SAND2010-6789C

Cloud Computing Security

October 7, 2010

Dongwan Shin
Bill Claycomb
Vince Urias

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

Agenda

- **Introduction to Cloud Computing Security**
- Cloud Computing Implementation Considerations
- Role Based Access Control and Demo
- Cloud Computing Forensics
- The Future of Cloud Computing Security

Cloud Deployment Models

- Public
- Private
  - May be managed and hosted on-site or off-site
  - May be part of a public cloud, a *virtual private cloud*
- Community
  - Shared by several organizations
- Hybrid
  - Composition of two or more clouds
  - Data and application portability exists between clouds

Cloud Services Models

- **Software as a Service (SaaS)**
  - **Service provider delivers everything**
    - **Applications provided by the cloud**
  - **Google Apps**
- Platform as a Service (PaaS)
  - Customer deploys applications on to a platform provided by the cloud
  - Google App Engine
- Infrastructure as a Service (IaaS)
  - Only fundamental resources are provided
  - Customer responsible for deployment and management
  - Amazon EC2

Software as a Service (SaaS)

Returning user? Sign in here | La

Reliable, secure online applications wherever you work

Google Apps reduces IT costs and empowers today's employees. Gmail, Google Docs, Google Sites, and more - $50 per user per year. Try it free for 30 days.

Gmail for business 25GB storage, less spam, and a 99.9% uptime SLA, and enhanced email security.

Google Calendar Agenda management, scheduling, shared online calendars and mobile calendar sync.

Google Docs Documents, spreadsheets, and presentations. Work online without attachments.

Google Groups User-created groups providing mailing lists, easy content sharing, searchable archives.

Google Sites Secure, coding-free web pages for intranets and team managed sites.

Google Video Private, secure, hosted video sharing.

Switch to Google Apps

Learn how switching from Microsoft Exchange or Lotus Notes helps you save money and reduce IT hassles.

Estimate your cost savings.

Cloud Services Models

- Software as a Service (SaaS)
  - Service provider delivers everything
    - Applications provided by the cloud
  - Google Apps
- **Platform as a Service (PaaS)**
  - **Customer deploys applications on to a platform provided by the cloud**
  - **Google App Engine**
- Infrastructure as a Service (IaaS)
  - Only fundamental resources are provided
  - Customer responsible for deployment and management
  - Amazon EC2

Platform as a Service (PaaS)

Google code Search

e.g. "templates" or "datastore"

Google App Engine Home Docs FAQ

Run your web apps on Google's infrastructure.

Easy to build, easy to maintain, easy to scale.

Java™ Language Support

App Engine recently unveiled its second language: Java. This release includes our Java runtime, integration with Google Web Toolkit, and a Google Plugin for Eclipse, giving you an end-to-end Java solution for AJAX web applications. The Java runtime is now available for anyone to use, so please give it a try and send us your feedback.

- Get the full scoop in our blog post.
- Click over to YouTube to watch our Campfire One announcements.
- See our docs for other new features like cron support, database import, and access to firewalled data.

Get an overview of App Engine's new Java runtime and see a demo of a sample app from creation to deployment.
Watch Now

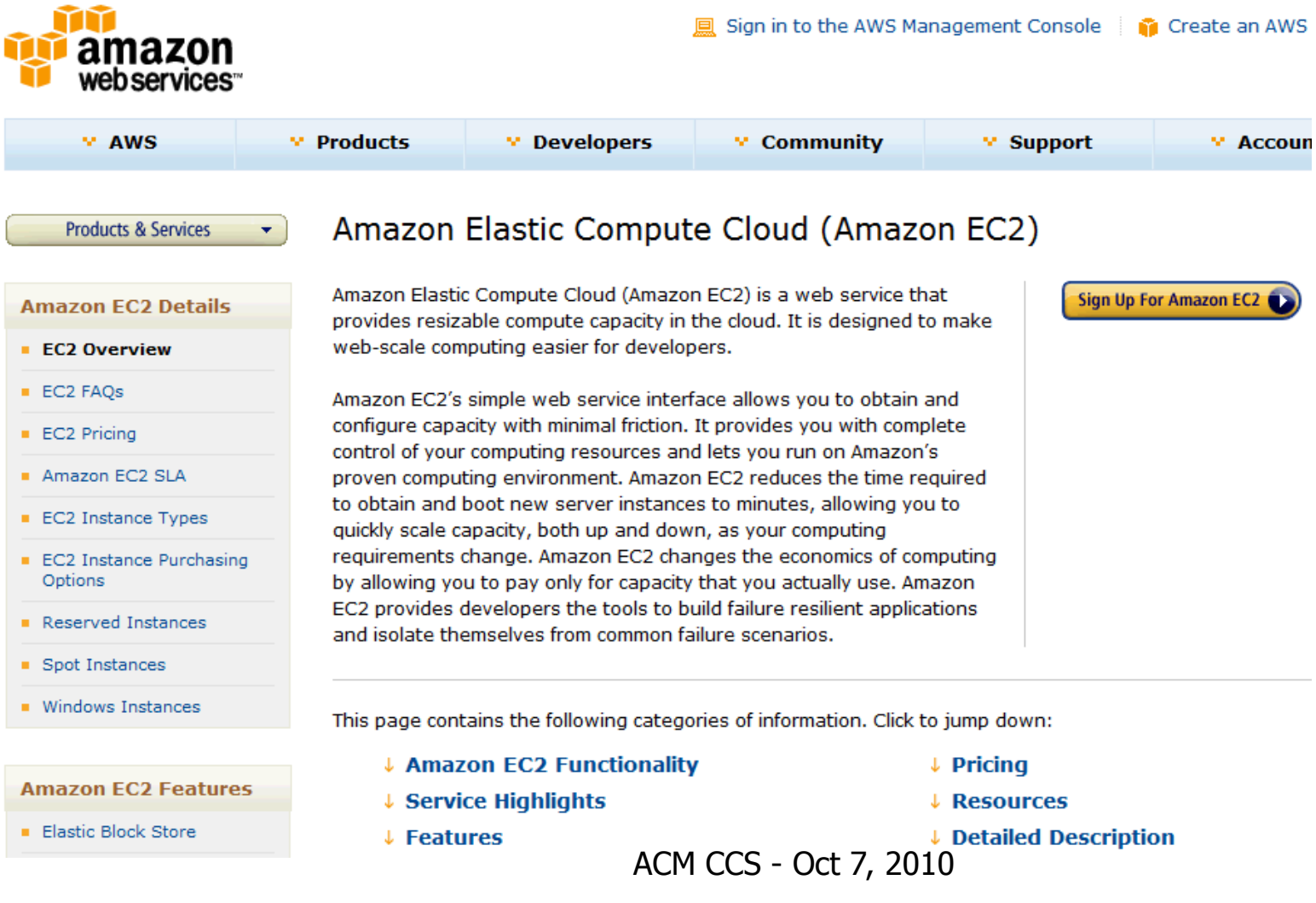
Grow Beyond The Free Quotas

App Engine developers can now purchase additional computing resources beyond the free quota limits. Scale your application to millions of users and pay only for what you use. App Engine will always be free to get started so you can try it out with no risk.

Cloud Services Models

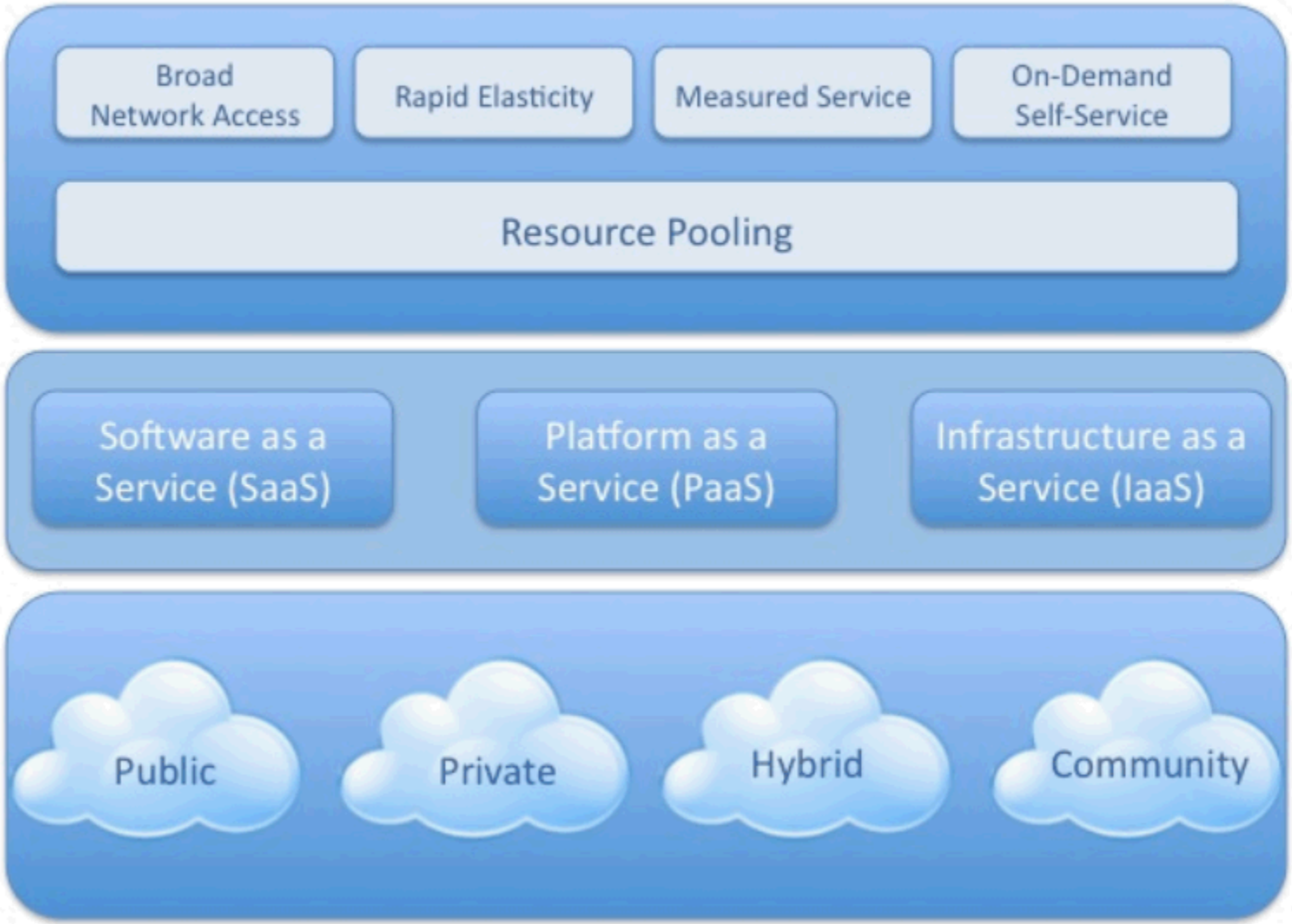
- Software as a Service (SaaS)
  - Service provider delivers everything
    - Applications provided by the cloud
  - Google Apps
- Platform as a Service (PaaS)
  - Customer deploys applications on to a platform provided by the cloud
  - Google App Engine
- **Infrastructure as a Service (IaaS)**
  - **Only fundamental resources are provided**
  - **Customer responsible for deployment and management**
  - **Amazon EC2**

Infrastructure as a Service (IaaS)

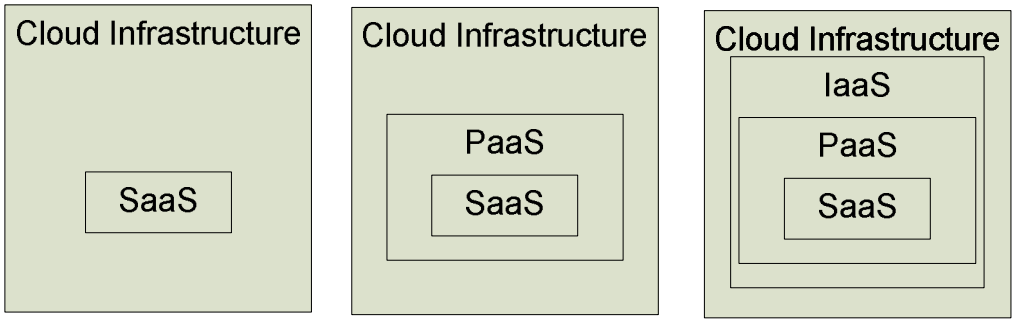


Visual Model Of NIST Working Definition Of Cloud Computing

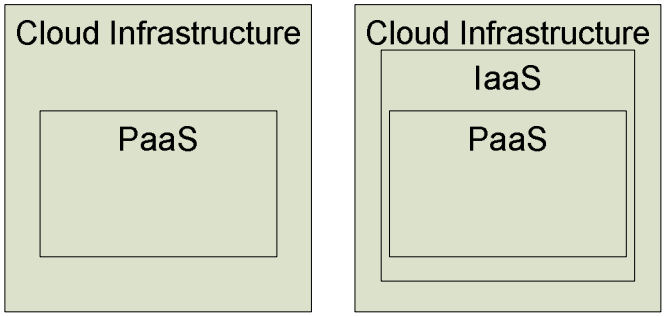
http://www.csrc.nist.gov/groups/SNS/cloud-computing/index.html



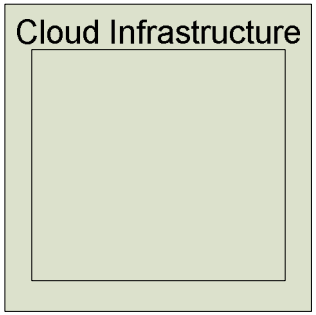
Service Model Architectures



Software as a Service (SaaS) Architectures



Platform as a Service (PaaS) Architectures



Infrastructure as a Service (IaaS) Architectures

Cloud Security Threats*

1. Abuse and Nefarious Use of Cloud Computing
2. Insecure Application Programming Interfaces
3. Malicious Insiders
4. Shared Technology Vulnerabilities
5. Data Loss/Leakage
6. Account, Service, and Traffic Hijacking
7. Unknown Risk Profile

* As Identified by the Cloud Security Alliance, 2010

Abuse and Nefarious Use

- Password and key cracking
- DDOS
- Launching dynamic attack points
- Hosting malicious data
- Botnet command and control
- Building rainbow tables
- CAPTCHA solving

- Exploits exist already

Prevention

Insecure Interfaces and APIs

- Could expose more functionality than intended
- Policy could be circumvented
- Credentials may need to be passed – is the interface secure?

Prevention

Malicious Insiders

- Particularly poignant for cloud computing
- Little risk of detection
- System administrator qualifications and vetting process for cloud services provider may be different that that of the data owner

Prevention

Shared Technology Issues

- Underlying architecture (CPU cache, GPU, etc.) not intended to offer strong isolation properties
- Virtualization hypervisor used to mediate access between guest OS and physical resources

- Exploits exist (Blue Pill, Red Pill)

Prevention

Data Loss or Leakage

- Data is outside the owner's control
- Data can be deleted or decoupled (lost)
- Encryption keys can be lost
- Unauthorized parties may gain access

- Caused by
  - Insufficient authentication, authorization, and access controls
  - Persistence and remanance
  - Poor disposal procedures
  - Poor data center reliability

Prevention

Account or Service Hijacking

- Exploits phishing attacks, fraud, or software vulnerabilities
- Credential reuse

Prevention

Unknown Risk Profile

- How well is the cloud being maintained?
  - Many companies are unwilling to release details
- Is the infrastructure up to date
  - Patches
  - Firmware
- Does the combination of different service providers create previously unseen vulnerabilities?

Prevention

Agenda

- Introduction to Cloud Computing Security
- **Cloud Computing Implementation Considerations**
- Role Based Access Control and Demo
- Cloud Computing Forensics
- The Future of Cloud Computing Security

Implementation Considerations

- **Demo**

Agenda

- Introduction to Cloud Computing Security
- Cloud Computing Implementation Considerations
- **Role Based Access Control and Demo**
- Cloud Computing Forensics
- The Future of Cloud Computing Security

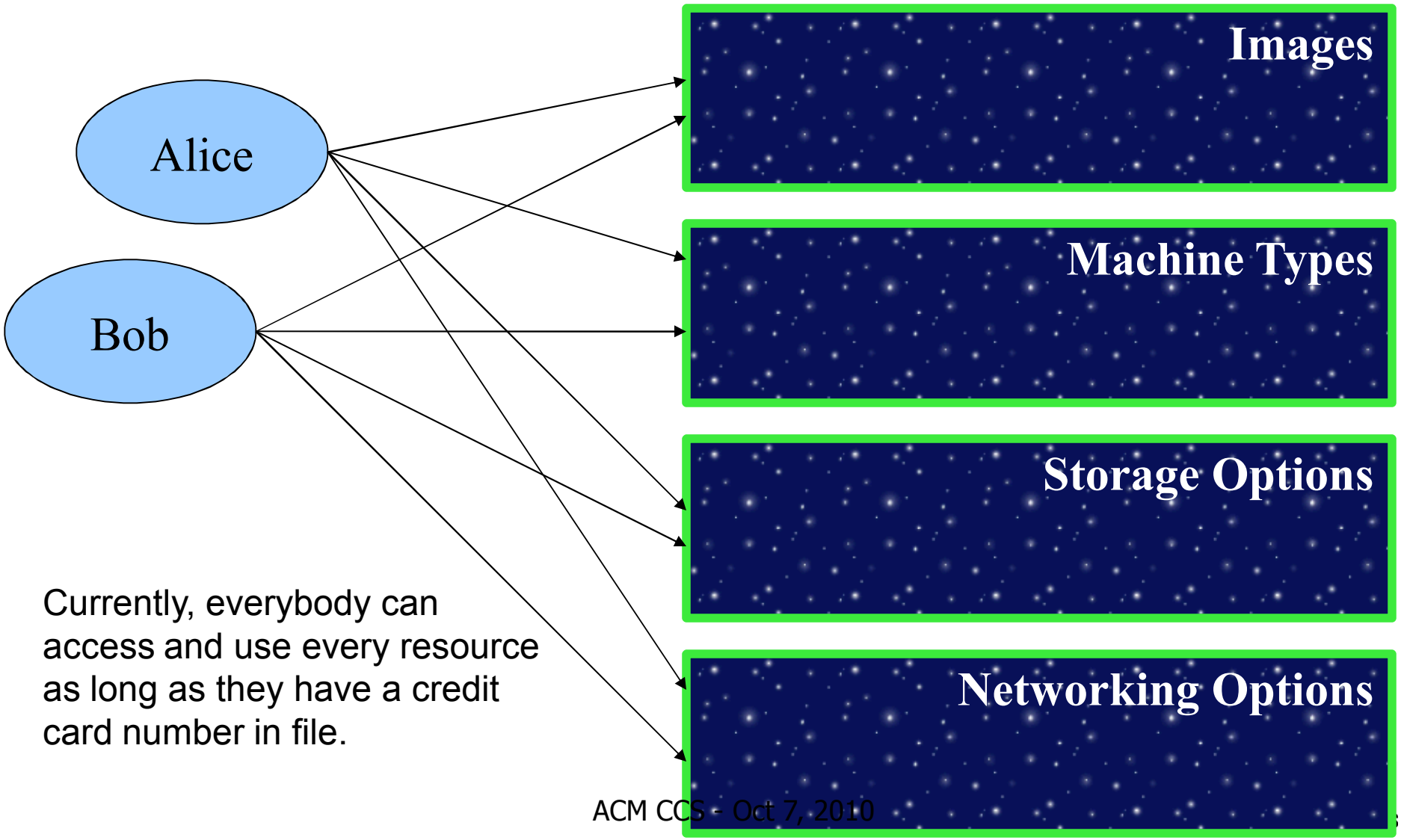
Motivation

- Access control in current IaaSes
  - User-resource direct mapping model
  - Utility (pay-as-you-go) computing
  - Very primitive access control support
  - Some of them only provide ACLs for images
- No organization-level security/governance policy support
- Inflexible pricing model for business

Motivating Example

- Assume that **SunnyTech.edu**, a fictitious university, decides to transition its university-wide information technology (IT) as well as department-wide service and research IT to the cloud services provided by **GoCloud.com**. The cloud services under consideration include not only software services such as email and HR applications, but also infrastructure services such as virtual machines with operating systems (OSes) having pre-configured application images installed. In addition, **SunnyTech.edu** decides to control access to, and thus limit the usage of, the virtualized resources based on the roles of the members of the university. Lastly, the university wants to keep access logs based on its audit policy as well as to implement other policies related to governance.

GoCloud.com



Currently, everybody can access and use every resource as long as they have a credit card number in file.

Motivating Example

- How to support an advanced access control such as role-based access control?
- How to implement organizational security and governance policies?
- How to support different pricing for individual users vs. businesses?

Our Objective

- We propose a domain-based framework for provisioning and managing users and resources in IaaS
  - Introduce the notion of domain to the user-resource direct mapping
  - Can address the three problems
- Provide a proof-of-concept implementation using Eucalyptus

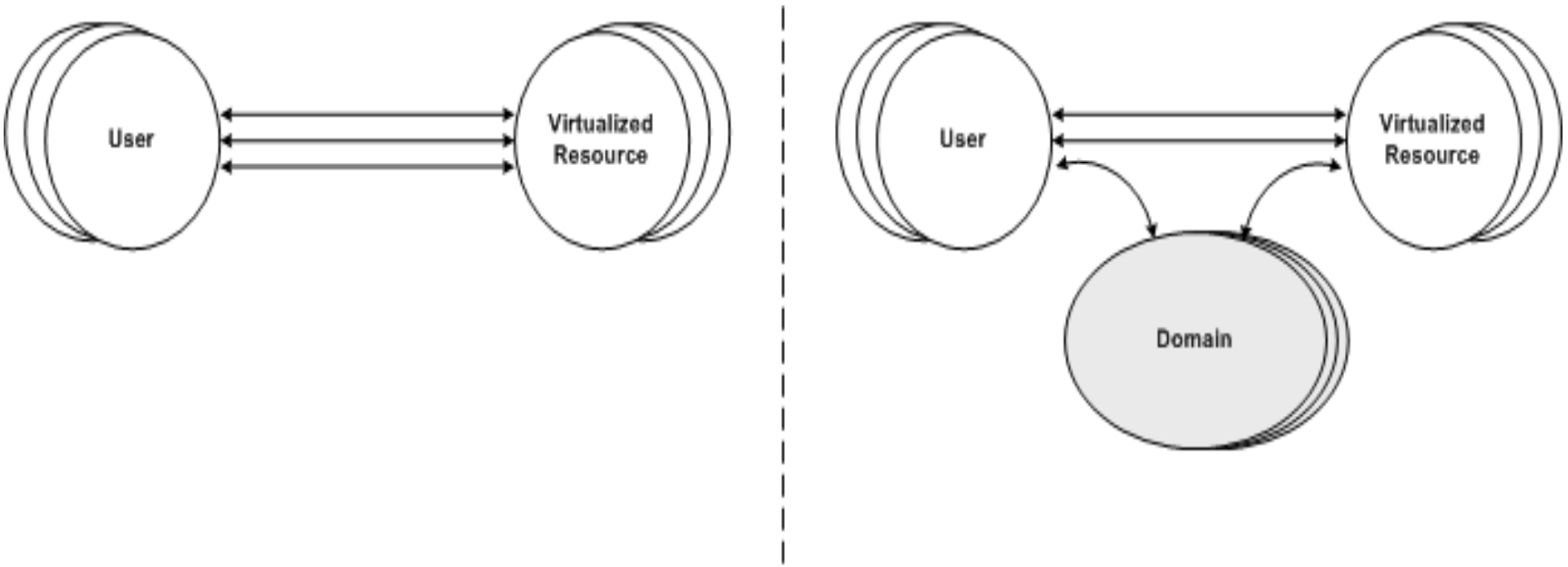
Background – Role Support in Clouds

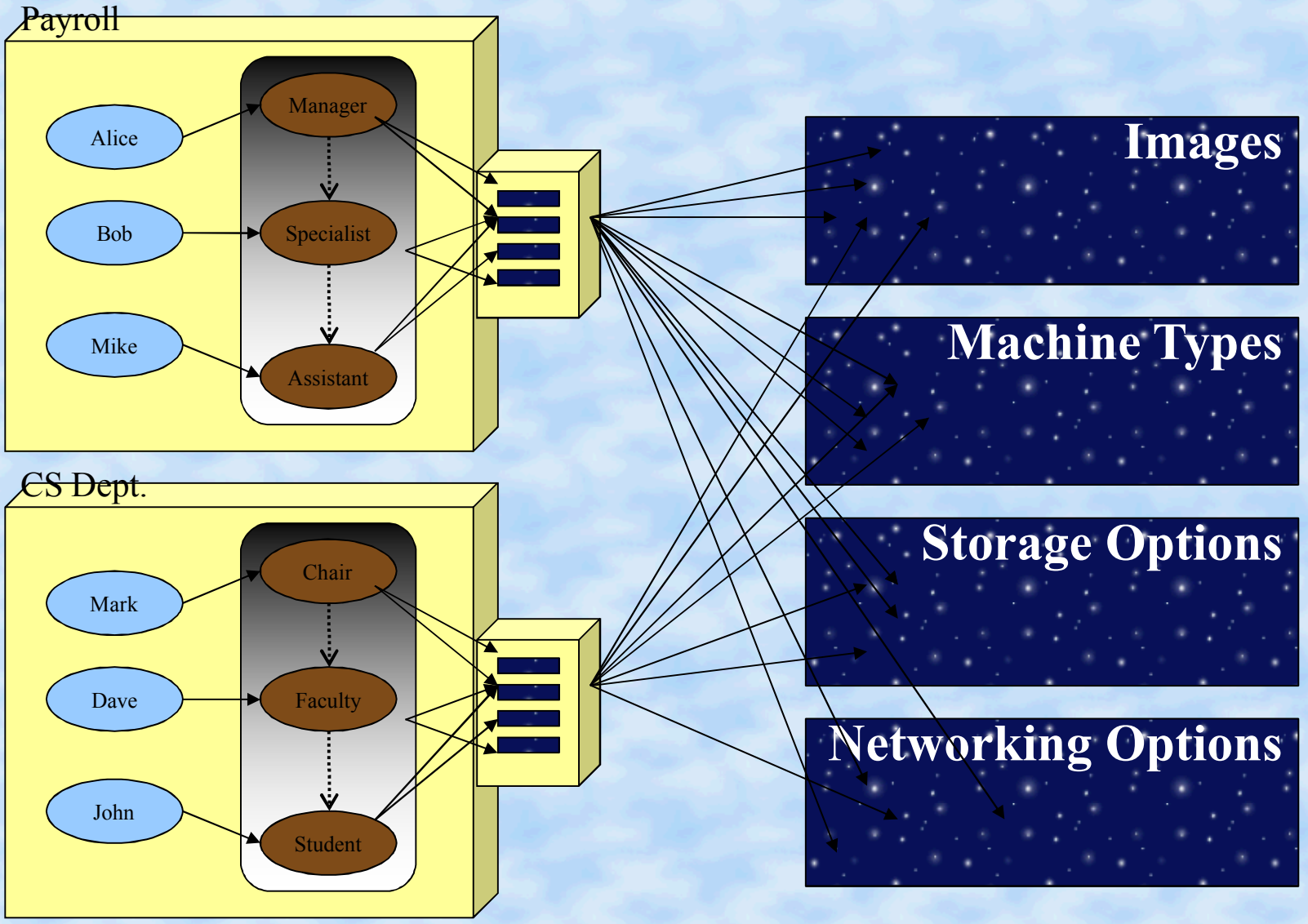
- Most of existing IaaSes do not support the notion of grouping
  - Amazon EC2
  - Nasa's Nebula
  - Windows Azure

Our Approach

- We want to add an access control mechanism so that the all-or-nothing approach will not be supported any more
- We want to group users and permissions so that all users belonging to the same group have exactly the same permissions
- Furthermore, we want to organize groups into a hierarchy
- Furthermore, we want to have separate administration domains

Our Approach




Payroll
Alice
Bob
Mike
Manager
Specialist
Assistant
CS Dept.
Mark
Dave
John
Chair
Faculty
Student
Images
Machine Types
Storage Options
Networking Options

Design

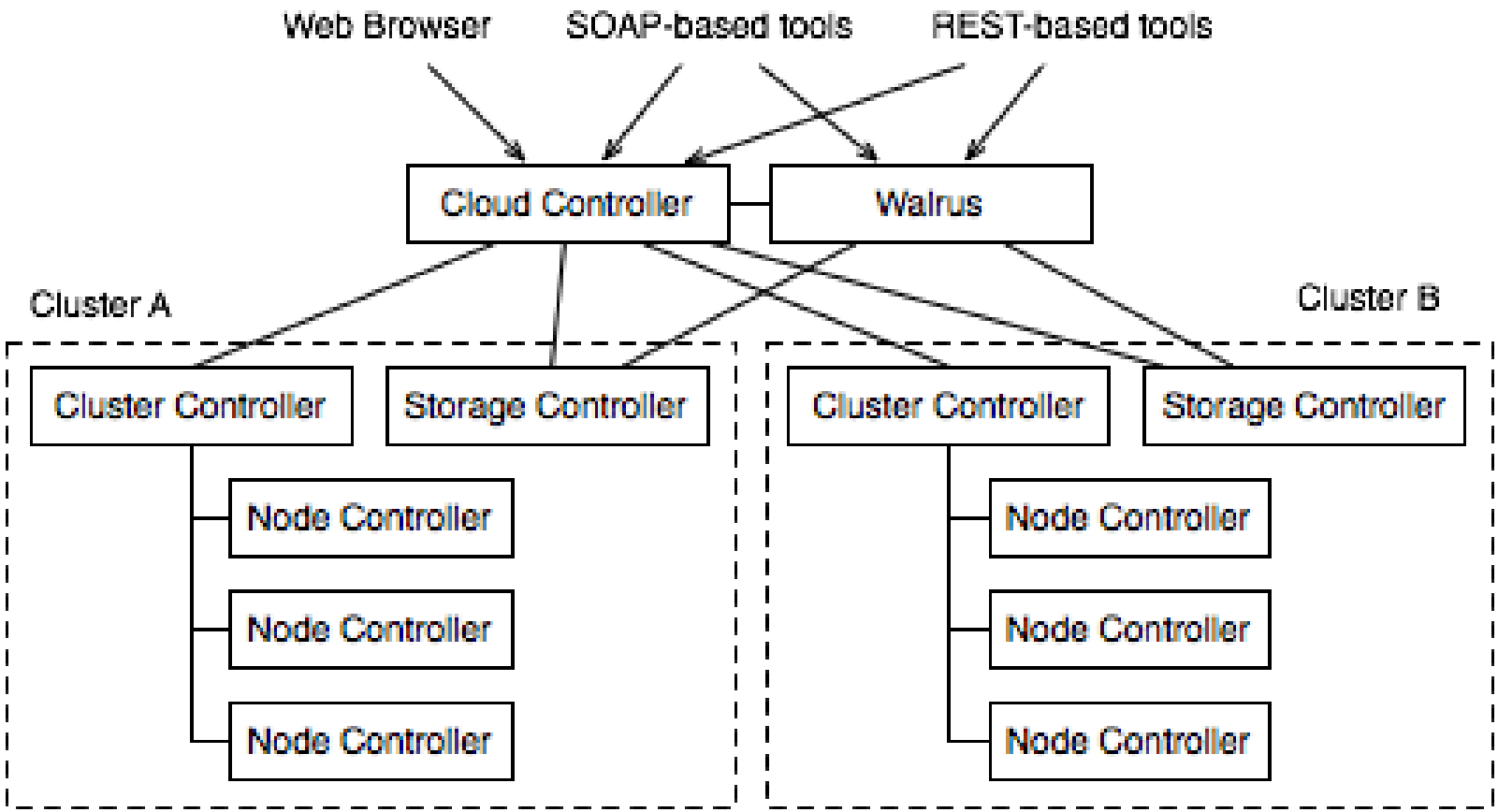
Cloud vs. Domain Administrators

- Domains create more indirection
  - Less messy mapping
- Cloud administrator:
  - Creates a domain
  - Assigns permissions to the domain
  - Creates a domain administrator
  - Manages the cloud
- Domain administrator:
  - Creates roles and builds a hierarchy
  - Creates users and assigns them to roles
  - Manages the domain

Implementation

- Open source project Eucalyptus v1.6.2
  - Designed to be exact clone of Amazon EC2
- Uses other open source Linux tools
  - Xen/KVM hypervisors
  - DHCP/Iptables/VLAN for network management
  - AoE for storage devices

Eucalyptus Architecture

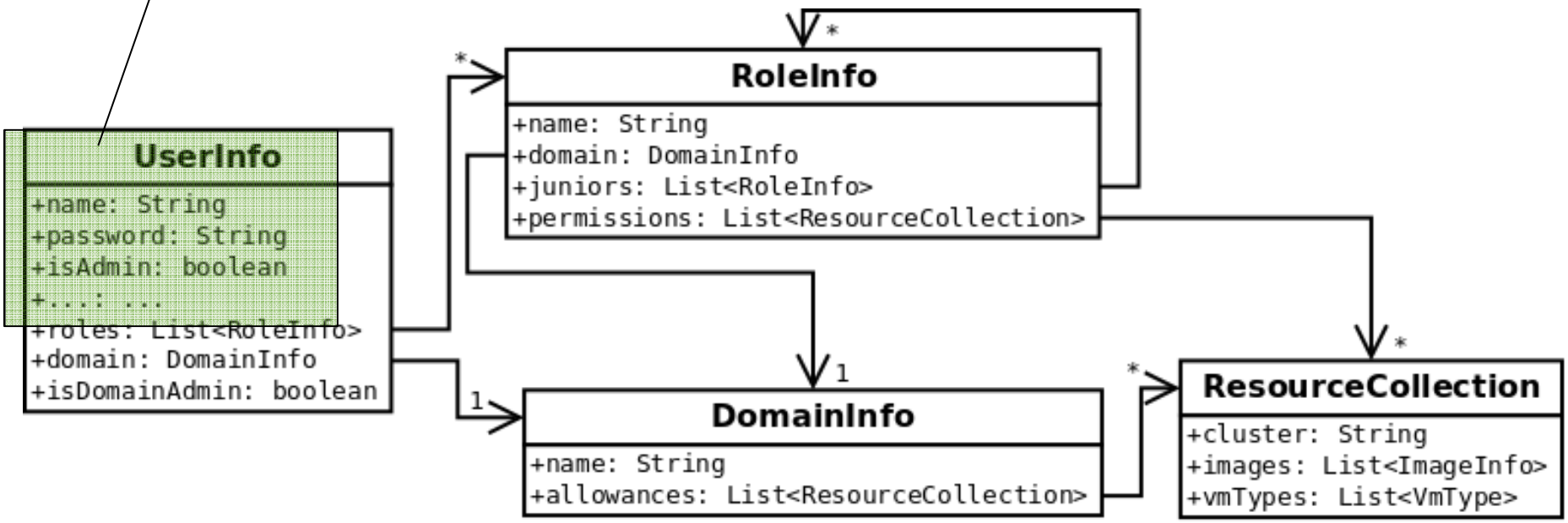


Implementation

- Our additions are only in the Cloud Controller
- Created objects that represent roles, domains, and permissions
- Modified the web administration interface for domain, role, and user management
- Added a reference monitor that makes the access control decisions

Implementing Objects

Original UserInfo object and attributes

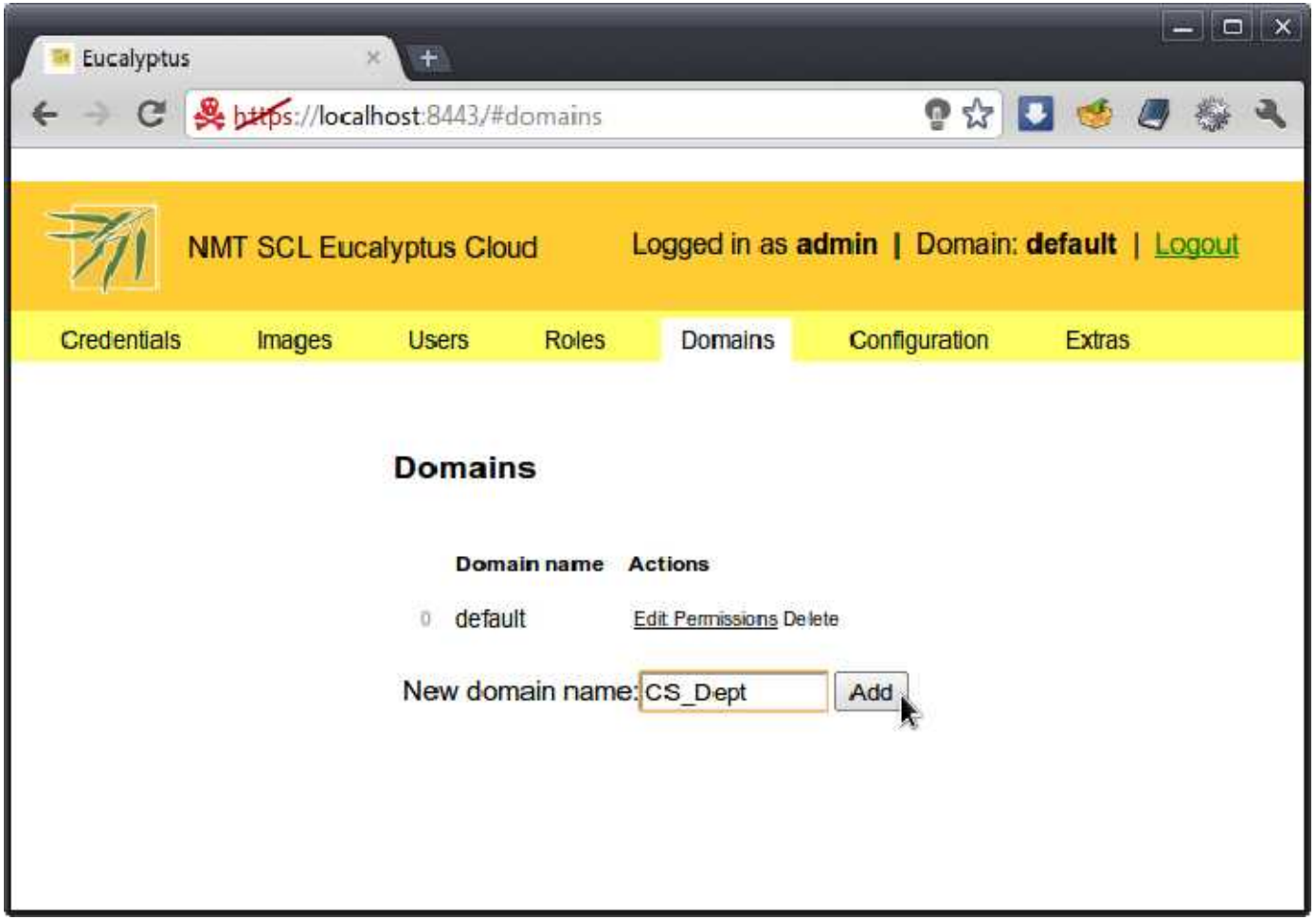


Representing Permissions

- **A ResourceCollection object represents all the images and virtual machine types that the domain/role is allowed to use in the specified cluster:**

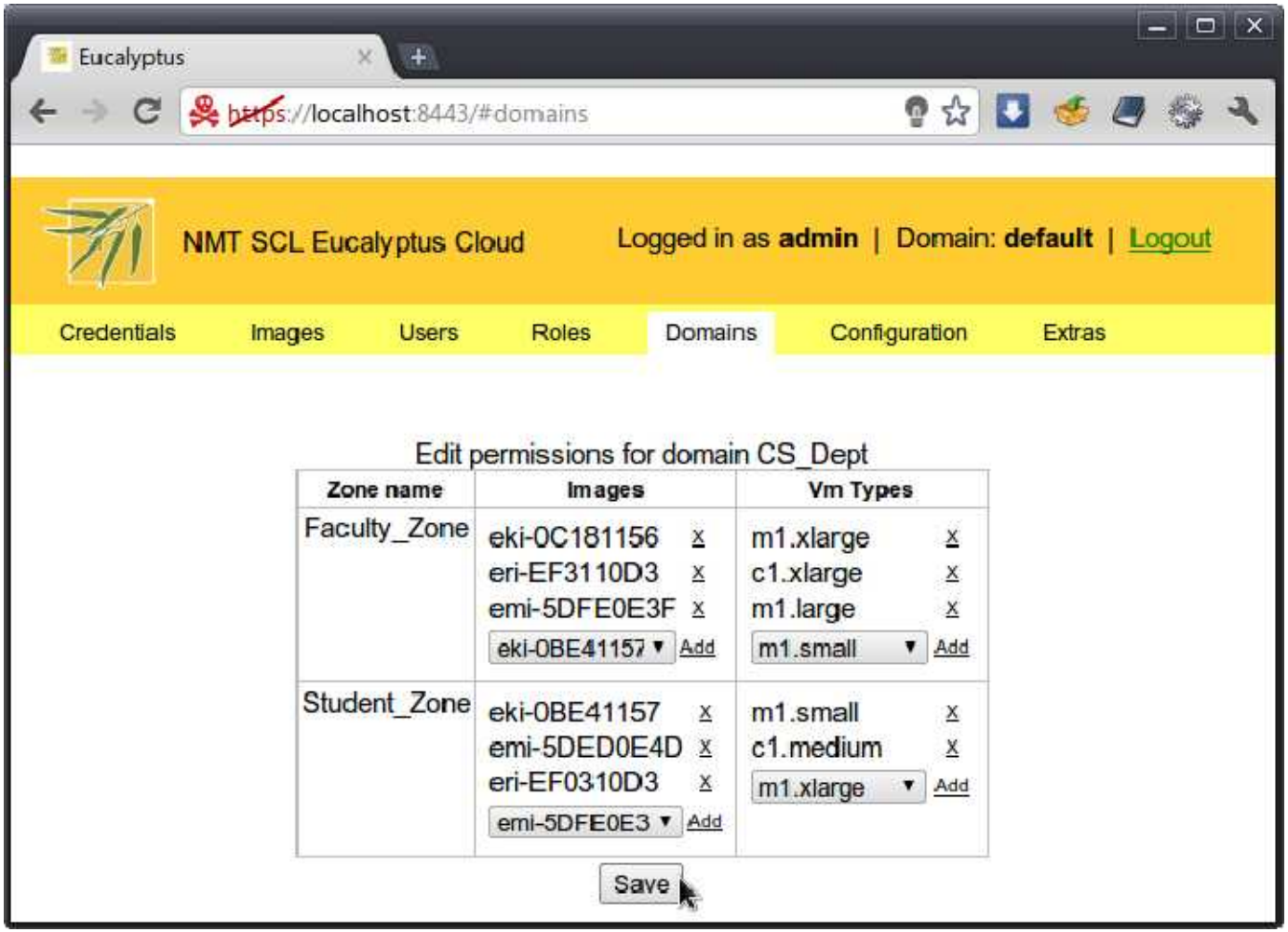
```
RC = {
cluster=”ZoneA”,
images=[“emi-AAAAAA”,”eri-BBBBBB”],
vmTypes=[“m1.small”, “m1.medium”]
}
```

Creating a New Domain

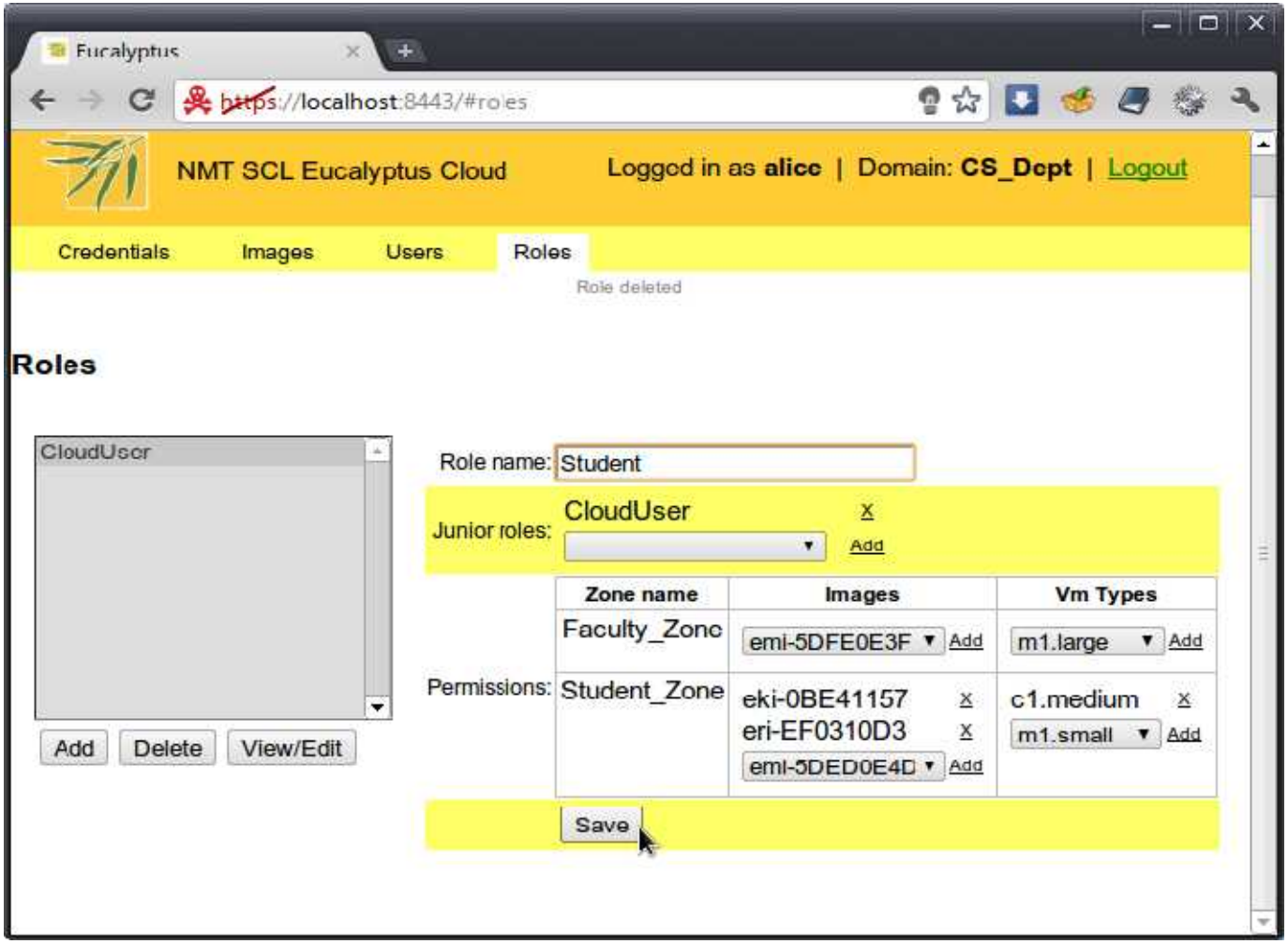


Sandia National Laboratories

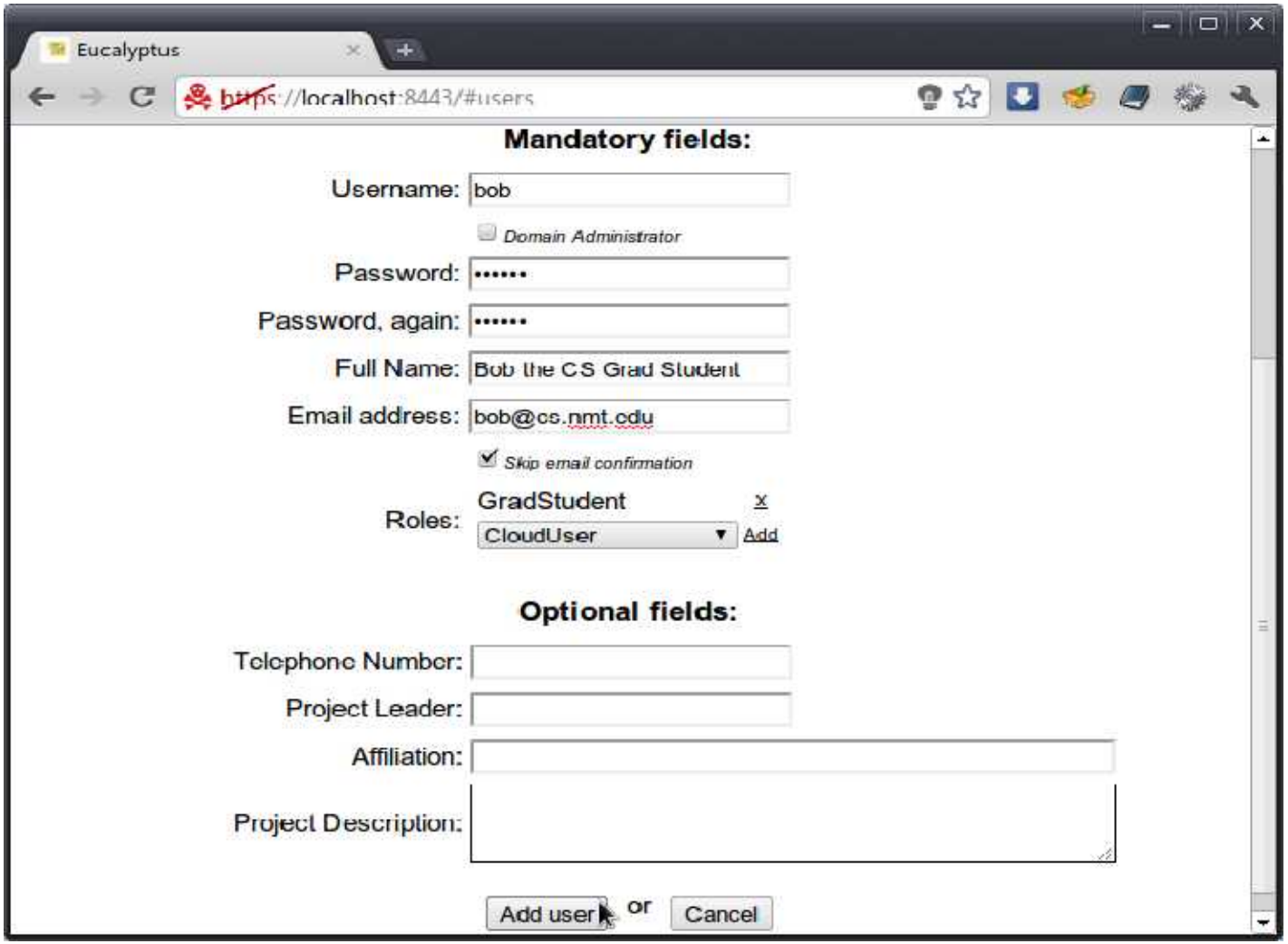
Domain Permission Management



Domain Role/Hierarchy Management



Domain User Management



Demo

Agenda

- Introduction to Cloud Computing Security
- Cloud Computing Implementation Considerations
- Role Based Access Control and Demo
- **Cloud Computing Forensics**
- The Future of Cloud Computing Security

Road Map

- Background
  - What is digital forensics?
- Problem statement
  - Focus on cyber infrastructure
- Approach
  - Analysis methodology
- Proposed solutions
- Case study
- Conclusion

Introduction/Problem

- Cloud computing is a disruptive technology for the forensics field
- I hypothesized that several types of standard techniques would no longer work in the cloud
- Certain Disk forensics techniques maybe come obsolete
  - But which ones? And why?
- As of right now, nobody has published anything; no white papers, reports, etc on the forensic implication on the cloud
  - Talked to vendors, they laughed
  - Little discussion on forums and blogs
- Goals:
  - Begin to identify the challenges to digital forensics and propose some solutions to the problem

What is Digital Forensics?

- Definition: “Tools and techniques to recover, preserve, and examine digital evidence on or transmitted by digital devices”
- What are digital devices?
  - Computers
  - Personal Digital Assistants (PDAs)
  - Storage Devices (SD cards, CD/DVD, Zip Disks, etc)
  - Cell Phones
  - Videogame Consoles (Xbox, Playstation, Wii, ..)
  - Printers

Digital Forensics Process

Legal Implications: Balance of need to investigate vs. privacy

What are We Looking For?

Inappropriate Emails

- Death threats
- Blackmailing
- Corporate espionage

Illicit Contraband

- Child pornography
- Explosives diagrams
- Software piracy
- Malicious software

Digital Communications

- Artifacts of network connections
- Cell phone messages

THE ROLE OF DIGITAL FORENSICS IN THE CLOUD

Sandia National Laboratories

Cyber Infrastructure

- **Clouds sit on top of a cyber infrastructure**
  - ***The comprehensive infrastructure needed to capitalize on dramatic advances in information technology has been termed cyberinfrastructure***
- **Cyberinfrastructure consists of a variety of things, one of which can be virtual technologies to quickly deploy reconfigurable environments…**
  - **VMware ESX, ESXi, vSphere**
  - **Zen**
  - **Citrix**

Industry Is Accepting Cloud Computing

- As the industry becomes increasingly popularized, several leaders are at the forefront of the technology
  - Microsoft, Google, Amazon, and IBM
- Showing their support through the promotion, encouragement, adoption, and leadership of cloud computing
- Security postures, architectures, and auditing methodologies have yet to be developed for the cloud; other areas, such as digital forensics, are also still up for exploration.

As Acceptance Grows, So Does Risk

- With more and more users starting to use services like EC2, there is increased risk
  - Need to define deficiencies
  - Need to start finding solutions earlier
- Don't want to be like other areas of Digital Forensics
  - Databases, Live Forensics
  - Need to stop being reactive and being to be proactive
- Has lots of opportunity for positive change

RESEARCH APPROACH AND ISSUES

Current Lack of Research/Methods

- No research or white papers on digital forensics in the clouds
- No foundational work on how to extend method
- Research Methods
  - Started by creating own research method
    - Looked at last 15 years of digital forensics
    - Found best/most flexible methodology
  - Used primary source data
  - Used experimental evidence

Approach

- To test my theory, these three methods were used
  - Combing Internet blogs and forums
  - Creating a matrix to compare every forensic step and brainstorm issues
  - Looked at the Grid and other predecessors to cloud computing
  - Conducting a small experiment
- Focused on two products
  - EC2 and VMware Products
- Using four phases of forensics
  - Identification
  - Acquisition
  - Analysis
  - Reporting
- Added Legal Issues
- Proposed Solution

FOUR PHASES OF FORENSICS

Identification

Easier

- No longer have physical media to identify
- Don’t have to worry about media compatibility in the lab

Harder

- Amazon Machine Images
- Amazon EC2 enables you to increase or decrease capacity within minutes
- Users have complete control of your instances
- Elastic IP address is associated with your account not a particular instance

Acquisition

Easier

- Reduced need for proprietary hardware
- VMs are just containers so they are portable
- EC2 Hashing
  - "Unimpeachable record"
  - Smart acquisition
- Reduced business down time
  - No longer need live forensics
  - VMotion
- Smaller 'devices' sizes

Harder

- No longer have write-blockers
- Focused solely on network acquisitions
  - Increased chance of failure
  - If network goes down?
- Increased dependence on infrastructure
  - Network speeds, etc
  - Failed network card on machine
- Verification of original VM
  - Does the service provider have to keep the device ?

Analysis

Easier

- Typical operations are supported
- Time stamping is improved
- Hashing is improved
- New, third party logger
  - Well documented that time on the cloud can be logged and validated
  - EC2 has the capacity

Harder

- Ephemeral nature of cloud computing
- Loss of artifacts
- Loss of intent and history of intent
- Virtual machine issues
  - Snapshots
  - Amazon Machine Image (AMI) in the EC2 environment: image is reverted to an original stored state in which the entire session data is removed and lost

Analysis Cont…

Harder

- Virtual hardware
  - Virtual Discs
  - Unallocated space
    - How long before its reclaimed?

Harder

- "You're kind of [at a loss] because temp files and registry settings are virtualized and deleted when the program exits."
- Recovering this data will be a harder challenge, in which data, logs, and services are spread across a multitude of servers and datacenters that could easily span different states or countries

Reporting

- Presentation of the analysis is the same
- Only change may be new visualizations of evidence on the cloud
- Other wise, it's the same..

LEGAL CONCERNS

Legal Issues

- Notion of liability
  - Who owns the data?
    - What if there is a hardware exploit and EC2 doesn't patch?
  - Is EC2 responsible for operating systems that are managed and operated by a user?
  - Does EC2 have to maintain and secure the data while mitigating risks to a person's virtual machine?

Legal Issues

- Cyber crimes
  - What if a crime is committed by a user on the cloud?
  - Is EC2 responsible for that crime by aiding them and by providing resources to commit the crime?
- What if the users were to start launching DDoS from EC2 equipment?
  - Does EC2 have the authority to stop the action?
    - If so, do they have a responsibility to prevent the action?
  - Should they actively scan for malware, etc?
- What about cyber warfare?
  - "[the US] would not rule out a kinetic response to a cyber attack"
  - What if they used EC2 as a cyber platform?

SOLUTIONS

Solution: Hypervisor Forensics

- Creation of a new sub domain in forensics
- Need to monitor and tap things at the hypervisor level
- Log changes to the VM
    - Virtual disk creations and removals, etc
- Creation of virtual network taps
- General logging and state of health of VMs using the hypervisor

Solution: Network Forensics

- Growing research area
  - 10+ papers
  - No real-world implementation
- Speculate a trend from physical disk forensics to network forensics
  - More network artifacts
  - Can externally validate
- Need a distributed forensic solution
  - Use the cloud to process the data?
    - Hadoop, Map Reduce

CASE STUDY

Case Study: Background

- Wanted proof of theories and ideas
- EC2 isn't free
- Erected a cloud-like infrastructure
  - ESXi is a widely accepted cloud infrastructure utility that provides the ability to provision and instantiate virtual environments.
- ESX is the foundation of many other cloud utilities, such as VDI or the more recent vSphere, all of which extend virtualization infrastructure to the cloud

Case Study: Construction

- Started by creating a small enterprise network
  - Realized it wasn't needed
- Started by using the tool
  - VMware's ESXi virtualization technology
  - Used as foundation for many cloud infrastructures
    - VDI, vSphere ..
- Tested 5 areas
  - Snapshot artifacts
  - Acquisition
  - Virtual devices
  - Typical forensics operations

Results

- Disappointing
  - Snapshots break everything
    - artifacts disappear, even unallocated space
  - Acquisitions breaks
    - Tools don't support vms/network acquisitions
  - Virtual drives
    - go undetected (their removal or addition)
  - Forensics operations
    - Once the acquisitions is resolved, work the same, better since VM are traditional smaller than normal devices

Agenda

- Introduction to Cloud Computing Security
- Cloud Computing Implementation Considerations
- Role Based Access Control and Demo
- Cloud Computing Forensics
- **The Future of Cloud Computing Security**

Future Directions

Questions

http://www.sandia.gov

http://scl.cs.nmt.edu

doshin@nmt.edu

wrclayc@sandia.gov

veuria@sandia.gov