Cloud Computing with Nimbus

XtreemOS Summer School 2009

Oxford, September 2009

Kate Keahey

keahey@mcs.anl.gov

Nimbus project lead

University of Chicago

Argonne National Laboratory

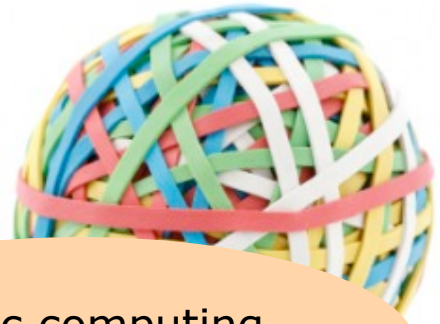
What Is a Cloud?

- The cloud:
  - Eliminates cost
  - Is Elastic
  - Removes Undifferentiated "Heavy Lifting"

(from W. Vogel's blog)

Benefits to Consumers

Eliminate expense and headaches of acquiring, managing and operating hardware



Elastic computing
Pay-as-you-go model

capital expense

operational expense

Benefits to Providers

Economies of scale to amortize the costs of buying and operating resources

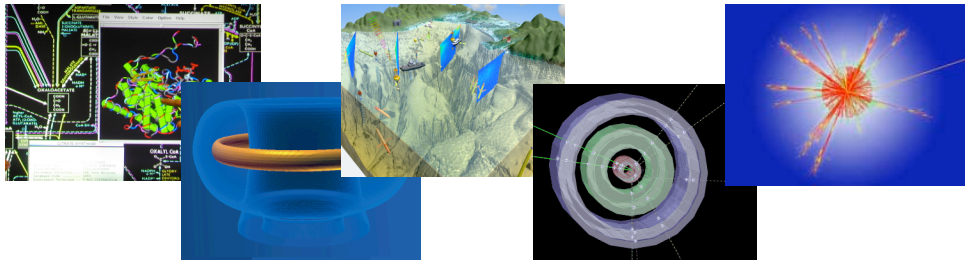
Avoid cost and complexity of managing multiple customer-specific environments and applications

Streamline and specialize

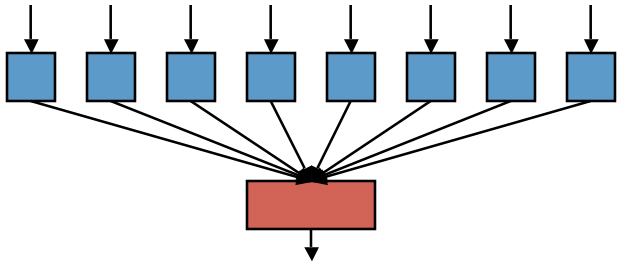
Unclouding the Cloud

Software-as-a-Service (SaaS)

Community-specific applications and portals



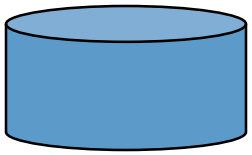
Platform-as-a-Service (PaaS)



Microsoft

Google

Infrastructure-as-a-Service (IaaS)

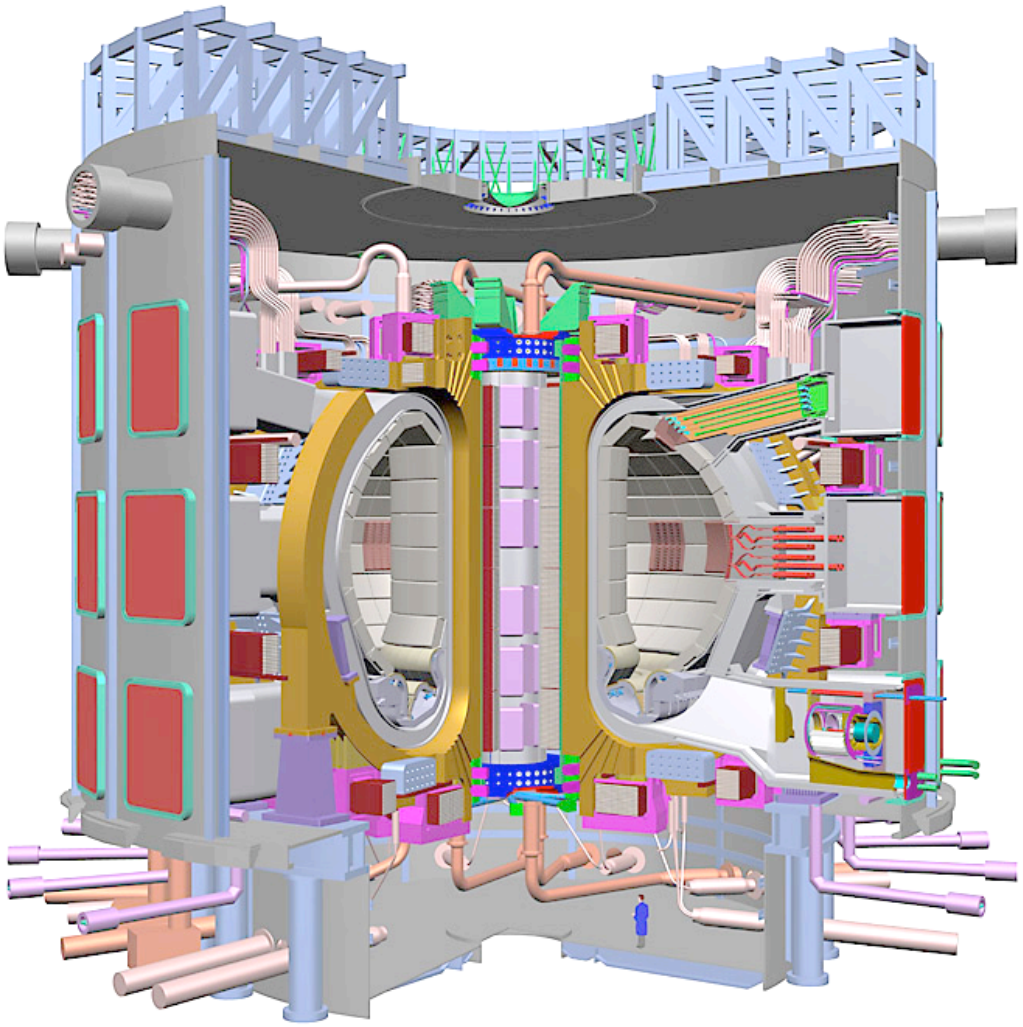


amazon web services™

flexiscale™

GOGRID beta
A ServePath Company

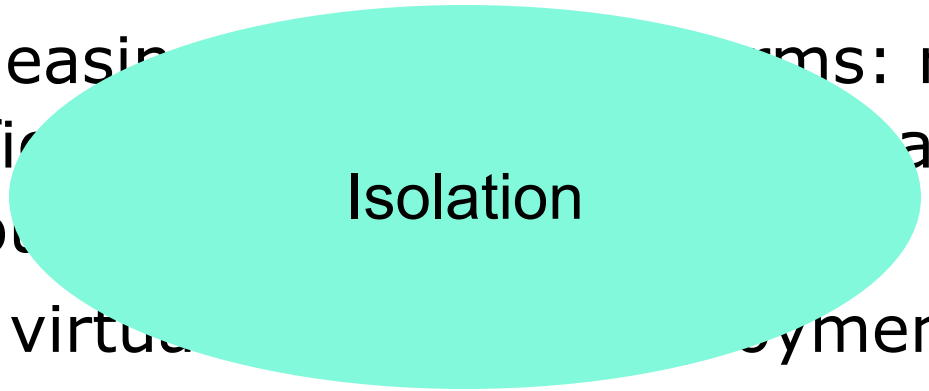
Cloud Computing for Science



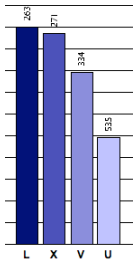
- Complex environments
- Need for control

"Workspaces"

- Dynamically provisioned environments
  - Environment control
  - Resource control
- Implementations
  - Via leasi ms: reimaging, confi amic accou
  - Via virtu yment



Grids to Clouds: a Personal Perspective



Xen released

EC2 goes online

STAR ☆

First STAR production run on EC2

Nimbus Cloud comes online

2003

2006

2009

"A Case for Grid Computing on VMs"
In-Vigo, VIOLIN, DVEs, Dynamic accounts
Policy-driven negotiation

First WSRF Workspace Service release

EC2 gateway available

Context Broker release

Support for EC2 interfaces

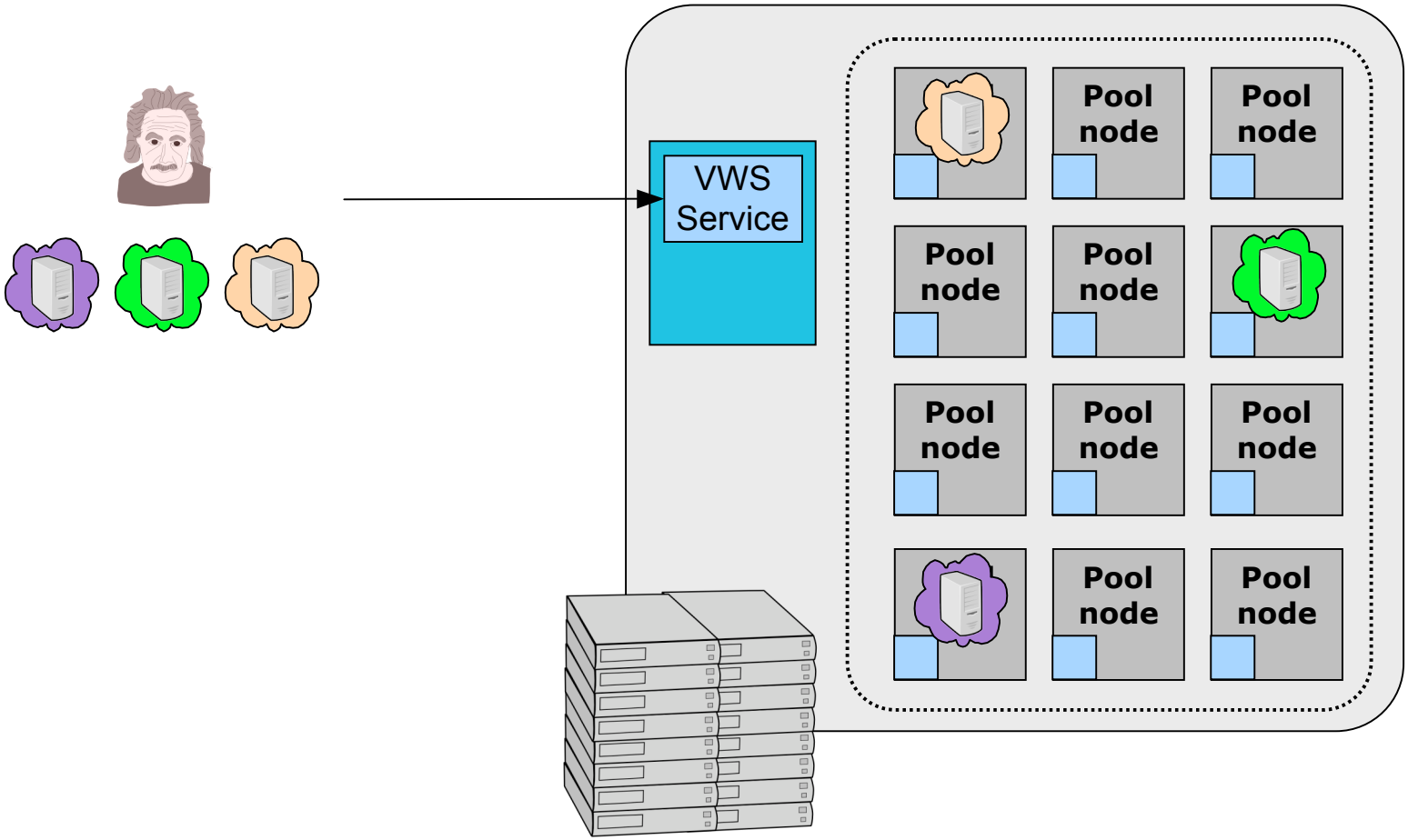
A Tour of Nimbus

Nimbus Goals

- Allow providers to build clouds
  - Private&shared (privacy, expense considerations)
  - Workspace Service: open source EC2 implementation
- Allow users to use cloud computing
  - Do whatever it takes to enable scientists to use IaaS
  - Context Broker: turnkey virtual clusters,
  - Also: protocol adapters, account managers, scaling tools…
- Allow developers to experiment with Nimbus
  - For research or usability/performance improvements
  - Community extensions and contributions: UVIC (monitoring), IU (EBS), Technical University of Vienna (privacy, research)
- Nimbus: http://workspace.globus.org

The Workspace Service

The Workspace Service

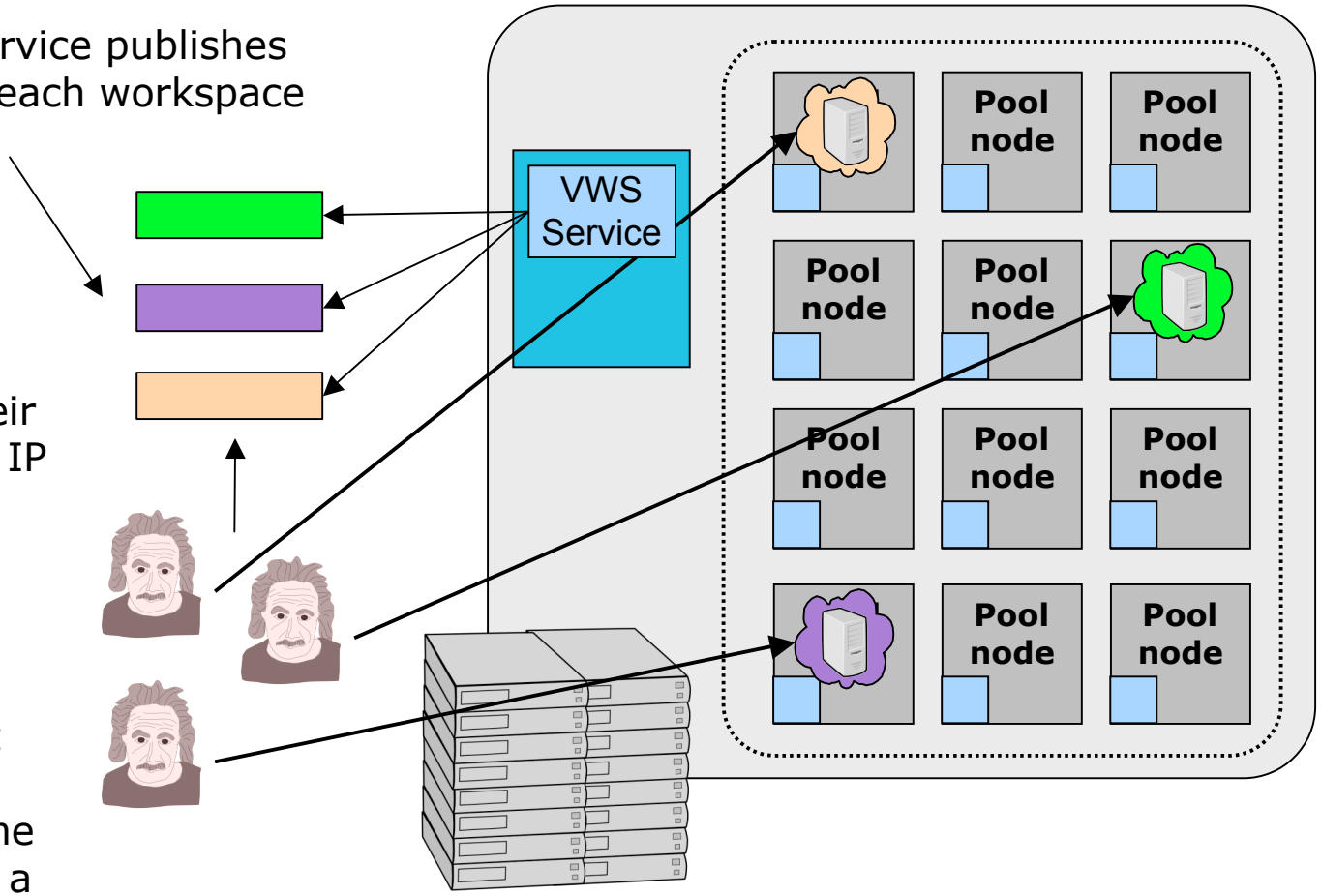


The Workspace Service

The workspace service publishes information about each workspace

Users can find out information about their workspace (e.g. what IP the workspace was bound to)

Users can interact directly with their workspaces the same way the would with a physical machine.



Workspace Service: Interfaces and Clients

- Web Services based
- Web Service Resource Framework (WSRF)
  - WS + state management (WS-Notification)
- Elastic Computing Cloud (EC2)
  - Compatible with EC2 clients
  - Supported: ec2-describe-images, ec2-run-instances, ec2-describe-instances, ec2-terminate-instances, ec2-reboot-instances, ec2-add-keypair, ec2-delete-keypair
  - Unsupported: availability zones, security groups, elastic IP assignment, REST
- Protocol adapter, moving towards messaging

Workspace Service: Security

- GSI authentication and authorization
  - PKI-based
  - VOMS, Shibboleth (via GridShib), custom PDPs
- Secure access to VMs
  - EC2 key generation
  - Accessed from .ssh
- Validating images and image data
  - Extensions from Vienna University of Technology
  - *Paper: Descher et al., Retaining Data Control in Infrastructure Clouds, ARES (the International Dependability Conference), 2009.*

Workspace Service: Networking

- Network configuration
  - External: public IPs or private IPs (via VPN)
  - Internal: private network via a local cluster network
- Each VM can specify multiple NICs mixing private and public networks (WSRF only)
  - E.g., cluster worker nodes on a private network, headnode on both public and private network

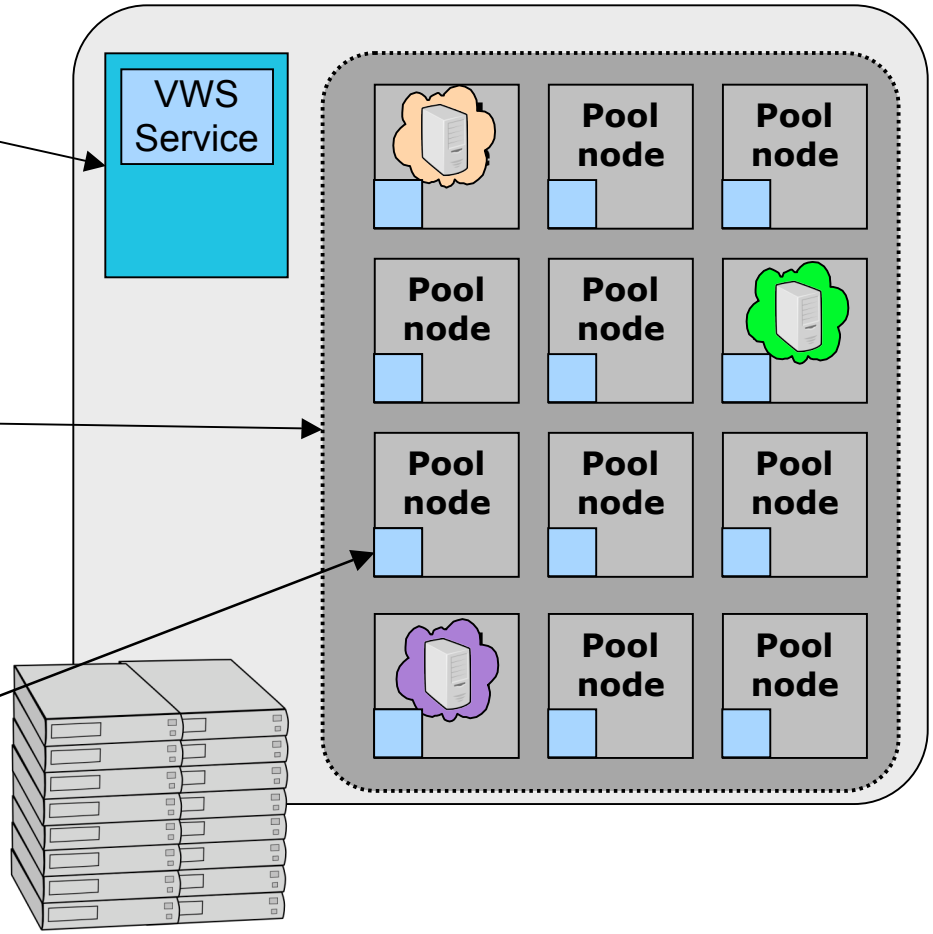
The Back Story

Workspace **WSRF front-end** that allows clients to deploy and manage virtual workspaces

Workspace back-end:

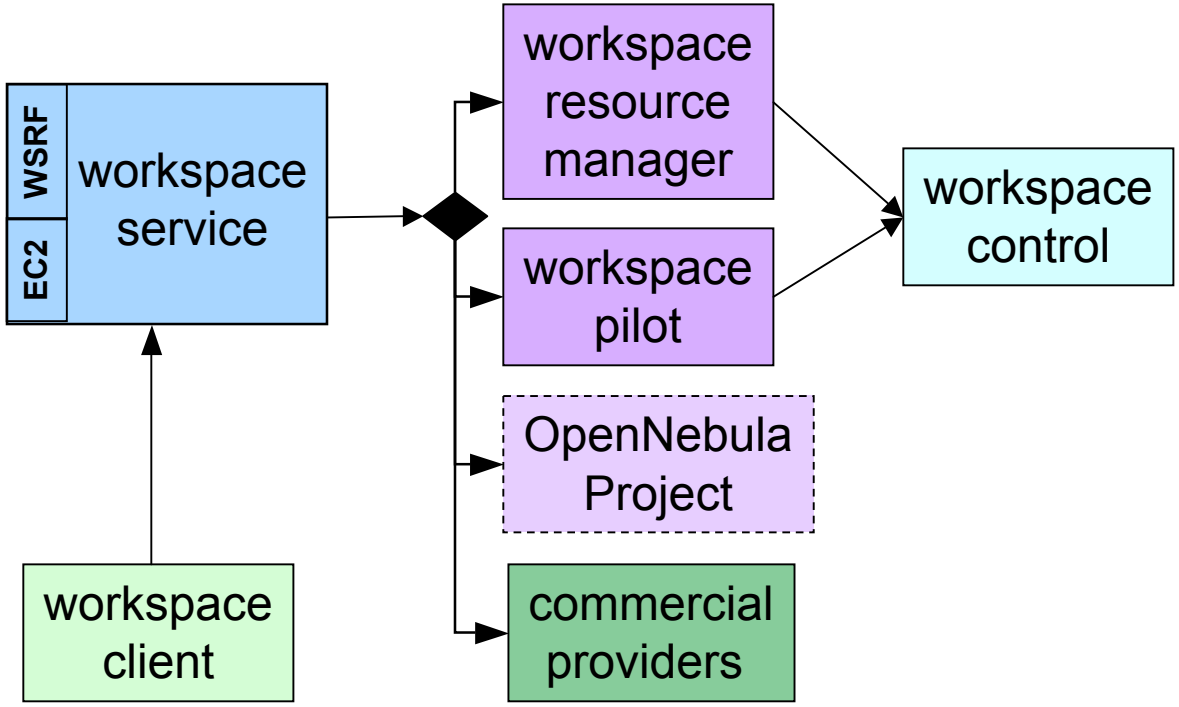
Resource manager for a pool of physical nodes Deploys and manages Workspaces on the nodes

Each node must have a VMM (Xen) installed, as well as the **workspace control** program that manages individual nodes



Trusted Computing Base (TCB)

Workspace Components



The Workspace Resource Manager

- Basic slot fitting
- Implements "immediate leases"
- Extensible vehicle to experiment with different leases
- Open source resource manager for multiple different VMMs
- Datacenter technology equivalent
  - Can be replaced by OpenNebula or other datacenter technologies
- Deployment
  - University of Chicago, University of Florida, Purdue, Masaryk University and all the other Science Cloud sites

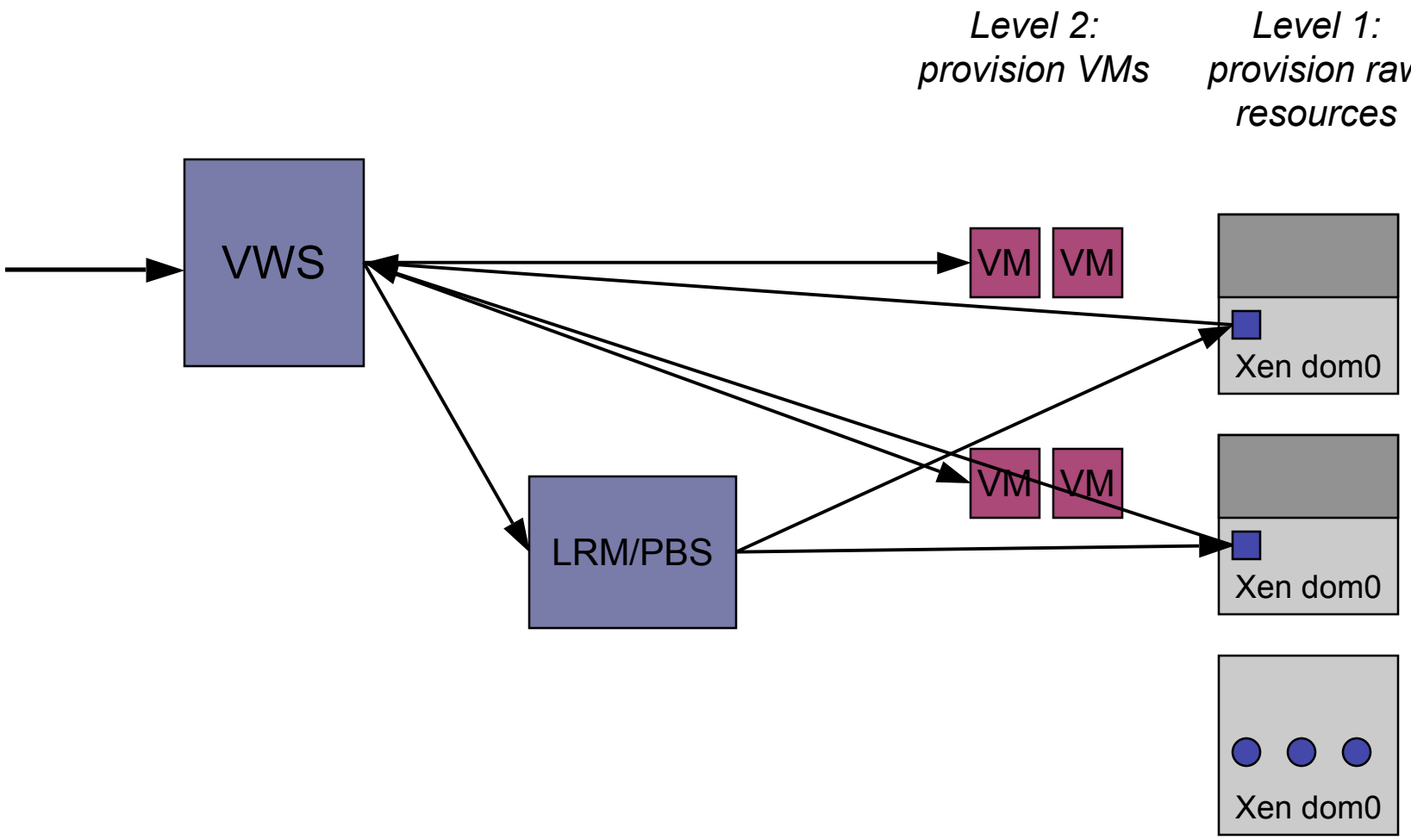
For research see papers: http://workspace.globus.org/papers/
"Combining Batch Execution and Leasing Using Virtual Machines", HPDC08

The Workspace Pilot

- Challenge: how can I provide a virtualization solution without significantly changing the current operation model of my cluster?
- The Workspace Pilot
  - Integrates with popular LRMs (such as PBS, SGE)
  - Glidein approach: submits a "pilot" program that claims a resource slot
  - Implements "best effort" leases
  - Handles signals
  - Includes administrator tools

Paper: "Simple Leases with Workspace Pilot", EuroPar08

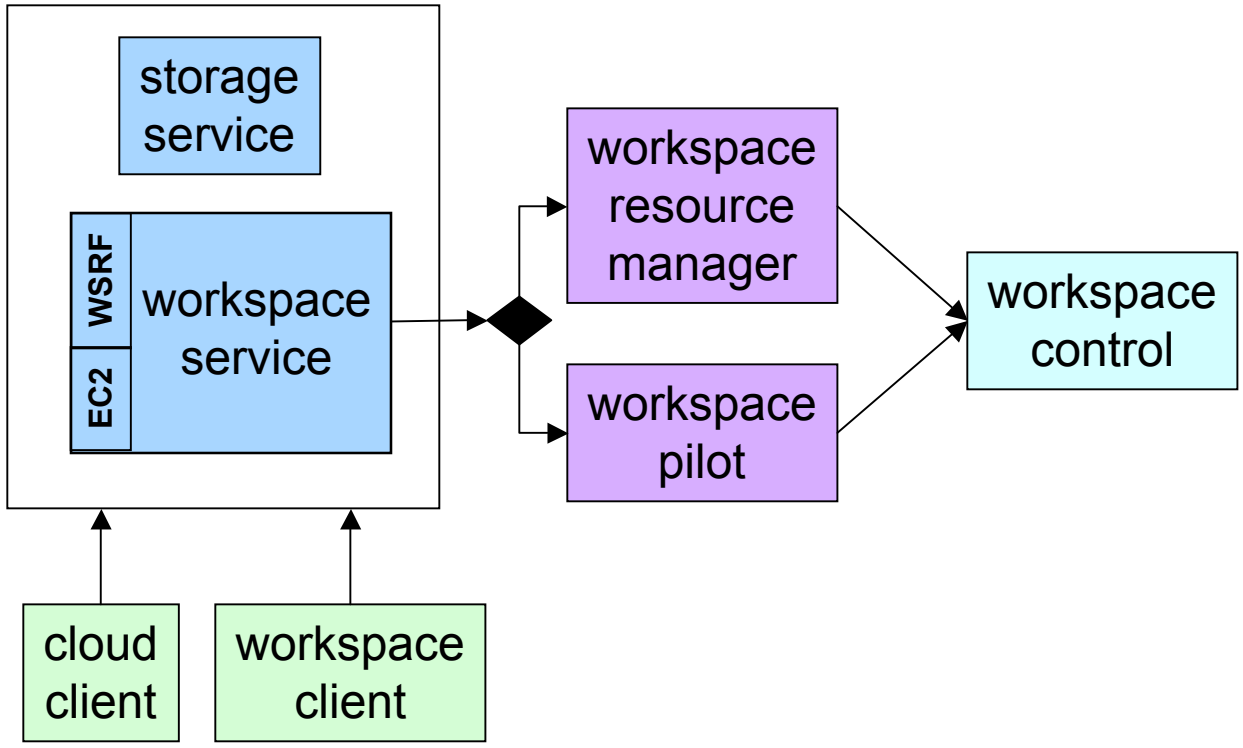
The Workspace Pilot



Workspace Control

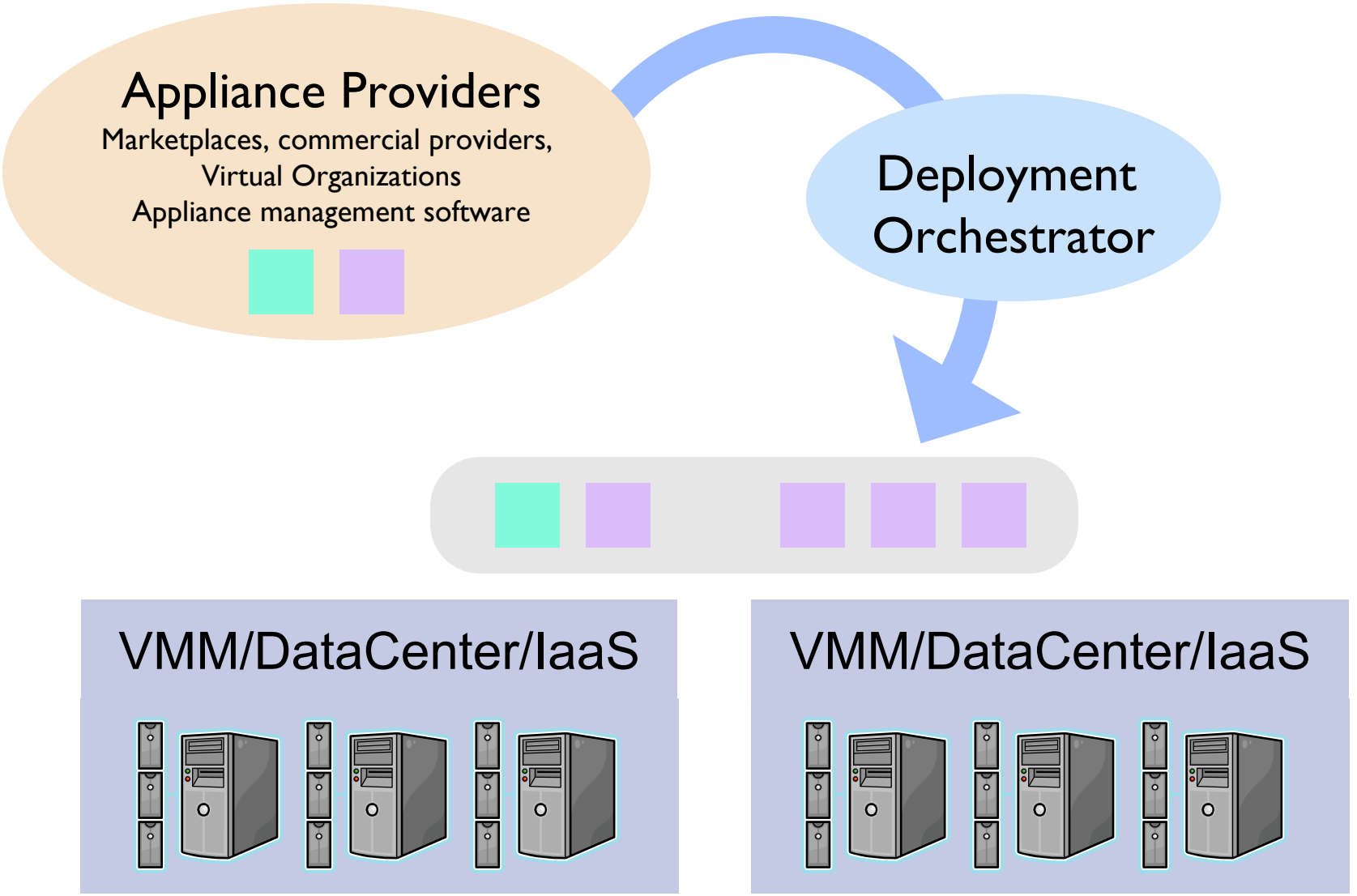
- VM image propagation
- Image management and reconstruction
  - Creating blank partitions, sharing partitions
- VM control
  - Starting, stopping, pausing, etc.
- Integrating a VM into the network
  - Assigning MAC addresses and IP addresses
  - DHCP delivery tool
  - Building up a trusted (non-spoofable) networking layer
- Contextualization information management
- Talks to the workspace service via ssh
- Standalone component
- Implementations in Xen and KVM (non-production)
- Moving towards a libvirt implementation

Cloud Capabilities

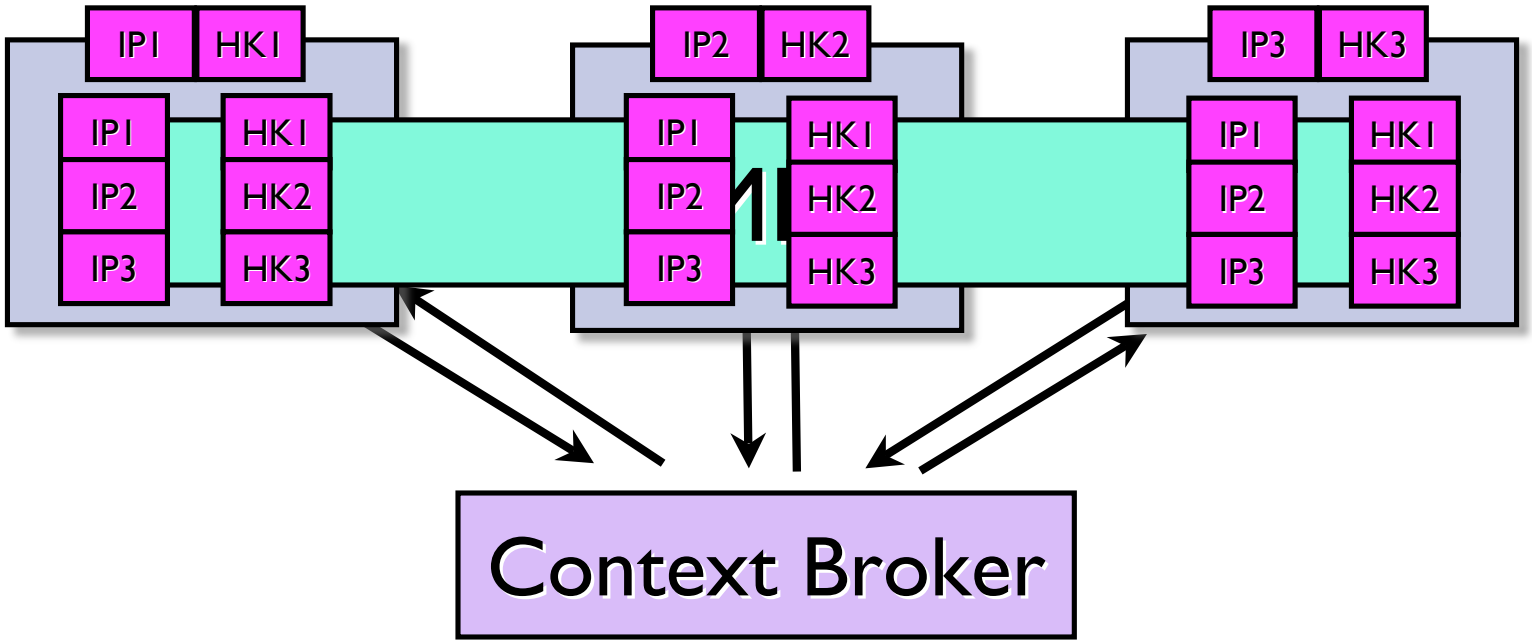


The Context Broker

Context Broker Background



Turnkey Virtual Clusters

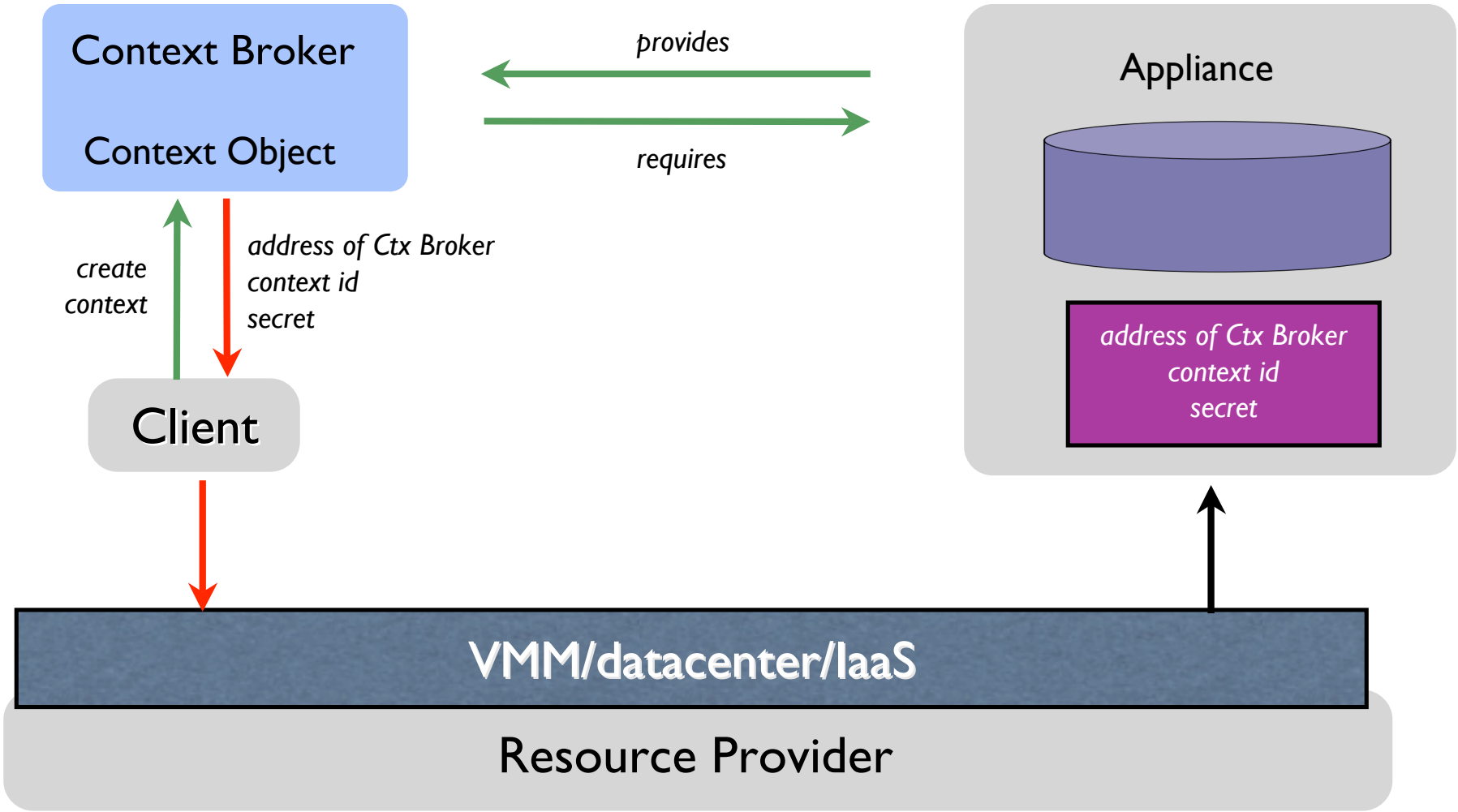


- Turnkey, tightly-coupled cluster
  - Shared trust/security context
  - Shared configuration/context information

Context Broker Goals

- Can work with every appliance
  - Appliance schema, can be implemented in terms of many configuration systems
- Can work with every cloud provider
  - Simple and minimal conditions on generic context delivery
- Can work across multiple cloud providers, in a distributed environment

Context Broker

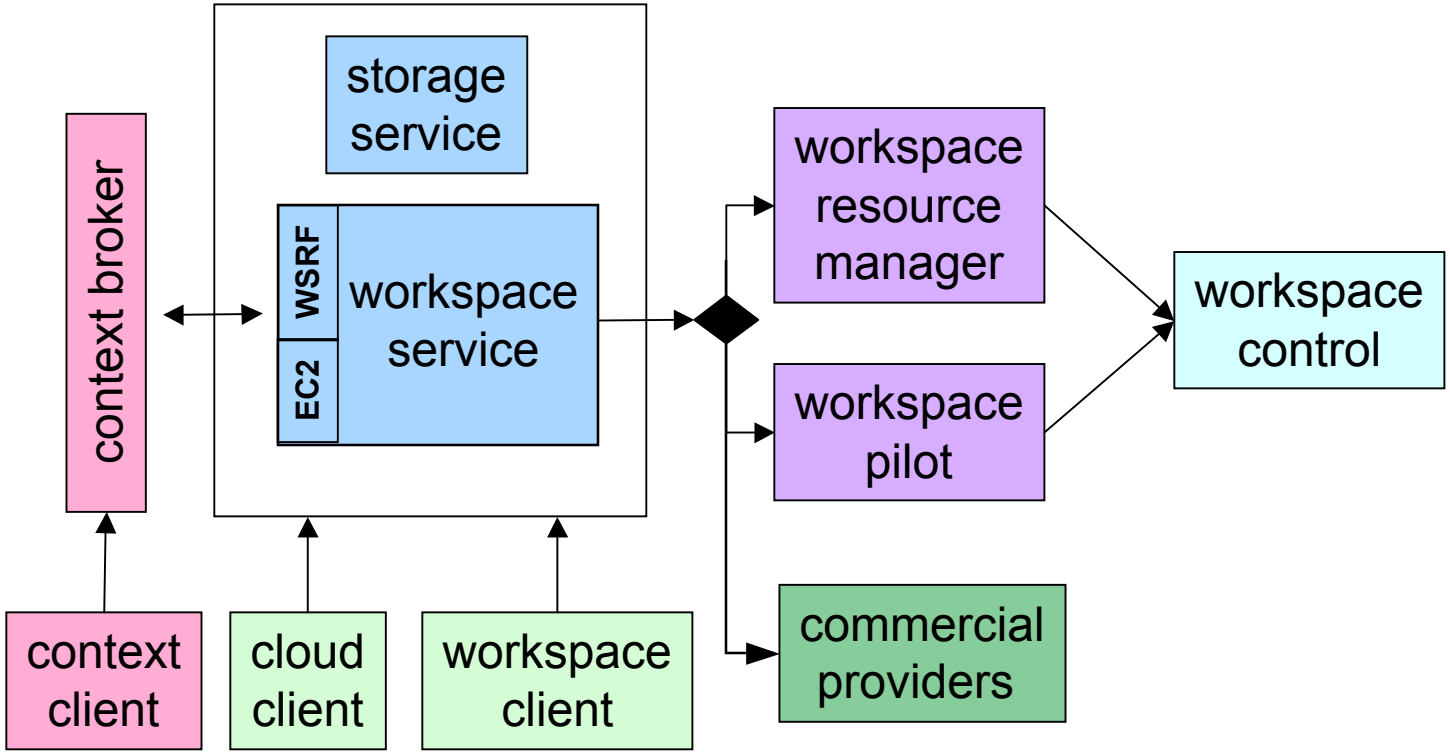


Context Broker Status

- Releases
  - In alpha since 08/07, first release 06/08, update 01/09
- Used to contextualize cluster composed of 100s of virtual nodes for multiple production apps
- Contextualized images on workspace marketplace
- Working with rPath to make contextualization easier for the user

Paper: Keahey&Freeman, Contextualization: Providing One-Click Virtual Clusters, eScience 2008

End of Nimbus Tour



Nimbus: Friends and Family

- Nimbus core team:
  - UC/ANL: Kate Keahey, Tim Freeman, David LaBissoniere
  - UVIC: Ian Gable & team
- Other efforts;
  - Cumulus: Raj Kettimuthu and John Bresnahan (ANL/UC)
  - EBS: Marlon Pierce, Xiaoming Gao, Mike Lowe (IU)
  - ViNe: Mauricio Tsugawa, Jose Fortes (UFL)
  - Others:
    - Descher et al (Technical U of Vienna): privacy extensions

Scientific Cloud Resources and Applications

Science Clouds

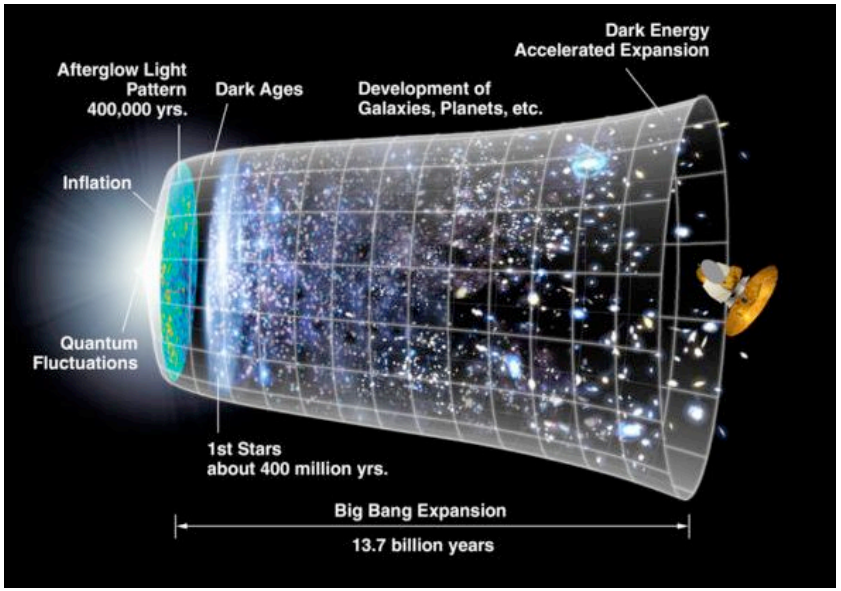
- Goals
  - Enable experimentation with IaaS
  - Evolve software in response to user needs
  - Exploration of cloud interoperability issues
- Available to all scientific projects
- Come and run:
  - http://workspace.globus.org/clouds

Science Clouds

- Participants
  - University of Chicago (since 03/08)
  - University of Florida (05/08, access via VPN)
  - Masaryk University, Brno, Czech Republic (08/08)
  - Wispy @ Purdue (09/08)
  - Other efforts in progress
  - Using EC2 for large runs
- Science Clouds Marketplace
  - OSG cluster, Hadoop cluster, test images etc.
- 100s of users, many diverse projects ranging across science, CS research, build&test, education, etc.

STAR experiment

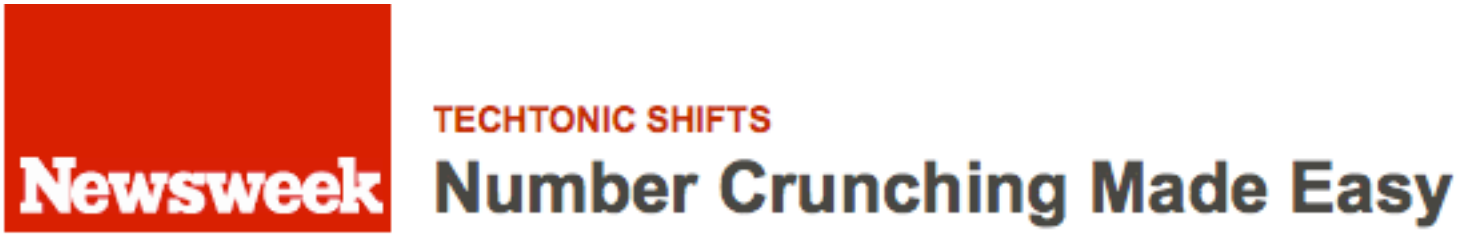
- STAR: a nuclear physics experiment at Brookhaven National Laboratory
- Studies fundamental properties of nuclear matter
- Problem: computations require complex and consistently configured environments that are hard to find in existing grids



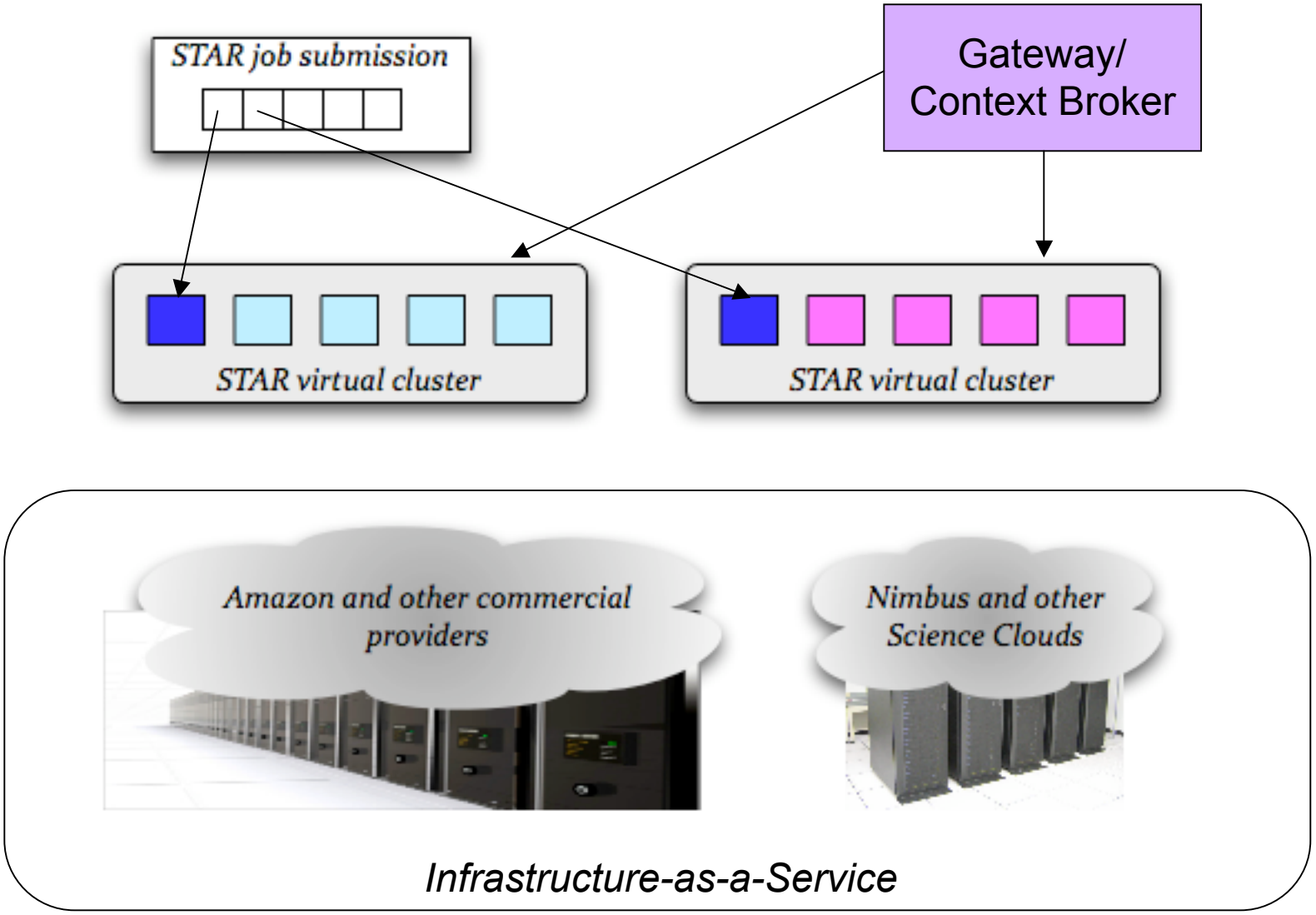
STAR Virtual Clusters

Work by Jerome Lauret, Leve Hajdu, Lidia Didenko (BNL), Doug Olson (LBNL)

- Virtual resources
  - A virtual OSG STAR cluster: OSG headnode (gridmapfiles, host certificates, NFS, Torque), worker nodes: SL4 + STAR
  - One-click virtual cluster deployment via Nimbus Context Broker
- From Science Clouds to EC2 runs
- Running production codes since 2007
- The Quark Matter run: producing just-in-time results for a conference: http://www.isgtw.org/?pid=1001735



STAR Quark Matter Run



Priceless?

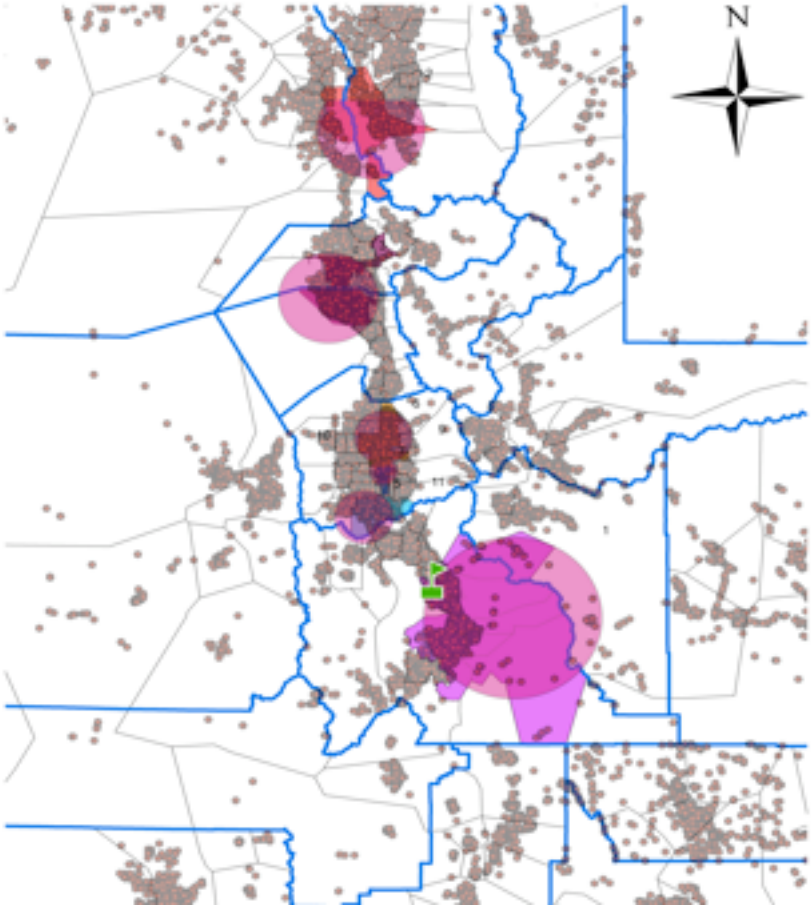
- Compute costs: $ 5,630.30
  - 300+ nodes over ~10 days,
  - Instances, 32-bit, 1.7 GB memory:
    - EC2 default: 1 EC2 CPU unit
    - High-CPU Medium Instances: 5 EC2 CPU units (2 cores)
  - ~36,000 compute hours total
- Data transfer costs: $ 136.38
  - Small I/O needs : moved <1TB of data over duration
- Storage costs: $ 4.69
  - Images only, all data transferred at run-time
- Producing the result before the deadline…

…$ 5,771.37

Modeling the Development of Epidemics

Work by Ron Price and others, University of Utah

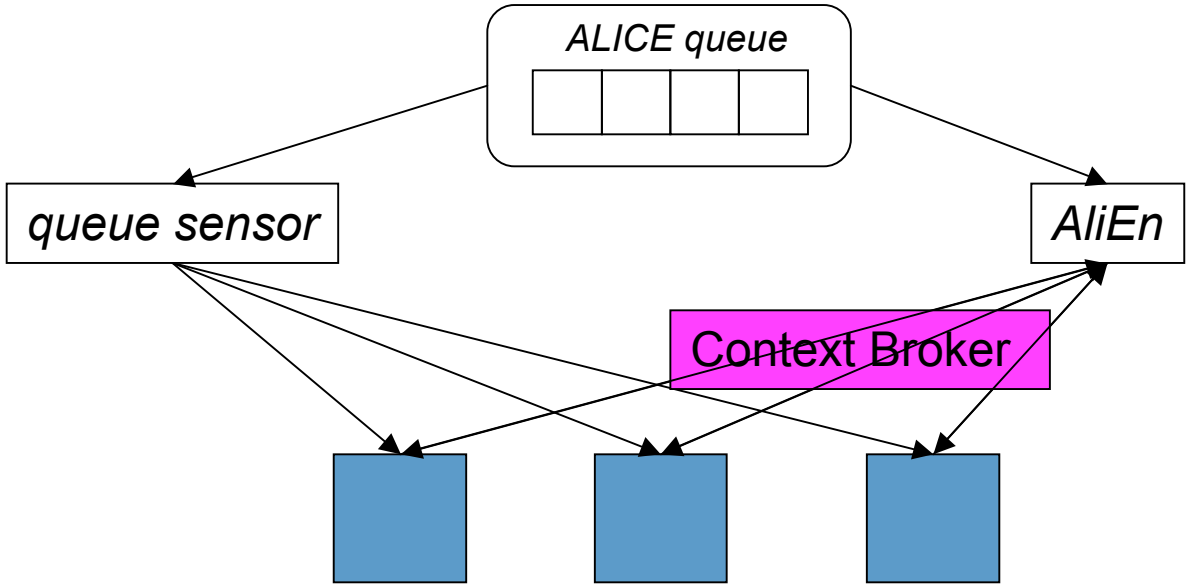
- Can we use clouds to acquire on-demand resources for modeling the progression of epidemics?
  - Monte-Carlo simulations
- What is the efficiency of simulations in the cloud?
  - Compare execution on:
    - a physical machine
    - 10 VMs on the cloud
    - The Nimbus cloud only
  - 2.5 hrs versus 17 minutes
  - Speedup = 8.81
  - 9 times faster

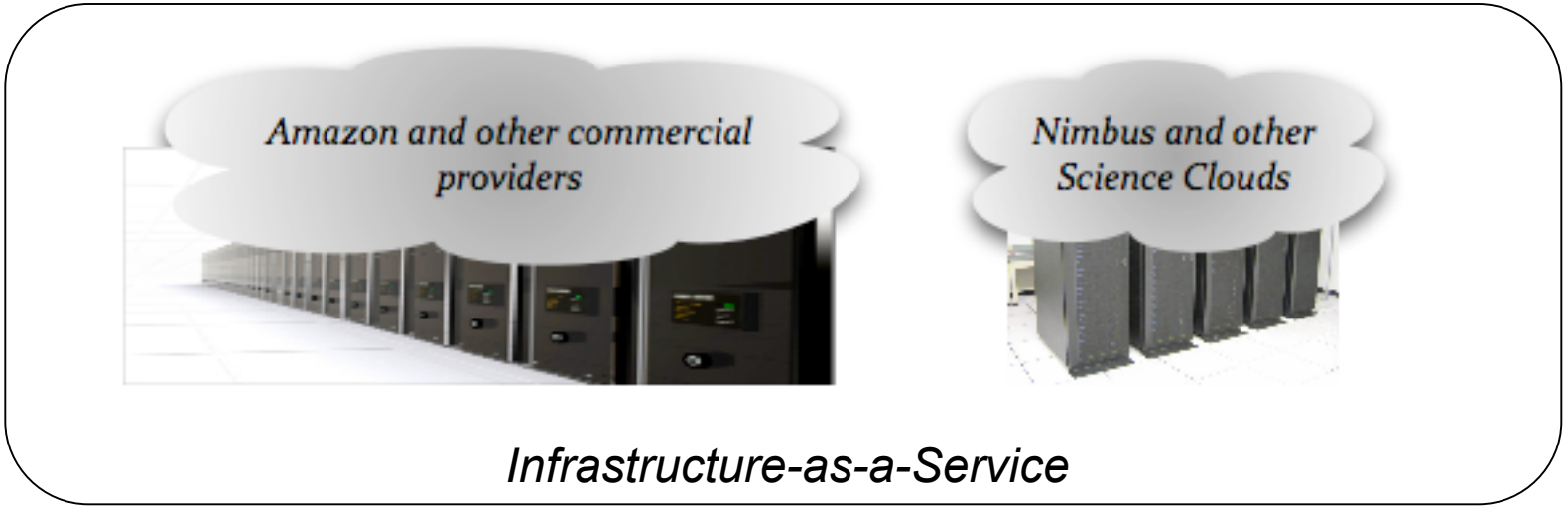


A Large Ion Collider Experiment (ALICE)

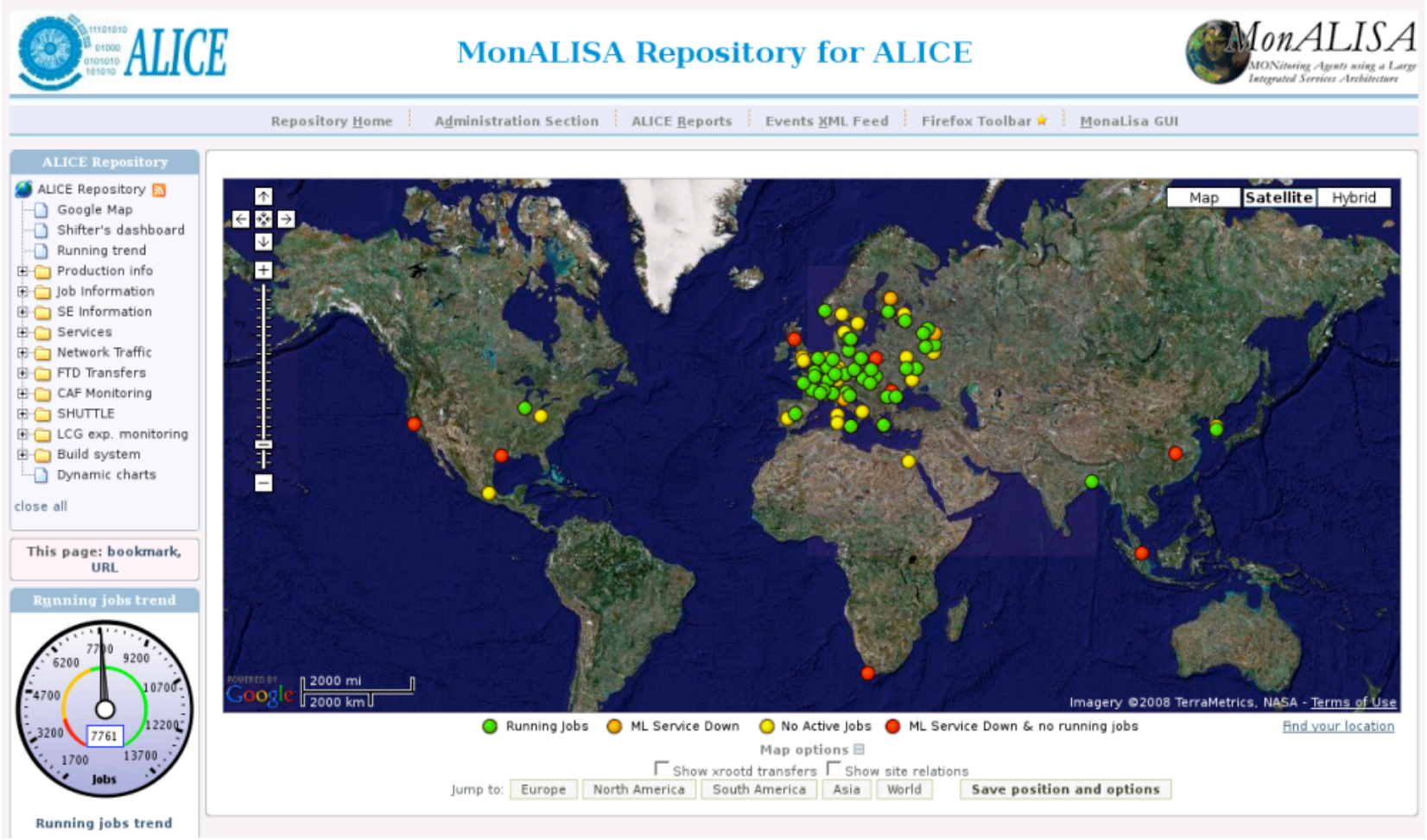
- Heavy ion simulations at CERN
- Problem: integrate elastic computing into current infrastructure
- Collaboration with CernVM project
- With Artem Harutyunyan and Predrag Buncic

Elastic Provisioning for ALICE HEP





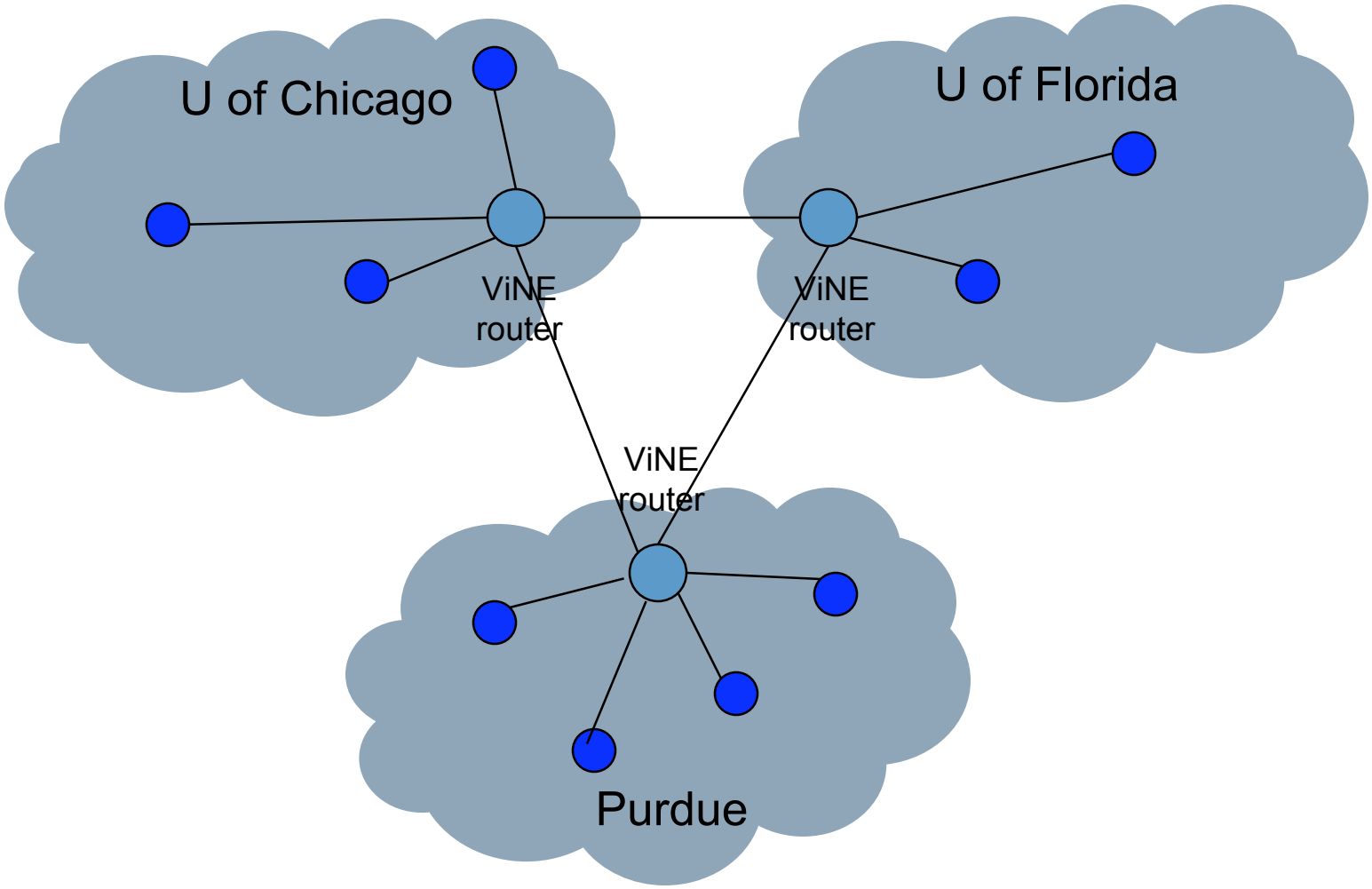
Elastically Provisioned Resources



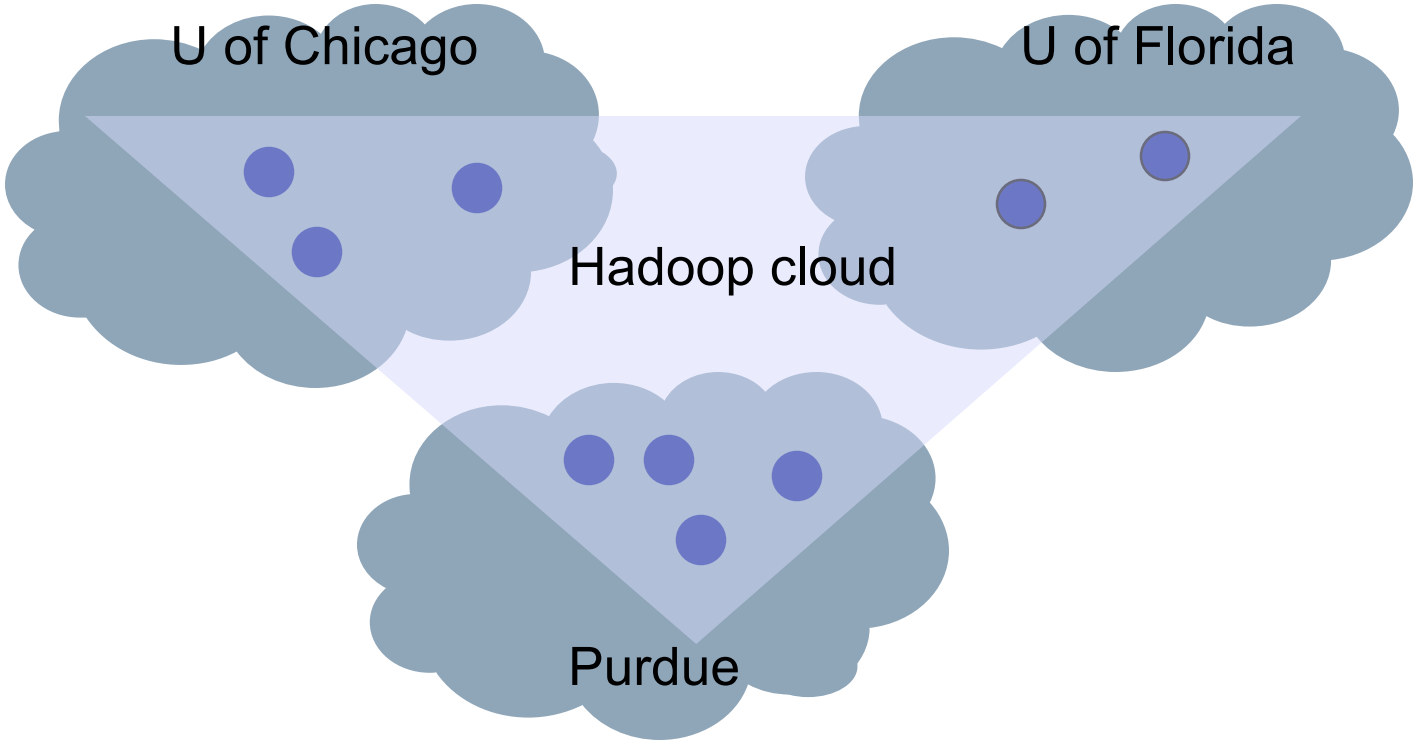
- *CHEP09 paper, Harutyunyan et al.*
- *Elastic resource base: ElasticSite, ATLAS, and others*

Sky Computing Environment

Work by A. Matsunaga, M. Tsugawa, University of Florida



Hadoop in the Science Clouds



- *Papers:*
  - *"CloudBLAST: Combining MapReduce and Virtualization on Distributed Resources for Bioinformatics Applications" by A. Matsunaga, M. Tsugawa and J. Fortes. eScience 2008.*
  - *"Sky Computing", by K. Keahey, A. Matsunaga, M. Tsugawa, J. Fortes, to appear in IEEE Internet Computing, September 2009*

Cloud Computing for Science: Issues and Challenges

Building the Ecosystem

- Configuring and maintaining appliances
  - Not just VMs, a variety of formats
  - CernVM, rBuilder (rPath)
- Licenses
  - Still vendor-specific approaches
- Getting used to dynamic sites
  - Host certificates and keys, community visibility, failure processing, etc.
- Infrastructure and leveraging

Security and Privacy Issues

- Security: new technology = new attacks
  - VMM issues: VM escape, drivers for smart NICs
  - Cloud infrastructure: IP spoofing?
  - Usage: is your VM up-to-date? are there any secrets on it? are there incentives to protect against attacks? Accepted "security" practices…
  - Attacks happen: e.g., VAServ
- Lack of features
  - Fine-grained authorization
  - *Paper: Palankar et al., Amazon S3 for Science Grids: a Viable Solution?*
- Data privacy
  - *Paper: Descher et al., Retaining Data Control in Infrastructure Clouds, ARES (the International Dependability Conference), 2009.*

Performance

- Difficult to track in a virtualized environment
  - I/O can be an issue
  - Tradeoffs between CPU power and throughput
  - Paravirtualized drivers
- Studies of cloud performance
  - *E.g., Walker, Benchmarking Amazon EC2 for high-performance scientific computing*
  - Low bandwidth from existing providers:
    - On the order of: 2-5 MB/sec, 17/21 MB/sec, 30MB/sec
  - Generally speaking, the existing cloud providers do not offer a very high-end computer… yet

Price

- Price for what?
  - Experimenting with business models
  - Estimating the cost is hard
- Price of Base Services for AWS:
  - Computation / EC2
    - On-demand: starting at $0.1 per hour
    - Reserved: starting at $227.50 per year for $0.03 per hour
  - Data / S3
    - Storage: $0.15 per GB/month,
    - Transfer: $0.17 per GB
    - AWS import/export for bulk
- Hosting Scientific datasets for free
  - Free on AWS for frequently used datasets

Service Levels

- Service levels
  - Computation: immediate, advance reservations, best-effort, periodic
  - Data: durability, high/low availability, access performance
  - Cross-cutting concern: security and privacy
- Different price points for different availability

Parting Thoughts

- IaaS cloud computing is science-motivated
  - Scientific applications are successfully using the existing infrastructure for production runs
- We are just at the beginning of the cloud revolution
  - *"Even though we keep rolling out new services and features, and several existing AWS services are already very successful, this is still Day One." (W. Vogels)*
- Project for the next few years: solve them!