

Datacenter Virtualization: Global-scale software switches

Hakim Weatherspoon

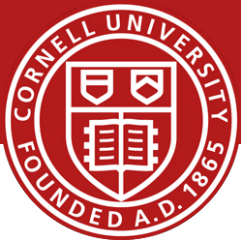
Assistant Professor, Dept of Computer Science

CS 5413: High Performance Systems and Networking

May 8, 2017

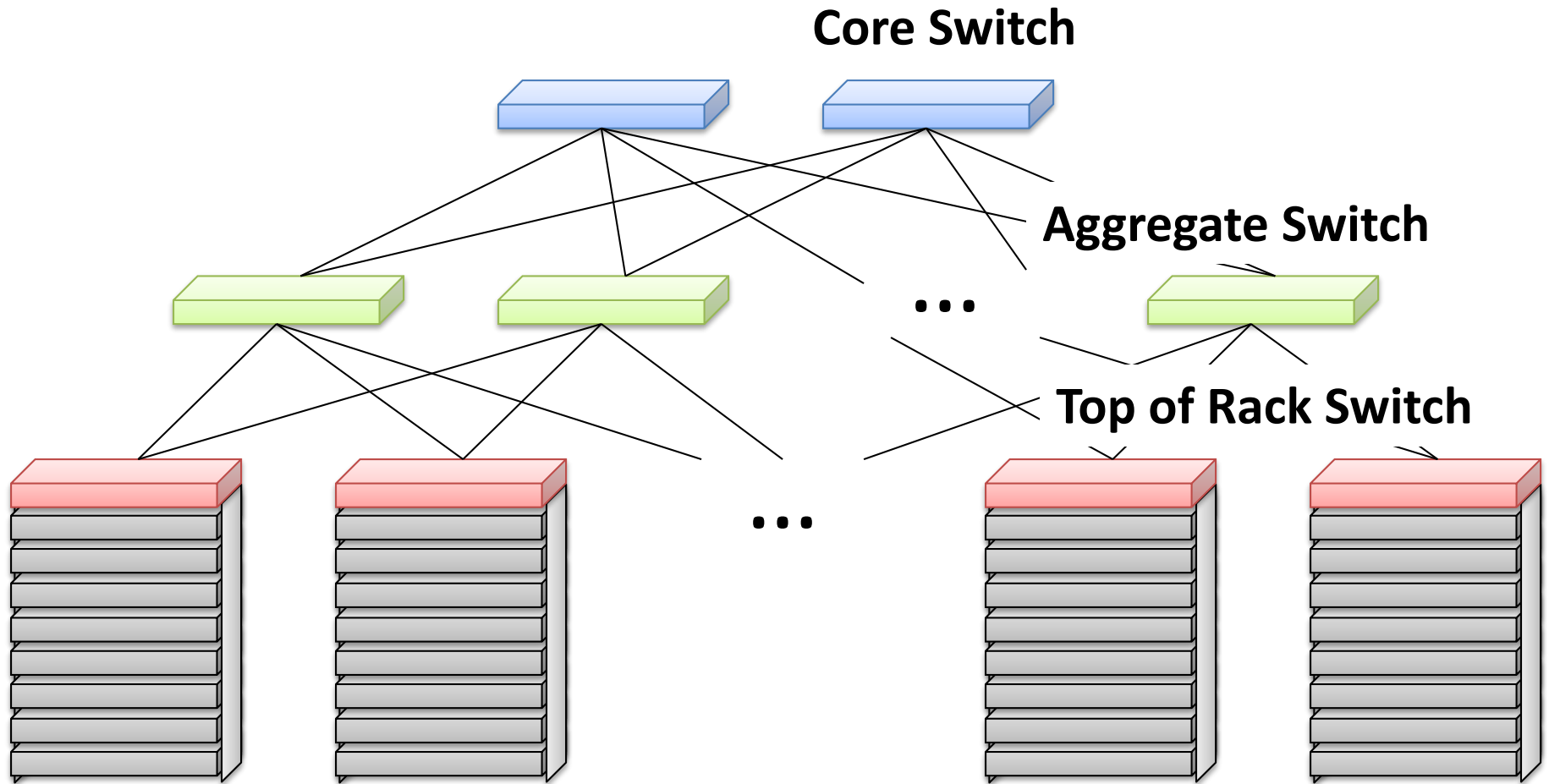
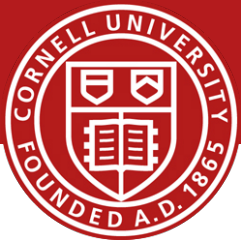
Slides from “Follow the Sun through the Clouds: Application Migration for Geographically Shifting Workloads “, Z. Shen, Q. Jia, E. Sela, B. Rainero, W. Song, R. van Renesse, H. Weatherspoon. In Proceedings of the ACM Symposium on Cloud Computing (SoCC), October 2016.

Where are we in the semester?

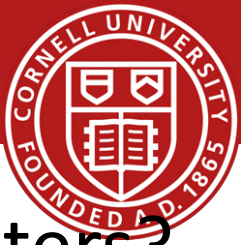


- Interested Topics:
 - SDN and programmable data planes
 - Disaggregated datacenters and rack-scale computers
 - Alternative switch technologies
 - Datacenter topologies
 - Datacenter transports
 - Advanced topics

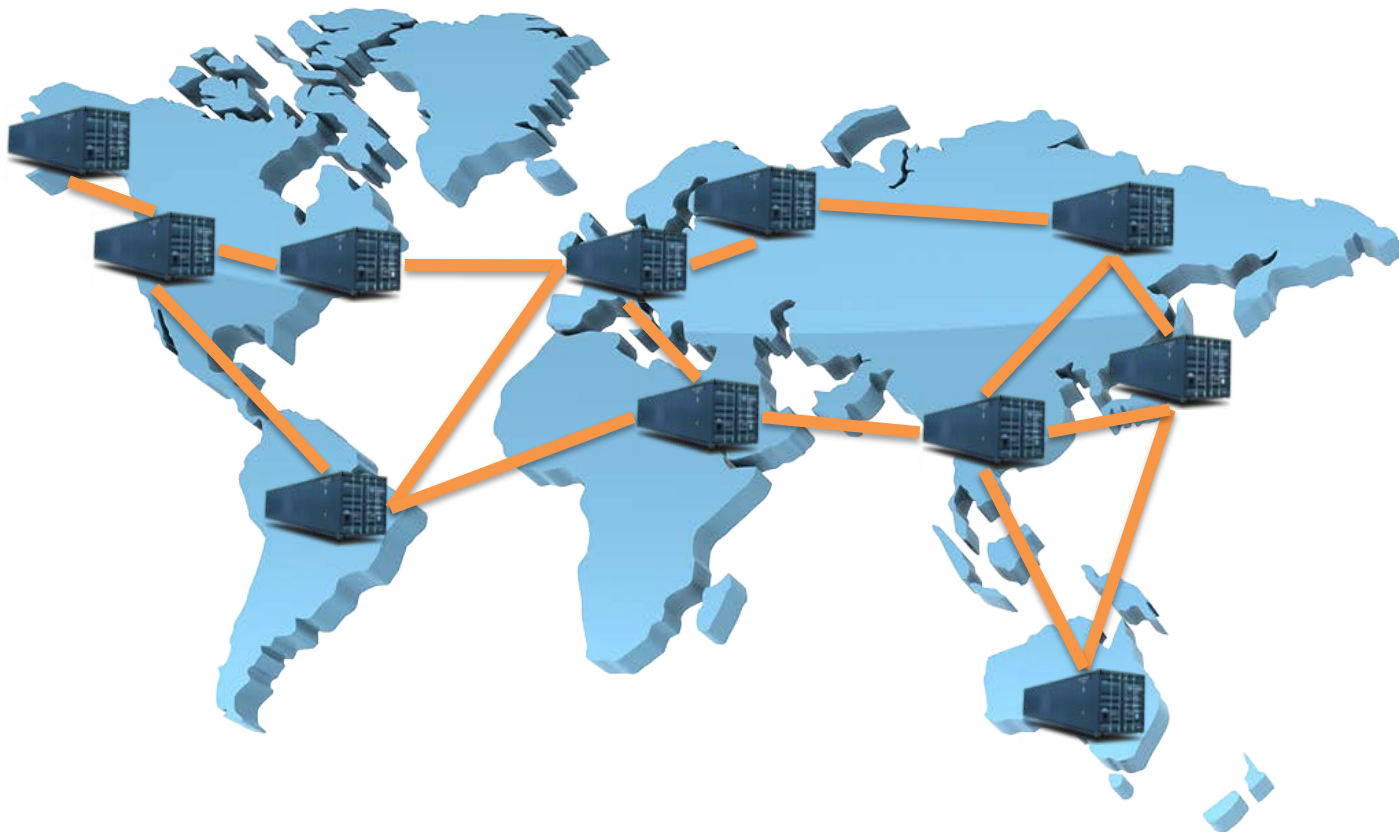
High Performance Systems and Networks



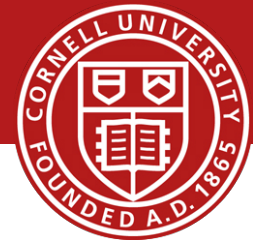
High Performance Systems and Networks



- How to optimize a global network of data centers?
- And, across multiple heterogeneous clouds

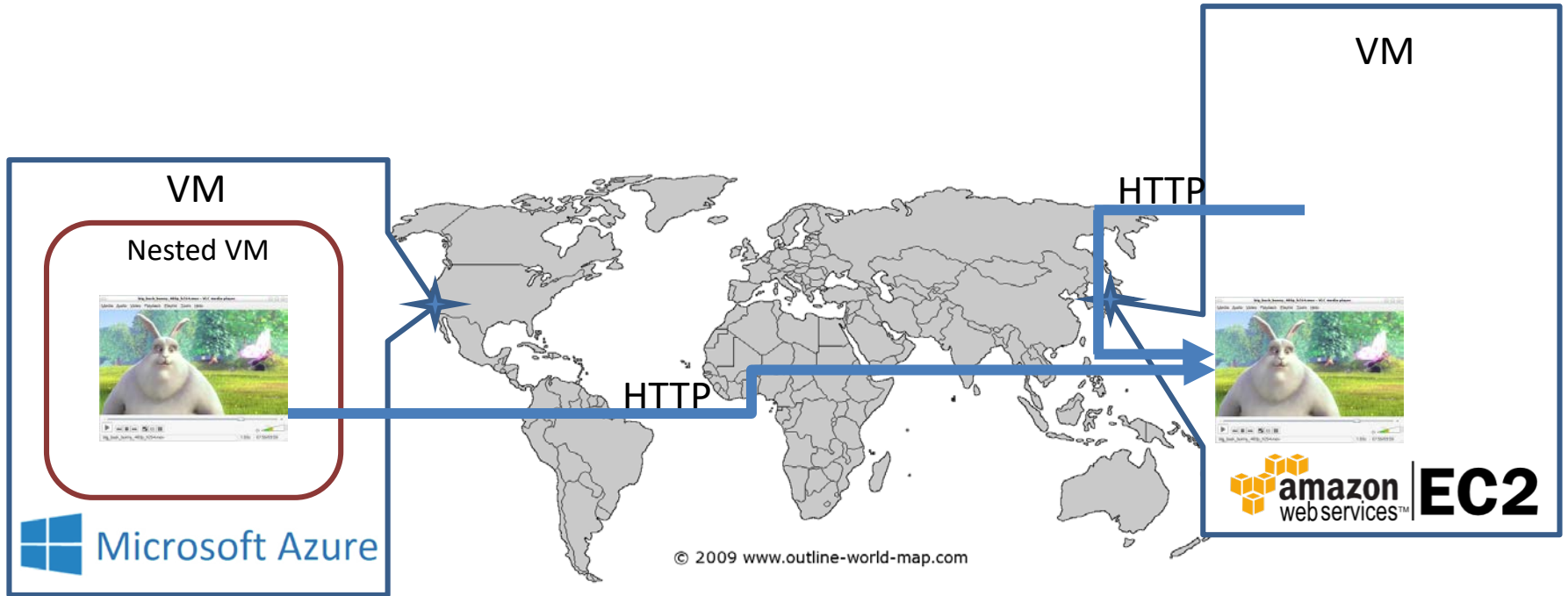
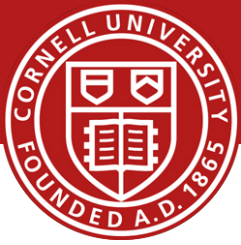


Goals for Today

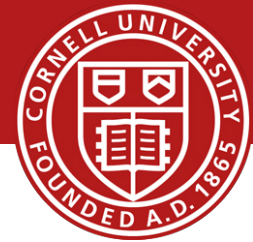


- Plug into the Supercloud
 - D. Williams, H. Jamjoom, and H. Weatherspoon. IEEE Internet Computing, Vol. 17, No 2, March/April 2013, pages 28-34
- Follow the Sun through the Clouds: Application Migration for Geographically Shifting Workloads
 - Z. Shen, Q. Jia, E. Sela, B. Rainero, W. Song, R. van Renesse, H. Weatherspoon. In the ACM Symposium on Cloud Computing (SoCC), October 2016.

Supercloud Demo



Highlights



- Automatic VM placement and migration
- Migrated VMs are LIVE
- IP addresses are not changed
- TCP connections are not broken

Demo



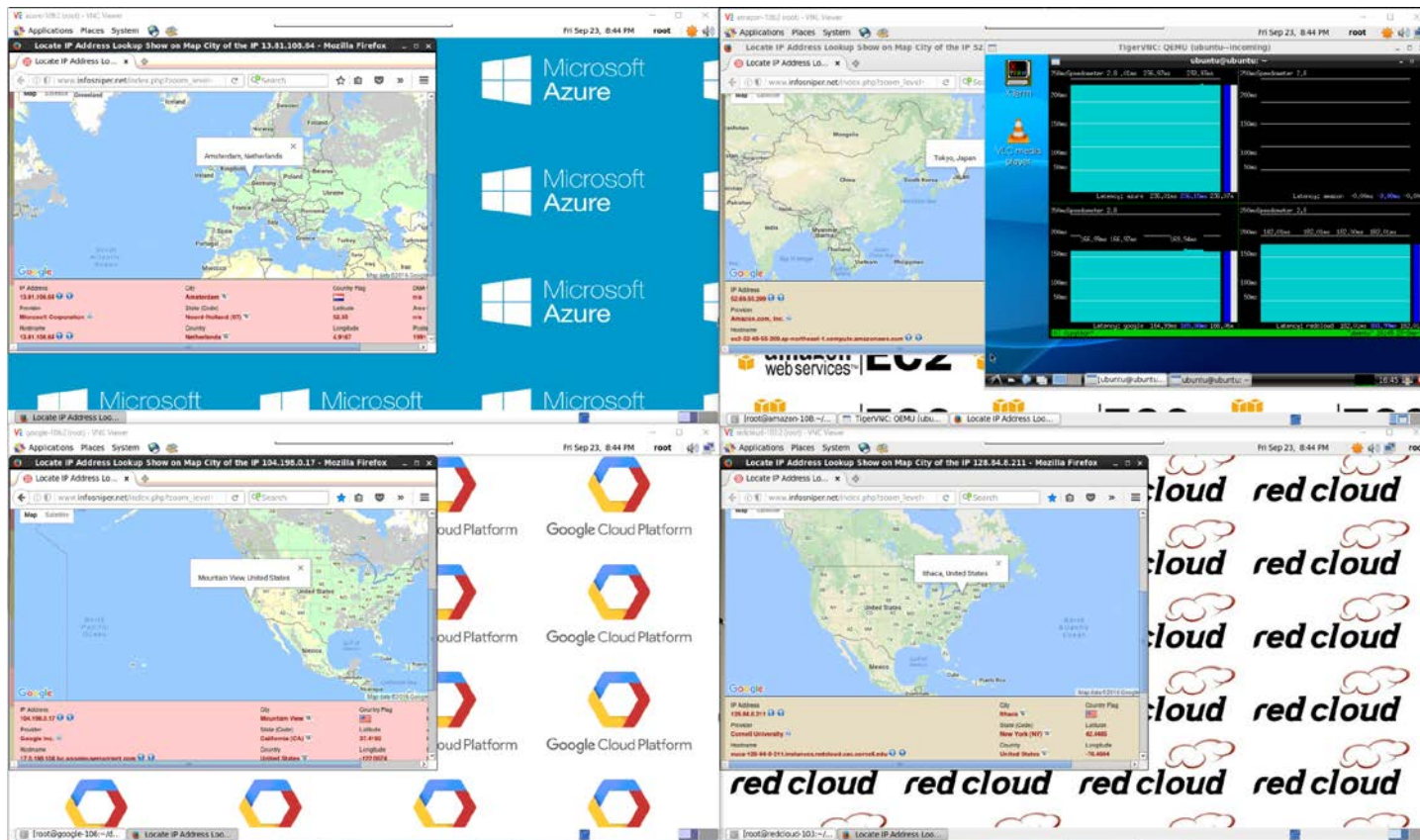
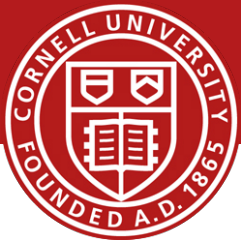
(Full video available at <http://supercloud.cs.cornell.edu>)

A screenshot of a multi-monitor VNC environment. The top-left monitor displays a Windows desktop with a blue background and multiple instances of the Microsoft Azure logo. The top-right monitor displays a Linux desktop with a blue background and the text "amazon EC2" repeated vertically. The bottom-left monitor shows a terminal window with the following text:

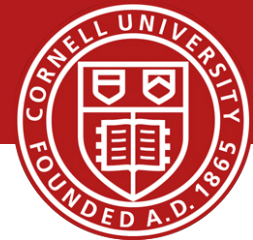
```
[root@centos6 scheduler]#  
[root@centos6 scheduler]#  
[root@centos6 scheduler]#  
[root@centos6 scheduler]#
```

The bottom-right monitor displays a slide titled "Demo" with a world map. The map has callouts for "Microsoft Azure" in North America and "amazon EC2" in Asia. The slide also includes the Cornell University logo and the text "© 2009 www.outline-world-map.com". A "RECORDED WITH SCREENCAST-MATIC" watermark is visible in the bottom-left corner of the terminal window.

Full Demo (<http://supercloud.cs.cornell.edu>)



Highlights



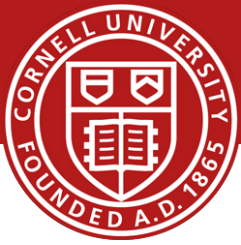
- Automatic VM placement and migration
- Migrated VMs are LIVE
- IP addresses are not changed
- TCP connections are not broken

- Appears as a unified private cloud that spans all clouds
- Controlled by the user!

Research Challenges

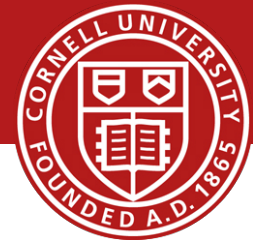


- How to migrate across incompatible virtualization platforms?
- How to keep IP addresses unchanged and TCP connections unbroken?
- How to decide when and where to migrate?
- How to make the system efficient?

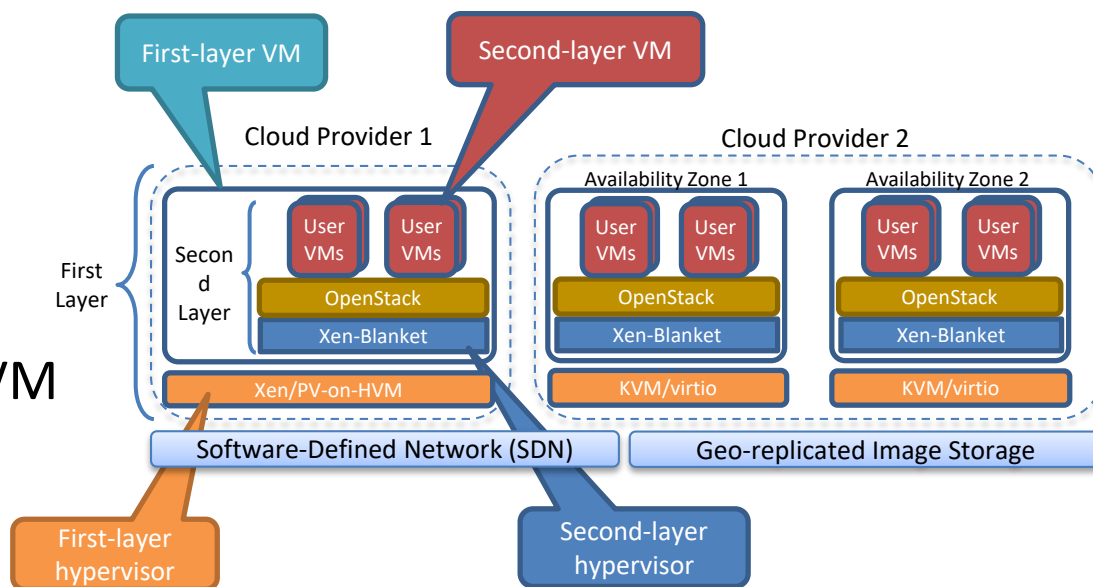


Supercloud is the first system that supports
automatic,
efficient,
and live
VM migration
across heterogeneous cloud providers
without changing IP addresses
or breaking TCP connections.

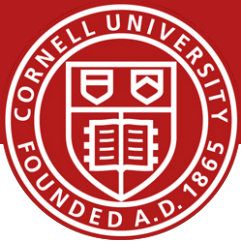
Supercloud Architecture



- Computation
 - Nested hypervisor: Xen-Blanket
 - Support all major platforms
- Network
 - SDN overlay
 - Support migration with public IP
- Storage:
 - Geo-replicated storage
 - Optimized for serving VM images
- Resource management
 - OpenStack platform
 - Automatic scheduling framework

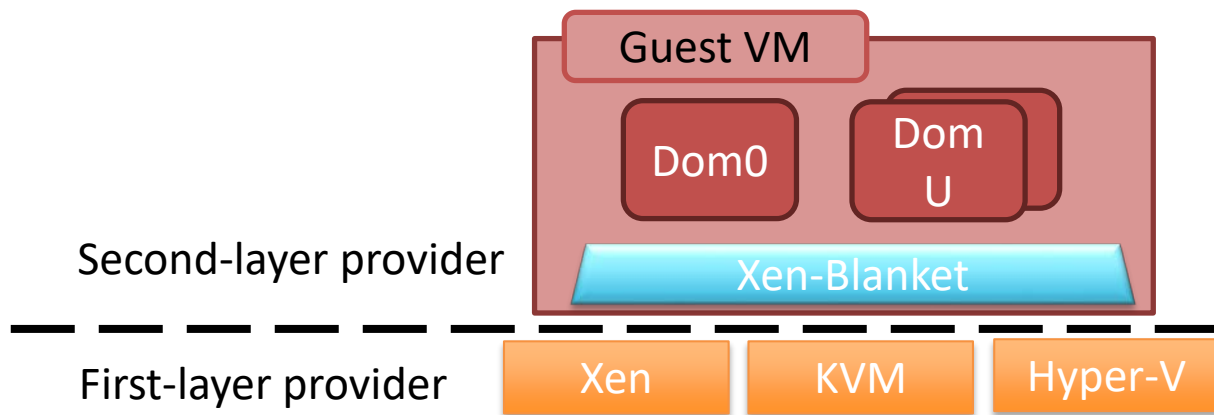


Supercloud Compute Instances



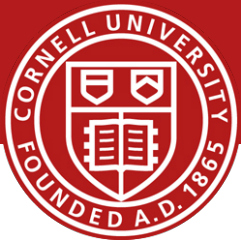
Xen-Blanket

- Second Layer Hypervisor
- Uniformity





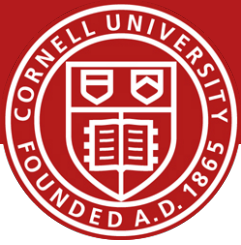
Supercloud Compute Instances



- Move computation via Virtualization
 - Virtualize processor Instruction Set Architecture
 - Full Virtualization vs Paravirtualization (of hardware)
 - VMWare vs (Original) Xen

- Migrate computation among different machines

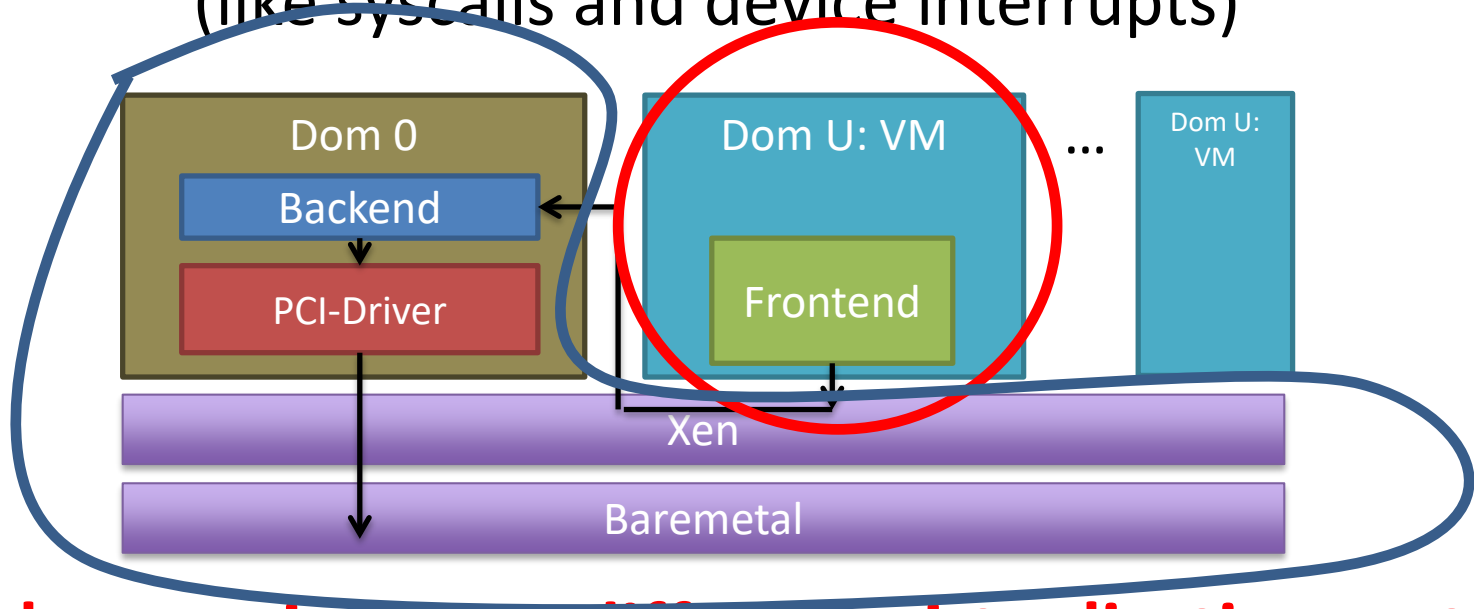
Supercloud Compute Instances



- Xen

- DomU hosts guest operating system in virtual machine
- Dom0 manages devices and guests
- Control Transfer: Hypercalls and Events

(like syscalls and device interrupts)

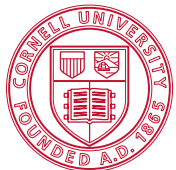
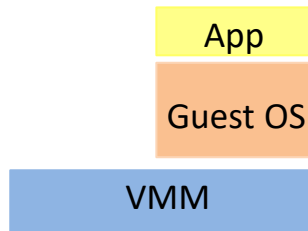
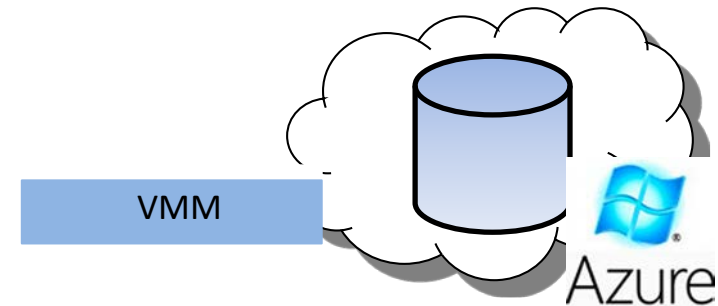
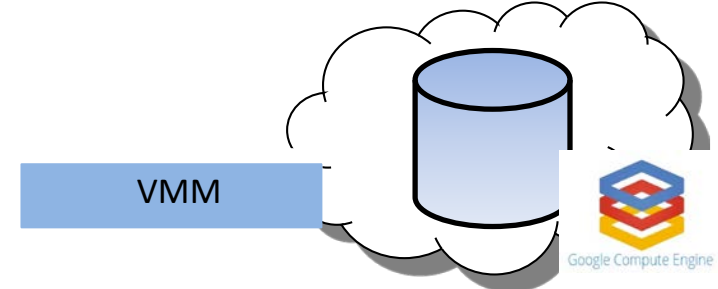
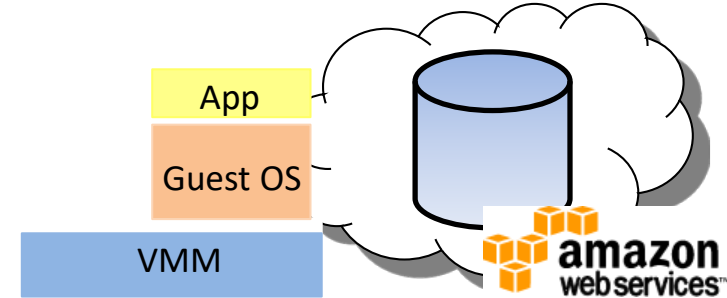
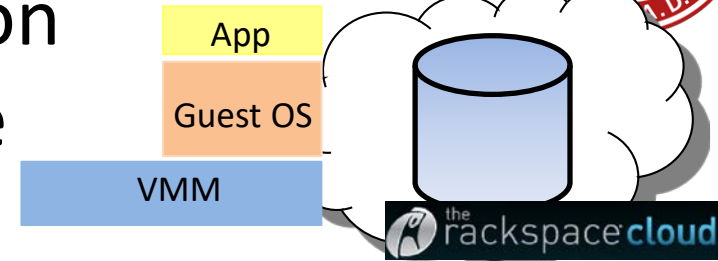


How do we migrate to different virtualization system?

Supercloud Compute Instances

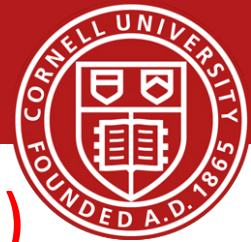


- Virtualization enables migration of computation across a single cloud

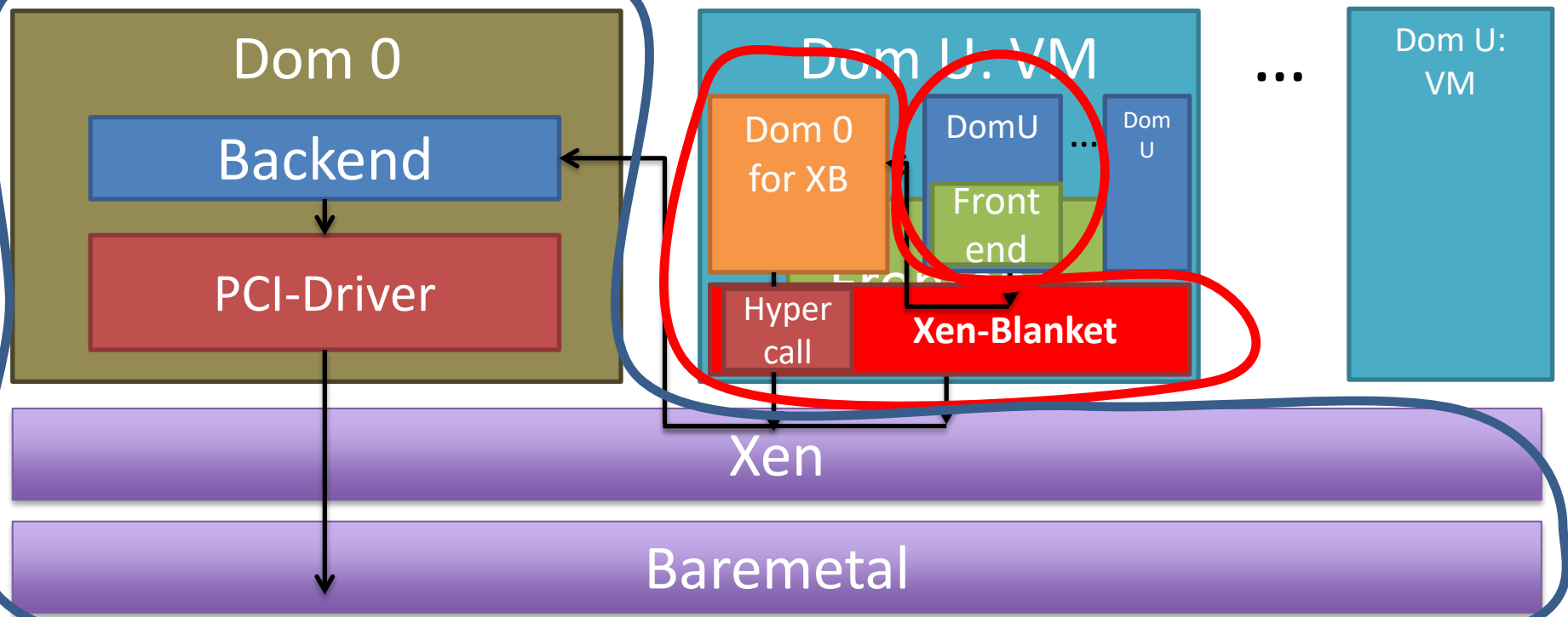


Cornell University

Supercloud Compute Instances



- Need another layer of virtualization (indirection)

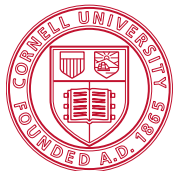
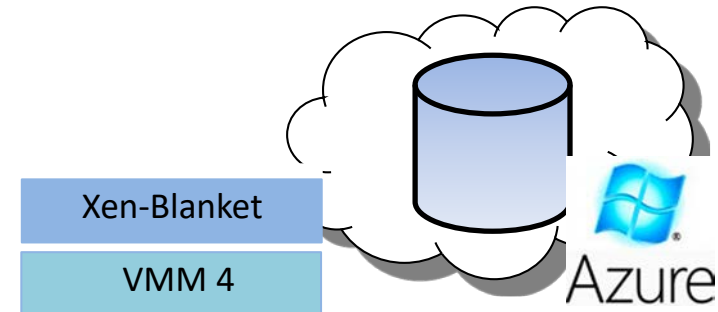
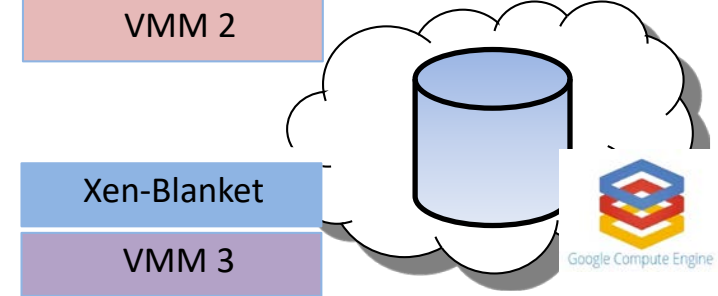
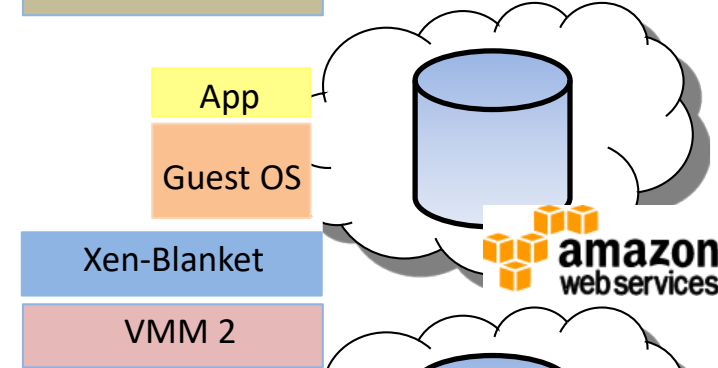
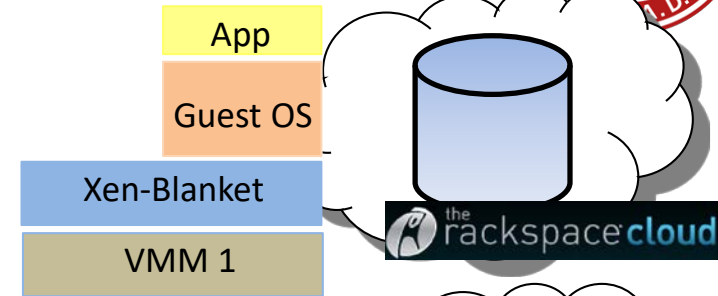


- Blanket drivers enable access to baremetal
- Blanket drivers support memory address translation
- Hypercalls provides privilege support for HVM

Supercloud Compute Instances



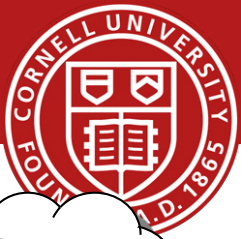
- Can create your own *Cloud-within-a-Cloud* aka a “Supercloud”



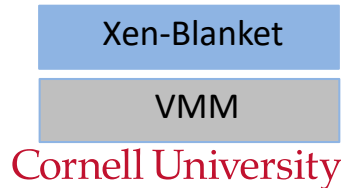
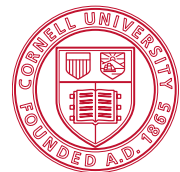
Cornell University

- Migrate computation among different cloud providers

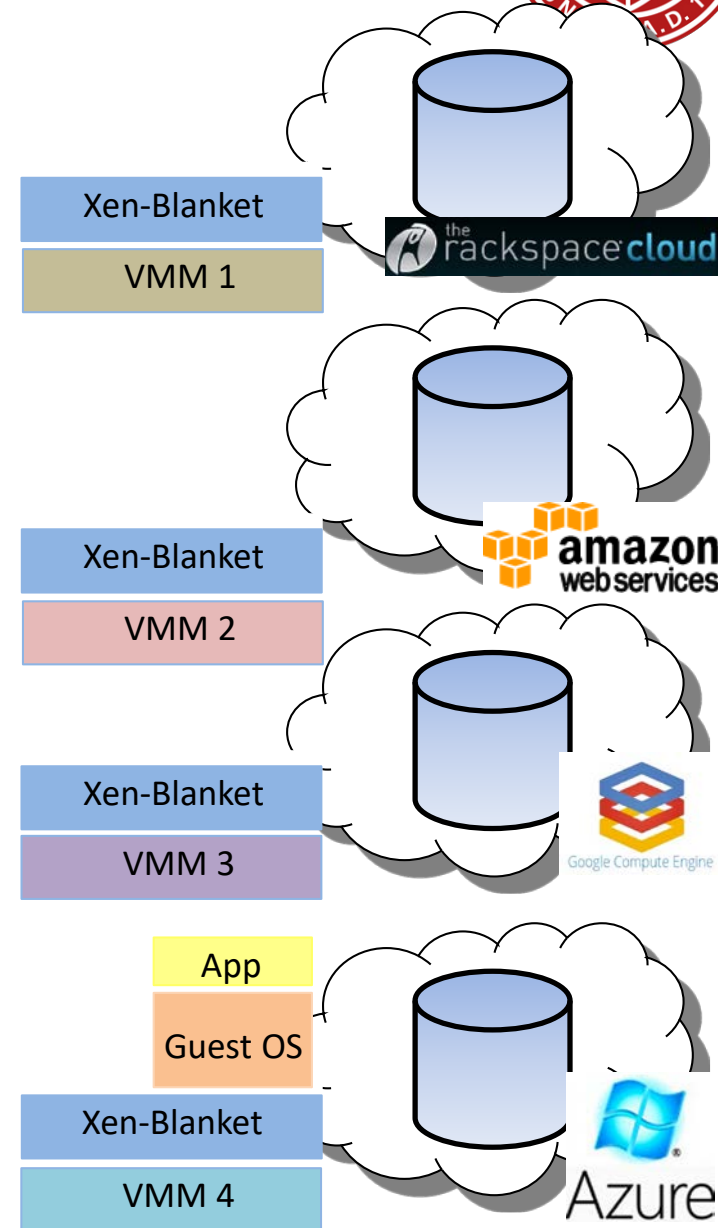
Supercloud Compute Instances



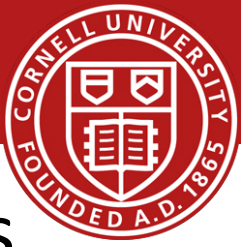
- Migration of computation among different cloud providers
- Small overhead
- Migration: 1 second downtime



- *Is VM migration sufficient?*

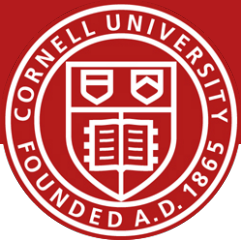


Supercloud Networking

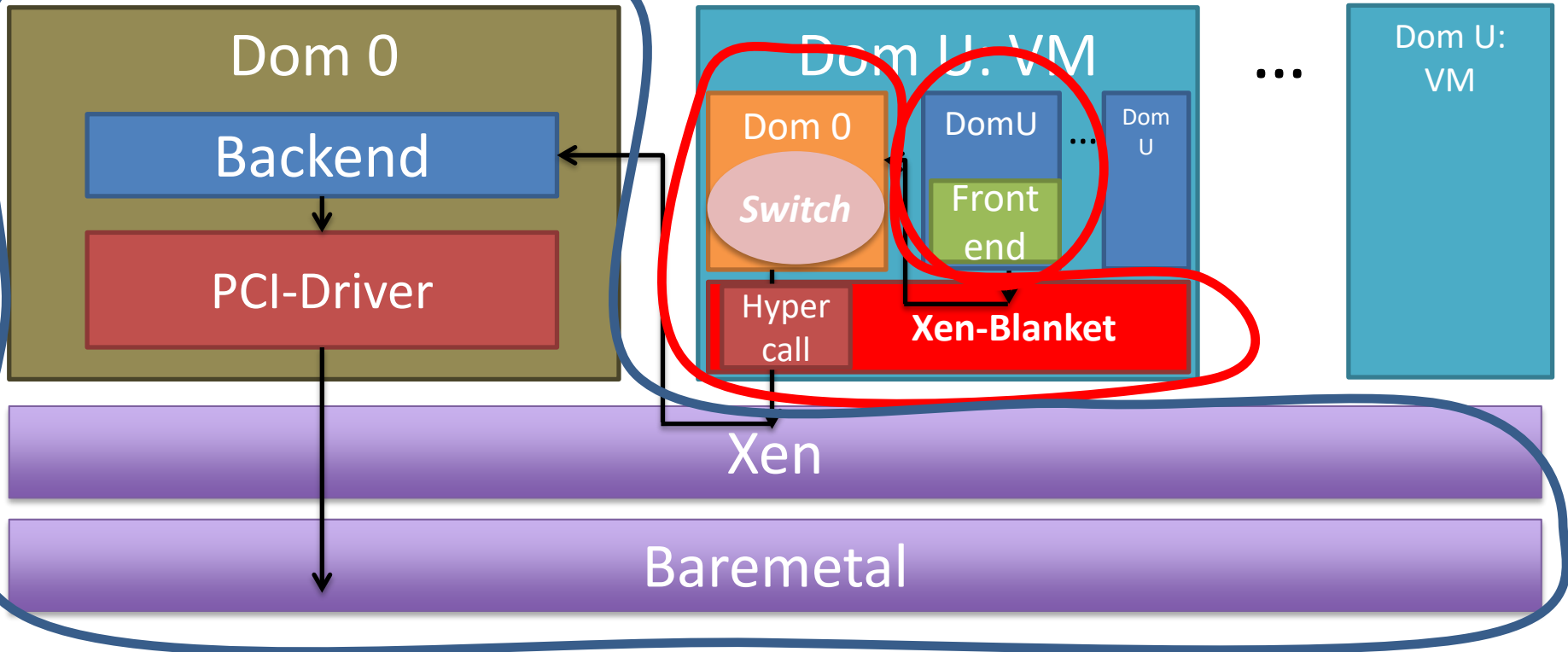


- VM migration is not enough for large systems
 - Single application includes an ecosystem
 - E.g. load balancer, app logic, database, and nw policies
- Network configuration is necessary
 - Policy based routing (e.g. firewall, load balancer, etc.)
 - Broadcast
 - Multicast
- Cloud providers' support is limited
- We need **user centric virtual wire; i.e. SDN!**

Supercloud Networking

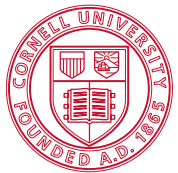
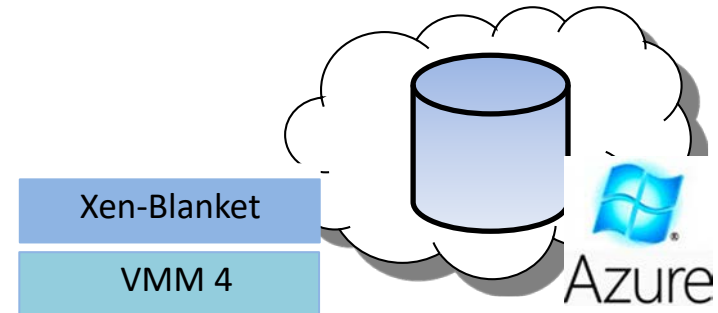
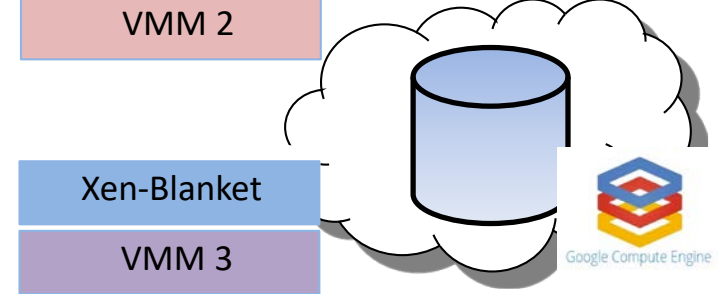
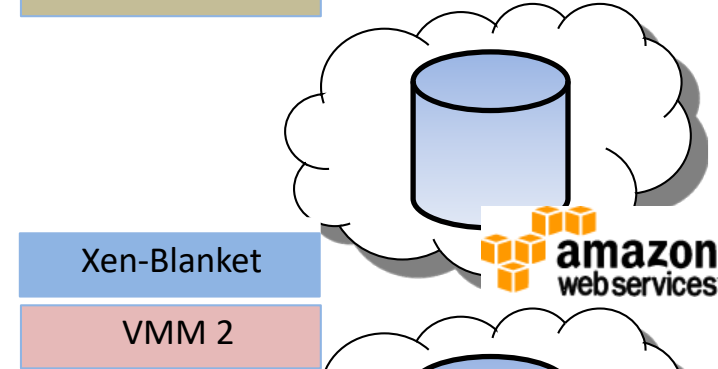
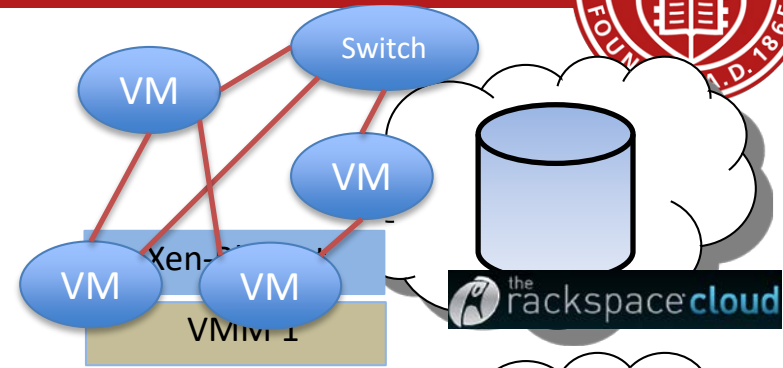


- How to build a “Supercloud” SDN?



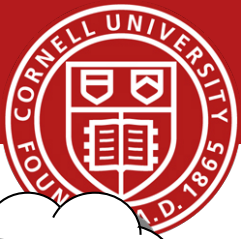
- Blanket drivers enable access to baremetal
- Blanket drivers support memory address translation
- Hypercalls provides privilege support for HVM

Supercloud Networking

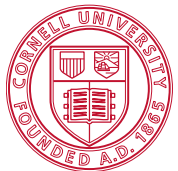
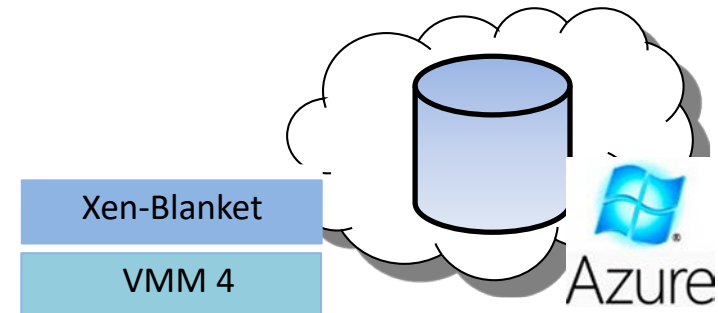
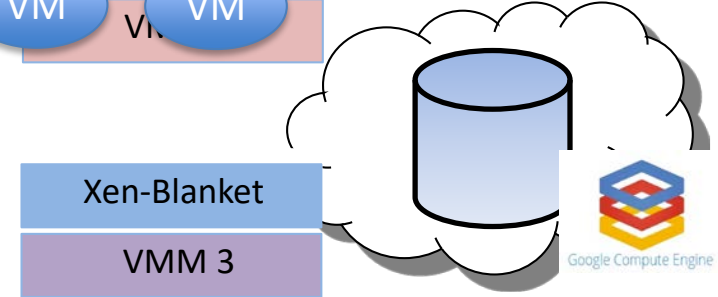
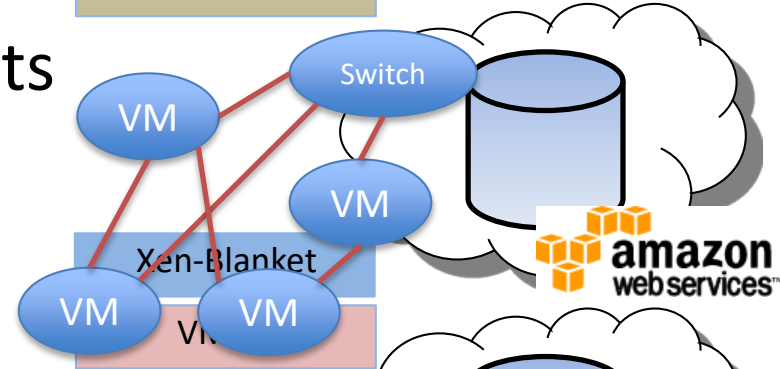
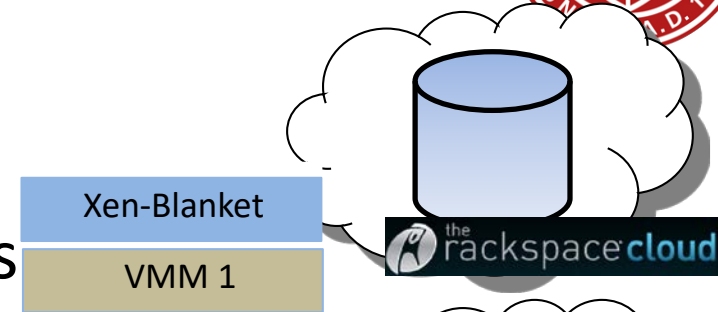


Xen-Blanket
VMM
Cornell University

Supercloud Networking



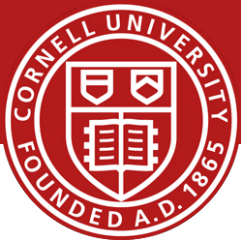
- Virtual Wires via SDN
 - Uses virtual switches
 - Installation of virtual connectors
 - Encapsulates and sends packets



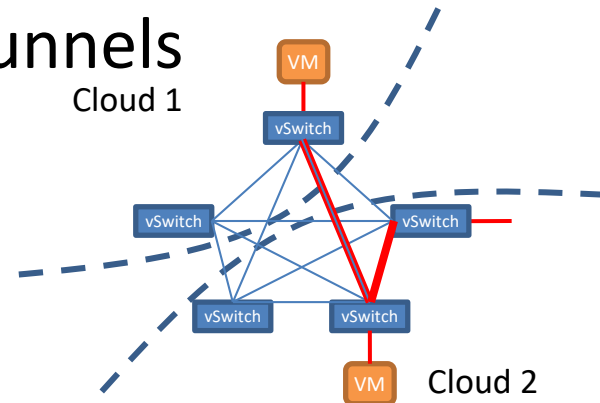
Xen-Blanket
VMM
Cornell University

- Migration of Network Configs
 - Migrates network topology
 - Obviates administrative configs

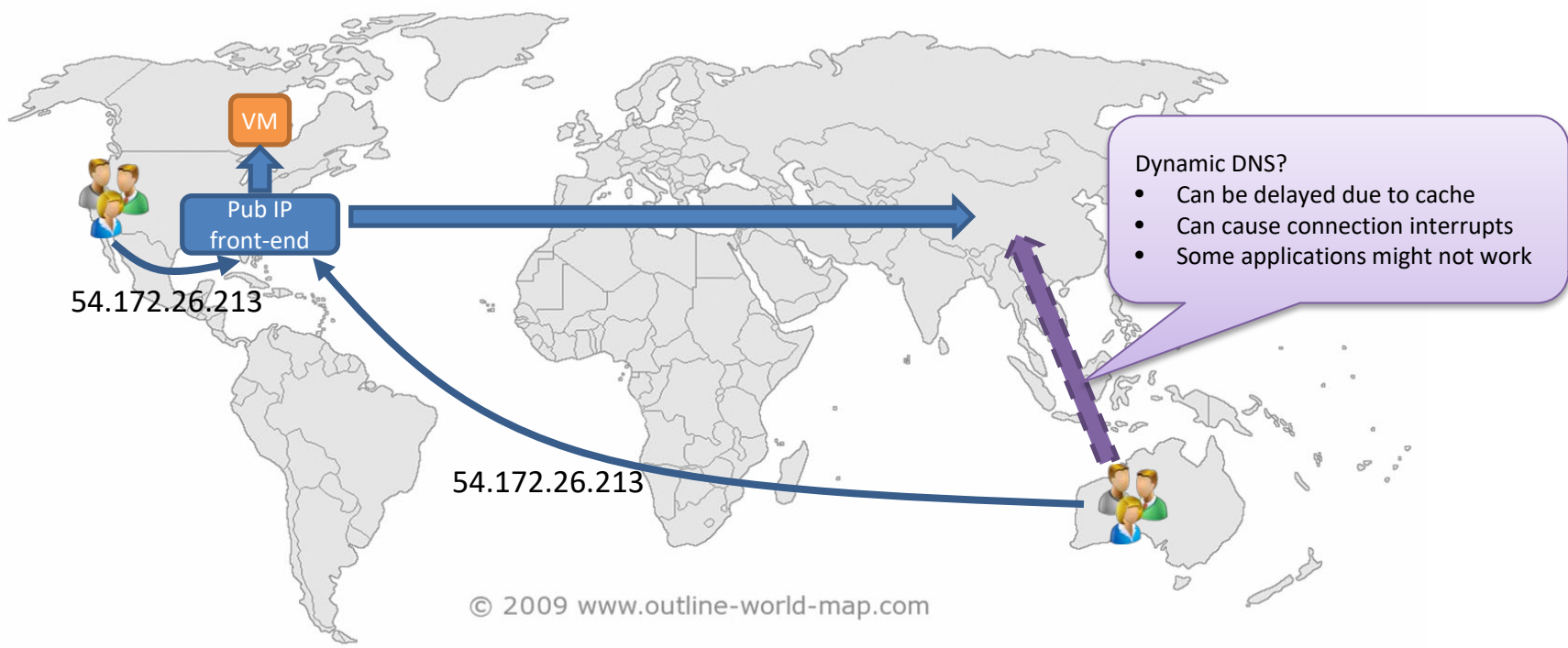
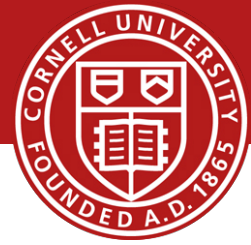
Supercloud Networking



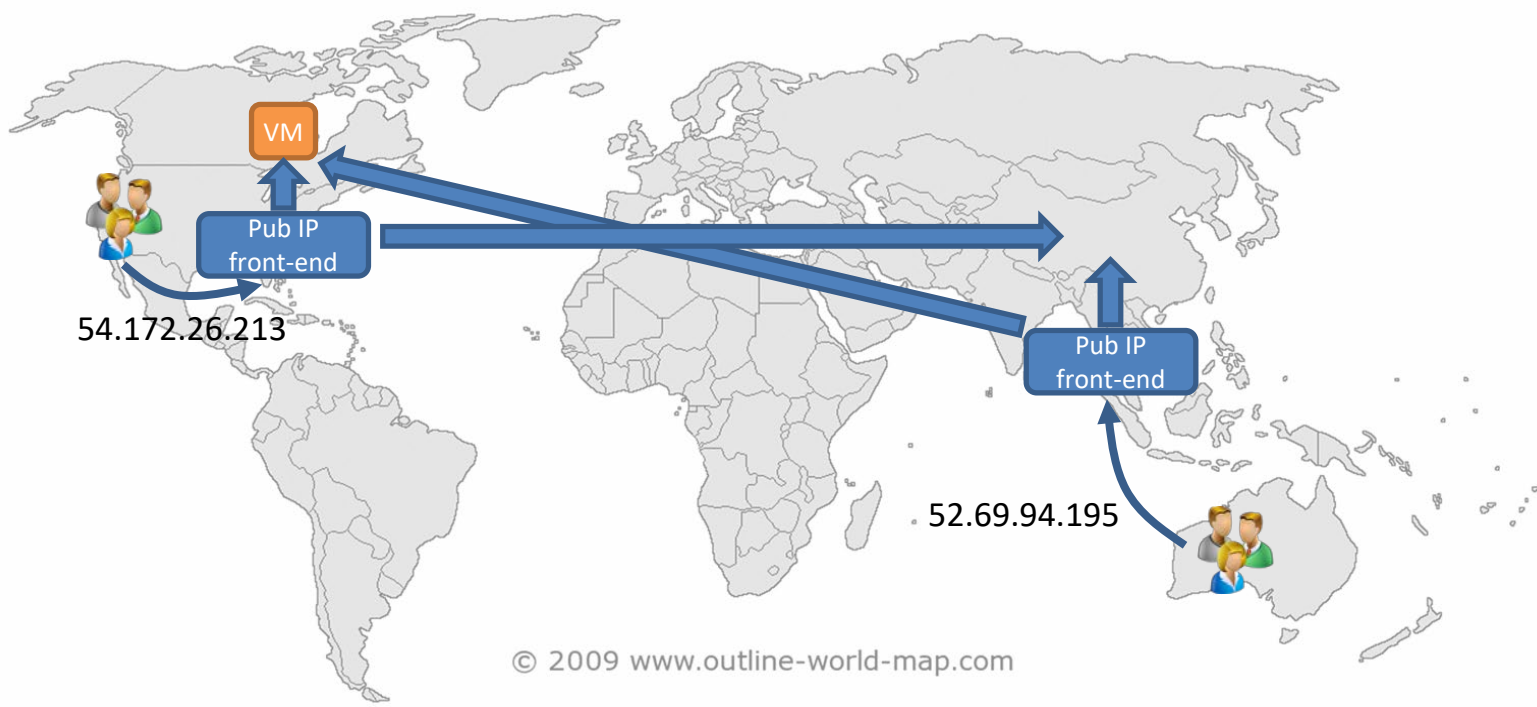
- Challenges:
 - Optimal routing without extra forwarding
 - Migration without changing IP addresses
- Solution:
 - VPN overlay with full-mesh tunnels
 - Frenetic SDN controller



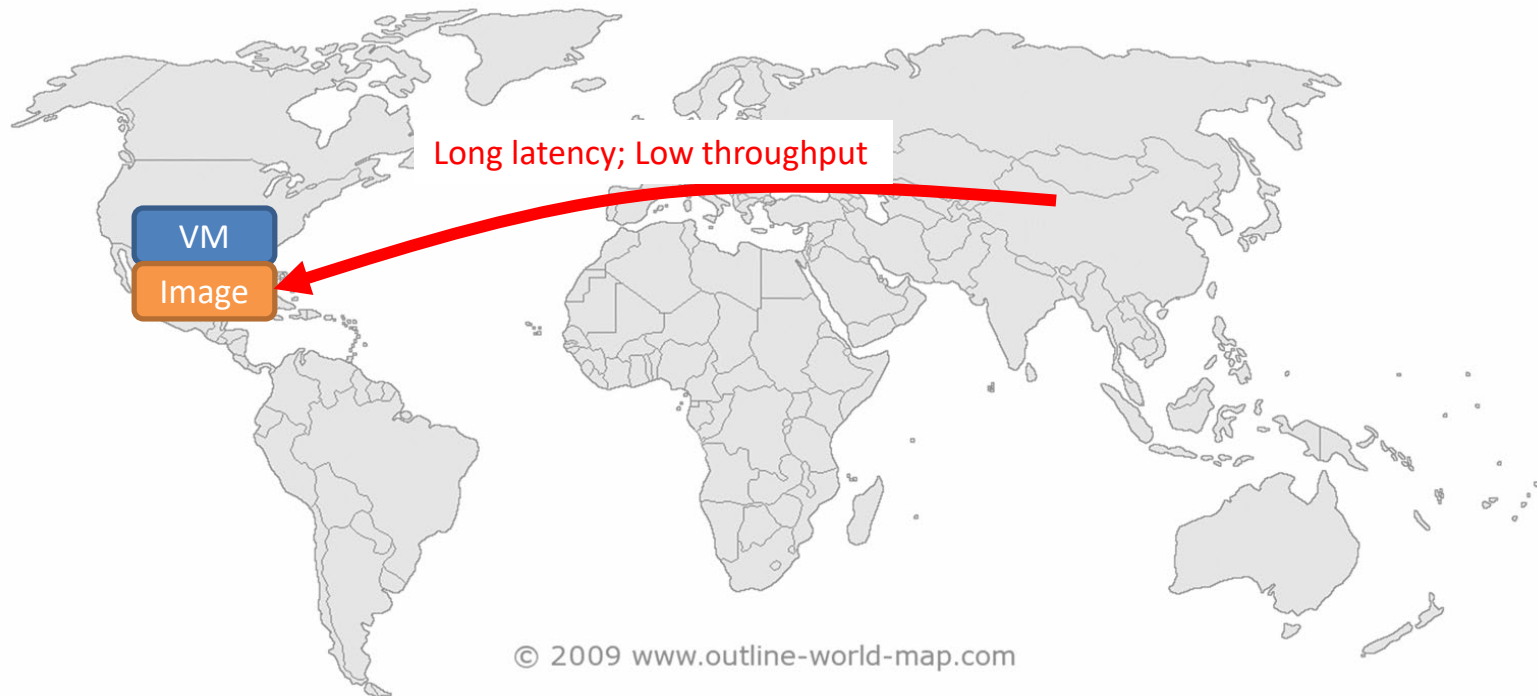
VM Migration with Public IP Address



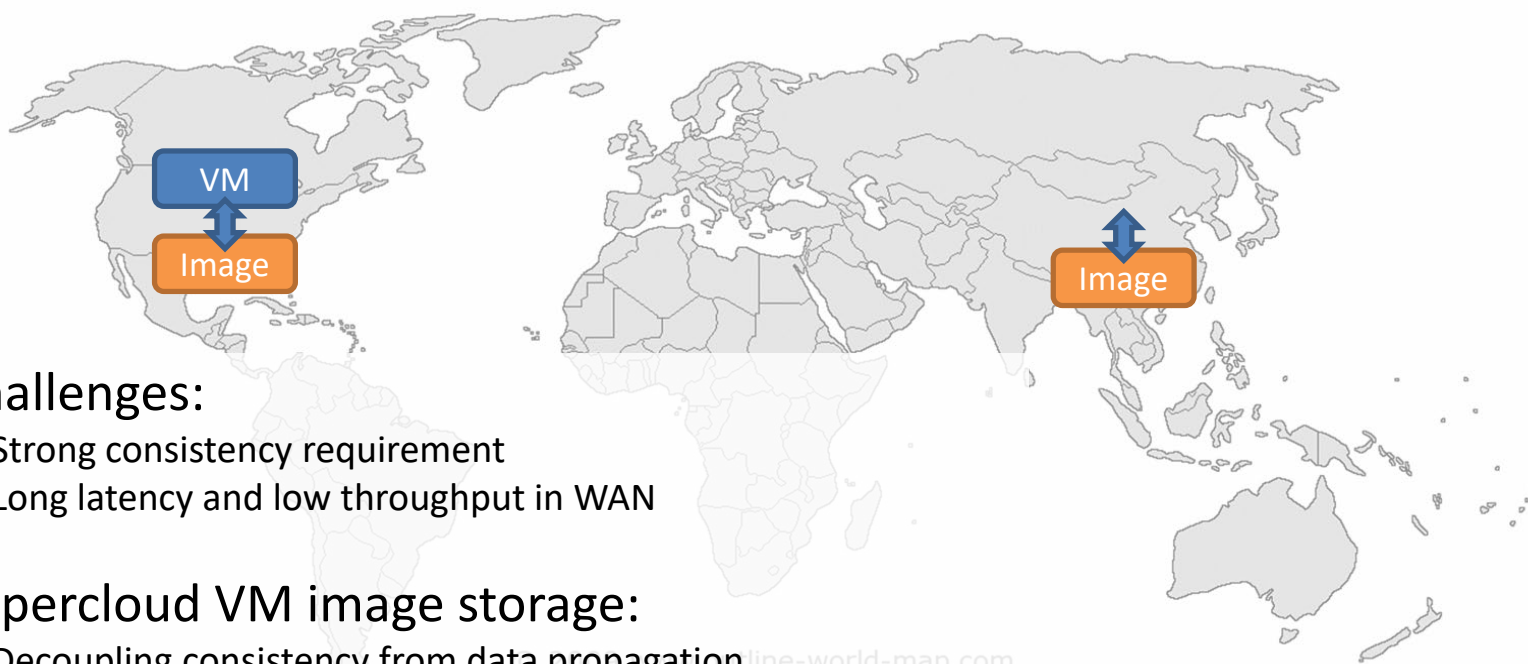
VM Migration with Public IP Address



Centralized VM Image Storage



Supercloud Geo-Replicated VM Image Storage



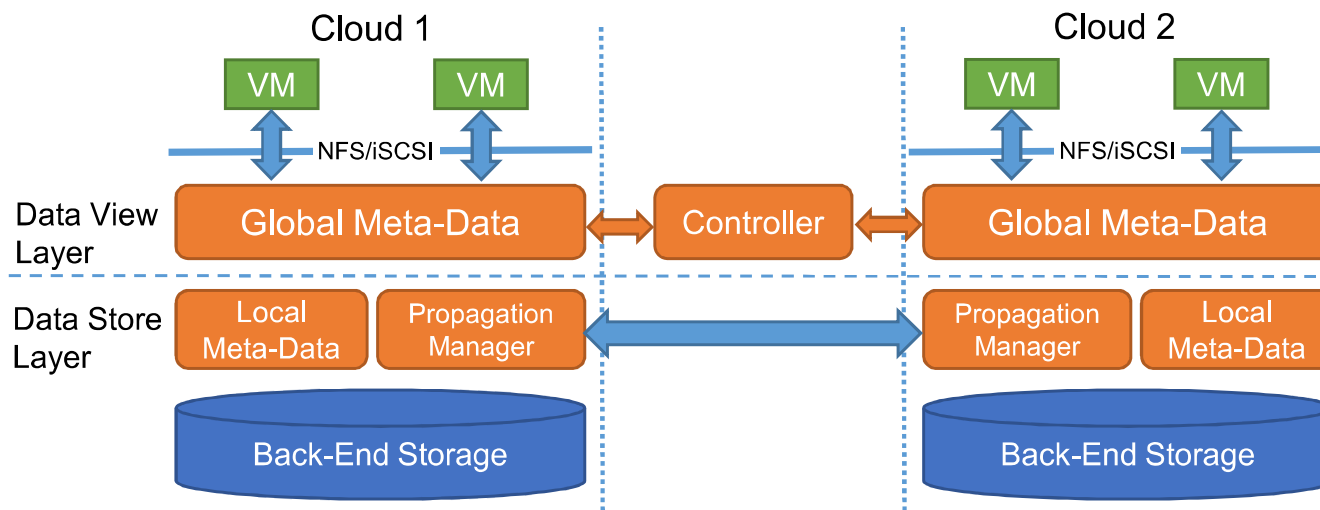
Challenges:

- Strong consistency requirement
- Long latency and low throughput in WAN

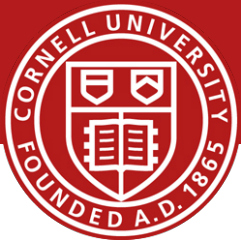
Supercloud VM image storage:

- Decoupling consistency from data propagation. ine-world-map.com
- Propagating data according to disk access patterns.

Supercloud Geo-Replicated VM Image Storage

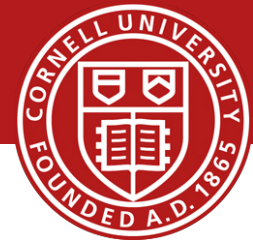


Scheduling Framework



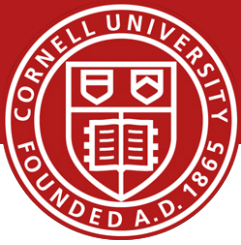
- Application provides:
 - A placement evaluation function f
 - $Score = f(Placement, Workload, Latency)$
 - A threshold T
 - Minimal score change for triggering migration
- The scheduler searches for a placement that:
 - Maximizes the score
 - Outperforms the current placement by at least T

Benefits of the Supercloud



- Case studies
 - Follow the Sun
 - Smart Spot Market
 - Smart Dynamic Resource Scheduler (SDRS)

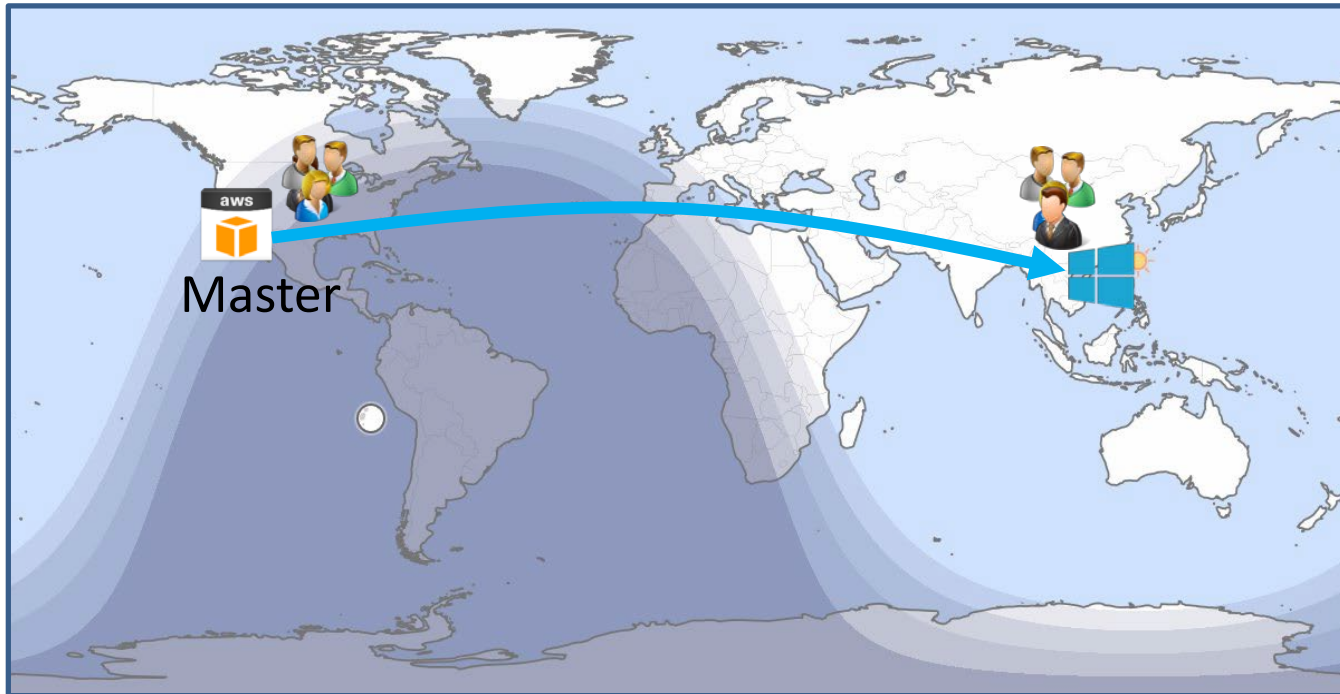
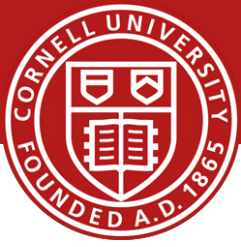
Benefits of the Supercloud



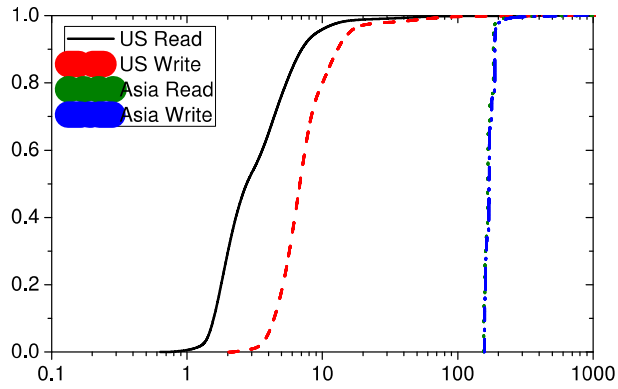
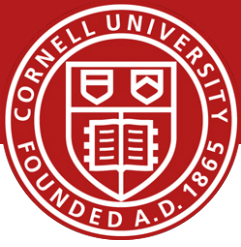
- By leveraging cloud diversity, users can:
 - Reduce cost
 - Reduce energy
 - Handle bursts
 - ✓ – Improve availability
 - ✓ – Reduce service response time
 - ✓ – Improve performance
 - Improve security
 - And more...



Follow the Sun with Zookeeper



Follow the Sun with Zookeeper



ZooKeeper Master in US

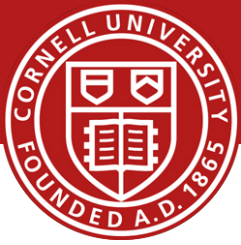
ZooKeeper Master migrates

Benefits of Exploiting Diversity

- By leveraging cloud diversity, users can:
 - Reduce cost
 - Reduce energy
 - ✓ Handle bursts
 - Improve availability
 - Reduce service response time
 - Improve performance
 - Improve security
 - And more...



Spot Instance



- Amazon Spot Instance

- Dramatic Price Change

Prices can jump up

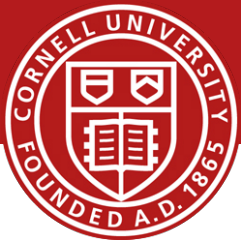
- Usually Cheap

- Charges hourly

Only stateless jobs!

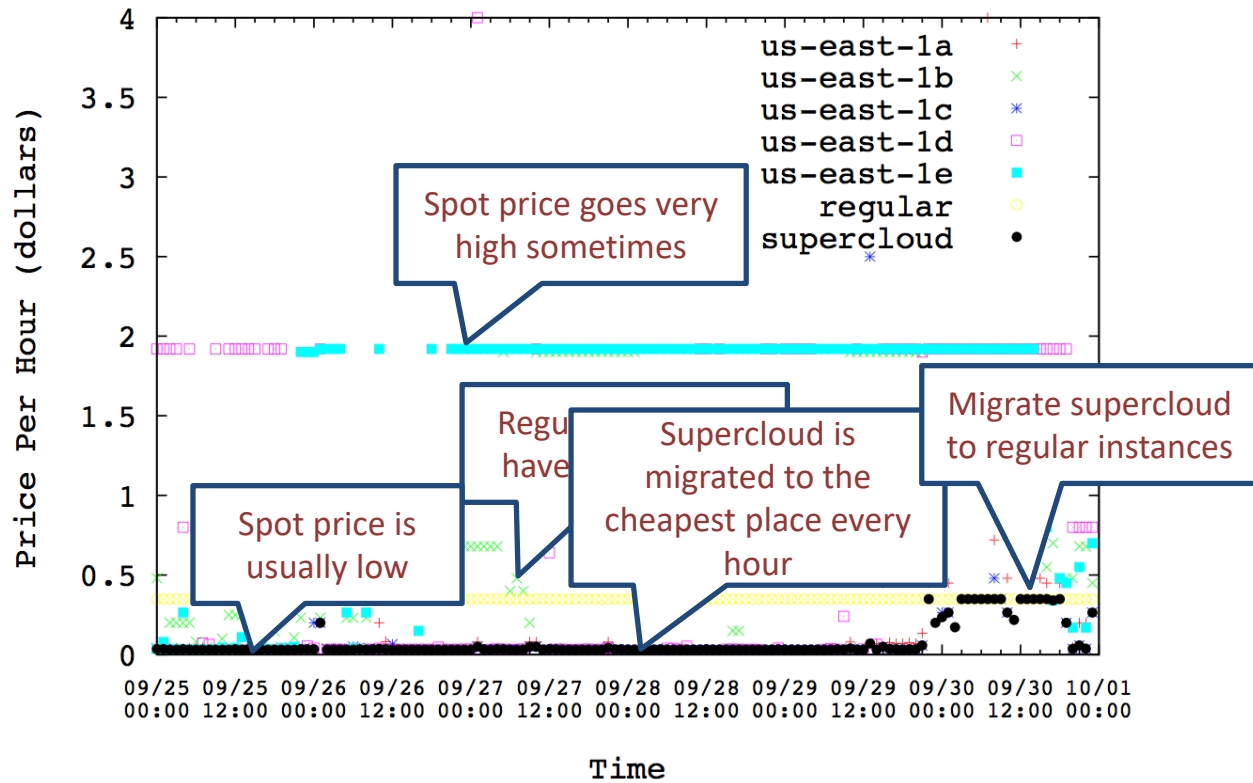
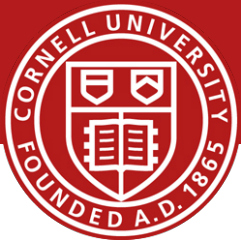
- Ready to be terminated

Smart Spot Instance

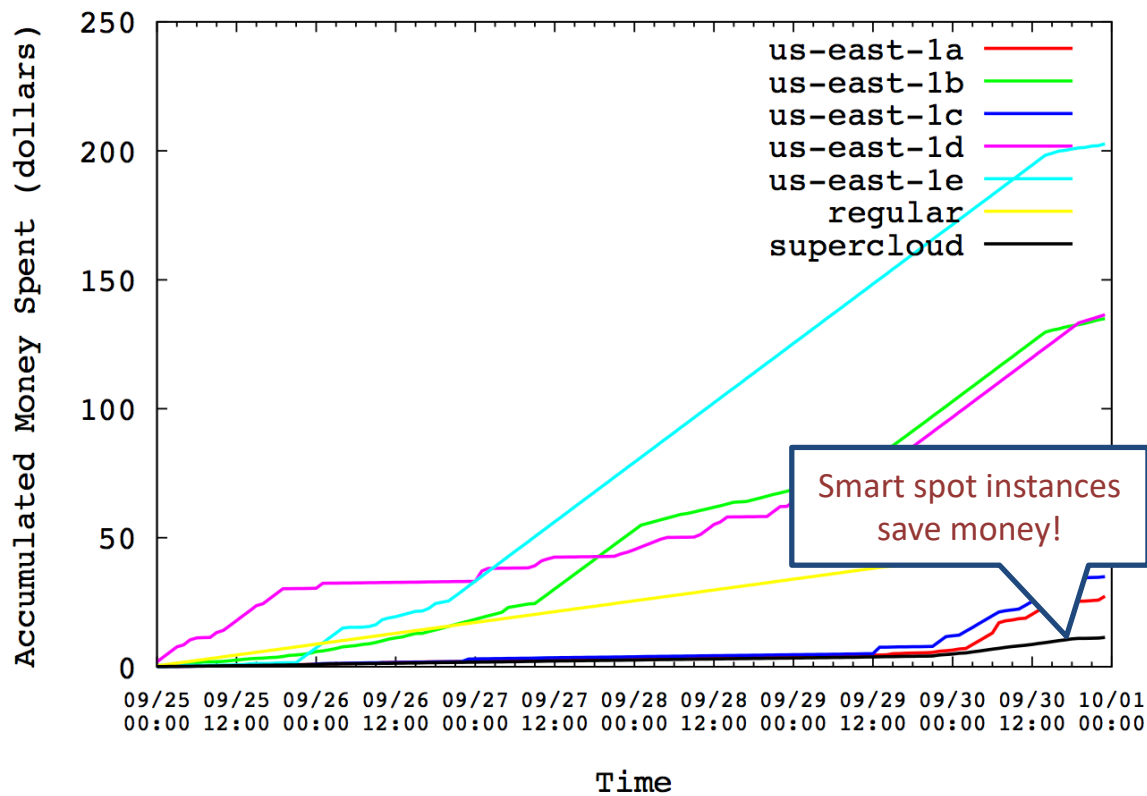
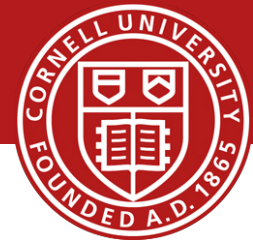


- Supercloud on Spot Instances
 - Migrate instances before being terminated
 - Migrate to the cheapest location before starting new billing hour
- Benefits
 - No termination
 - Reduced cost

Spot Instance Price History



Accumulated Price

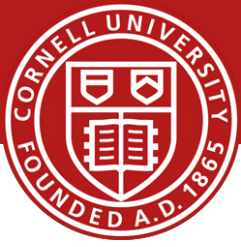


Benefits of Exploiting Diversity

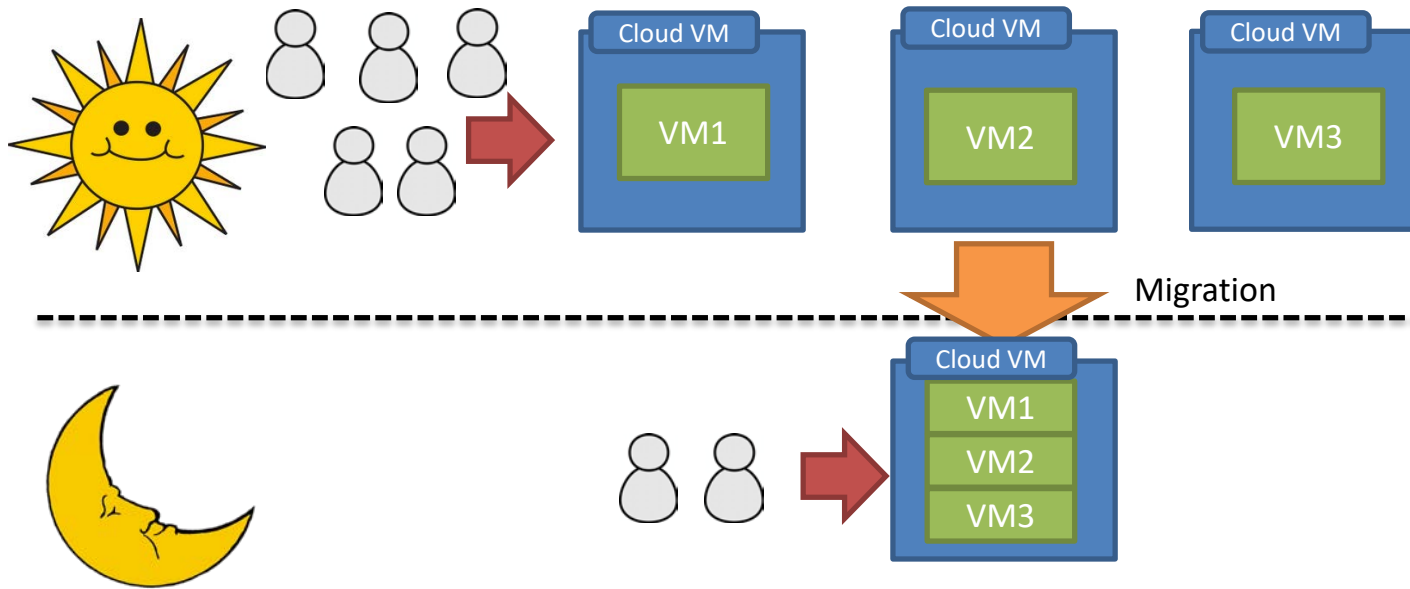
- By leveraging cloud diversity, users can:
 - Reduce cost
 - Reduce energy
 - ✓ Handle bursts
 - ✓ Improve availability
 - Reduce service response time
 - Improve performance
 - Improve security
 - And more...



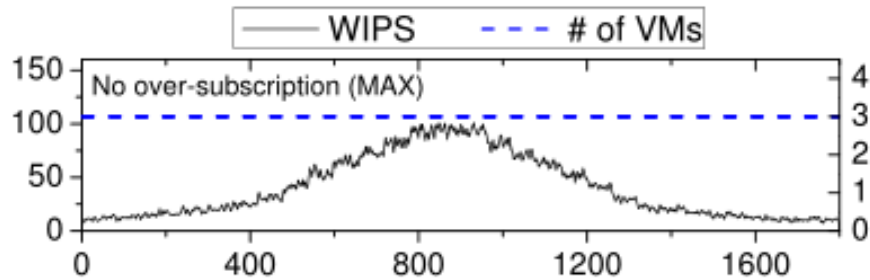
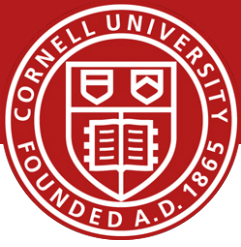
Resource Oversubscription



- Pack VMs when loads becomes light

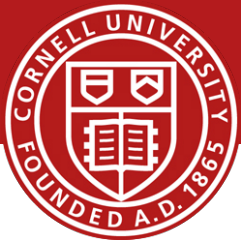


Resource Oversubscription



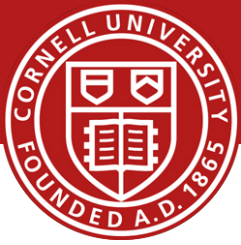
29% cost saving with 1.5% performance degradation.

In the Paper



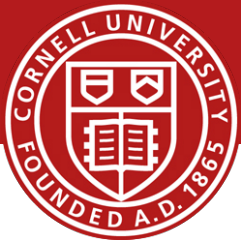
- Comparison with application-level migration
- Placement policies for different types of applications
- Detail design of the image storage
- Hierarchical network topology
- Evaluations

Perspective



- Supercloud: application migration for geographically shifting workloads
 - Crossing heterogeneous cloud providers
 - Automatic placement and migration
 - Geo-replicated image storage
 - Wide-area SDN
- A unified private cloud that spans all clouds
- Controlled by the user!

Paper Trail



- Supercloud Multicloud Deployment: SoCC 2016
- Supercloud Economics: CrossCloud 2016
- Supercloud: Opportunities/Challenges in OSR-2015
- Plug into the Supercloud in IEEE Internet Computing-2013
- Supercloud/Xen-Blanket in EuroSys-2012
- Supercloud in HotCloud-2011/2014
- Overdriver in VEE-2011
- RACS in SOCC-2010
- See also, Storage: Gecko in FAST 2013 / HotStorage 2012,
Antiquity in EuroSys 2007 and SMFS in FAST 2009
- Networking: SoNIC in SIGCOMM 2016, NSDI 2013/2014, Maelstrom NSDI 2008
Wireless DC in ANCS 2012, and NetSlice in ANCS 2012
- More at <http://supercloud.cs.cornell.edu>
and also <http://fireless.cs.cornell.edu>
- Email: hweather@cs.cornell.edu