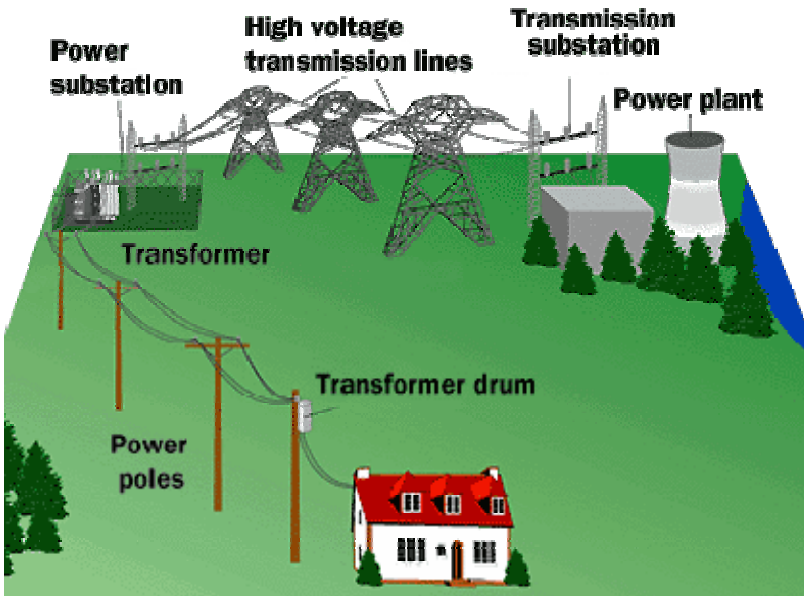
Working Spaces: Virtual Machines in the Grid

Kate Keahey
keahey@mcs.anl.gov
Argonne National Laboratory
Tim Freeman, Frank Siebenlist
{tfreeman,franks}@mcs.anl.gov

Towards Realizing the Grid Vision



- Quality of Service
  - Dynamically controlled enforcement of various qualities
  - Not just per-process enforcement
- Quality of Life
  - Being able to find the right configuration on the Grid

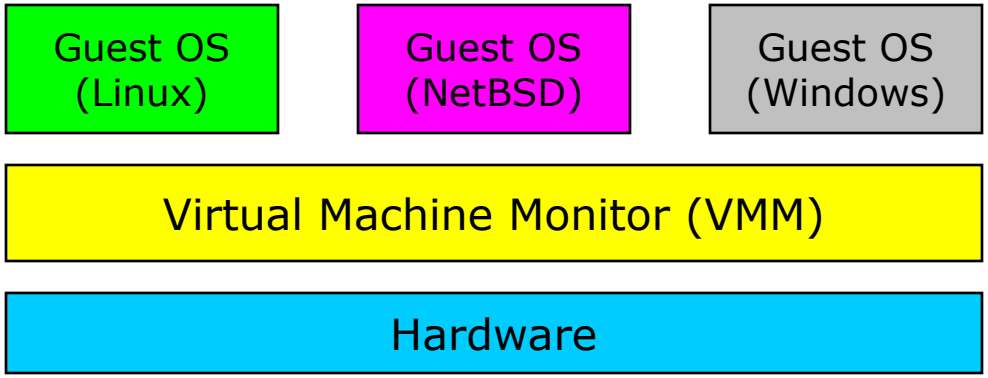
We Need a Workspace!

- A configurable execution environment, container
  - Good isolation properties
  - Good enforcement potential
  - Customizable software configuration
    - Library signature, OS, maybe even 64/32-bit architectures
- We need to be able to create, manage and deploy it
  - We need to be able to negotiate/renegotiate an environment shape with a VO
  - A broker would then be able to negotiate resource allocations and map these workspaces onto

Virtual Machines (VMs)

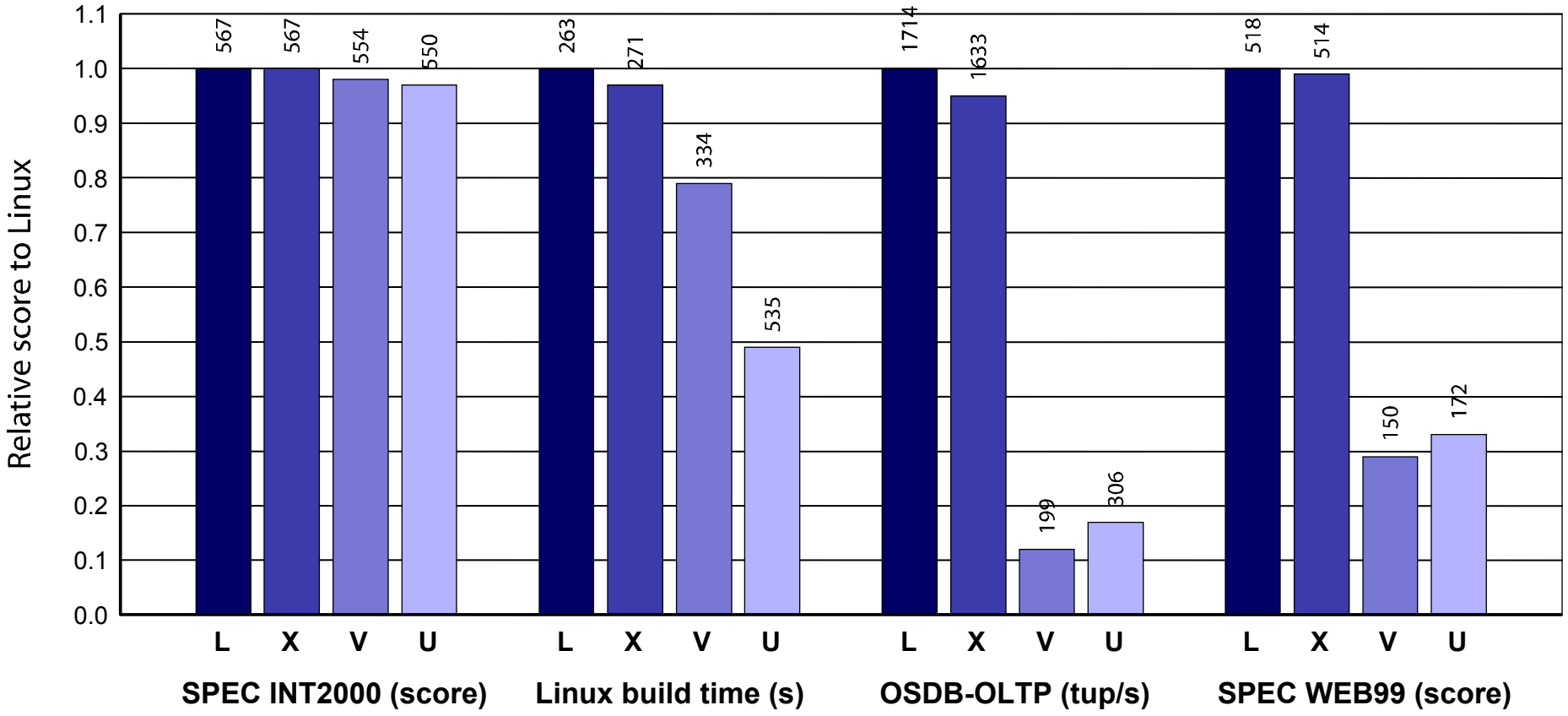
- VMs happen to have:
  - Good isolation properties
    - Generally enhanced security, audit forensics
  - Good enforcement potential
  - Customizable software configuration
    - Library signature, OS, maybe even 64/32-bit architectures
  - Serialization property
    - VM images (include RAM), can be copied
  - The ability to pause and resume computations
    - Allow migration
- Common concern:
  - Overhead: application, startup, resource usage

Virtual Machines Primer



- Different types of virtual machines
  - Full virtualization (VMware)
    - Run multiple unmodified guest OSs
  - Para-virtualization (Xen, UML, Denali)
    - Run multiple guest OSs ported to a special architecture
  - Single OS image (Vserver)
- Paper: "*From Sandbox to Playground: Dynamic Virtual Environments in the Grid*", Grid 2004

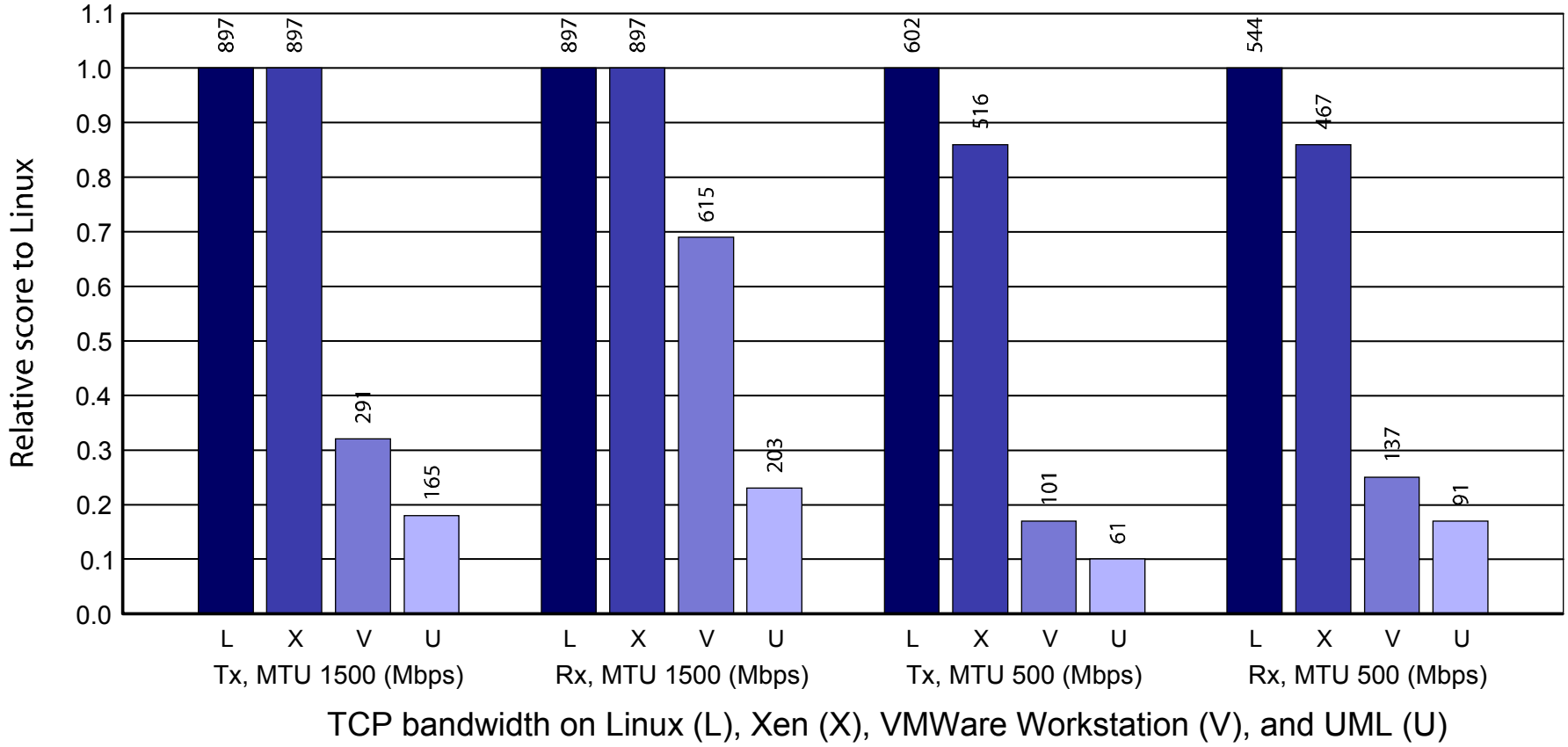
The Need for Speed



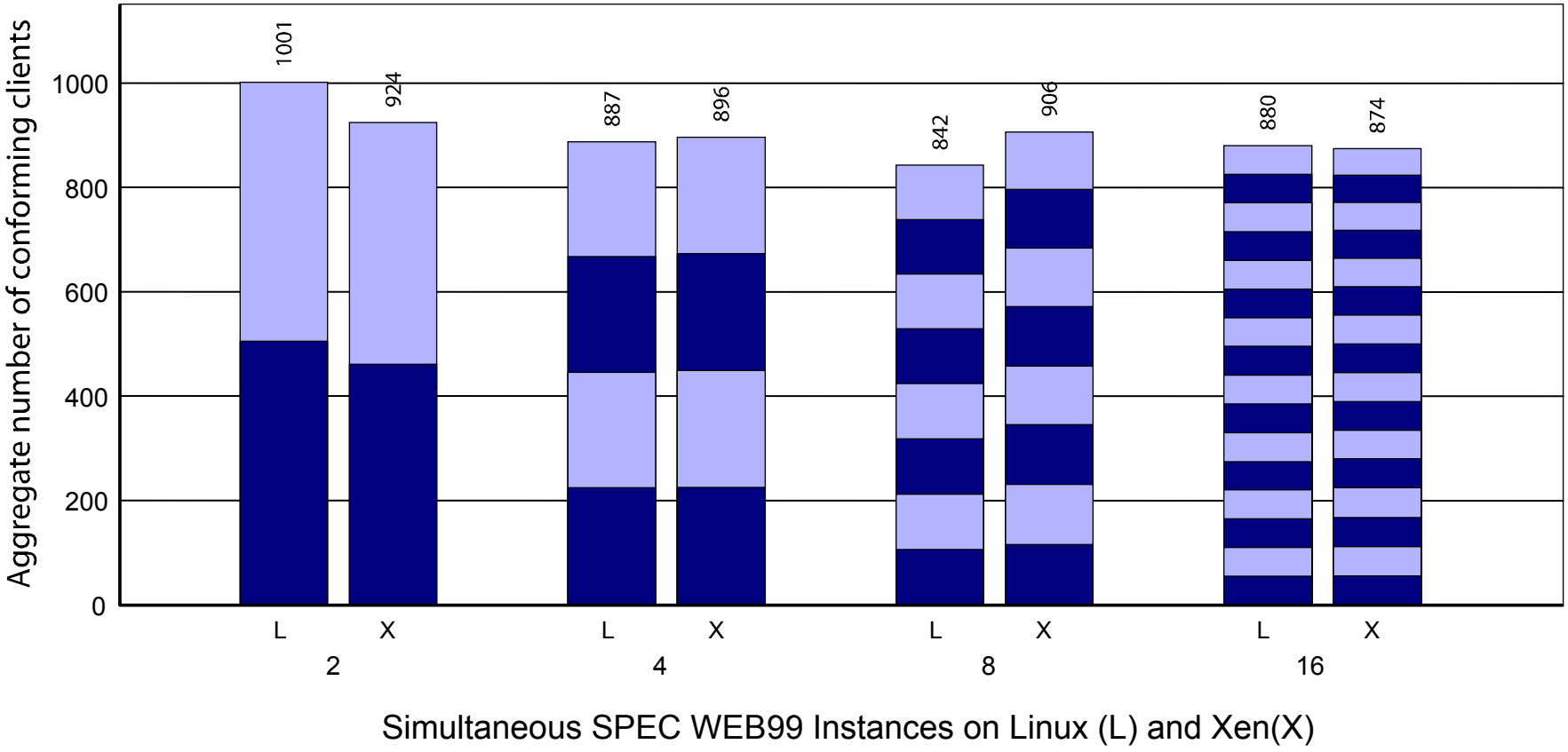
Benchmark suite running on Linux (L), Xen (X), VMware Workstation (V), and UML (U)

Paper: "Xen and the Art of Virtualization", SOSP 2003

TCP results


Relative score to Linux
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
897
897
291
165
897
897
615
203
602
516
101
61
544
467
137
91
L
X
V
U
Tx, MTU 1500 (Mbps)
Rx, MTU 1500 (Mbps)
Tx, MTU 500 (Mbps)
Rx, MTU 500 (Mbps)
TCP bandwidth on Linux (L), Xen (X), VMWare Workstation (V), and UML (U)

Scalability



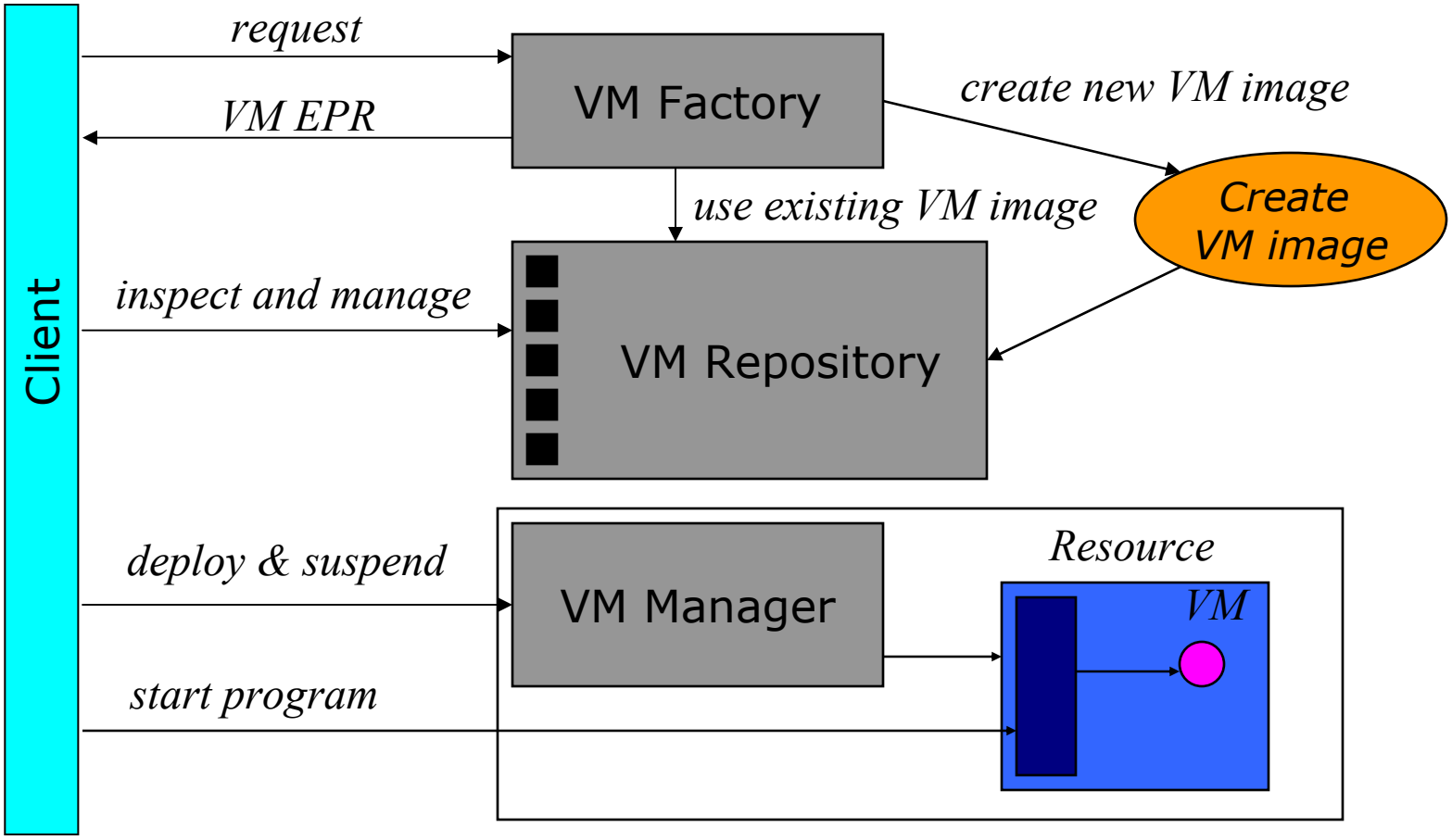
Other Concerns

- License
  - Open source (Xen, UML)
    - Visible effects of open source community work
  - Commercial (VMware)
    - Also, XenSource
- Distribution/Installation
  - Para-virtualization requires kernel modifications
    - Yes, but … everything else stays the same
    - Xen is soon to be part of the Linux kernel, Fedora Core 4 (May '05), is in Debian unstable, unofficial: Gentoo, Mandrake and SUSE distributions
  - Privilege
    - Xen (root, patch kernel, domain 0 privileges setup)
    - VMware Workstation (root, installation only)

New Technology, New Challenges

- How can we leverage the benefits of VMs in Grid technology?
- How efficient will be this new technology?
- How can we ensure a secure environment under these new assumptions?
- How well will it all work in practice?
- What new scenarios will this enable?
- How will it change Grid computing?
- What new problems will it create?

Integrating VMs into the Grid Architecture



Supporting Services

- Factory
  - Creates VM images
    - Eventually it may have to support negotiation
  - Images may be created based on an already existing image
- VM Repository
  - Access to state describing a VM
  - Allows inspection, management, termination, potentially renegotiation, etc.
- VM Manager
  - Service deploying VMs on nodes
  - Operations: stage, start, pause, stop, checkin
- Once deployed, jobs may be executed in the virtual machines in a variety of ways

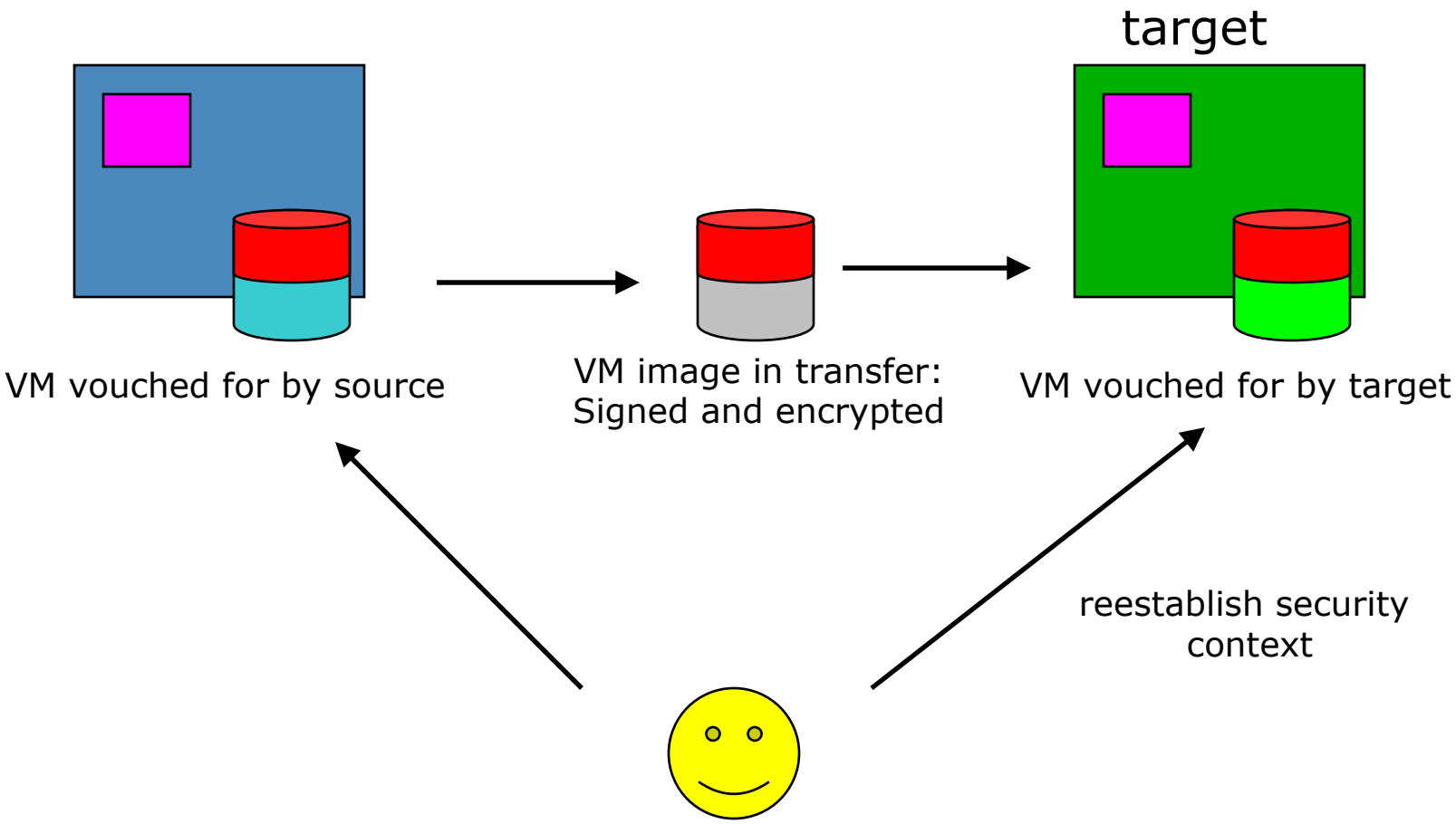
Workspace Structure

- Elements of a workspace
  - Workspace description (meta-data)
    - EPR/name, category, state, etc., also hardware, network, software configuration, etc.
  - Workspace implementation (VM image)
- Workspace "types"
  - Workspaces conforming to set configurations
  - Provenance of VM images
- Workspace instances
  - Workspace meta-data contains a name
  - Instance equality
  - Copying operation: copy image, meta-data, create a new name
  - Signing instances

Security: New Opportunities, New Problems

- VMs introduce a new layer of trust
  - VM monitor needs to be trusted as a technology
  - Trusted computing
- VMs can be serialized and transferred as data
  - The integrity of a VM image needs to be protected (signing)
  - Private information on a VM image need to be protected (encryption)
    - VM private key: should a VM be able to assert its identity?
    - Application private keys, security context
- VMs can be migrated (source --- > target)
  - Trust management: a VM image may be moved between parties that don't trust each other
    - A popular problem
  - Key management: target VM needs to verify the integrity and identity of a VM image in ways acceptable to the client: key management
  - Security context has to be preserved or renegotiated

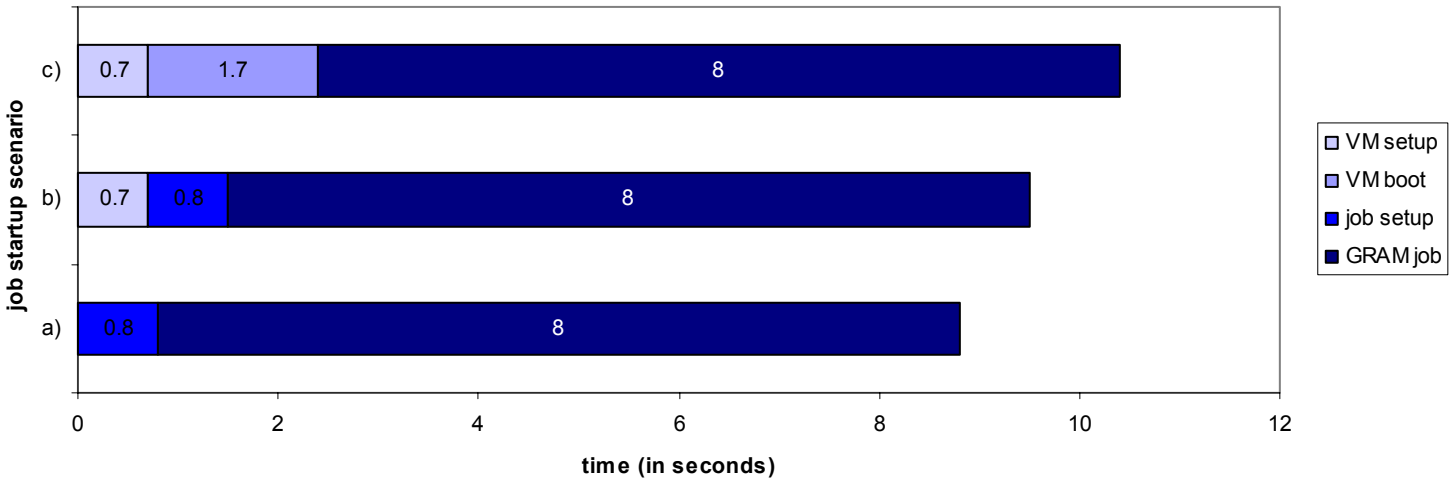
Migrating Securely



Security: New Twist on Old Problems

- Deployment problem
  - Protecting the VM from the world
    - VMs are only as secure as the software they run
    - Who maintains all those VMs? Local administrators would have to maintain too many images yet need to protect against vulnerabilities
  - Protecting the world from the VM
    - One could use one's privileges as root on a VM (for example to generate harmful network traffic)
    - Although audit works great by the time the damage is done it is too late!
- Deployment solutions
  - VO certification and authorization
    - Certification Authority to certify VM image
    - Site policies may require VMs to conform to site policies
  - Detection: Intrusion Detection Systems
  - Actions
    - Restricting network traffic: putting the good guys in jail
    - Right to take action

Job Startup



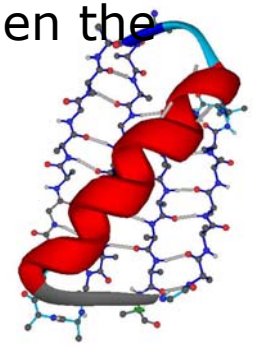
- a) job startup using a pre-configured job
- b) job startup using an unpaused VM
- c) job startup using GRAM

Things We Didn't Talk About

- Security
  - Processing of encryption and signing
- Moving images
  - Image size: 1MB, more typically 200 MB, and upwards
  - Image diff (Rosenblum)
  - Proximity of image as matching criterion (VMPlant)
  - Mounting partitions (general on-site assembly)
- Scalability
  - Distribute processes among existing VMs
  - Lightweight VMs: Denali
- Clusters
  - Currently in progress: work on virtual cluster, collaboration with the COD team at Duke

How does it work in practice?

- Recent project: combining VMs and Grids to create a platform for bioinformatics applications
- Some of the conclusions:
  - Use of virtual machines can significantly broaden the resource base
  - Saves installation time
    - EMBOSS installation: ~45 minutes
    - Deploying a 2GB VM image: ~6.5 minutes
    - Peace of mind: priceless!
  - Enforcement capabilities
    - Depend on the implementation but are generally better than what we have now



- SC04 poster:
  - *"Quality of Life in the Grids: VMs meet Bioinformatics Applications", D. Galron et al*

Virtual Playgrounds

- Define a virtual Grid in terms of requirements
  - Virtual workspaces
  - Networking requirements, virtual network
  - Storage and other requirements
- Provide mechanisms to create a Grid
- Provide services for the deployment of such "virtual playgrounds" on real resources
- Ephemeral Grids built for a special purpose:
  - Scientists getting up a Grid for the purposes of a specific experiment run
  - A scientific simulation that gets discarded or interrupted but can potentially be restored later

Conclusions

- For Grids to scale we need a way to create and manage remote environments in the dynamically and effortlessly
- Virtual is the new Real!
  - VMs present a very compelling solution
    - Efficiency, flexibility, migration, etc.
  - …and introduce some new challenges
    - New services, different models of sharing, security, etc.
- Watch out for emergent properties and behaviors!
- We need to work with the VM community to fine-tune requirements and features
  - Open Source helps!
- A growing role for Virtual Organizations
- Policy, Policy, Policy…
  - Policy of resource owners, VOs, users…
- Closer to the dream on seamlessly negotiating, provisioning, renegotiating…
- Status: a GT4 prototype that we keep evolving!

Related Efforts

- Condor
  - ClassAds, glide-ins
- In-Vigo (VMPlant)
- Virtuoso (VNET)
- VIOLIN (UML + private networks)
- Cluster on Demand
- Workspaces Project
  - www.mcs.anl.gov/workspace

Credits

- Actively working:
  - Tim Freeman
  - Frank Siebenlist
  - Xuehai Zhang
- Past contributions:
  - Karl Doering (UCSD)
  - Daniel Galron (OSU)