

Cloud Programming and Software Environments: Amazon AWS

Ioan Raicu (Presented by Alex Orhean)

Computer Science Department
Illinois Institute of Technology

CS 553: Cloud Computing
April 18th, 2018

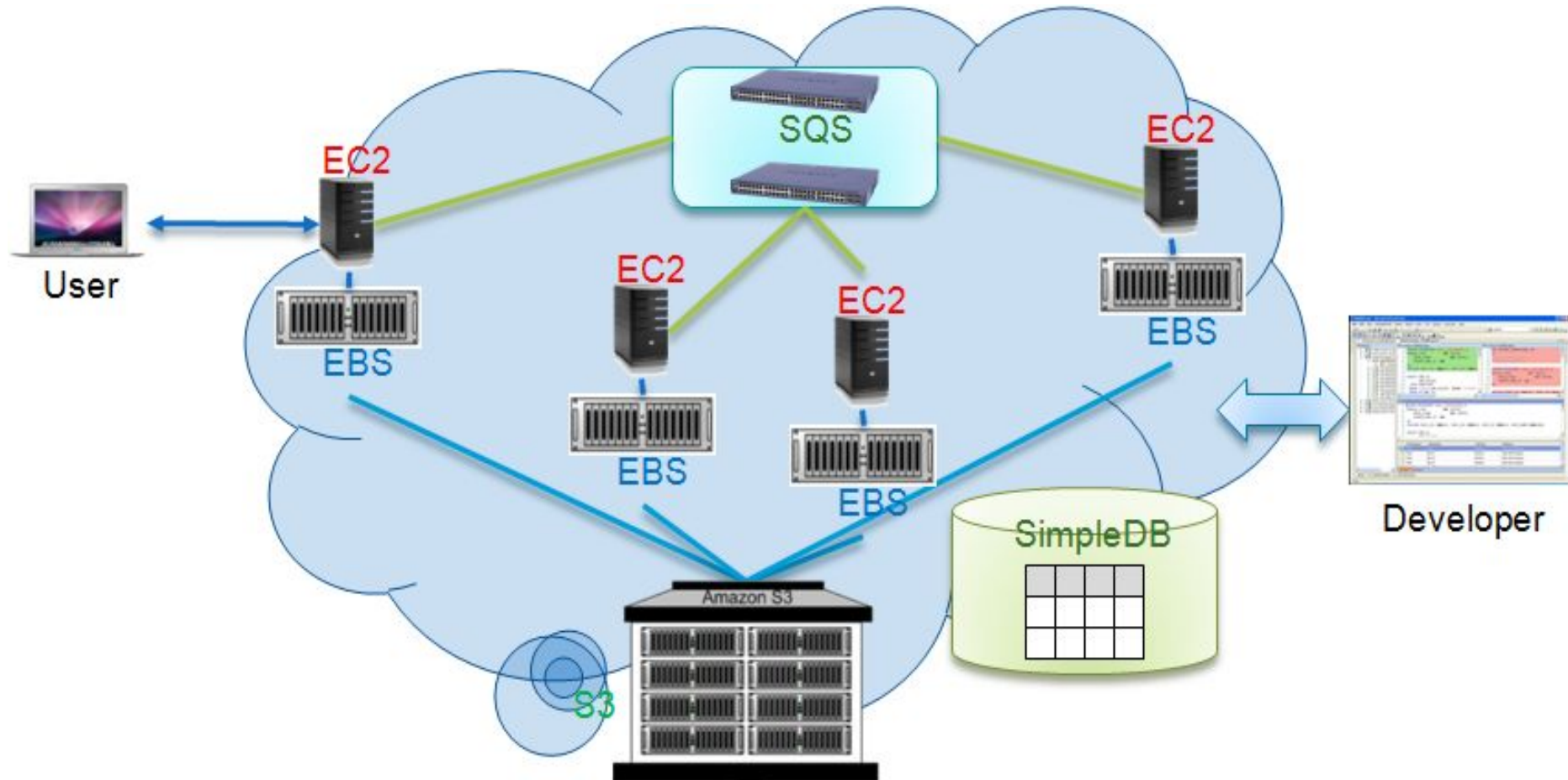
digitalblasphemy.com

Amazon Web Services (AWS)

- Amazon has been a leader in providing public cloud services
- Amazon applies the IaaS model in providing its services
- Different from Google, Amazon provides a more flexible cloud computing platform for developers to build cloud applications.
- Elastic Computing Cloud (EC2) provides the virtualized platforms to the host VMs where the cloud application can run
- S3 (Simple Storage Service) provides the object-oriented storage service for users
- EBS (Elastic Block Service) provides the block storage interface which can be used to support traditional applications
- SQS stands for Simple Queue Service, and its job is to ensure a reliable message service between two processes
 - The message can be kept reliably even when the receiver processes are not running
 - Users can access their objects through SOAP with either browsers or other client programs which support the SOAP standard.

Amazon Web Services (AWS)

- Amazon cloud computing infrastructure



Amazon Web Services (AWS)

- AWS offerings in 2011

Service Area	Service Modules and Abbreviated Names
Compute	Elastic Compute Cloud (EC2), Elastic MapReduce, Auto Scaling
Messaging	Simple Queue Service (SQS), Simple Notification Service (SNS)
Storage	Simple Storage Service (S3), Elastic Block Storage (EBS), AWS Import/Export
Content Delivery	Amazon CloudFront
Monitoring	Amazon CloudWatch
Support	AWS Premium Support
Database	Amazon SimpleDB, Relational Database Service (RDS)
Networking	Virtual Private Cloud (VPC) (Example 4.6), Elastic Load Balancing
Web Traffic	Alexa Web Information Service, Alexa Web Sites
E-Commerce	Fulfillment Web Service (FWS)
Payments and Billing	Flexible Payments Service (FPS), Amazon DevPay
Workforce	Amazon Mechanical Turk

Amazon Web Services (AWS) Technologies

- *Relational Database Service (RDS)*
- Elastic MapReduce capability is equivalent to Hadoop running on the basic EC2 offering
- AWS Import/Export allows one to ship large volumes of data to and from EC2 by shipping physical disks
 - Offers the highest bandwidth connection between geographically distant systems
- CloudFront implements a content distribution network
- MPI clusters and cluster compute instances
 - AWS cluster compute instances use hardware-assisted virtualization instead of the para-virtualization used by other instance types and requires booting from the EBS

Amazon Web Services (AWS) Technologies

- CloudWatch is a Web service that provides monitoring for AWS cloud resources
 - It provides customers with visibility into resource utilization, operational performance, and overall demand patterns, including metrics such as CPU utilization, disk reads and writes, and network traffic
- ELB automatically distributes incoming application traffic across multiple Amazon EC2 instances
 - Both auto-scaling and ELB are enabled by CloudWatch which monitors running instances

Amazon Web Services (AWS)

Billing and Accounting

- DevPay is a simple-to-use online billing and account management service that makes it easy for businesses to sell applications that are built into or run on top of AWS
- FPS provides developers of commercial systems on AWS with a convenient way to charge Amazon's customers that use such services built on AWS
 - Customers can pay using the same login credentials, shipping address, and payment information they already have on file with Amazon
- FWS allows merchants to access Amazon's fulfillment capabilities through a simple Web service interface
 - Merchants can send order information to Amazon to fulfill customer orders on their behalf

Programming Support in AWS

- EC2
- S3
- SimpleDB
 - NOSQL support
- Relational Database Service (RDS)
- Elastic MapReduce
 - Capability is equivalent to Hadoop running on the basic EC2 offering.
- Simple Queue Service (SQS)
- Simple Notification Service (SNS)

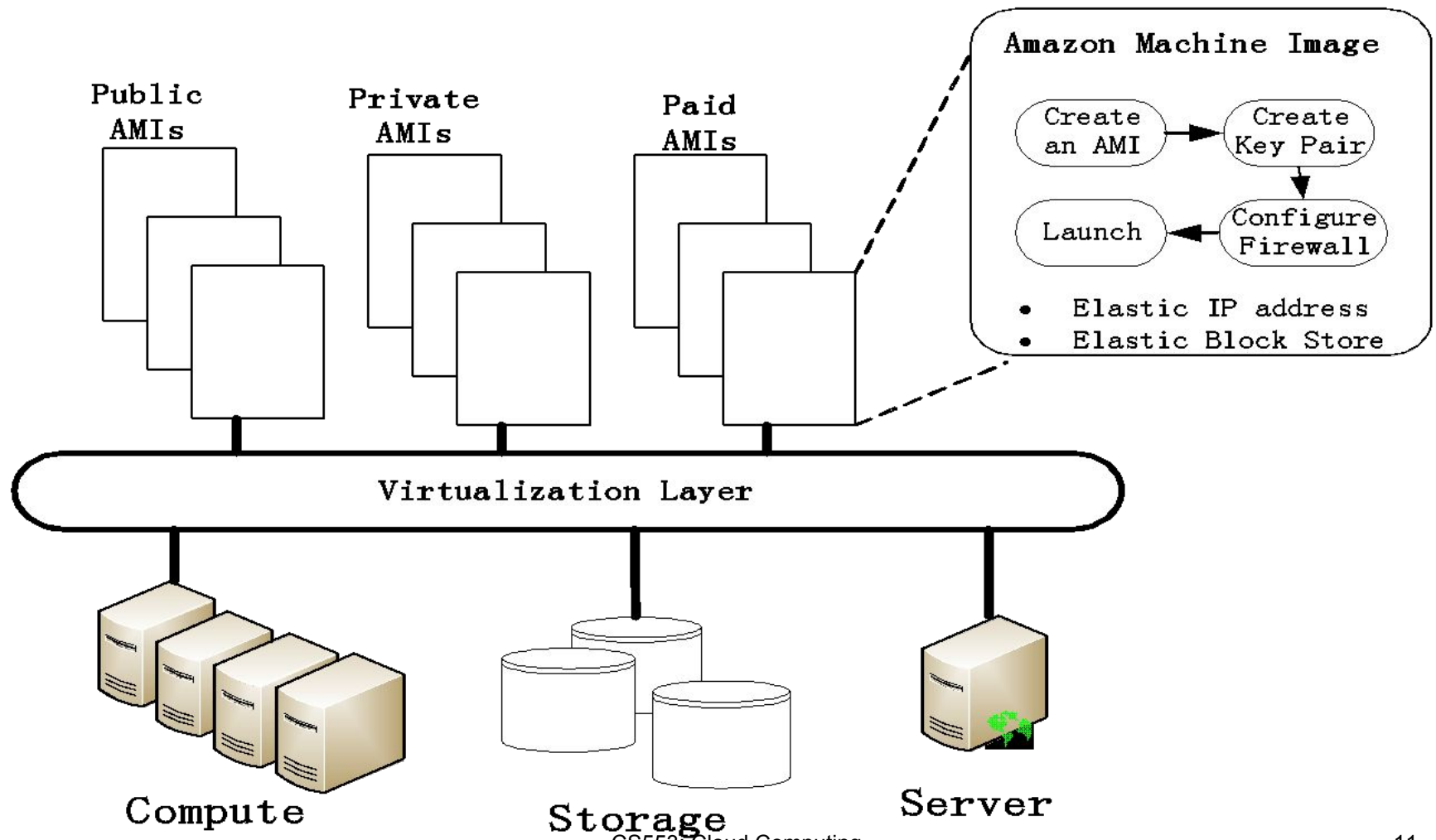
Programming Support in AWS

- **Auto-scaling**
 - Auto-scaling enables you to automatically scale your Amazon EC2 capacity up or down according to conditions that you define
 - With auto-scaling, you can ensure that the number of Amazon EC2 instances you're using scales up seamlessly during demand spikes to maintain performance, and scales down automatically during demand lulls to minimize cost
- **Elastic load balancing services**
 - Elastic load balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances and allows you to avoid nonoperating nodes and to equalize load on functioning images
- **Both auto-scaling and elastic load balancing are enabled by CloudWatch which monitors running instances**
 - CloudWatch is a Web service that provides monitoring for AWS cloud resources, starting with Amazon EC2
 - It provides customers with visibility into resource utilization, operational performance, and overall demand patterns—including metrics such as CPU utilization, disk reads and writes, and network traffic

Programming on Amazon EC2

- Amazon was the first company to introduce VMs in application hosting
- Customers can rent VMs instead of physical machines to run their own applications
- By using VMs, customers can load any software of their choice
- The elastic feature of such a service is that a customer can create, launch, and terminate server instances as needed, paying by the hour for active servers
- Amazon provides several types of preinstalled VMs
- Instances are often called *Amazon Machine Images (AMIs)* which are preconfigured with operating systems based on Linux or Windows, and additional software

Amazon EC2 Execution Environment



Amazon EC2 Execution Environment

- Three types of AMI

Image Type	AMI Definition
Private AMI	Images created by you, which are private by default. You can grant access to other users to launch your private images.
Public AMI	Images created by users and released to the AWS community, so anyone can launch instances based on them and use them any way they like. AWS lists all public images at http://developer.amazonwebservices.com/connect/kbcategory.jspa?categoryID=171 .
Paid QAMI	You can create images providing specific functions that can be launched by anyone willing to pay you per each hour of usage on top of Amazon's charges.

Amazon EC2 Execution Environment

- Types of instances
 - **Standard instances** are well suited for most applications.
 - **Micro instances** provide a small number of consistent CPU resources and allow you to burst CPU capacity when additional cycles are available. They are well suited for lower throughput applications and Web sites that consume significant compute cycles periodically.
 - **High-memory instances** offer large memory sizes for high-throughput applications, including database and memory caching applications.
 - **High-CPU instances** have proportionally more CPU resources than memory (RAM) and are well suited for compute-intensive applications.
 - **Cluster compute instances** provide proportionally high CPU resources with increased network performance and are well suited for high-performance computing (HPC) applications and other demanding network-bound applications. They use 10 Gigabit Ethernet interconnections.

Amazon EC2 Execution Environment

- Instance Types Available on Amazon EC2 (October 6, 2010)

Compute Instance	Memory GB	ECU or EC2 Units	Virtual Cores	Storage GB	32/64 Bit
Standard: small	1.7	1	1	160	32
Standard: large	7.5	4	2	850	64
Standard: extra large	15	8	4	1690	64
Micro	0.613	Up to 2		Only EBS	32 or 64
High-memory	17.1	6.5	2	420	64
High-memory: double	34.2	13	4	850	64
High-memory: quadruple	68.4	26	8	1690	64
High-CPU: medium	1.7	5	2	350	32
High-CPU: extra large	7	20	8	1690	64
Cluster compute	23	33.5	8	1690	64

Amazon EC2 Execution Environment

- **Cost of Amazon On-Demand VM Instant Types (October 6, 2010)**

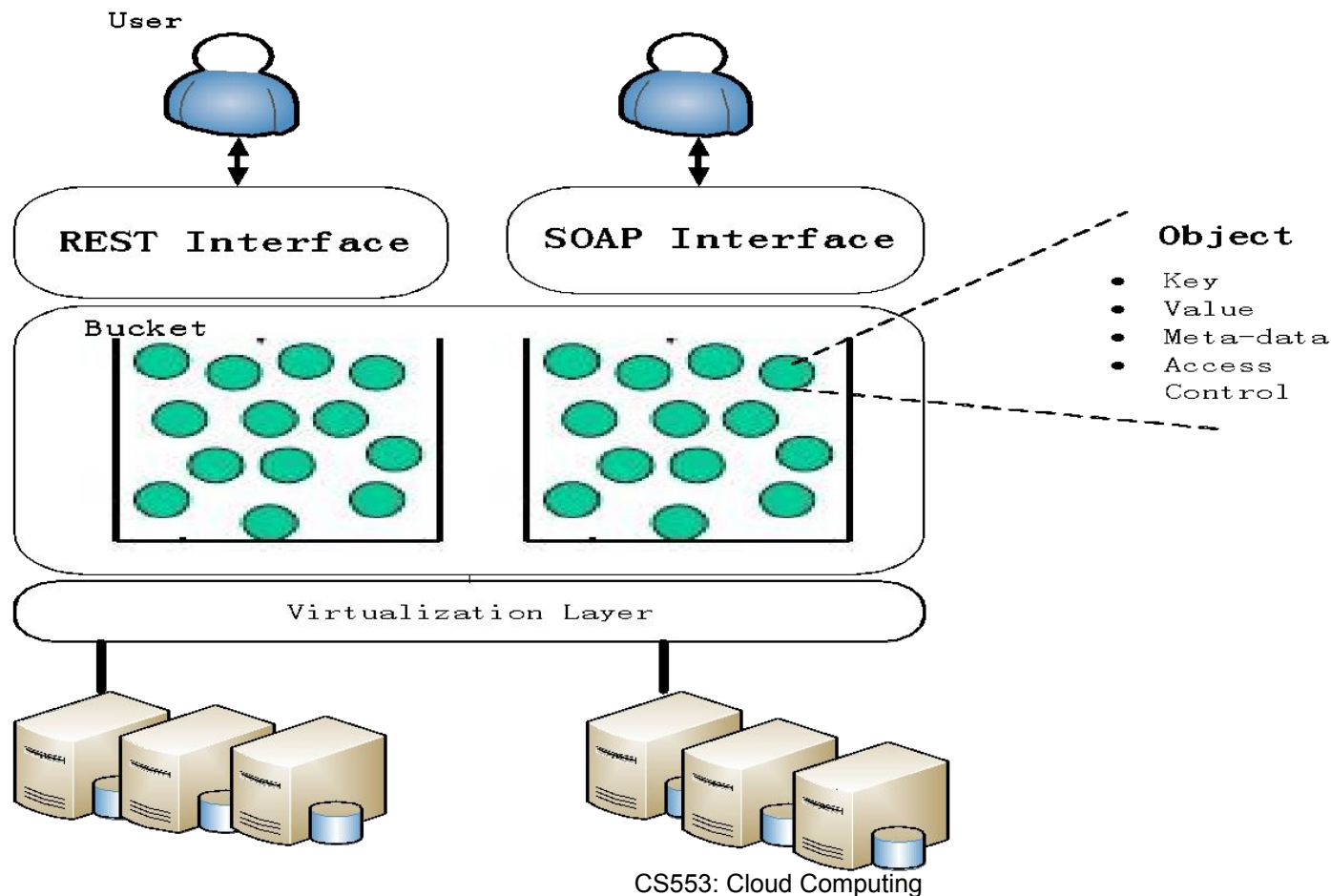
VM Instance Type	Size	Linux/UNIX Usage	Windows Usage
Standard instances	Small (default)	\$0.085 per hour	\$0.12 per hour
	Large	\$0.34 per hour	\$0.48 per hour
	Extra large	\$0.68 per hour	\$0.96 per hour
Micro instances	Micro	\$0.02 per hour	\$0.03 per hour
High-memory instances	Extra large	\$0.50 per 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
Cluster compute instances	Quadruple extra large	\$1.60 per hour	Not available

Amazon S3

- Amazon S3 provides a simple Web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the Web
- S3 provides the object-oriented storage service for users
- Users can access their objects through *Simple Object Access Protocol* (SOAP)
 - Supports both browsers and client programs which support SOAP
- SQS is responsible for ensuring a reliable message service between various processes, even if the receiver processes are not running

Amazon S3

- Amazon S3 execution environment



Amazon S3

- Fundamental operation unit of S3 → *object*
- Each object is stored in a *bucket* and retrieved via a unique, developer-assigned key
 - A bucket is the container of the object
 - Objects have other attributes such as values, metadata, and access control information
- S3 is essentially a key-value pair
 - Users can write, read, and delete objects containing from 1 byte to 5 gigabytes of data each
- Two types of Web service interfaces to access data stored in Amazon clouds
 - REST (Web 2.0)
 - SOAP

Amazon S3

- Key Features of S3:
 - Redundant through geographic dispersion
 - Designed to provide 99.999999999 percent durability and 99.99 percent availability of objects over a given year with cheaper reduced redundancy storage (RRS)
 - Authentication mechanisms to ensure that data is kept secure from unauthorized access
 - Objects can be made private or public, and rights can be granted to specific users
 - Per-object URLs and ACLs (access control lists)
 - Default download protocol of HTTP. A BitTorrent protocol interface is provided to lower costs for high-scale distribution
 - \$0.055 (more than 5,000 TB) to 0.15 per GB per month storage (depending on total amount)
 - First 1 GB per month input or output free and then \$.08 to \$0.15 per GB for transfers outside an S3 region
 - There is no data transfer charge for data transferred between Amazon EC2 and Amazon S3

Amazon Elastic Block Store (EBS)

- Traditional EC2 instances will be destroyed after use
- Note that S3 is “Storage as a Service” with a messaging interface
- The *Elastic Block Store* (EBS) provides the volume block interface for saving and restoring the virtual images of EC2 instances onto S3
 - The status of EC2 can now be saved in the EBS system after the machine is shut down
 - Users can use EBS to save persistent data and mount to the running instances of EC2
 - EBS is analogous to a distributed file system accessed by traditional OS disk access mechanisms
 - EBS allows you to create storage volumes from 1 GB to 1 TB that can be mounted as EC2 instances

Amazon Elastic Block Store (EBS)

- Multiple volumes can be mounted to the same instance
- These storage volumes behave like raw, unformatted block devices, with user-supplied device names and a block device interface
- You can create a file system on top of Amazon EBS volumes, or use them in any other way you would use a block device (like a hard drive)
- Snapshots are provided so that the data can be saved incrementally
 - This can improve performance when saving and restoring data
- Pricing → similar pay-per-use schema as EC2 and S3
 - Volume storage charges are based on the amount of storage users allocate until it is released, and is priced at \$0.10 per GB/month
 - EBS also charges \$0.10 per 1 million I/O requests made to the storage (as of October 6, 2010)
- The equivalent of EBS has been offered in open source clouds such as Nimbus

Amazon SimpleDB Service

- SimpleDB provides a simplified data model based on the relational database data model
 - Structured data from users must be organized into domains
 - Each domain can be considered a table
 - The items are the rows in the table
 - A cell in the table is recognized as the value for a specific attribute (column name) of the corresponding row
- SimpleDB is similar to a table in a relational database
 - However, it is possible to assign multiple values to a single cell in the table
 - This is not permitted in a traditional relational database which wants to maintain data consistency

Amazon SimpleDB Service

- Many developers simply want to quickly store, access, and query the stored data
- SimpleDB removes the requirement to maintain database schemas with strong consistency
- SimpleDB is priced at \$0.140 per Amazon SimpleDB Machine Hour consumed
 - First 25 Amazon SimpleDB Machine Hours consumed per month free (as of October 6, 2010).
- SimpleDB, like Azure Table, could be called “LittleTable,” as they are aimed at managing small amounts of information stored in a distributed table
- BigTable (from Google) is aimed at basic big data, whereas LittleTable is aimed at metadata
- Amazon Dynamo is an early research system along the lines of the production SimpleDB system

Next Semester:

CS554: Data-Intensive Computing

- Tour through various research topics in distributed data-intensive computing
 - Topics in cluster computing, grid computing, supercomputing, and cloud computing
 - Explore solutions and learn design principles for building large network-based computational systems to support data intensive computing
 - This course is a research focused course, with research papers as the reading material, and semester long group project as the primary basis of the evaluation
 - Last years website:

<http://www.cs.iit.edu/~iraicu/teaching/CS554-F13/>

Questions

