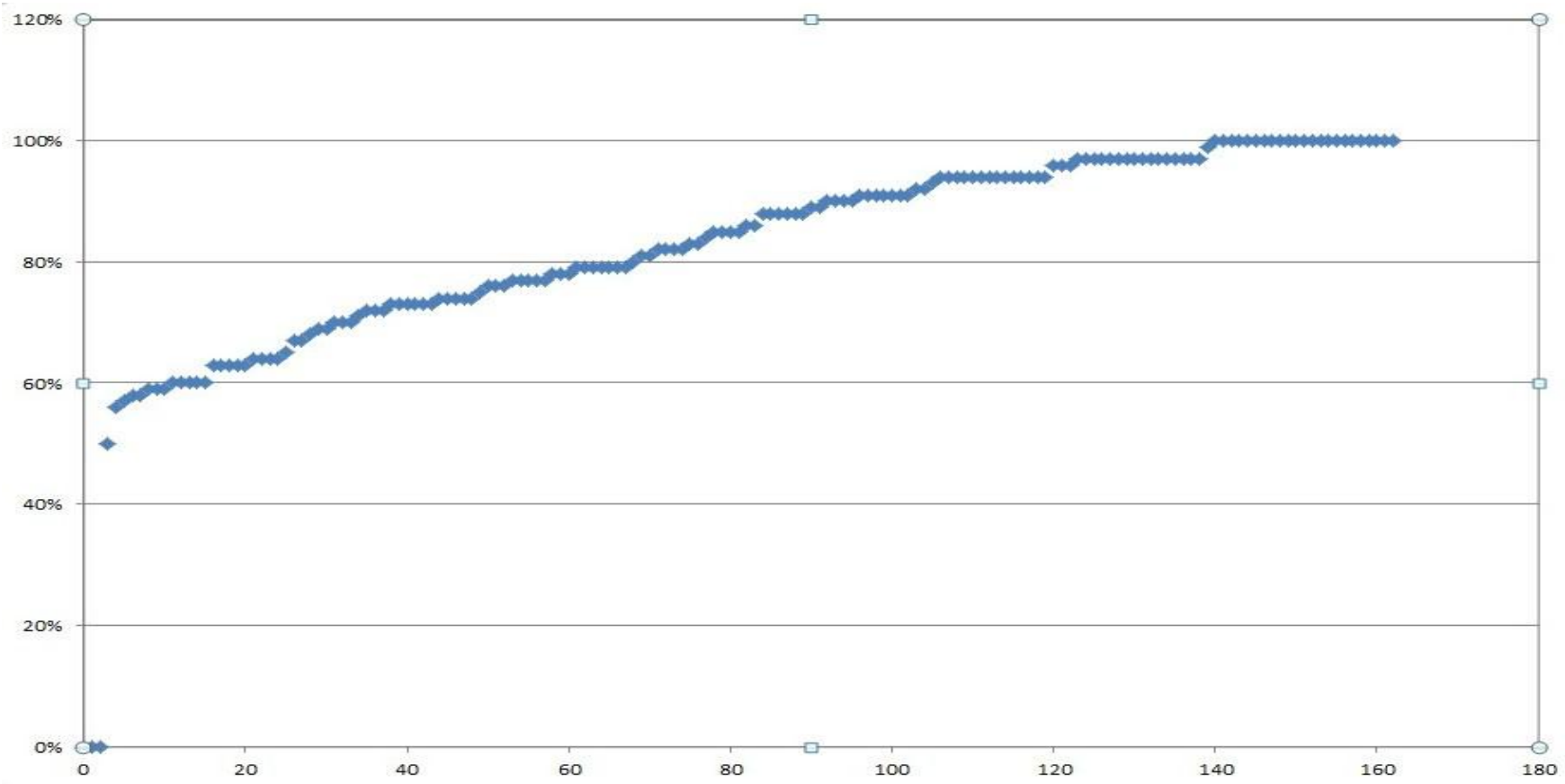
September 24th & 27th, 2013

Announcements

- Encounter a general bug:
  - Post on Piazza
- Encounter a grading bug:
  - Post Privately on Piazza
  - Don’t ask if my answer is correct
  - Search before posting
- Post feedback on OLI

UNIT 2: Checkpoint Quiz 2

- Student Performance:
  - Average is 80%



Project 2 Student Progress

- Introduction and APIs
  - Single Instance Benchmarks
- Elastic Load Balancing (2 modules due)
  - Elastic Load Balancer
  - Static Load Benchmarking

Available Modules

- UNIT 3: Virtualizing Resources for the Cloud
  - Module 6: Introduction and Motivation
  - Module 7: Virtualization
  - Module 8: Resource Virtualization - CPU
  - Module 9: Resource Virtualization - Memory
  - Module 10: Resource Virtualization – I/O
  - Module 11: Case Study
  - Quiz 3: Virtualizing Resources for the Cloud

Piazza Questions

- Question 2 of Checkpoint Project 2, first module

  We would manually grade the requests per second for each instance type.

- How to make sure instance is running?

  Launching instance has two stages: Running (h/w resources allocated) and Initializing (loading AMI image)

  Ec2-describe-instance-status to describe the current status.

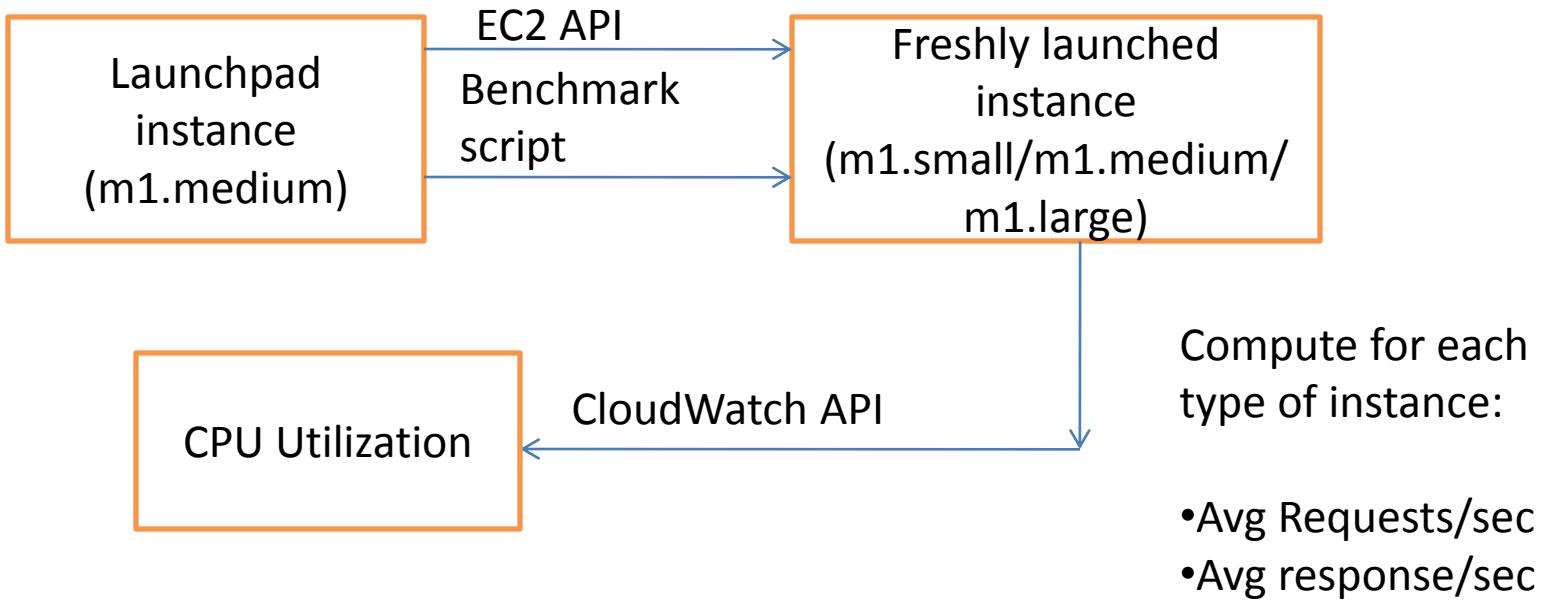
- Run the benchmark test on launchpad instance instead of test instance.

- Run the script/code on the launchpad instance instead of running it on your local machine.

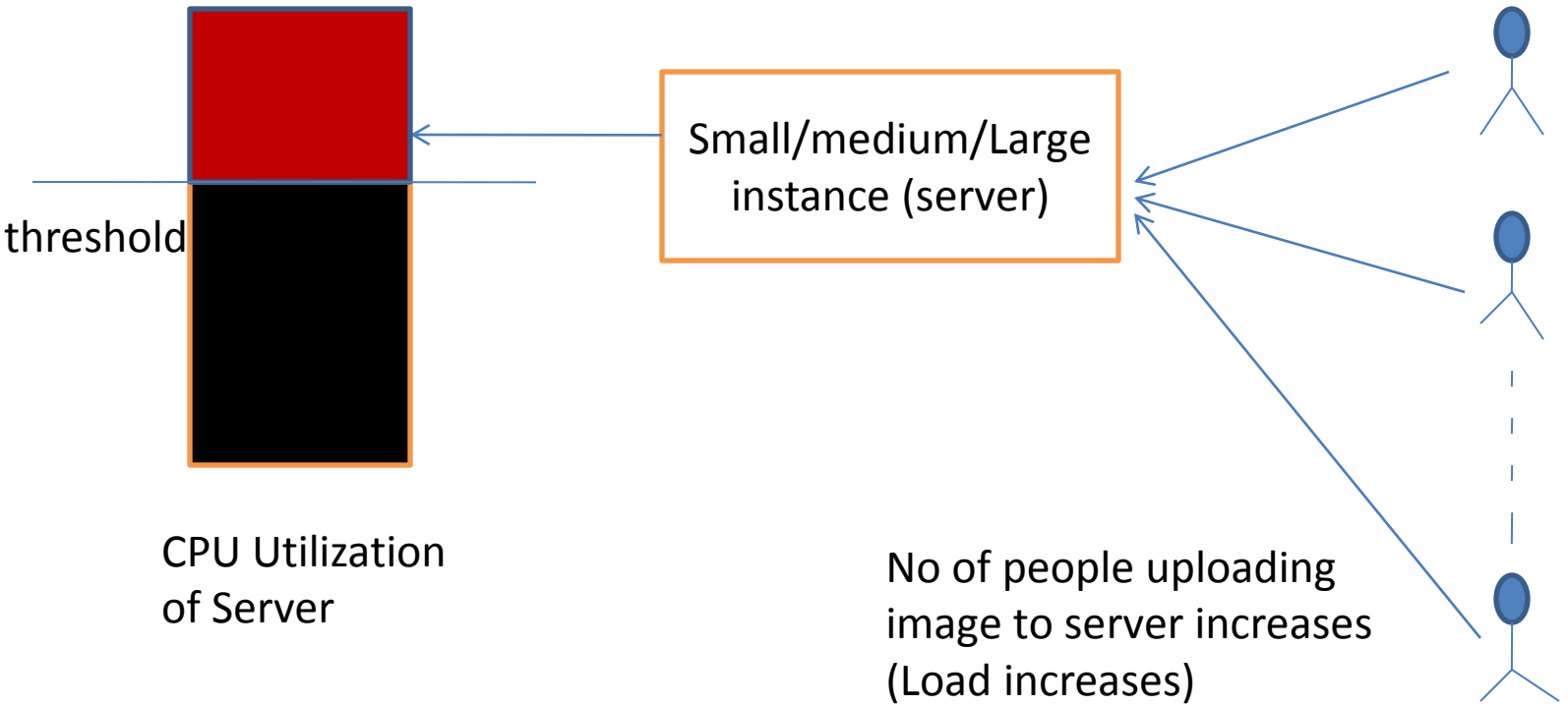
Vertical Scaling (First Module)

- Takeaways from previous checkpoint of Project 2

  Vertical scaling (increasing the capacity of web servers) improves bandwidth (average responses/sec)

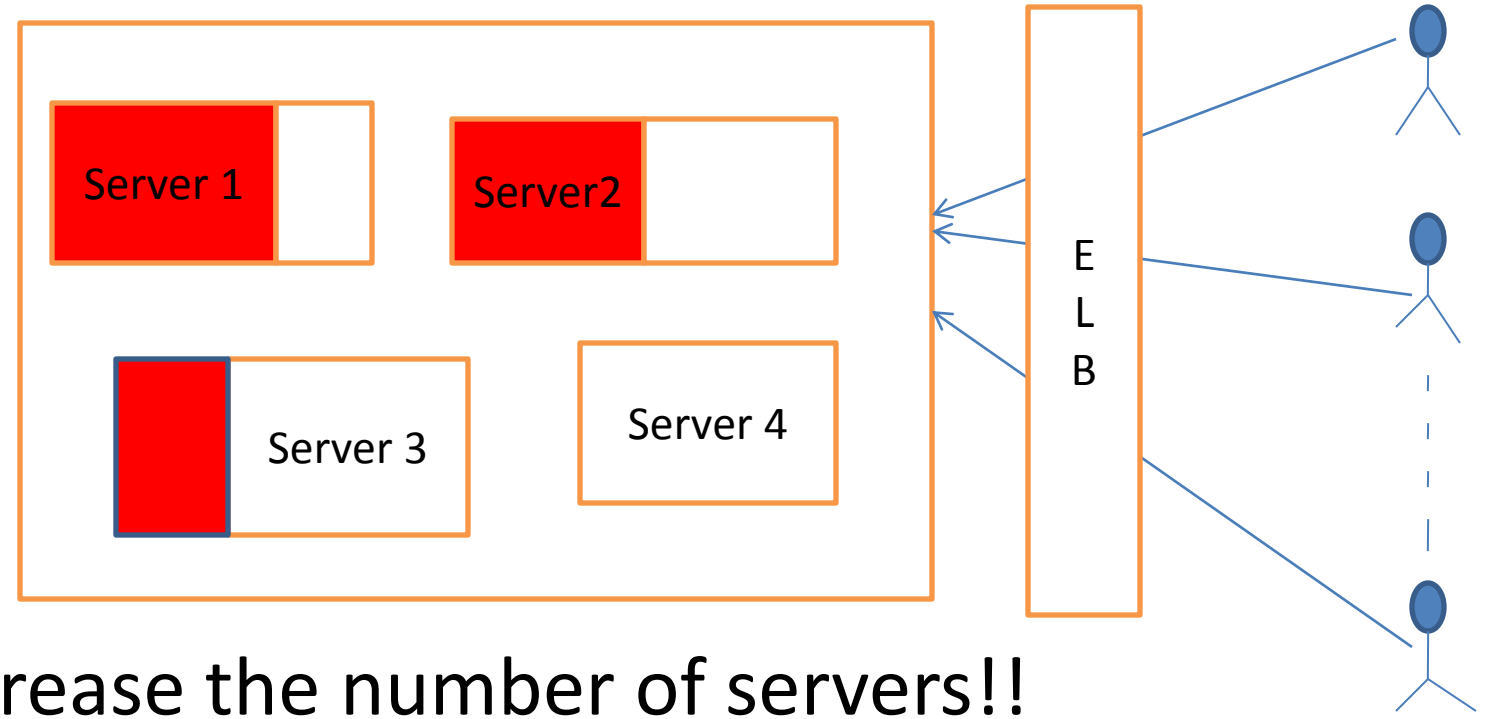


Checkpoint, Project 2, Second Module



What can be the solution to deal with the increasing load of requests????

Solution

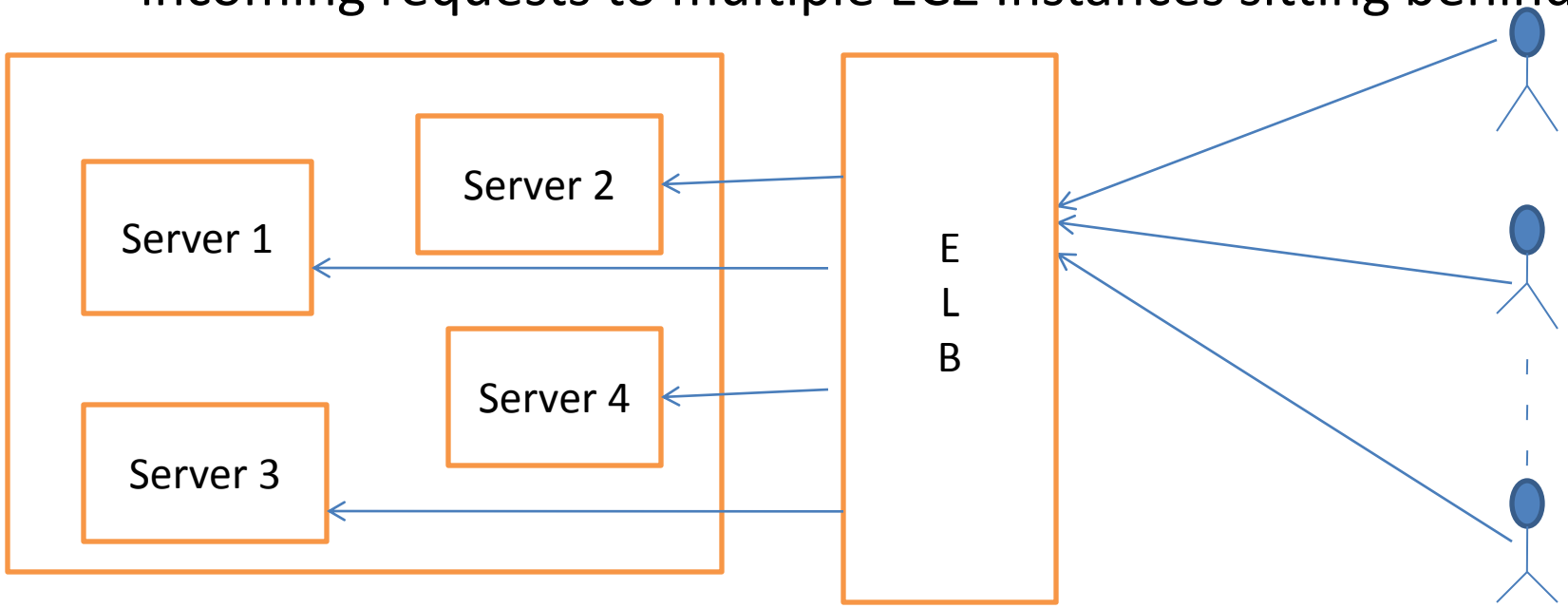


Increase the number of servers!!

How to best utilize all servers??

Elastic Load Balancer

- We need a way to redirect the requests from clients to upload the image to one of the servers.
- ELB is a gateway that acts as a router interface and sends incoming requests to multiple EC2 Instances sitting behind it.



Features of ELB

- Using ELB, you can distribute incoming traffic across your Amazon EC2 instances in a single Availability Zone or multiple Availability Zones.

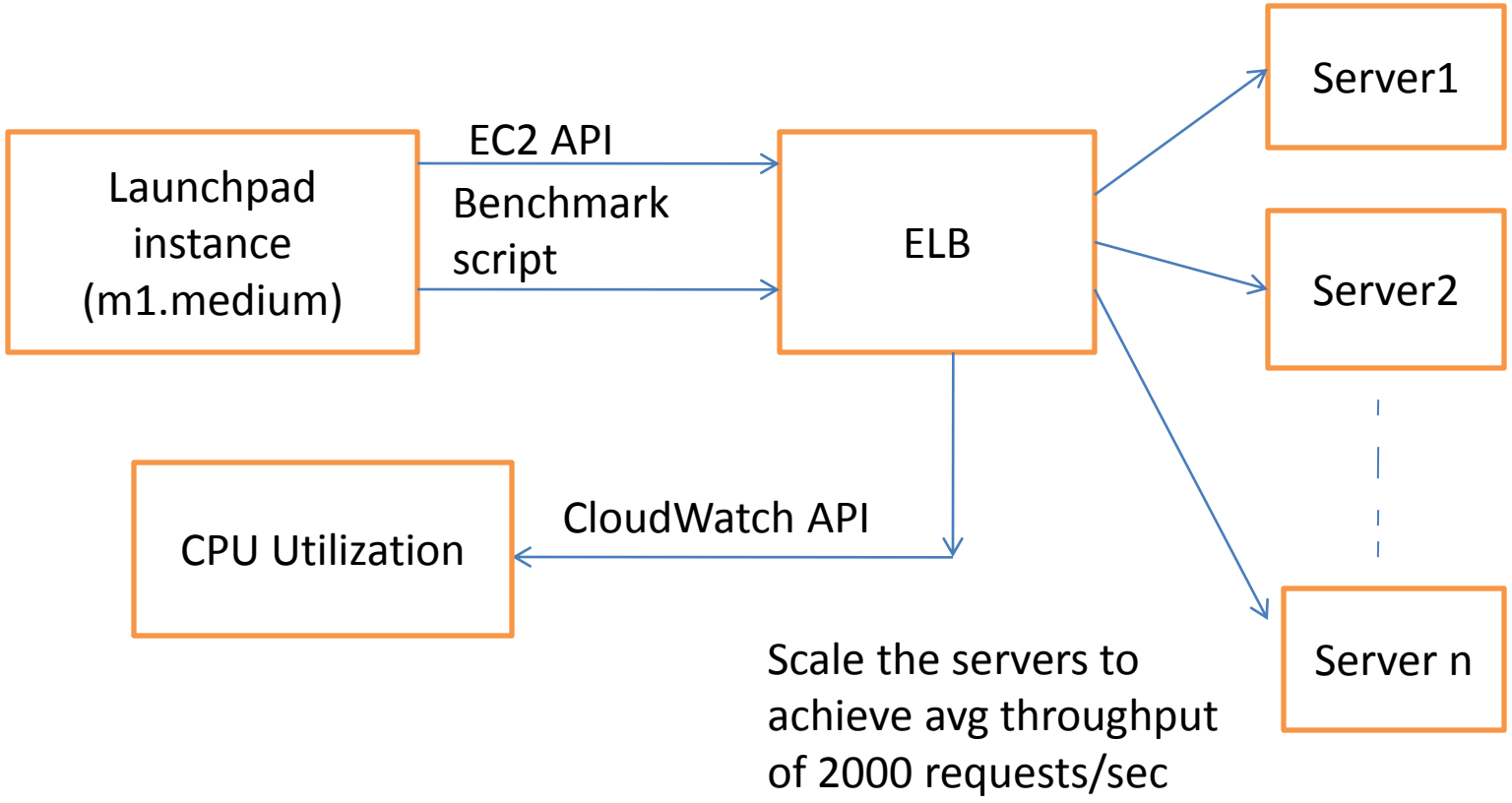
- Elastic Load Balancing can detect the health of Amazon EC2 instances. When it detects unhealthy instances, it spreads the load to other healthy instances.

Second Module (using Web Console)

- Deliverables:
- Launch 3 EC2 instances (1 launchpad + 2 servers)
- Create a new ELB
- Attach one of the provisioned instances/servers to ELB.
- Fire up Apache Benchmarking from launchpad to ELB this time.
- Repeat the same for second instance.
- Answer Checkpoint Quiz

Horizontal Scaling (Third Module)

- Develop a policy to increase the number of Web servers behind ELB to improve throughput.



Third Module (programmatically)

- Deliverables:
- Script that should:
  - Launch an instance and create ELB
  - Attach launched instance to ELB
  - Run benchmark from launchpad to ELB
  - Retrieve Req/sec. If its less than 2000, add another instance, attach it to ELB and benchmark again

Demos

- Create a new Load Balancer manually. (Project 2, Second Module)

- Create a new Load Balancer and attach/remove the instance to ELB programmatically using Python. Though you can use the programming language of your own choice. (Project 2, Third Module)

Upcoming Deadlines

- Project 2:

Project 2		
Introduction and APIs		
Single Instance Benchmarks	Checkpoint	Available Now Due 9/22/13 11:59 PM
Elastic Load Balancing		
Elastic Load Balancer	Checkpoint	Available Now Due 09/29/13 11:59 PM
Static Load Benchmarking	Checkpoint	Available Now Due 09/29/13 11:59 PM

- Unit 3:

UNIT 3: Virtualizing Resources for the Cloud
Module 6: Introduction and Motivation
Module 7: Virtualization