

# **Sky Computing: When Multiple Clouds Become One**

**José Fortes**

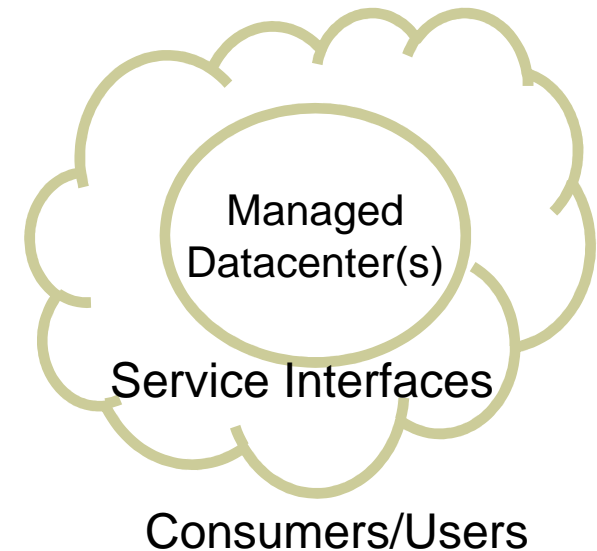
**Advanced Computing and Information Systems Lab**

and

**NSF Center for Autonomic Computing**

# Clouds

- Provider view
  - Economies of scale
  - Statistical multiplexing
  - Avoid customer-specific complexities
- Consumer view
  - No need to (over)provision
  - No operating costs
  - Pay per use
- Win-win decoupling
  - Virtualization in the large



# Cloud computing trivia for 2010

- 20+ conferences with “cloud” in the title
- 5/10 top search hits for “Cloud”
- 200+ cloud providers in the US
- Much more related research/work than what I can acknowledge in this talk



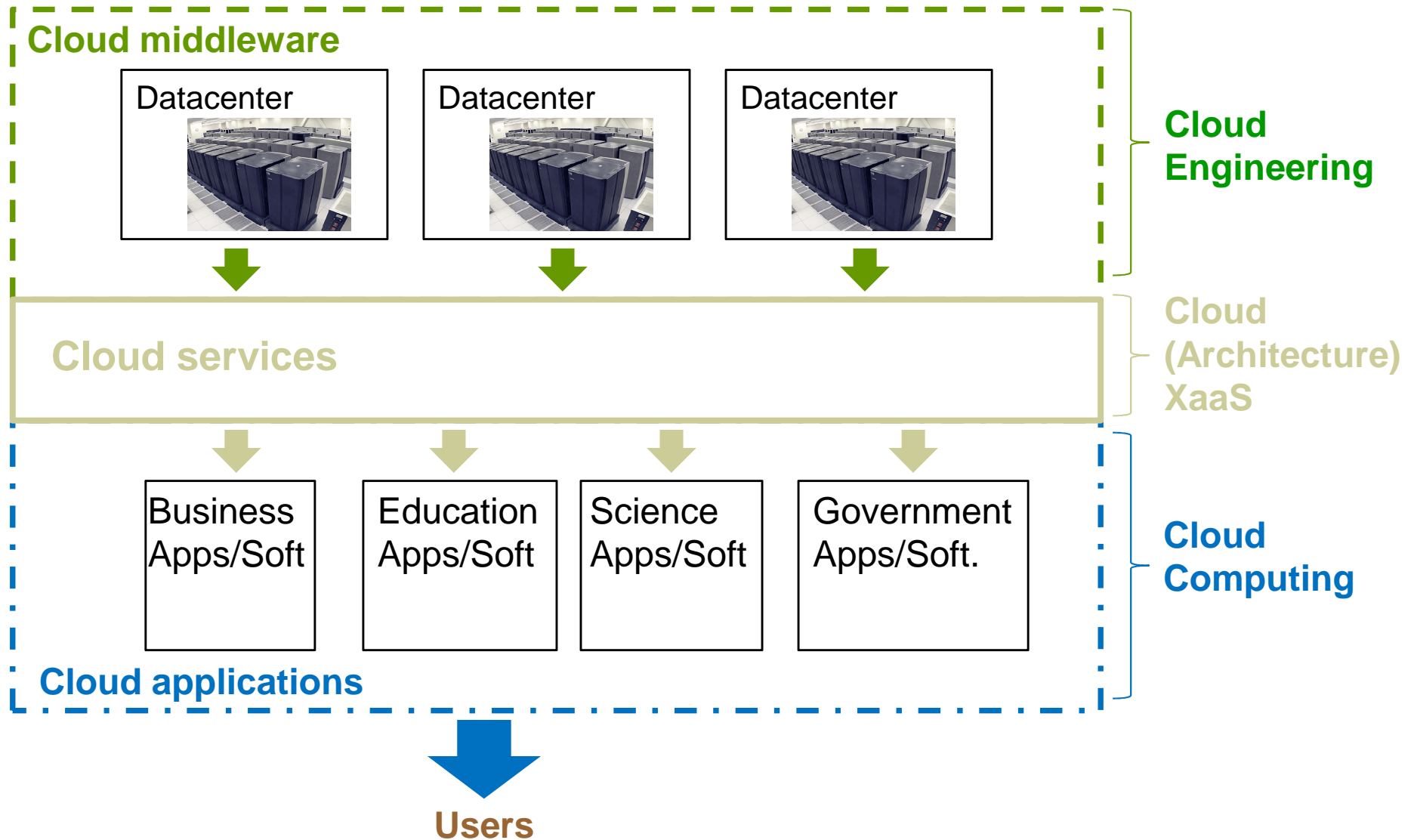
<http://oitblog.wordpress.com/2009/10/14/looking-ahead-in-the-clouds/>

# Outline

---

- Cloud world and ecosystem
- Sky computing
- Networking across clouds
- Other issues
- Conclusions

# The world of a cloud



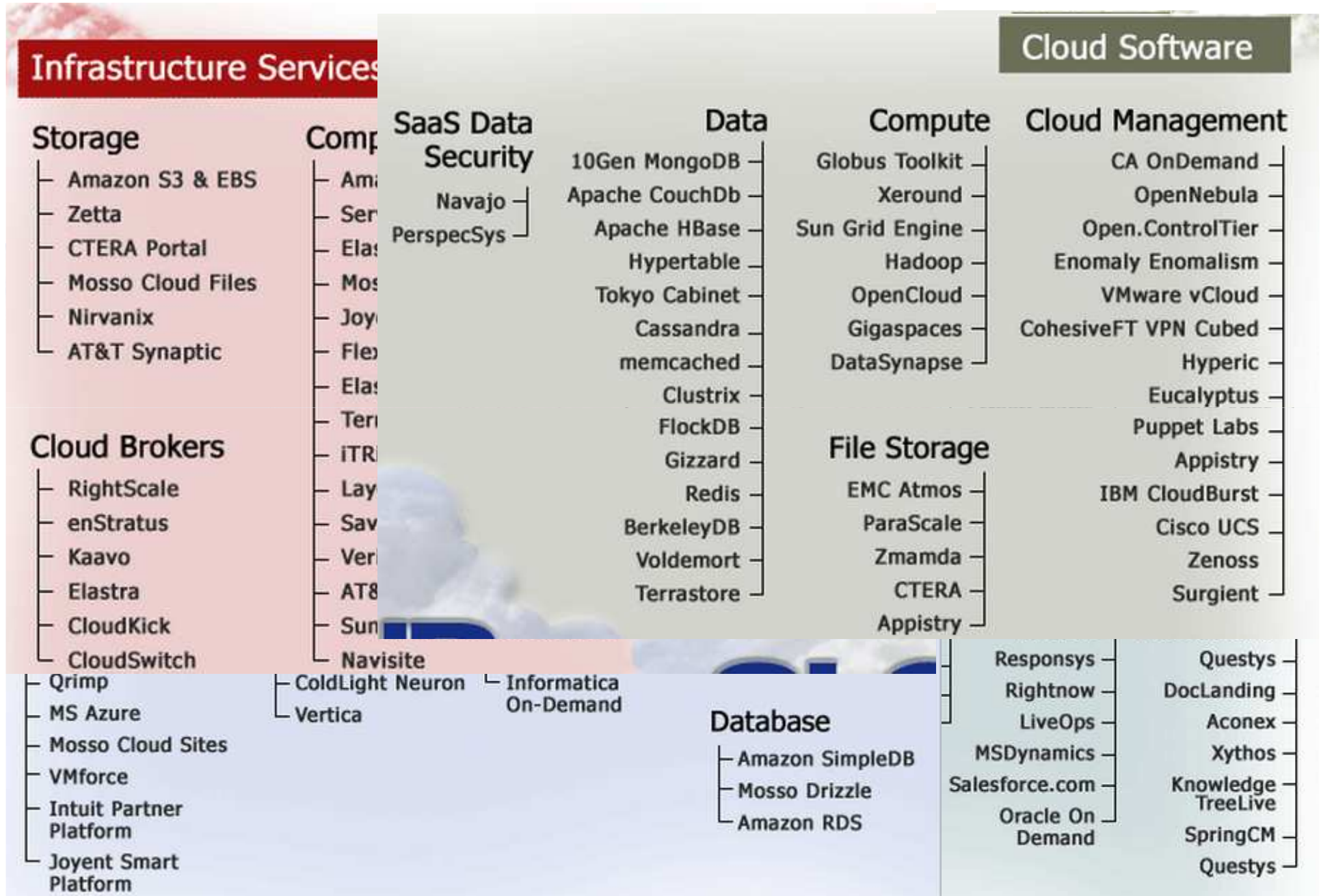
# Datacenters

- Building blocks in shipping containers
  - Servers come setup and wired
  - Large pluggable components ready to go once connected to power, network, cooling
  - Densely packed
- ~5-to-7 fold savings in operational costs (server admin, Mbps, GB)



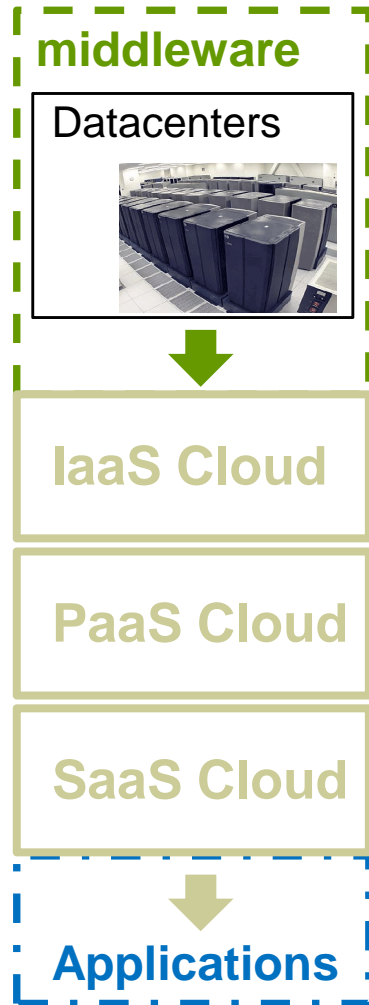


# Cloud middleware and services

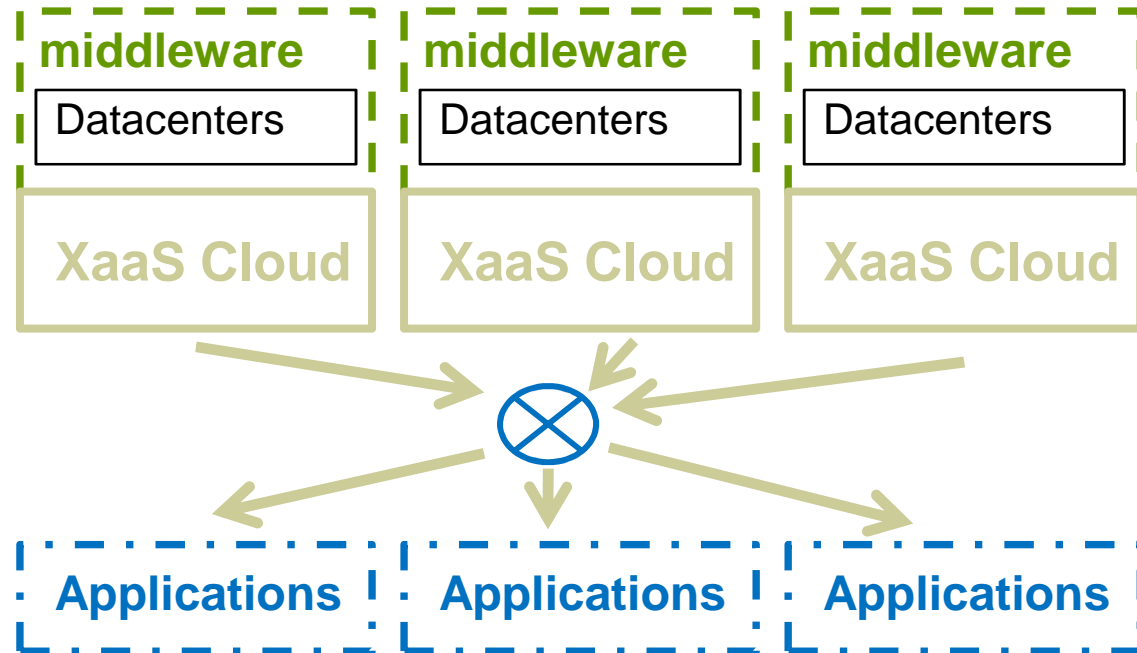


# Early multi-cloud consumer systems

## Nested



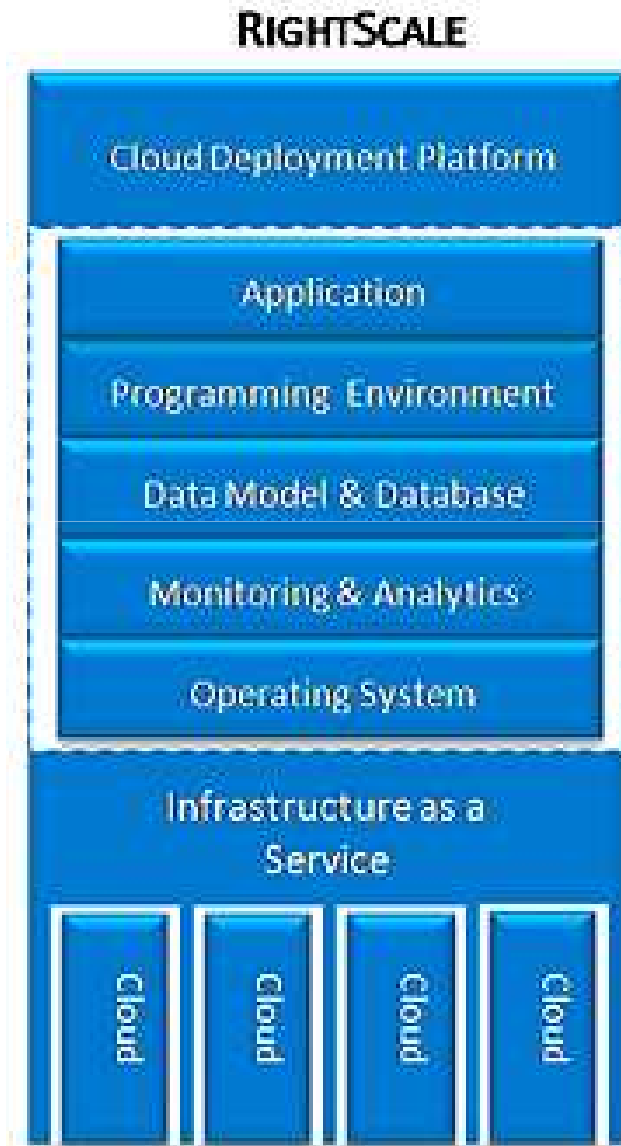
## Replicated or diversified



- Geographic factors
  - Markets, location-dependent services
- Dependability/continuity
  - 24/7, disaster recovery, diversity ...
- Provider independence
- On demand scale-out
- Differentiated services
- Different(iated) apps
- Hybrids



# Multi-cloud management tools



Dashboard - RightScale

https://my.rightscale.com/dashboard/overview

andrea@rightscale.com DEMO | Support | Feedback | Logout

**RIGHT SCALE**

- Dashboard
- Manage
  - Clouds
    - AWS Global
      - Queues
      - S3 Browser
      - Image locator
      - CloudFront
    - AWS US
      - Instances
      - Images
      - Security Groups
      - SSH Keys
      - Elastic IPs
      - EBS Volumes
      - EBS Snapshots
    - AWS EU
      - Instances
      - Images
      - Security Groups
      - SSH Keys
      - Elastic IPs
      - EBS Volumes
      - EBS Snapshots
    - Flexiscale
      - Instances
      - Images
      - Disks
      - Ip Blocks
      - Network Interfaces
      - Vlans
      - Packages
    - GoGrid
      - Instances
      - Images
      - IP Addresses
    - Slicehost
      - Instances
      - Images
      - Flavors
      - Zones
      - Records
  - Design
  - Report
  - Settings

**Dashboard**

Overview | Deployments | Monitoring

**Cloud Service Credential Validity**

Service	Validity
Amazon Web Services	Online
EC2	Online
S3	Online
SQS	Online
CloudFront	Online
Flexiscale	Online
GoGrid	Offline
Slicehost	Offline

**RightScale News**

- New Dashboard Release (April 21st, 2009) View [release notes](#).
- Dashboard UI redesign Please provide feedback on the UI redesign in this [forum thread](#)
- New to RightScale? Check out our [video tutorials](#). Set up a Production Deployment in RightScale.

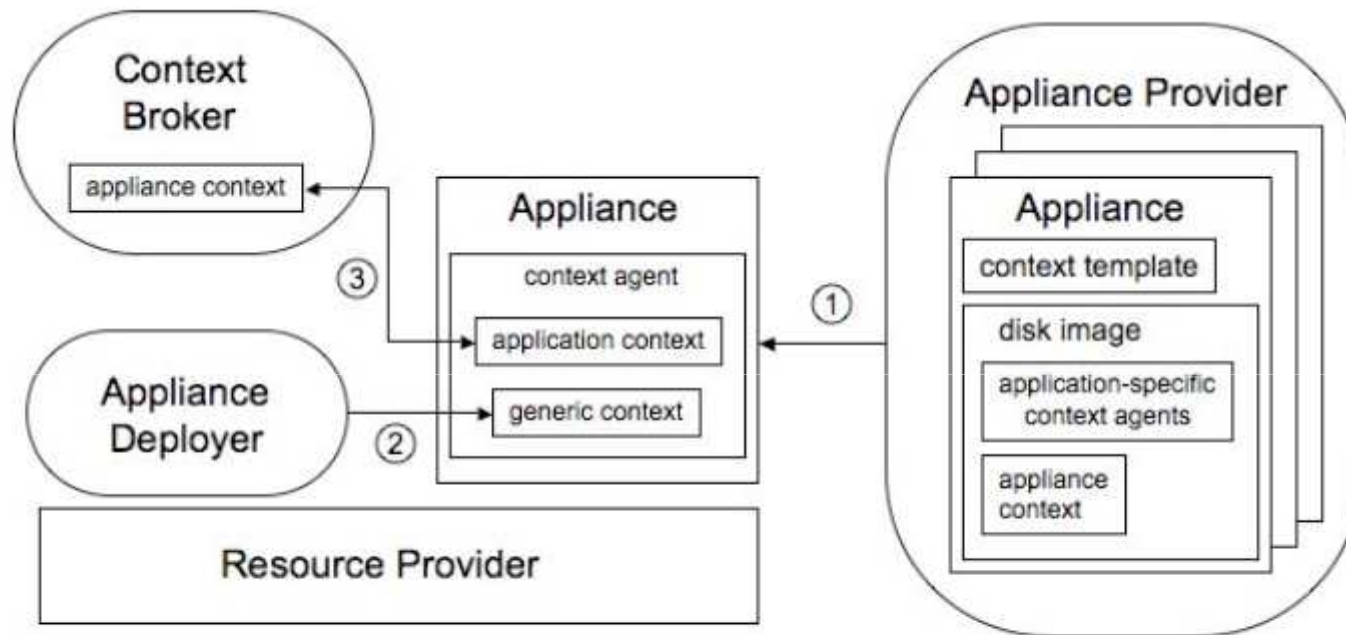
**Deployments Budget Estimate**

Deployment	Current Runrate (hour/day)*	14-day average*	Servers
<a href="#">AWS - Scalable Photo Site</a>	-	\$0.07	4
<a href="#">Bitnami Servers</a>	-	-	4
<a href="#">Default</a>	-	\$0.01	10
<a href="#">Demo Scheduler</a>	-	\$0.07	1
<a href="#">GRID Image Processing Demo</a>	-	\$0.24	1
<a href="#">PHP Scalable Website</a>	-	-	6
<a href="#">PHP-MySQL Cluster</a>	-	-	4
<a href="#">PHP-MySQL Dev Environment</a>	-	-	4

RECENT ACTIVITY 18:35

# Contextualization

- Nimbus

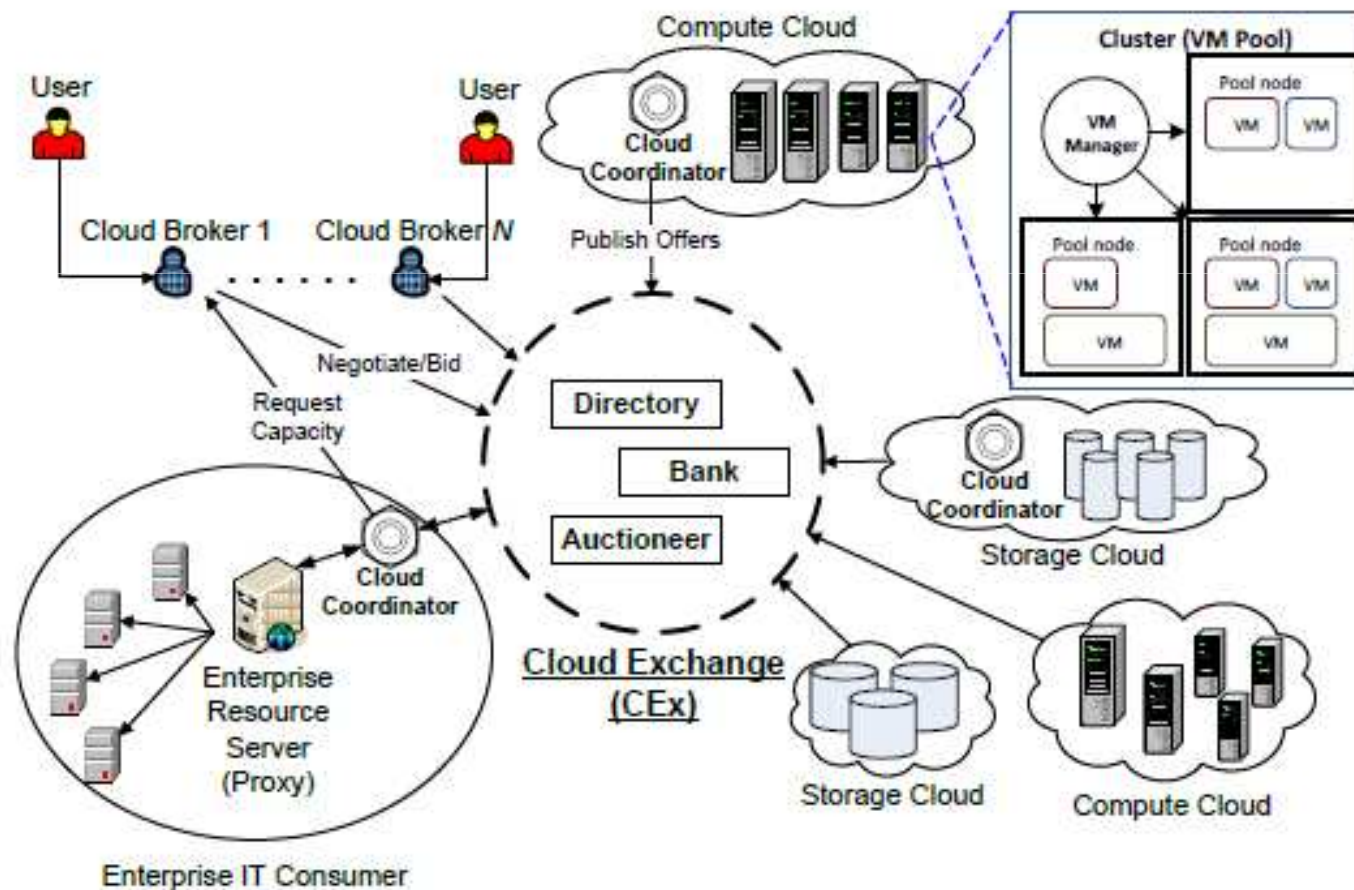


**Figure 2: Relationship between appliance provider, appliance deployer, and context broker.**

**Contextualization: Providing One-Click Virtual Clusters, Keahey, K., T. Freeman. eScience 2008, Indianapolis, IN. December 2008.**

# Cloud Federation with InterCloud

- Application scaling across multiple cloud providers
- Brokers, exchanges and coordinators

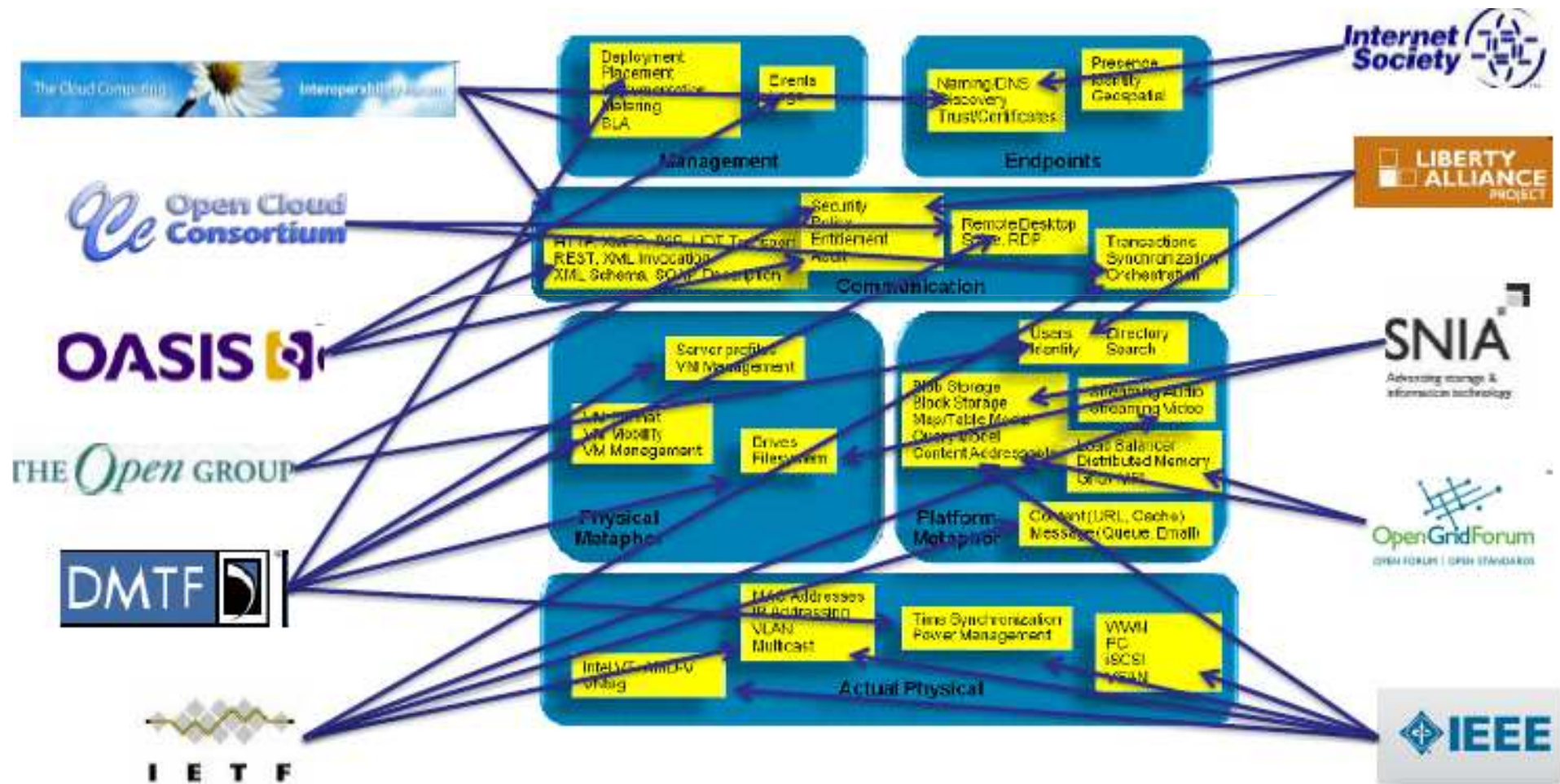


Buyya et al

CLOUDS lab @  
U. Melbourne

# Intercloud standards

- Protocols, formats and mechanisms for interoperability
- From David Bernstein, Hwawei Tech., [www.cloudstrategypartners.com](http://www.cloudstrategypartners.com)



[http://www.iaria.org/conferences2009/filesFUTURECOMPUTING09/DavidBernstein\\_Intro\\_to\\_Intercloud\\_V6.pdf](http://www.iaria.org/conferences2009/filesFUTURECOMPUTING09/DavidBernstein_Intro_to_Intercloud_V6.pdf)



# Combinatorial Innovation

---

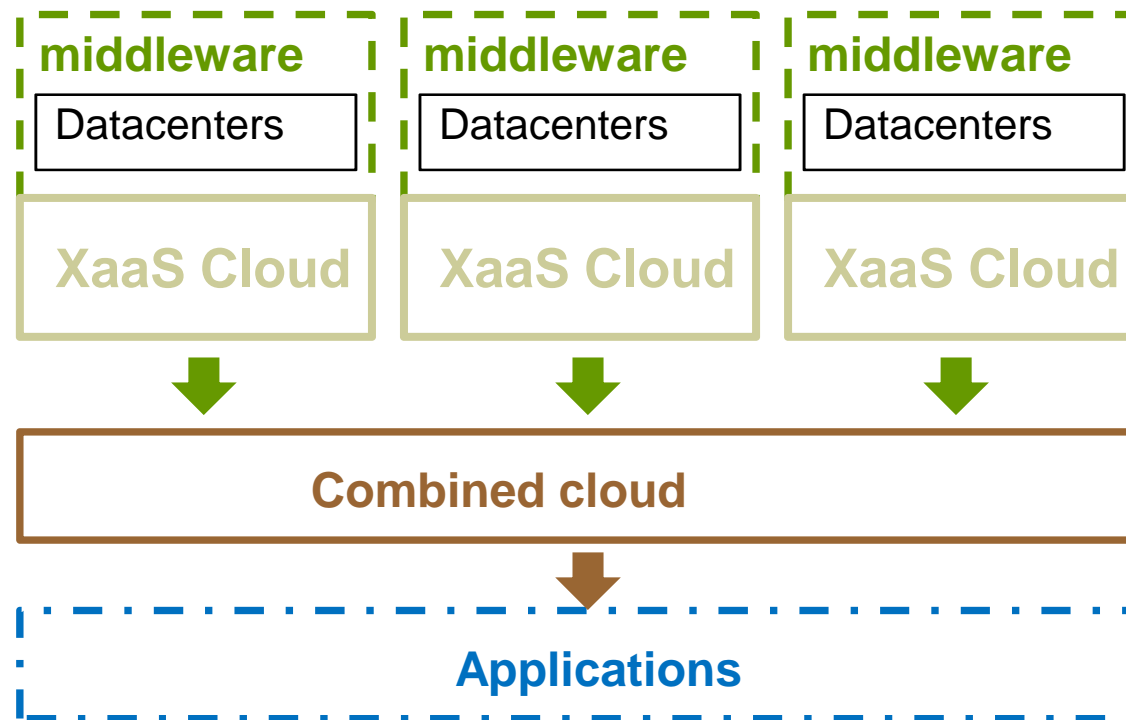
“... historically, you’ll find periods in history where there would be the availability of a different component parts that innovators could combine or recombine to create new inventions. In the 1800s, it was interchangeable parts. In 1920, it was electronics. In the 1970s, it was integrated circuits. Now what we see is a period where you have Internet components, where you have software, protocols, languages, and capabilities to combine these component parts in ways that create totally new innovations. The great thing about the current period is that component parts are all bits. That means you never run out of them. **You can reproduce them, you can duplicate them, you can spread them around the world, and you can have thousands and tens of thousands of innovators combining or recombining the same component parts to create new innovation.** So there’s no shortage. There are no inventory delays. It’s a situation where the components are available for everyone, and so we get this tremendous burst of innovation that we’re seeing.”

**Hal Varian, chief Google economist and professor at UC Berkeley**



# Combined clouds

- Combine: to bring into such close relationship as to obscure individual characteristics



- “Heterogeneous virtual cluster on a WAN” aaS
- “(Excel-based) geospatial market analytics” aaS
- “Personalized health from multiple providers” aaS

# Sky computing

- Combined use of multiple clouds
  - Resources/apps/platforms across independent clouds are used
  - Services other than those of each individual cloud
  - Transparency of multiple clouds - single-cloud like
  - Sky providers are consumers of cloud providers
  - “Virtual” datacenter-less dynamic clouds
- Many challenges and questions
  - Communication among resources in different clouds is of key importance

# Communication Problems

- Connectivity limitations due to the lack of publicly accessible addresses, firewalls, NATs ...
  - Grid computing solutions available (API-based, VPN, P2P, VNET, ViNe ...)
  - User-level network virtualization is a solution adopted by the majority of projects
- Dangers of VM privileged users on the cloud
  - change IP and/or MAC addresses
  - configure Network Interface Card in promiscuous mode
  - use raw sockets
  - attack network (spoofing, proxy ARP, flooding, ...)
- Cloud providers impose network restrictions that severely affect the ability of network virtualization techniques to work efficiently

# Network Restrictions in Clouds

---

- Internal routing and NAT
  - IP addresses (especially public) are not directly configured inside VMs, and NAT techniques are used
- Sandboxing
  - VMs are connected to host-only networks
  - VM-to-VM communication is enabled by a combination of NAT, routing and firewalling mechanisms
- Packet filtering (beyond usual)
  - VMs packets are inspected and only those packets containing valid addresses (IP and MAC assigned by the provider) are allowed

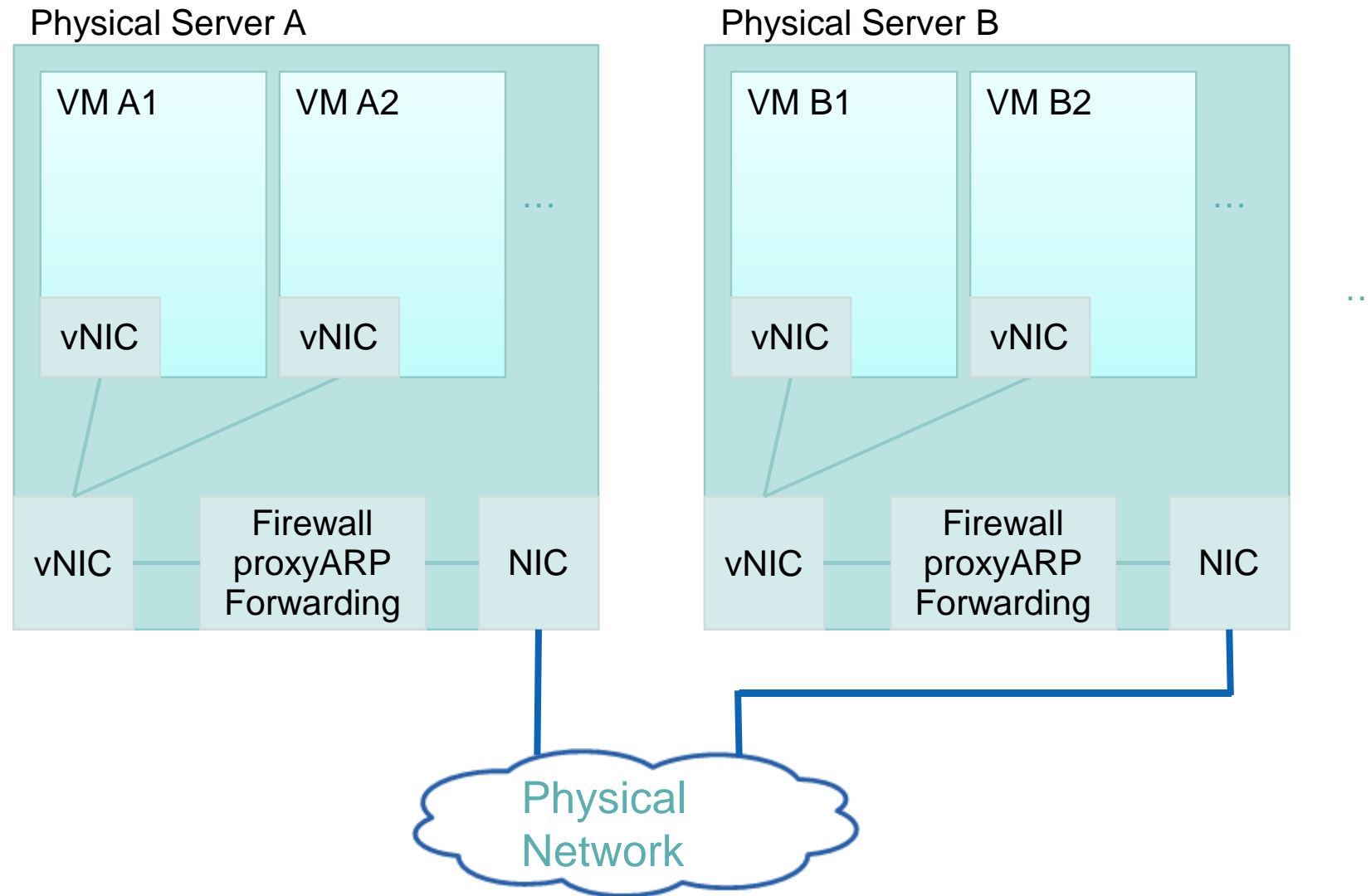
# Network Challenges in Clouds

---

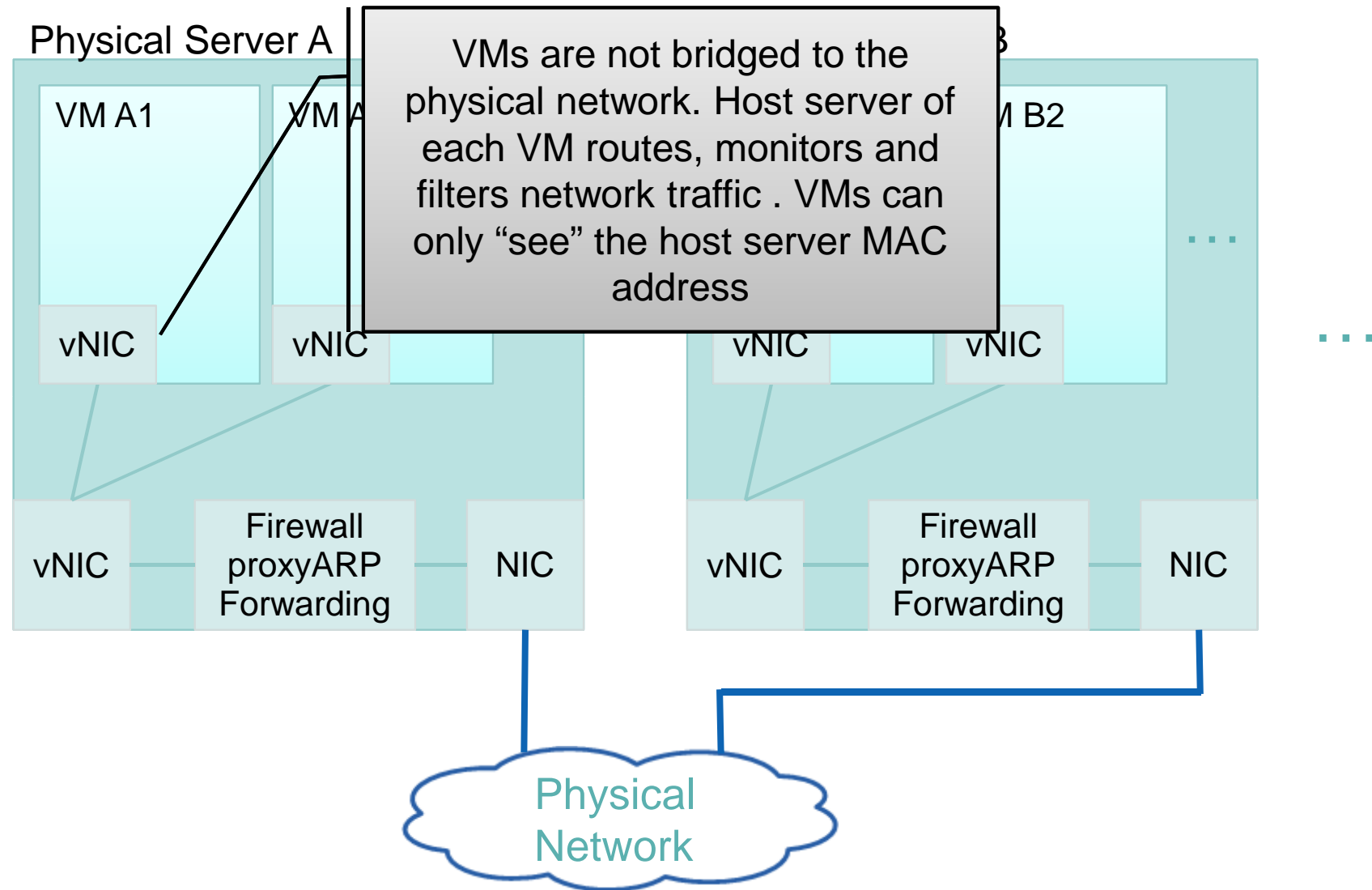
- Internal routing / NAT
  - High number of intermediate nodes (hops) in LAN communication (nodes on the same subnet, thus no hops in-between are expected)
    - EC2 public-to-public – 6+ hops
    - EC2 private-to-private – 3+ hops (better)
- Sandboxing
  - Disables direct datalink layer (L2) communication
  - Can't use VMs as routers
  - No node-to-gateway communication
- Packet filtering
  - Only allows packets w/ source IP address
  - Disables VM ability to act as a router
  - No gateway-to-node communication



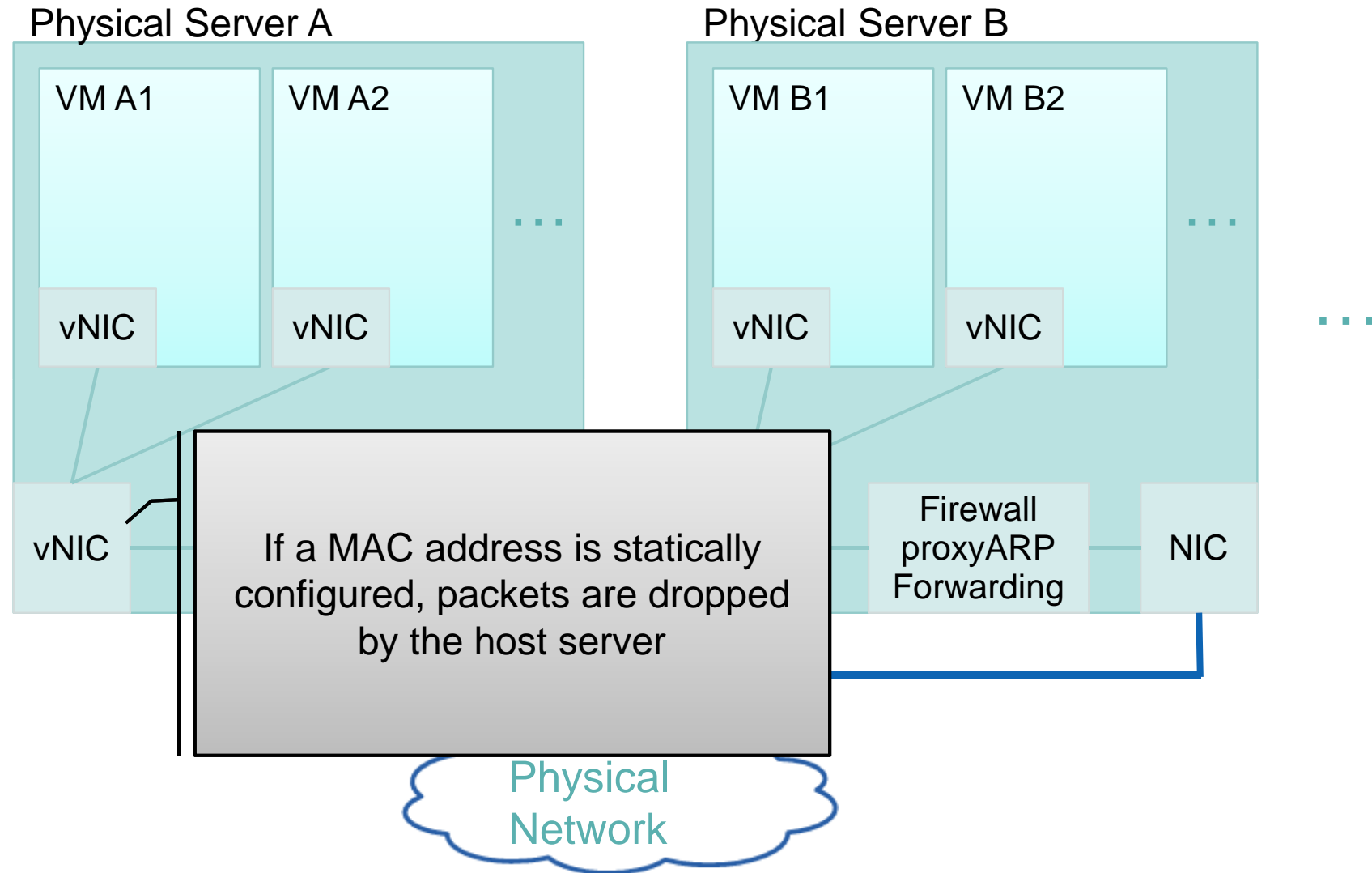
# Typical IaaS Network



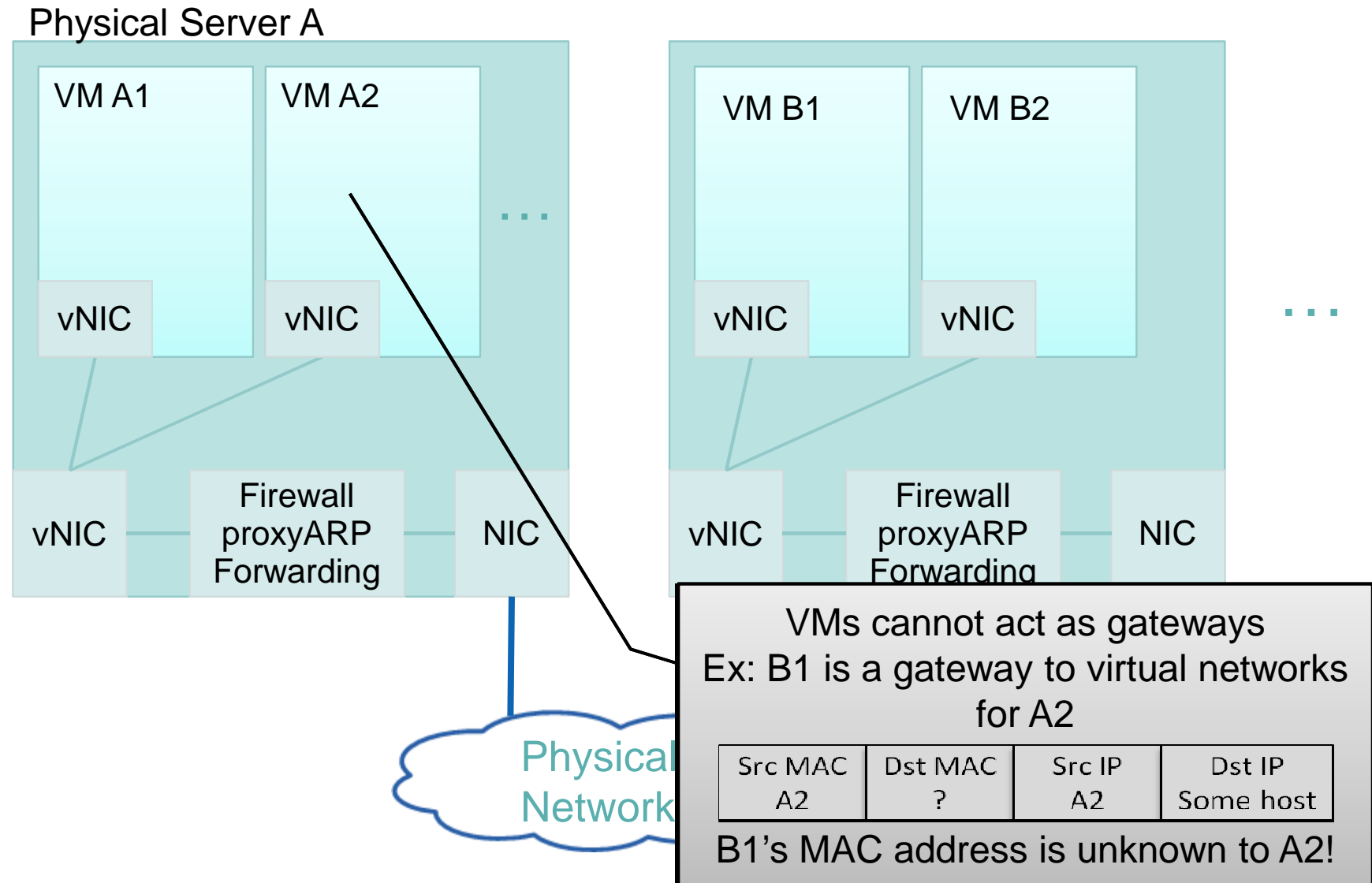
# Typical IaaS Network



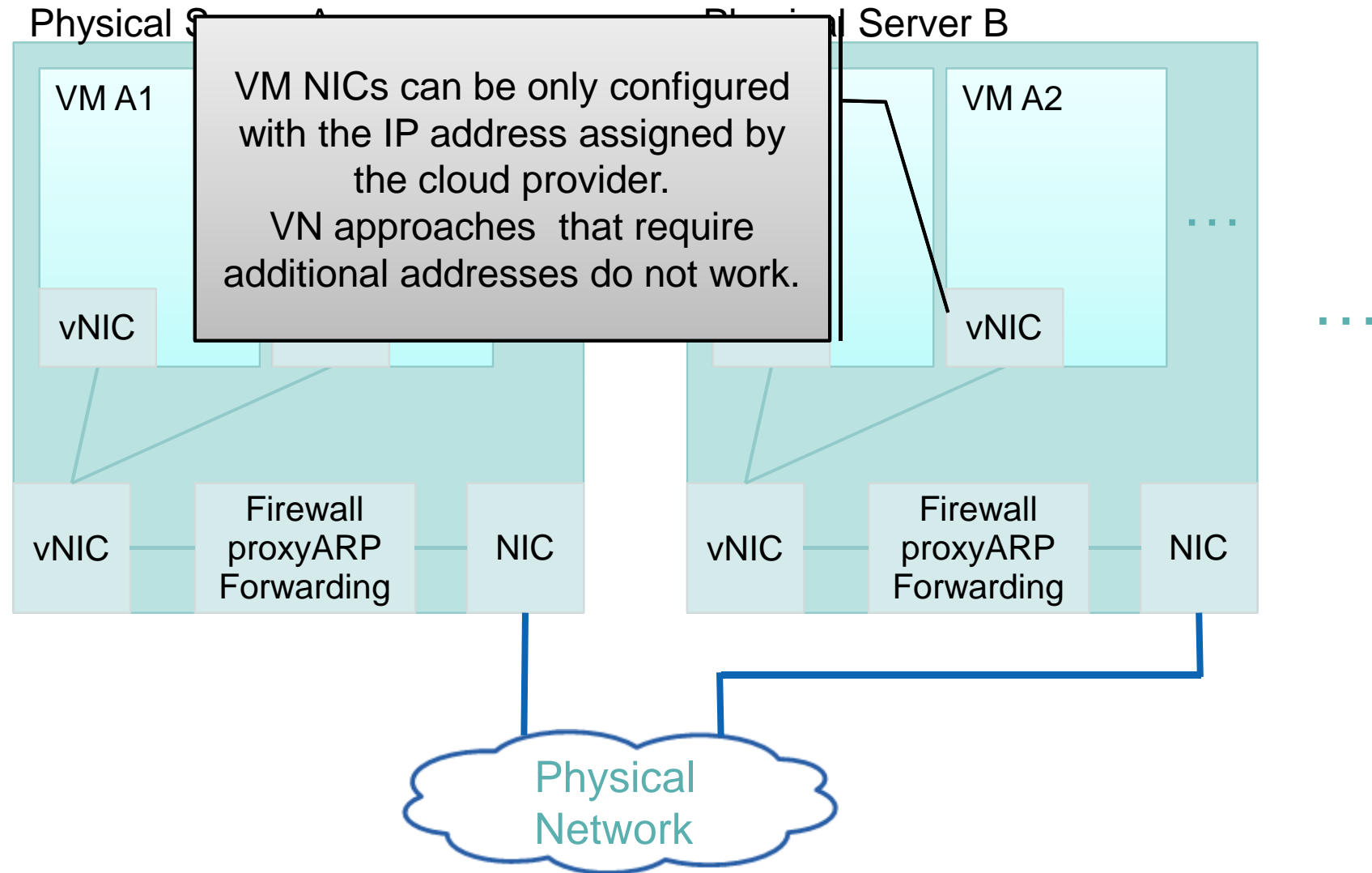
# Typical IaaS Network



# Typical IaaS Network

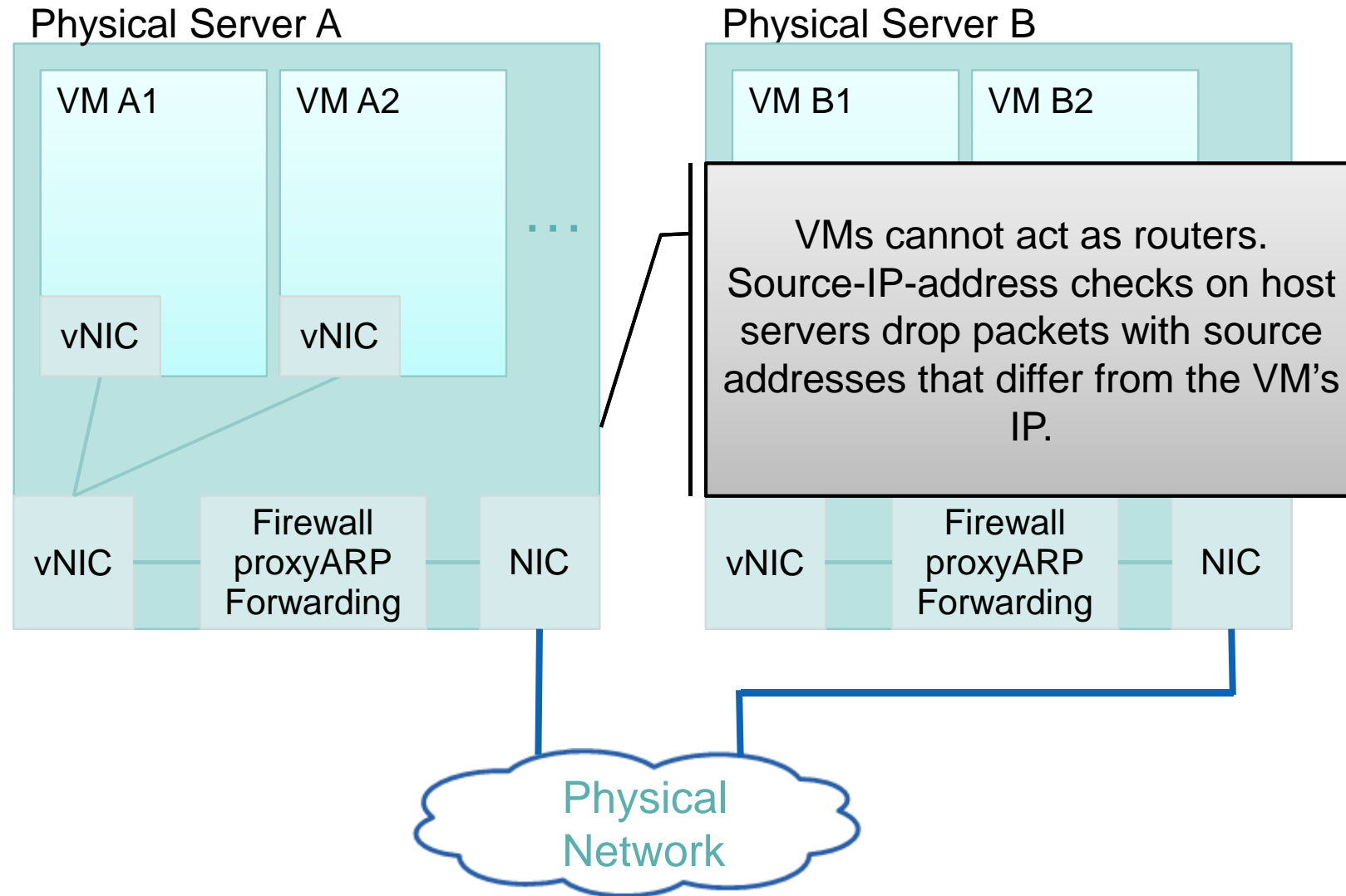


# Typical IaaS Network

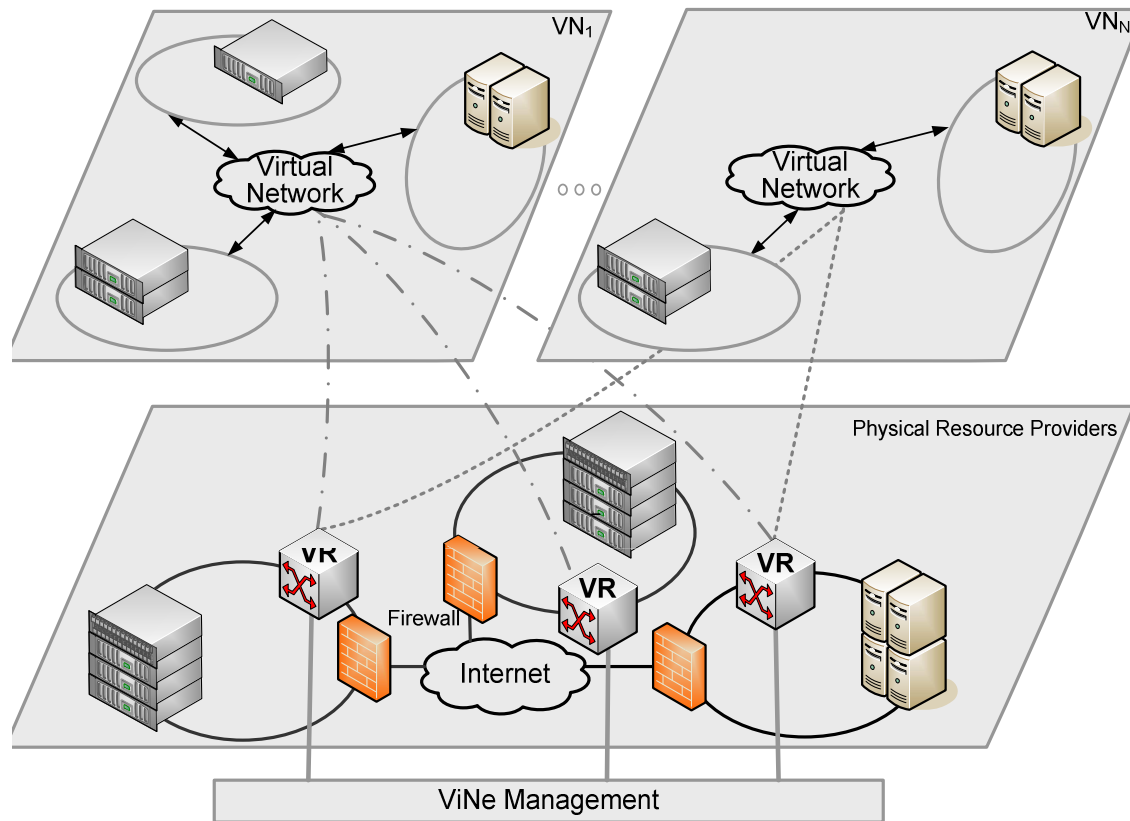




# Typical IaaS Network



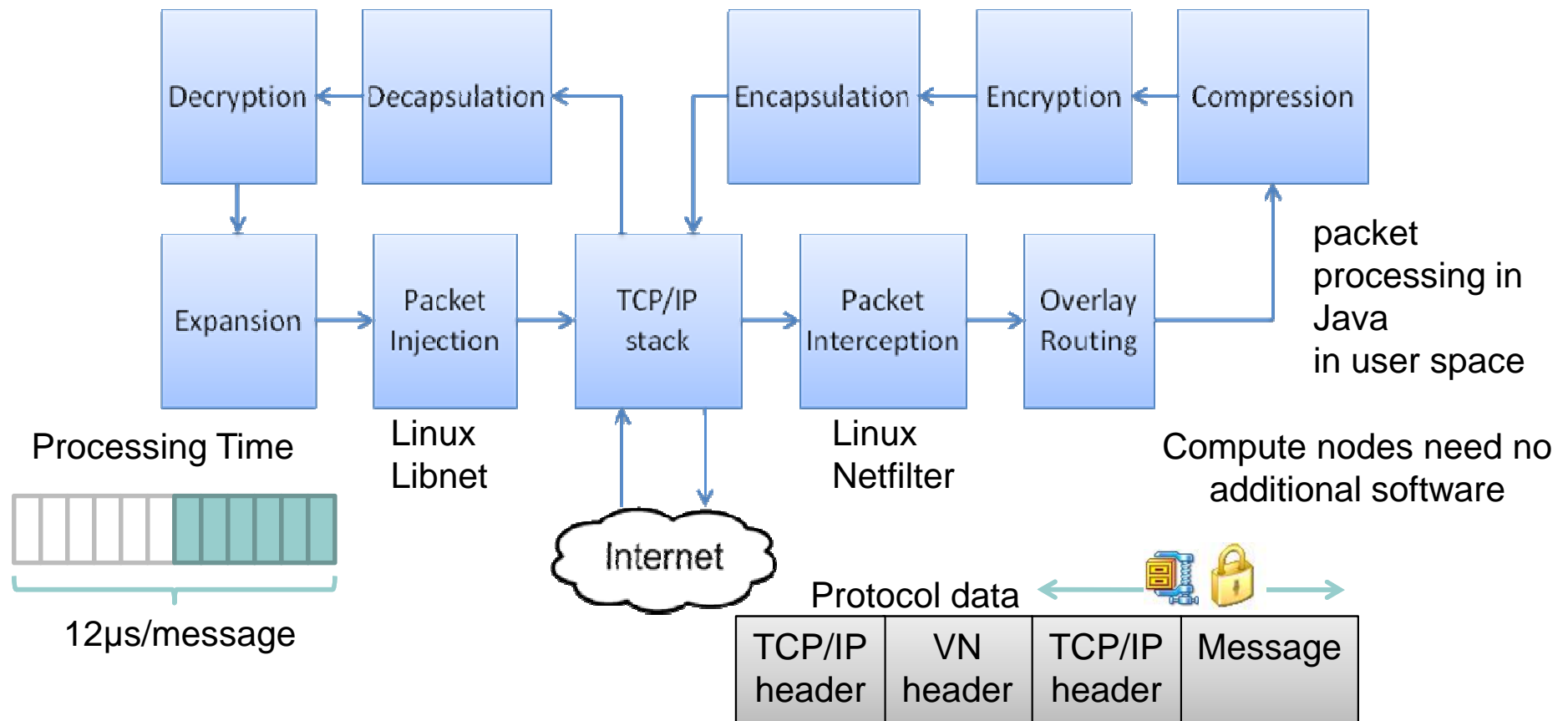
# User-level Virtual Network - ViNe



- General purpose overlay network solution
- Based on deployment of user-level virtual routers
- VRs are used as gateways by nodes that do not run ViNe software
- Applications run unmodified
- Best performance

- VRs ability to perform overlay routing affected by cloud network restrictions
  - Limited node-to-VR communication due to use of L2 communication

# ViNe Routing



- **Local Network Description Table (LNDT)**
  - Describes the VN membership of a node
- **Global Network Description Table (GNDT)**
  - Describes sub-networks for which a VR is responsible

# ViNe Routing

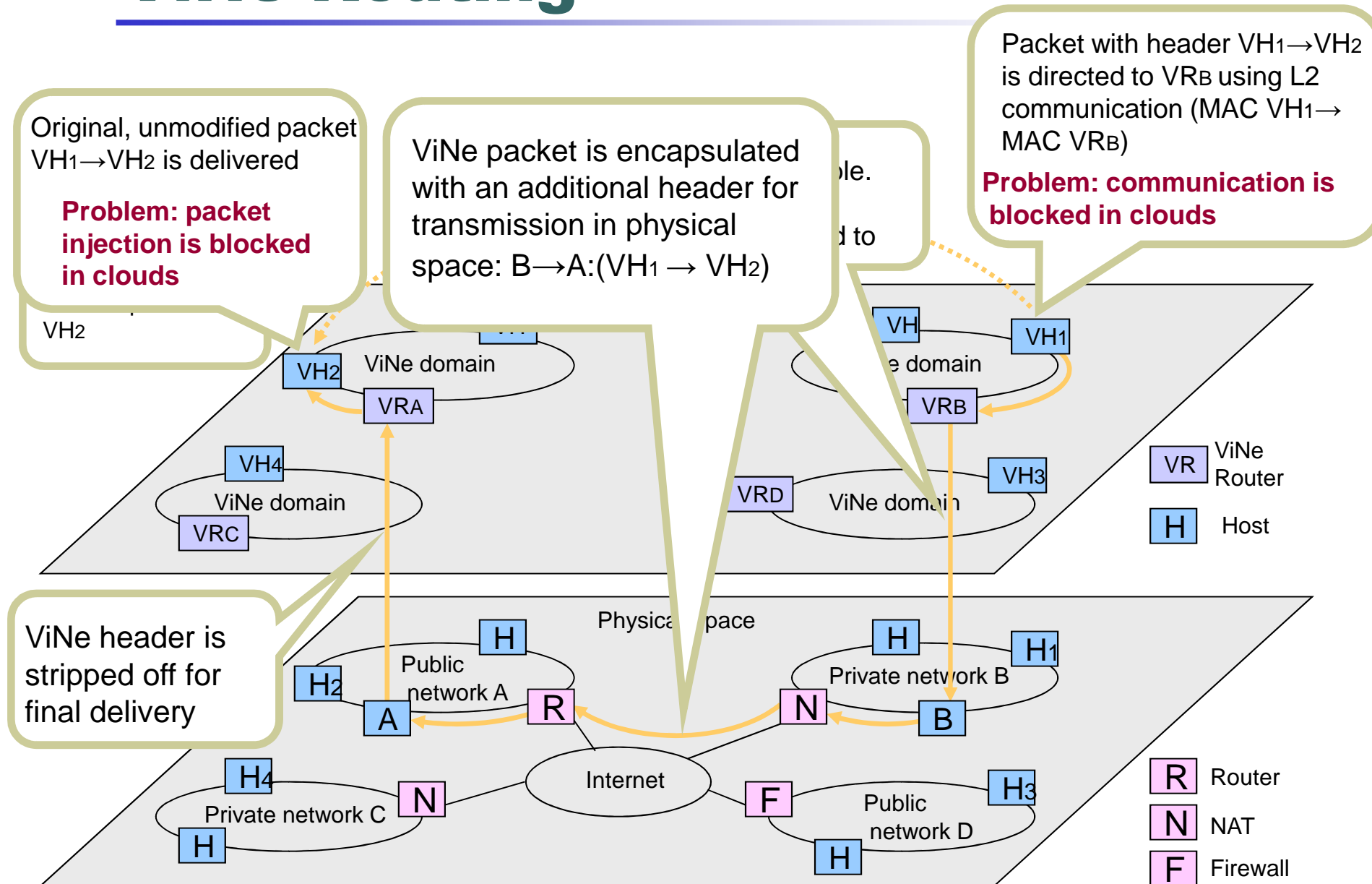
- Local Network Description Table (LNDT)
  - Describes the VN membership of a node
- Global Network Description Table (GNDT)
  - Describes the sub-networks that a VR is responsible
- Suppose that a VR with the following routing tables, received a packet from 172.16.0.10 destined to 172.16.10.90

LNDT	
Host	ViNe ID
172.16.0.10	1
172.16.0.11	2

GNDT – ViNe ID 1	
Network/Mask	Destination
172.16.0.0/24	VR-a
172.16.10.0/24	VR-b

GNDT – ViNe ID 2	
Network/Mask	Destination
172.16.0.0/24	VR-a
172.16.20.0/24	VR-c

# ViNe Routing





# Solution

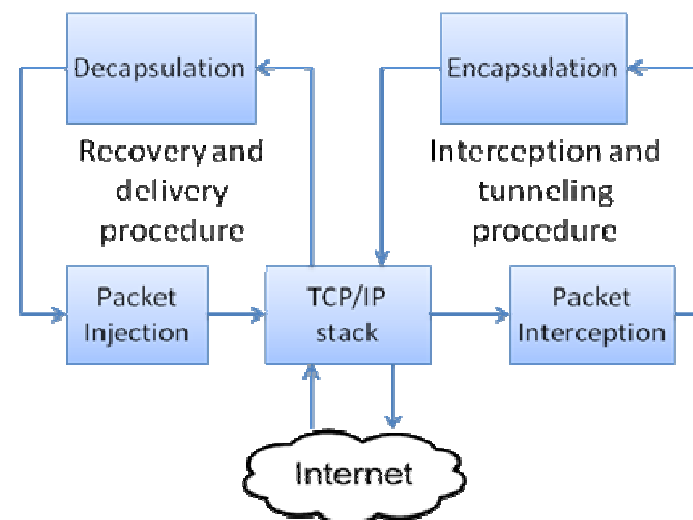
---

- Configure all nodes to work as VRs
  - No need for host-to-VR L2 communication
  - TCP or UDP based VR-to-VR communication circumvents the source address check restriction
- But...
  - Network virtualization software required in all nodes
  - Network virtualization overhead in inter- and intra-site communication
  - Complex configuration and operation
- TinyViNe
  - No need to implement complex network processing – leave it to specialized resources (i.e., full-VRs)
  - Keep it simple, lightweight, tiny
  - Use IP addresses as assigned by providers
  - Make it easy for end users to deploy

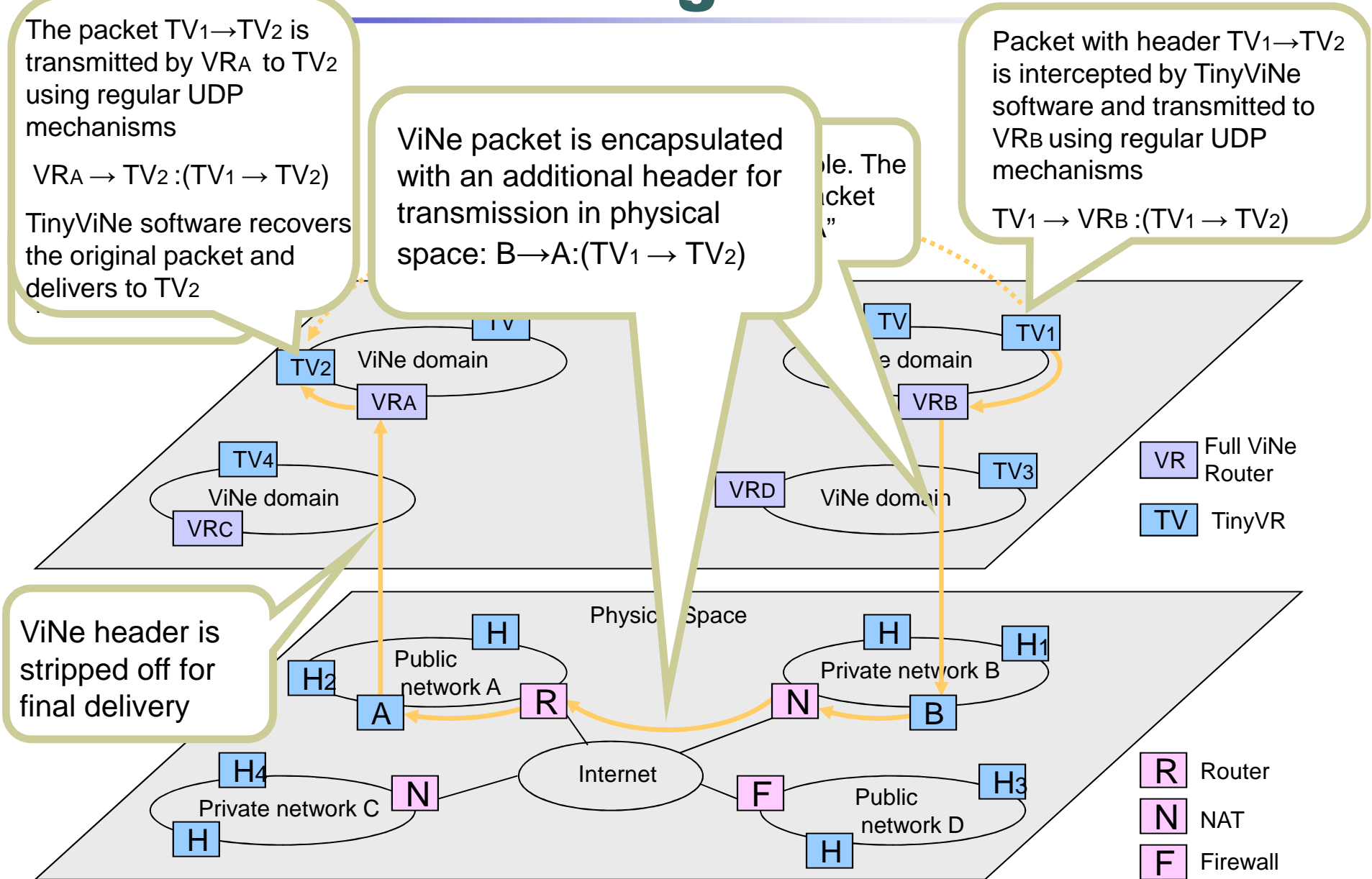
M. Tsugawa\* et al. “User-level Virtual Networks Support for Sky Computing”, e-Science, 12/09.

# TinyViNe

- TinyViNe software
  - Enables host-to-VR communication on clouds using UDP tunnels
  - TinyVR – nodes running TinyViNe software
- TinyVR processing
  - Intercept packets destined to full-VRs
  - Transmit the intercepted packets through UDP tunnels
  - Decapsulate incoming messages through UDP tunnels
  - Deliver the packets



# TinyViNe Routing

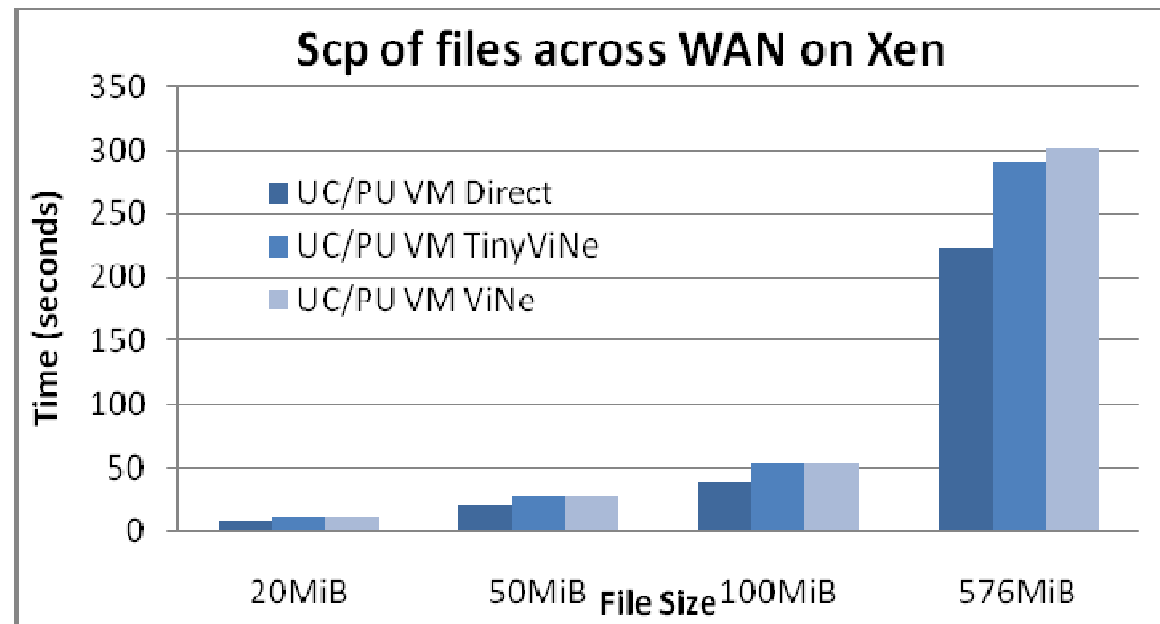


# TinyViNe Overheads

- Interference on other applications

CPU Utilization	Direct	TinyViNe	ViNe
Matrix multiplication	95.62%	92.74%	89.04%
Network application	1.44%	1.17%	1.04%
(Tiny)ViNe software	-	9.27%	11.98%

- Impact on communication-intensive application

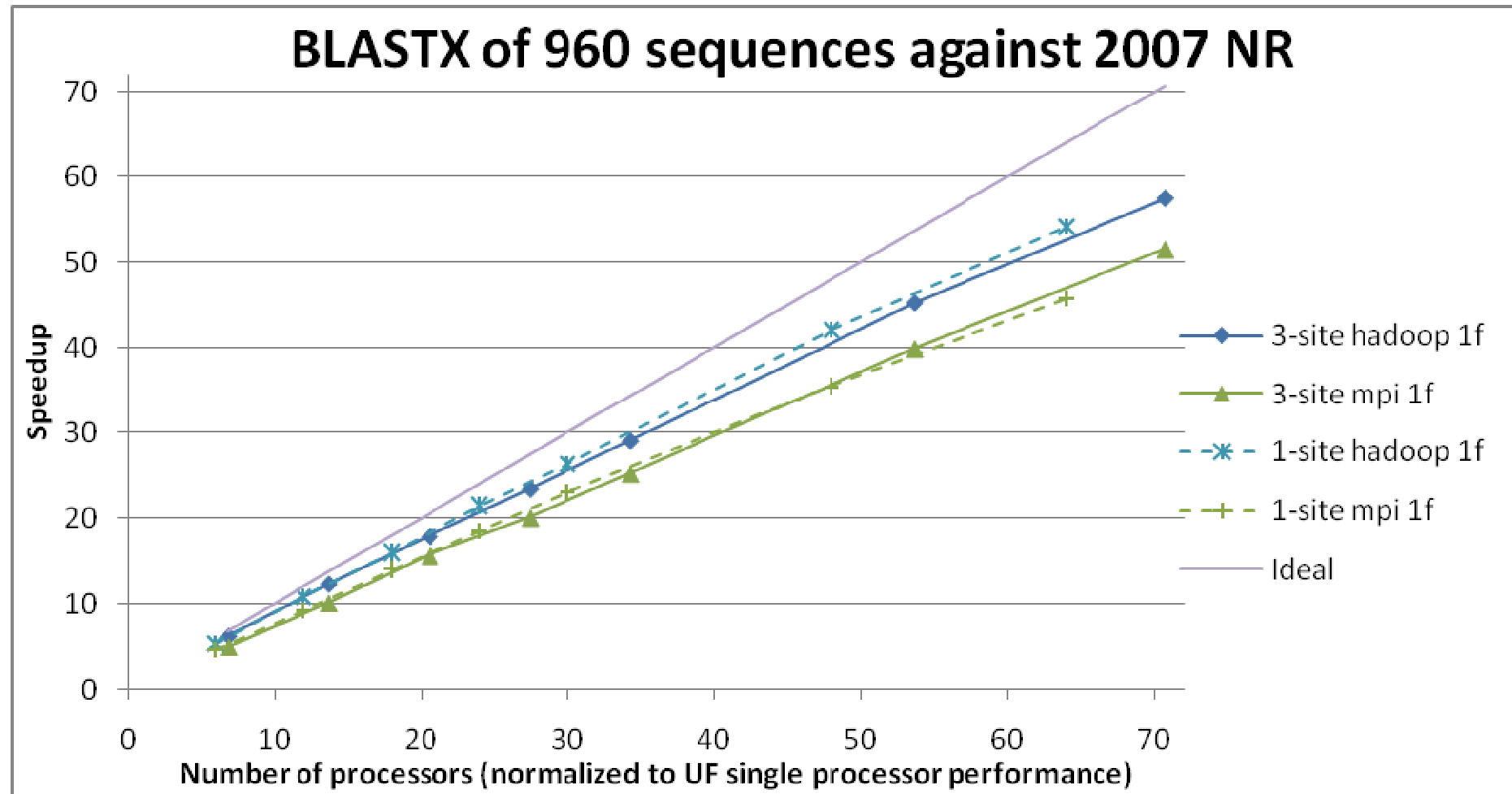


# Experiment on the sky

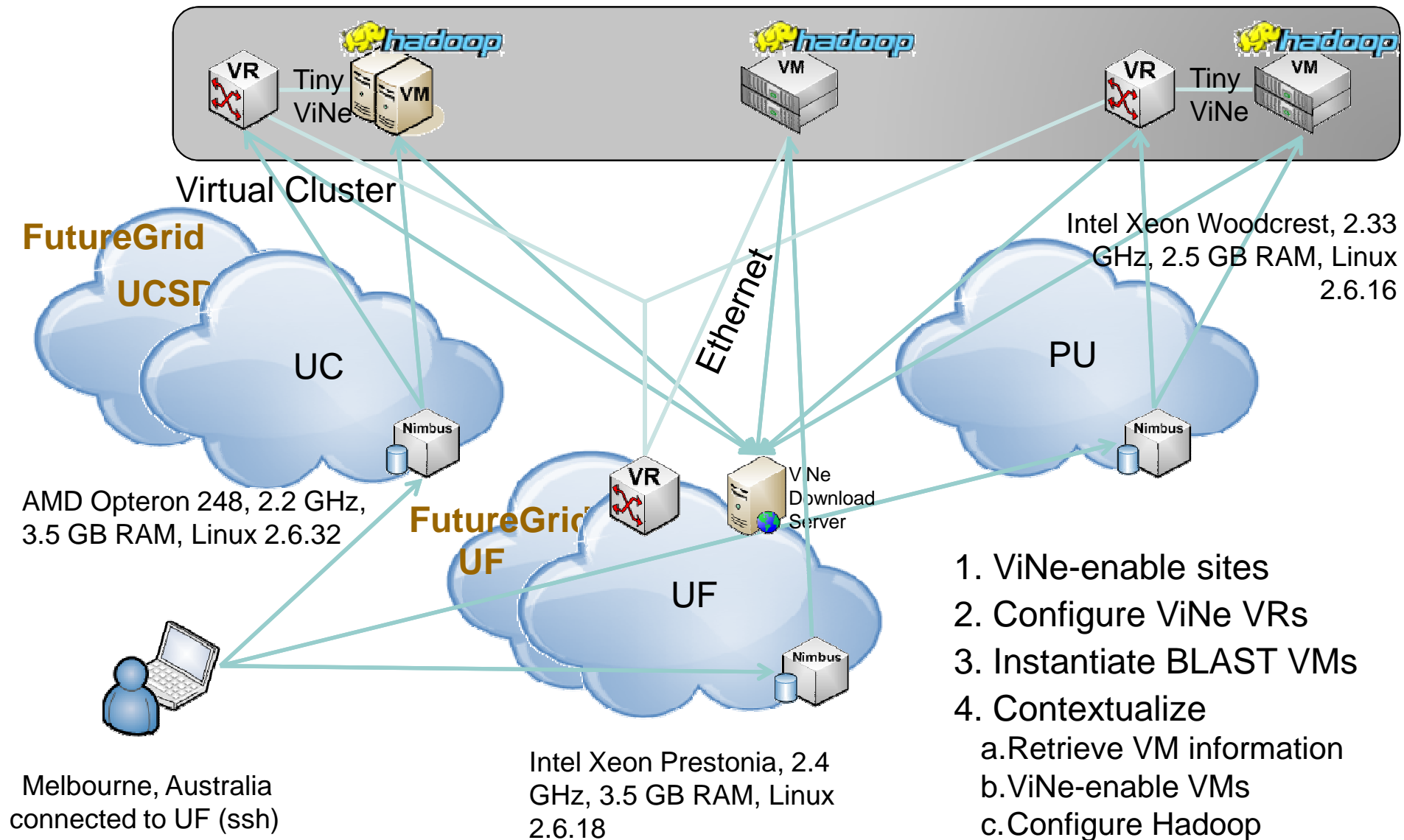
---

- 3 cloud providers at University of Florida, University of Chicago and Purdue University
- All 3 providers offer IaaS using Nimbus
  - Offers APIs to create and manage VMs
  - Contextualization: mechanisms by which VMs are “adapted” to their environments – e.g., on a virtual cluster deployment, each node can be configured with different “roles”
- TinyViNe deployed by adding 1 line to the Nimbus virtual cluster configuration file

# 3-sites experiment



# Sky Computing





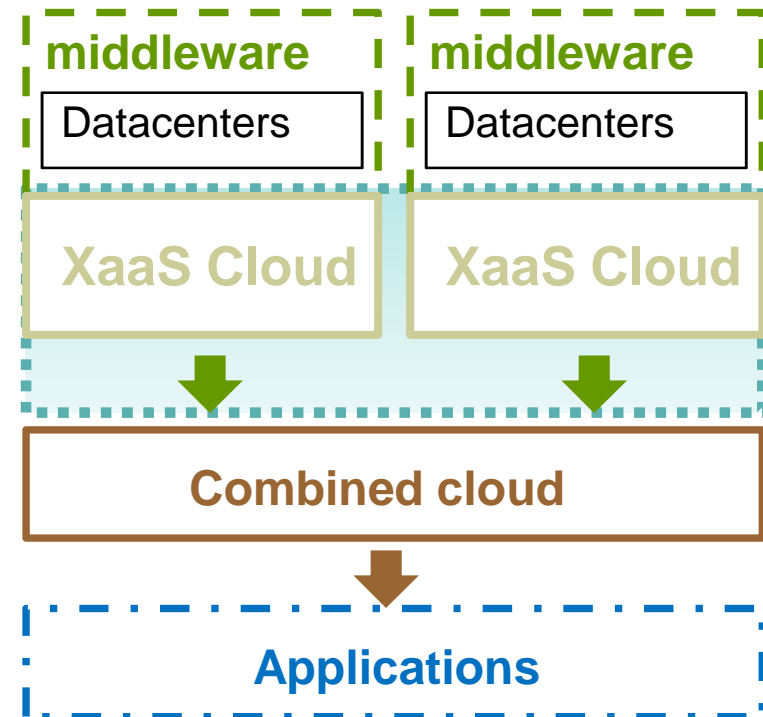
# Summary

---

- User-level overlay networks needed for inter-cloud communication
- Hard to deploy due to cloud-specific restrictions
  - Overcome via network-virtualization software in VMs
  - It is important to keep the software simple and light
- TinyViNe enables applications across clouds
  - Experiments with parallel bioinformatics applications show that it efficiently enables sky computing
  - Can be implemented as a service by a cloud provider, in the context broker of a sky provider or by the consumer
  - Being improved to enable autonomic networking

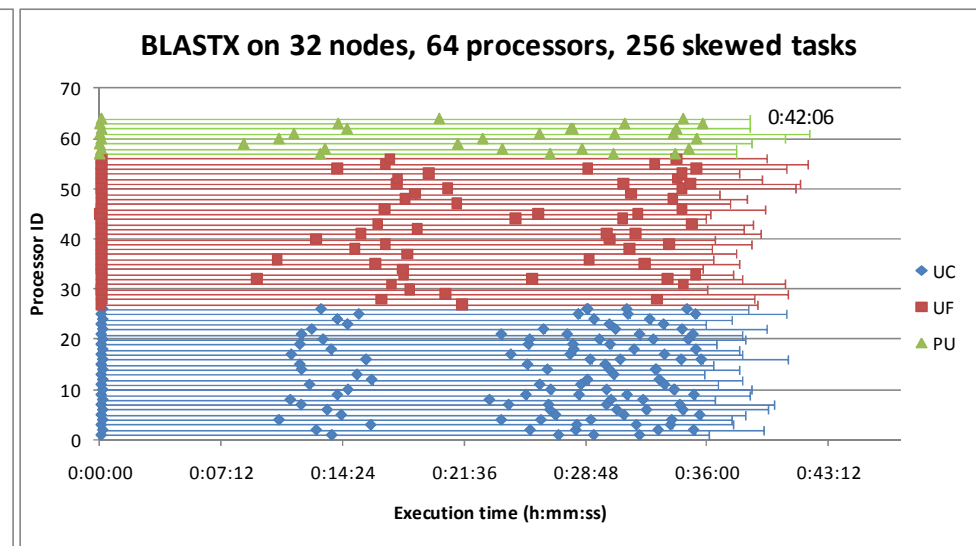
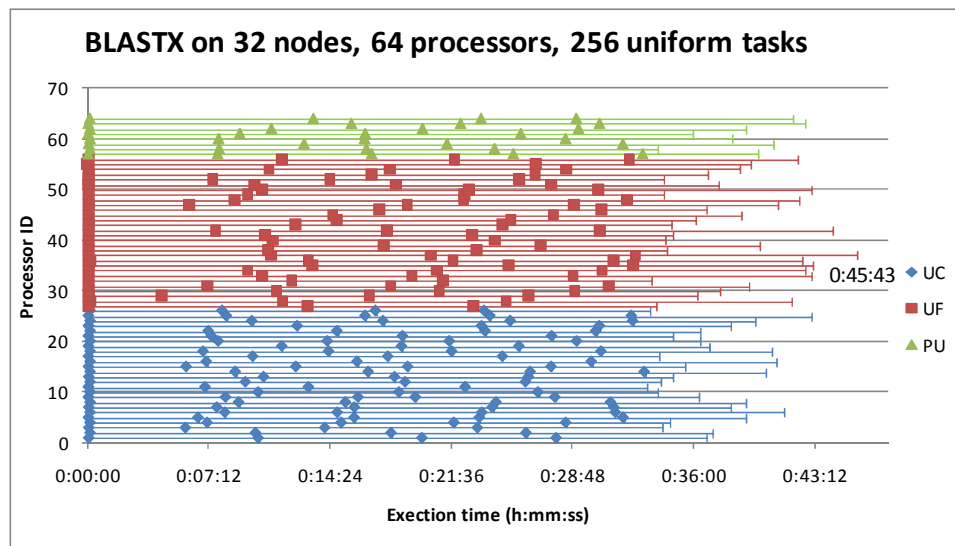
# Beyond communication

- XaaS = abstractions as a service
  - Which ones to use?
  - SLAs: what is in them and how to support?
  - Affect management of performance, complexity, dependability, ...
- Contextualization, coordination and management
- Modeling is essential
- Issues: security, privacy, business models...



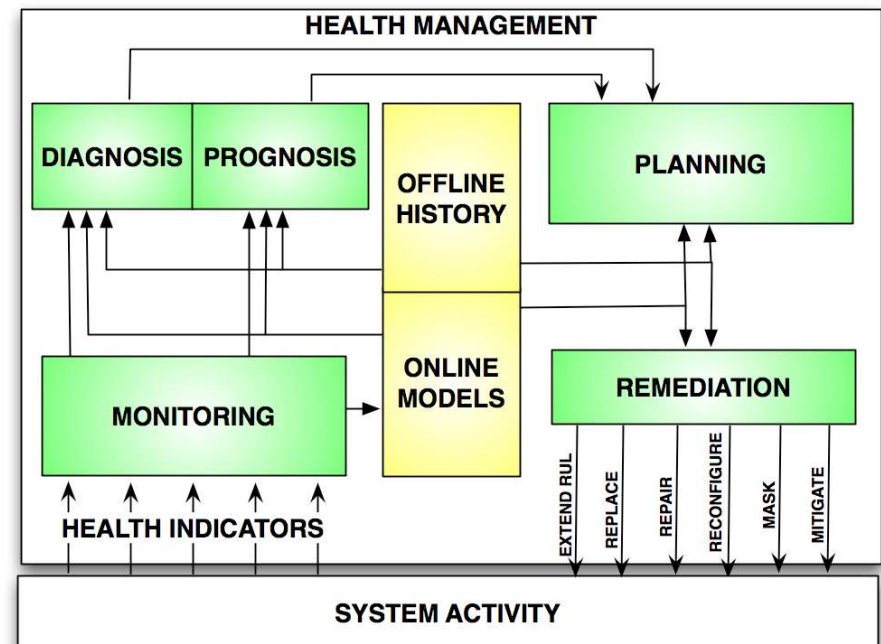
# Resource usage estimation

- Provider perspective
  - can improve resource utilization, as schedulers are able to fit more requests in the same resource
- Consumer perspective
  - to choose the most cost-effective cloud and resource configuration for a given problem



# Fault modeling

- If components are cloud services, what is a component fault?
  - SLA violation? User-defined condition? Unusual behavior?
  - E.g. resource-exhaustion faults
- How can the health of a sky system/app be managed?
- What/how are concerns separated?
  - E.g. virtual routers
  - “A distributed system is one in which the failure of a computer you didn't even know existed can render your own computer unusable”  
Leslie Lamport
- Similar issues for security, privacy, performance...
- Autonomic solutions desirable



# Conclusions

---

- Clouds provide the components for novel types of IT systems or novel implementations of familiar IT system architectures
  - Sky-computing refers to such systems and their use
  - In particular, combined clouds capable of providing environments, workflows, enterprise IT, etc as a service
- Design and management of combined clouds face challenges and need fundamental and system-oriented advances
  - A new area for IT research
  - Essential for standards and next generation of IT businesses

# Acknowledgments

- Sponsors



- National Science Foundation
- BellSouth Foundation
- Center for Autonomic Computing (UF site)
  - Citrix, IBM, Intel, Microsoft, Northrop-Grumman

- Collaborators

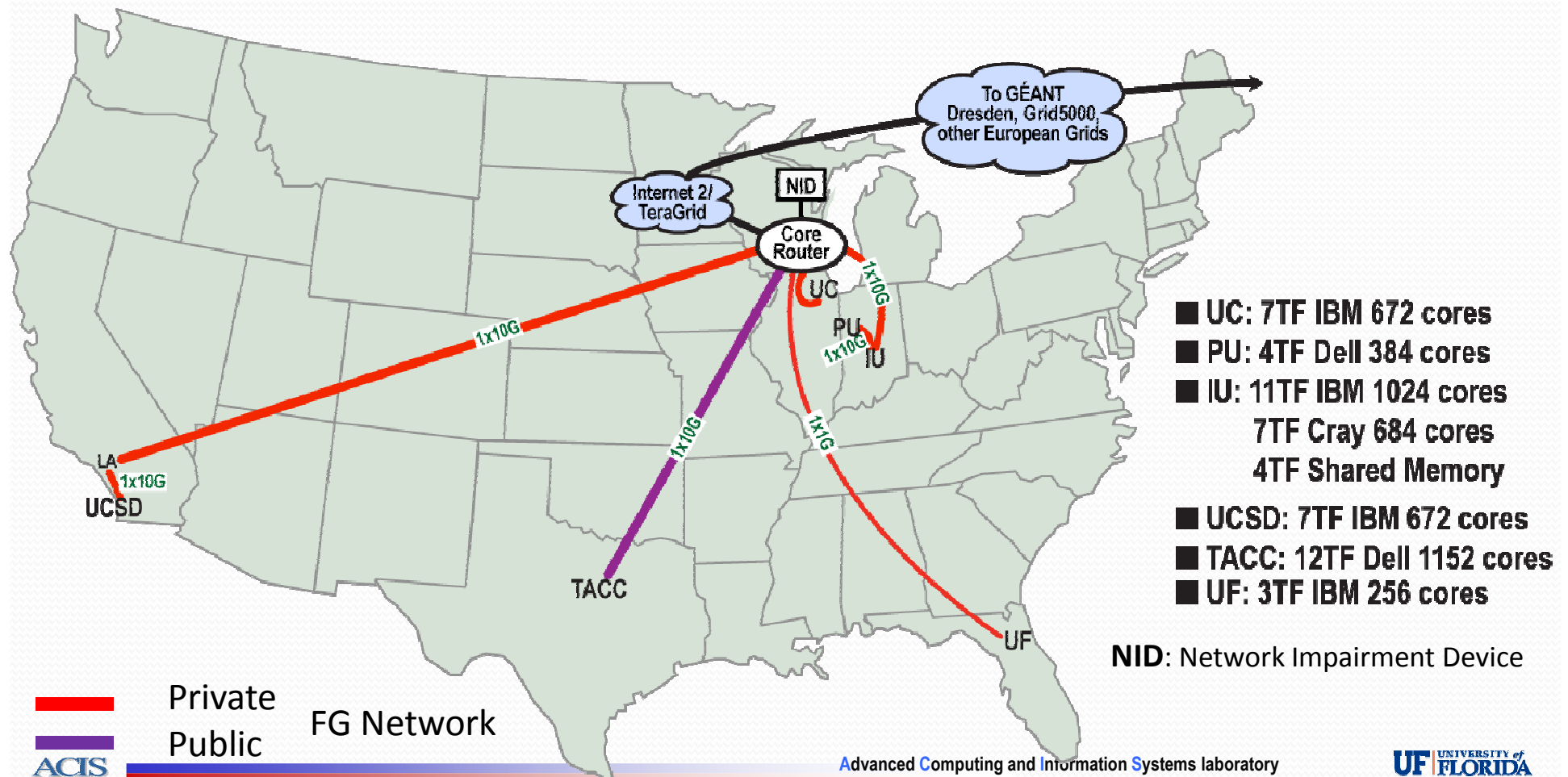
- Andrea Matsunaga and Mauricio Tsugawa
- Kate Keahey, Tim Freeman
- Renato Figueiredo and others at ACIS/CAC
- NSF FutureGrid team



# FutureGrid: a Grid Testbed



- **IU** Cray operational, **IU** IBM (iDataPlex) completed stability test May 6
- **UCSD** IBM operational, **UF** IBM stability test completes ~ May 12
- **Network**, **NID** and **PU** HTC system operational
- **UC** IBM stability test completes ~ May 27; **TACC** Dell awaiting delivery of components





# FutureGrid



- The goal of FutureGrid is to **support the research** on the future of distributed, grid, and cloud computing.
- FutureGrid will build a robustly managed simulation environment or testbed to support the development and early use in science of new technologies at all levels of the software stack: from **networking to middleware to scientific applications**.
- FutureGrid is a (small 5600 core) **Science/Computer Science Cloud** but it is more accurately a **virtual machine based simulation environment**



Browser window showing the Future Grid website. The address bar displays <http://futuregrid.org/about>. The browser's bookmark bar includes links to AfreshWiki, Clock, Customize Links, TWC weather, TeleCenter, my schedule, Webster, my page, ACIS, Google Groups, and Other bookmarks.

The website header features the Future Grid logo and a navigation bar with links to [About](#), [Register](#), and [Help & Support](#).

**Links**

- ▼ About
  - News
  - Events
  - Sponsors
  - Contact
- ▼ Hardware
  - Status
- Register
- Help & Support


**Internal Pages**

**Book navigation**

- Education and outreach

**Home**

## About

 This project provides a capability that makes it possible for researchers to tackle complex research challenges in computer science related to the use and security of grids and clouds. These include topics ranging from authentication, authorization, scheduling, virtualization, middleware design, interface design and cybersecurity, to the optimization of grid-enabled and cloud-enabled computational schemes for researchers in astronomy, chemistry, biology, engineering, atmospheric science and epidemiology. The project team will provide a significant new experimental computing grid and cloud test-bed, named FutureGrid, to the research community, together with user support for third-party researchers conducting experiments on FutureGrid.

The test-bed will make it possible for researchers to conduct experiments by submitting an experiment plan that is then executed via a sophisticated workflow engine, preserving the provenance and state information necessary to allow reproducibility.

The test-bed includes a geographically distributed set of heterogeneous computing systems, a data management system that will hold both metadata and a growing library of software images, and a dedicated network allowing isolatable, secure experiments. The test-bed will support virtual machine-based environments, as well as native operating systems for experiments aimed at minimizing overhead and maximizing performance. The project partners will integrate existing open-source software packages to create an easy-to-use software environment that supports the instantiation, execution and recording of grid and cloud computing experiments.

One of the goals of the project is to understand the behavior and utility of cloud computing approaches. Researchers will be able to measure the overhead of cloud technology by requesting linked experiments on both virtual and bare metal systems. FutureGrid

# FutureGrid Partners

- [Indiana University](#) (Architecture, core software, Support)
- [Purdue University](#) (HTC Hardware)
- [San Diego Supercomputer Center](#) at University of California San Diego (INCA, Monitoring)
- [University of Chicago](#)/Argonne National Labs (Nimbus)
- [University of Florida](#) (ViNE, Education and Outreach)
- University of Southern California Information Sciences Institute (Pegasus to manage experiments)
- University of Tennessee Knoxville (Benchmarking)
- [University of Texas at Austin](#)/Texas Advanced Computing Center (Portal)
- University of Virginia (OGF, Advisory Board and allocation)
- Center for Information Services and GWT-TUD from Technische Universität Dresden. (VAMPIR)

# Center for Autonomic Computing (nsfcac.org)



## Founding industry members

- collaborative partnership amongst industry, academe, and government;
- concepts, technologies and resources for industry-relevant autonomic computing research;
- interdisciplinary education on autonomic computing;

□ Industry and government agencies invited to join as members

**BAE SYSTEMS**

**CITRIX**

**EWA Systems**

**IBM**

**intel** Leap ahead™

**imaginestics**

**ISCA**

**Merrill Lynch**

**Microsoft**

**NORTHROP GRUMMAN**

DEFINING THE FUTURE

**NEC**

**Raytheon**

**xerox**



