Cloud Computing, Hands On

David J. Malan
dmalan@harvard.edu

Self-service, Prorated Super Computing Fun!

- “The New York Times has decided to make all the public domain articles from 1851-1922 available free of charge. These articles are all in the form of images scanned from the original paper. In fact from 1851-1980, all 11 million articles are available as images in PDF format. To generate a PDF version of the article takes quite a bit of work — each article is actually composed of numerous smaller TIFF images that need to be scaled and glued together in a coherent fashion. . . . I then began some rough calculations and determined that if I used only four machines, it could take some time to generate all 11 million article PDFs. But thanks to the swell people at Amazon, I got access to a few more machines and churned through all 11 million articles in just under 24 hours using 100 EC2 instances . . . . (In fact, it work so well that we ran it twice, since after we were done we noticed an error in the PDFs.)”

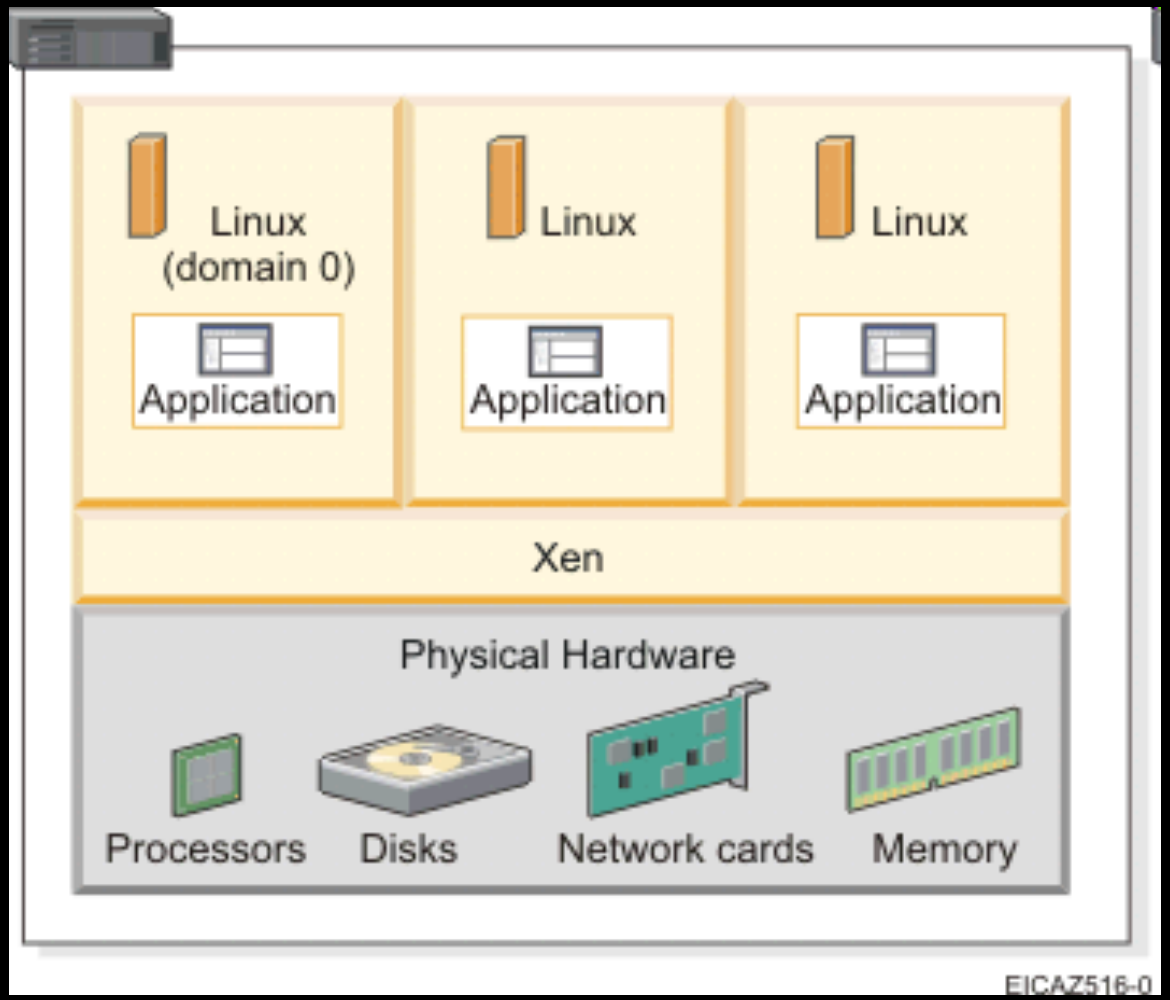
Virtualization

http://communication.howstuffworks.com/server-virtualization1.htm

Virtualization

- **Xen**
- KVM
- VMware
  - Workstation, Fusion, Server, vSphere
- Parallels
  - Desktop, Server, Virtuozzo, ...
- Microsoft
  - Hyper-V, Virtual PC
- Oracle
  - VirtualBox
- QEMU
- FreeVPS
- . . .

Xen



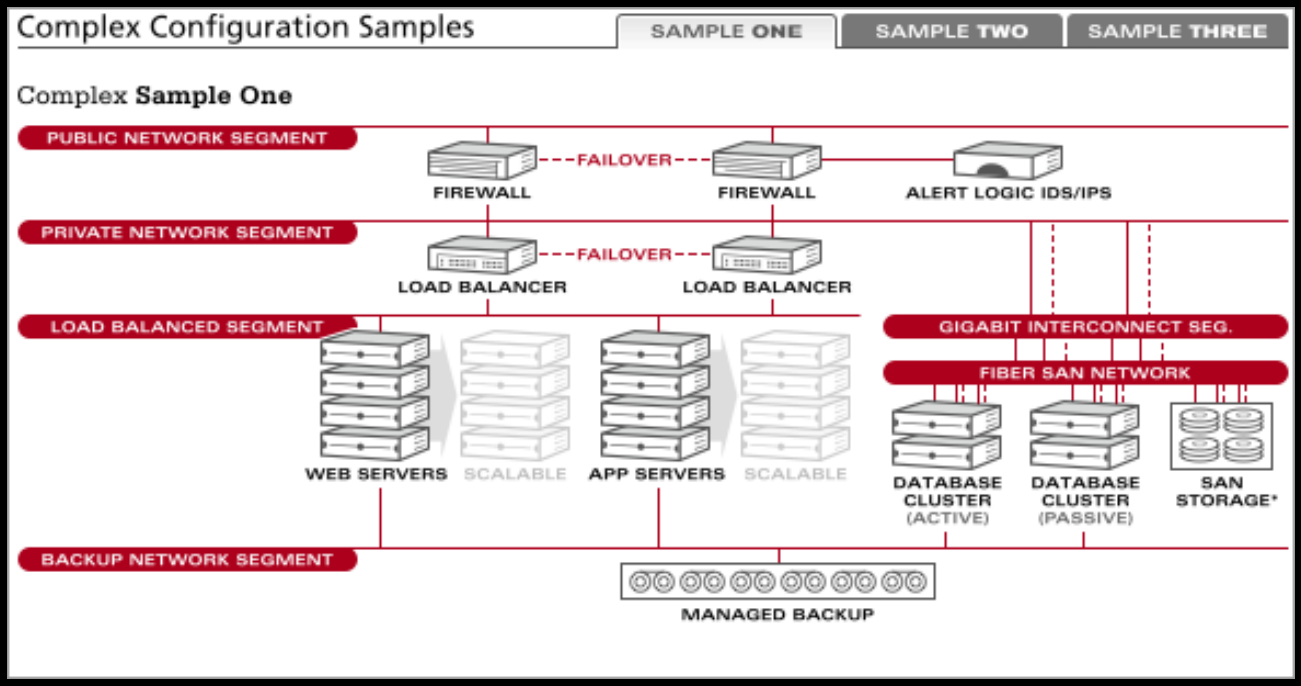
http://publib.boulder.ibm.com/infocenter/eserver/v1r2/topic/eicaz/eicaz516.gif

Virtual Private Servers (VPSes)

- Linode
- Rackspace
- ServInt
- Slicehost
- VPSLAND
- . . .

Managed Colocation

- Rackspace
- . . .



http://www.rackspace.com/solutions/managed_hosting/configurations/index.php

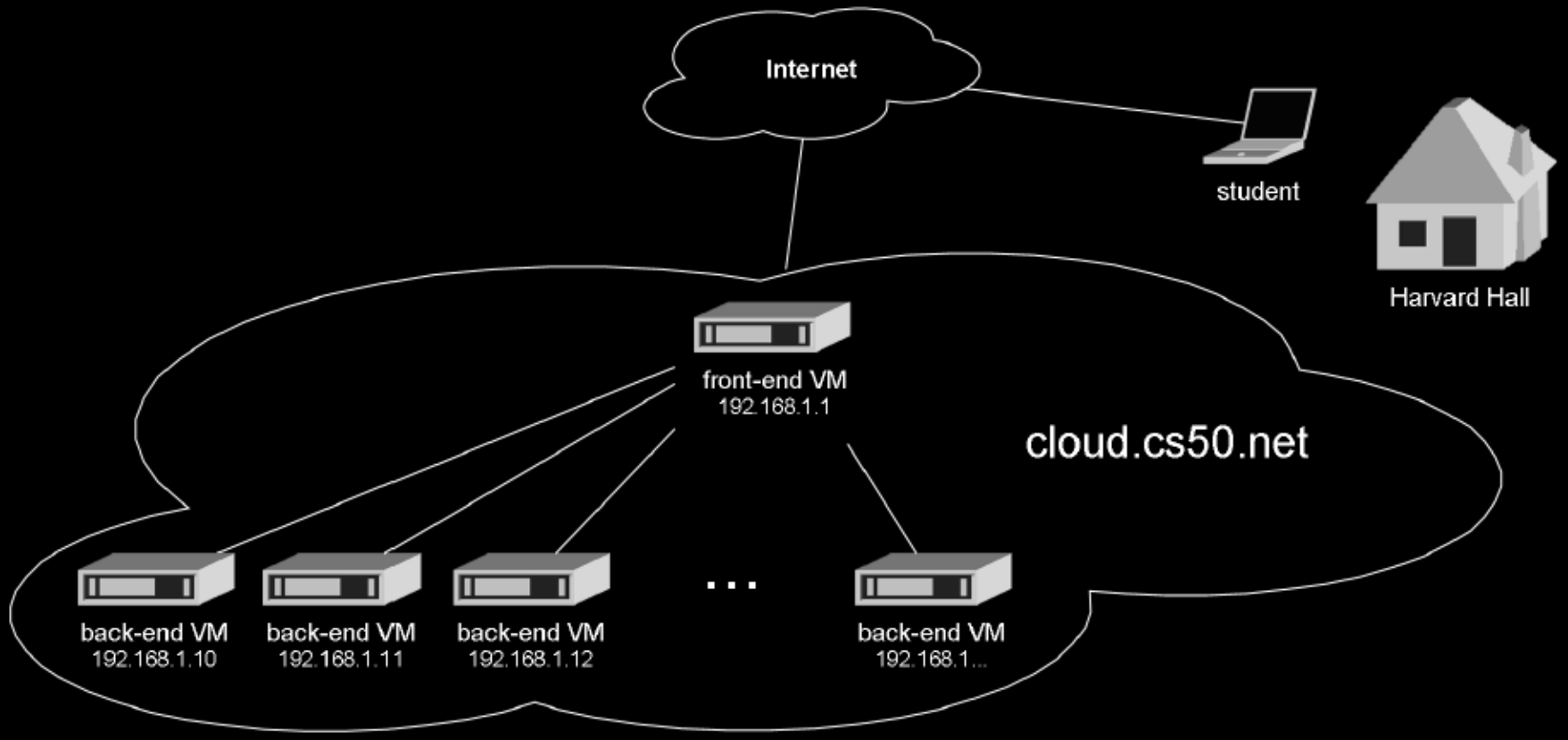
Clouds

- Amazon Elastic Compute Cloud (Amazon EC2)
- Google App Engine
- Microsoft Azure Services Platform
- . . .

Amazon Web Services (AWS)

- **Amazon Elastic Compute Cloud (Amazon EC2)**
- Amazon Simple Storage Service (Amazon S3)
- Amazon SimpleDB
- Amazon CloudFront
- Amazon Relational Database Service (RDS)
- Amazon Elastic MapReduce
- Amazon Virtual Private Cloud (Amazon VPC)
- Elastic Load Balancing
- . . .

cloud.cs50.net



Jargon

- AMI
- Instance
- EBS Volume
- . . .

AMIs

- Amazon Linux AMI
- Debian
- Fedora
- Gentoo Linux
- Red Hat Enterprise Linux
- OpenSolaris
- SUSE Linux Enterprise
- Oracle Enterprise Linux
- Ubuntu Linux
- Windows Server 2003/2008
- . . .

Standard Instances

- Small Instance (m1.small)
  - 1.7 GB memory
  - 1 EC2 Compute Unit
  - 160 GB instance storage
  - 32-bit platform
- Large Instance (m1.large)
  - 7.5 GB memory
  - 4 EC2 Compute Units
  - 850 GB instance storage
  - 64-bit platform
- Extra Large Instance (m1.xlarge)
  - 15 GB memory
  - 8 EC2 Compute Units
  - 1,690 GB instance storage
  - 64-bit platform

High-CPU Instances

- High-CPU Medium Instance (c1.medium)
  - 1.7 GB of memory
  - 5 EC2 Compute Units
  - 350 GB of instance storage
  - 32-bit platform
- High-CPU Extra Large Instance (c1.xlarge)
  - 7 GB of memory
  - 20 EC2 Compute Units
  - 1690 GB of instance storage
  - 64-bit platform

High-Memory Instances

- High-Memory Extra Large Instance (m2.xlarge)
  - 17.1 GB of memory
  - 6.5 EC2 Compute Units
  - 420 GB of instance storage
  - 64-bit platform
- High-Memory Double Extra Large Instance (m2.2xlarge)
  - 34.2 GB of memory
  - 13 EC2 Compute Units
  - 850 GB of instance storage
  - 64-bit platform
- High-Memory Quadruple Extra Large Instance (m2.4xlarge)
  - 68.4 GB of memory
  - 26 EC2 Compute Units
  - 1690 GB of instance storage
  - 64-bit platform

Regions + Availability Zones

- United States
  - California
  - Virginia
- Europe
  - Ireland
- Asia
  - Singapore

On-Demand Instances

US – N. Virginia | US – N. California | EU – Ireland | APAC – Singapore

Standard On-Demand Instances	Linux/UNIX Usage	Windows Usage
Small (Default)	$0.085 per hour	$0.12 per hour
Large	$0.34 per hour	$0.48 per hour
Extra Large	$0.68 per hour	$0.96 per hour
Micro On-Demand Instances		
Micro	$0.02 per hour	$0.03 per hour
High-Memory On-Demand Instances		
Extra Large	$0.50 per hour	$0.62 per hour
Double Extra Large	$1.00 per hour	$1.24 per hour
Quadruple Extra Large	$2.00 per hour	$2.48 per hour
High-CPU On-Demand Instances		
Medium	$0.17 per hour	$0.29 per hour
Extra Large	$0.68 per hour	$1.16 per hour
Cluster Compute Instances		
Quadruple Extra Large	$1.60 per hour	N/A*
Cluster GPU Instances		
Quadruple Extra Large	$2.10 per hour	N/A*
* Windows is not currently available for Cluster Compute or Cluster GPU Instances.		

http://aws.amazon.com/ec2/

Data Transfer

Data Transfer In	US & EU Regions	APAC Region
All Data Transfer	$0.10 per GB	$0.10 per GB

Data Transfer Out ***	US & EU Regions	APAC Region
First 1 GB per Month	$0.00 per GB	$0.00 per GB
Up to 10 TB per Month	$0.15 per GB	$0.19 per GB
Next 40 TB per Month	$0.11 per GB	$0.15 per GB
Next 100 TB per Month	$0.09 per GB	$0.13 per GB
Over 150 TB per Month	$0.08 per GB	$0.12 per GB

http://aws.amazon.com/ec2/

EBS Volumes + Snapshots

http://aws.amazon.com/ec2/

Cost Us Less than $15 per Student

	CPU	Disk	I/O Requests	Bandwidth
Sep	2,275 Hrs	125 GB	45,348	14 GB
Oct	3,425 Hrs	108 GB	93,257,314	191 GB
Nov	5,484 Hrs	199 GB	337,019,916	239 GB
Dec	5,206 Hrs	300 GB	427,639,962	52 GB
Jan	5,208 Hrs	300 GB	1,502,614,186	62 GB

AWS Simple Monthly Calculator

http://calculator.s3.amazonaws.com/calc5.html

Concerns

- Time
- PEBKAC
- “Cloud is wicked laggy”
- Bandwidth Costs
- “Lightning Strike Triggers Amazon EC2 Outage”
- “Outage Hits Amazon European Sites”
- …

Resources

- Getting Started Guide
  - http://docs.amazonwebservices.com/AWSEC2/latest/GettingStartedGuide/
- User Guide
  - http://docs.amazonwebservices.com/AWSEC2/latest/UserGuide/
- Developer Guide
  - http://docs.amazonwebservices.com/AWSEC2/latest/DeveloperGuide/
- . . .