Workspaces for CE Management

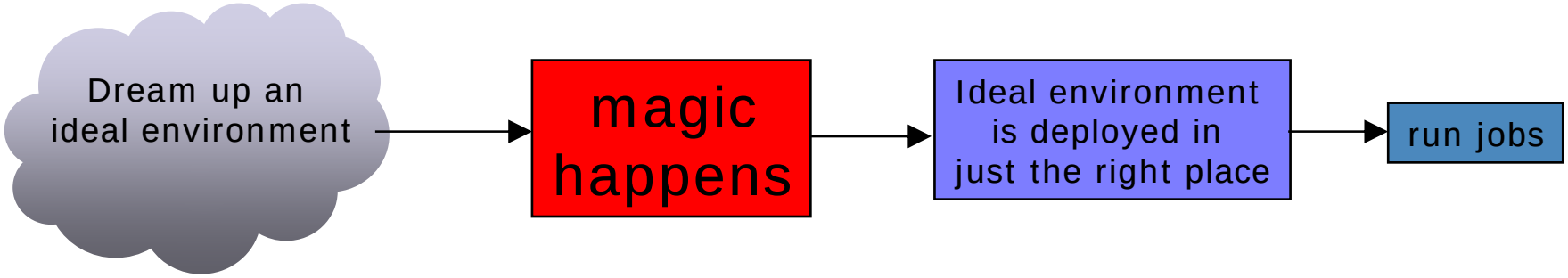
Kate Keahey

keahey@mcs.anl.gov

Argonne National Laboratory

Why Workspaces?

- We need to be able to dynamically create an execution environment on remote resources

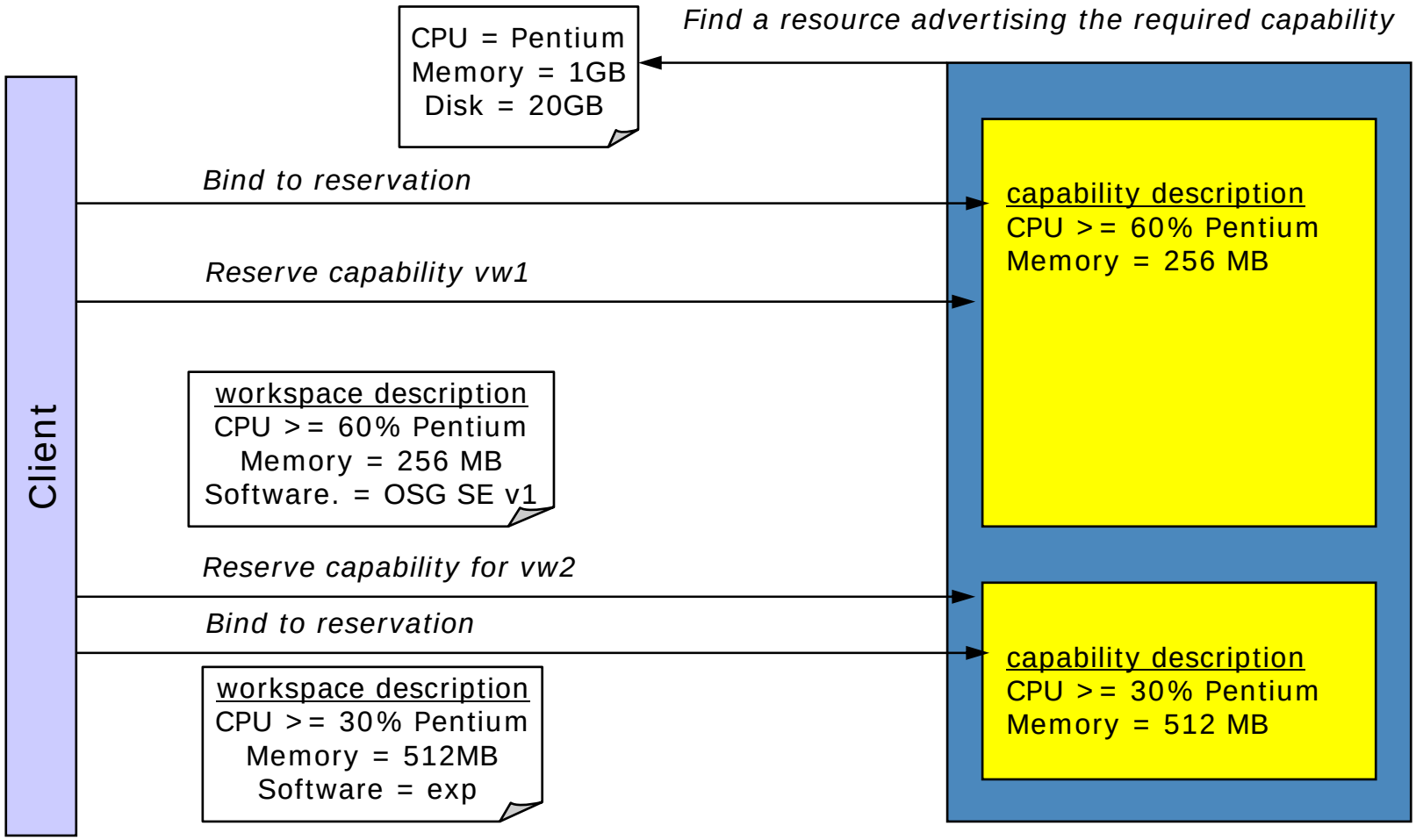


- The aspects of workspaces:
  - Quality of Service: isolation and enforcement
  - Quality of Life: providing the right configuration at the right time

What are Virtual Workspaces?

- A description of an execution environment
  - Software configuration requirements
    - OSG worker node, submit node for a Grid3 cluster
  - Resource allocation requirements
    - Use exactly X memory, at least Y disk space, Z bandwidth…
  - Sharing and isolation properties
    - Unix account, sandbox, various kinds of virtual machines …
  - And others…
- Basic workspace example : a Unix account on a remote machine
- Workspace can be managed and refined
  - In terms of lifetime, meta-data, access policies…
- A workspace can be deployed on a resource
- Jobs can be deployed in a workspace
- A workspace can have various implementations
  - Dynamic accounts & configuration tools
    - Pacman, SoftEnv, Softricity
  - Virtual Machines

Binding Wrokspaces to Resources

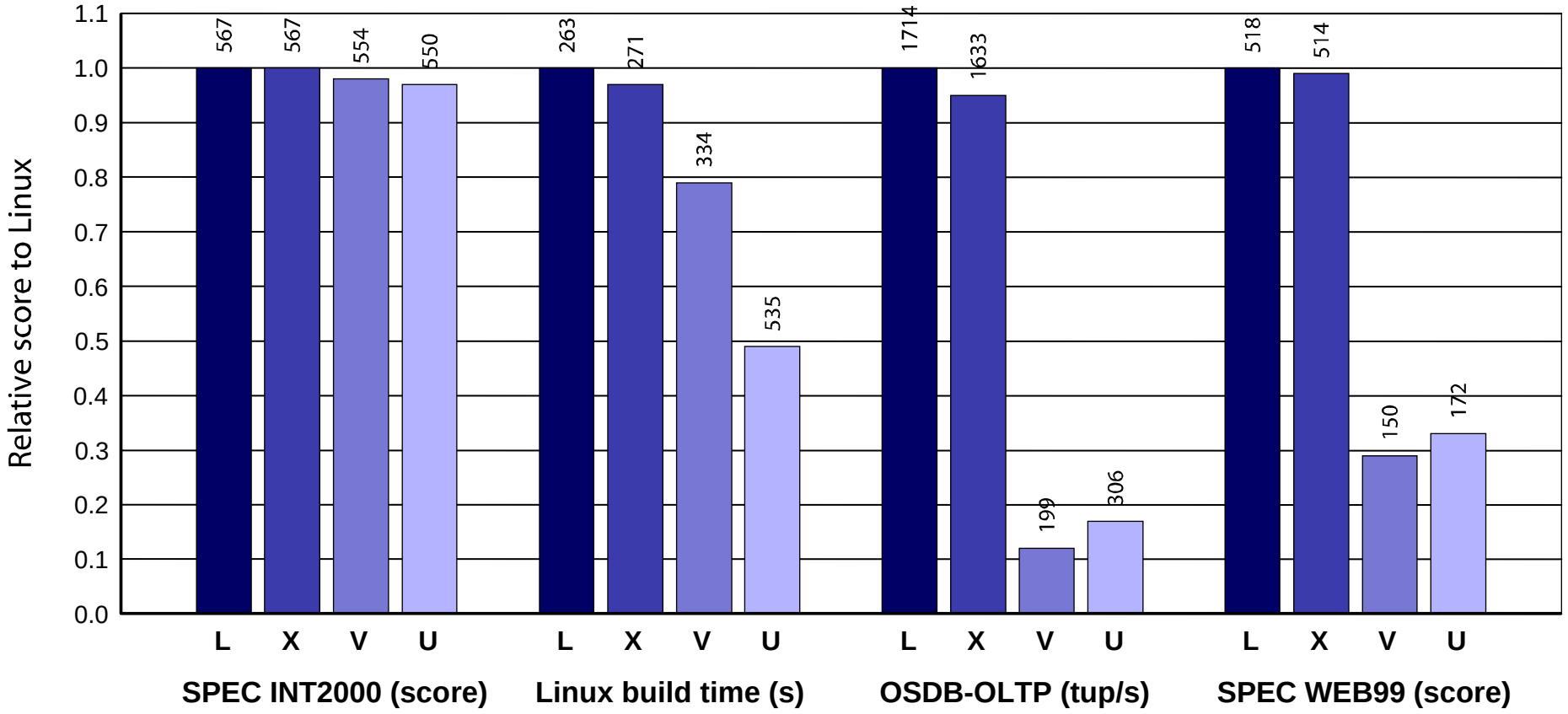


VW Implemenations: Virtual Machines

- Advantages
  - Customizable software configuration
    - Library signature, OS, 64/32-bit architectures
  - Excellent enforcement potential
    - Enforcement on a sandbox rather than process level
  - Excellent isolation
    - Generally enhanced security, audit forensics
  - Pausing, serialization, and migration
    - VM images (include RAM), can be copied
- Available implementations
  - Commercial (VMware)
  - Open source (Xen, UML)
    - Also support for Xen from XenSource and many Linux distributors
- Xen is rapidly emerging as the most popular implementation
  - The fastest, freeest, the most open source, the most backed…

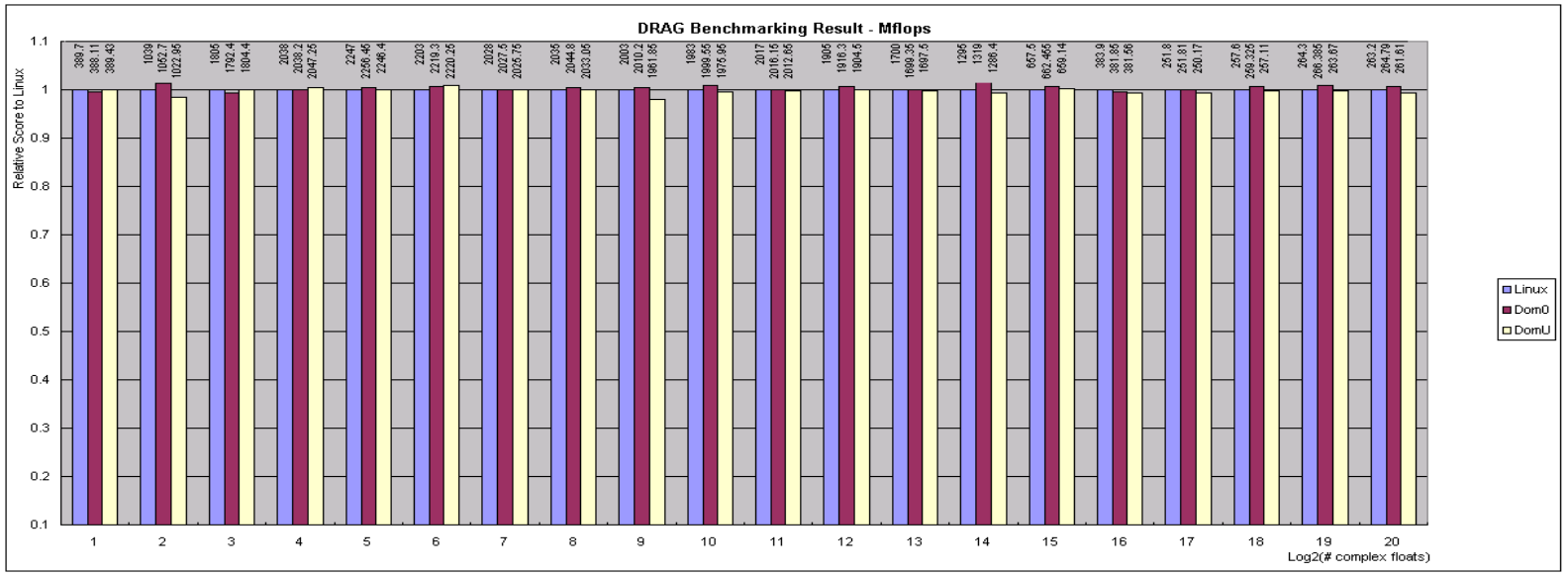
the globus alliance
www.globus.org

The Need for Speed



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

DRAG Benchmark Results



- DRAG suite: FFT-based benchmark
- Comparison (by Xuehai Zhang, UC):
  - Linux: machine runs native 2.6 Linux.
  - Dom0: machine runs Xen and domain 0.
  - DomU: machine runs Xen, domain 0 and a user domain.
- Similar performance as native Linux
  - <3% degradation, but sometimes actually better than native Linux
- More details at http://people.cs.uchicago.edu/~hai/vm1/drag/.

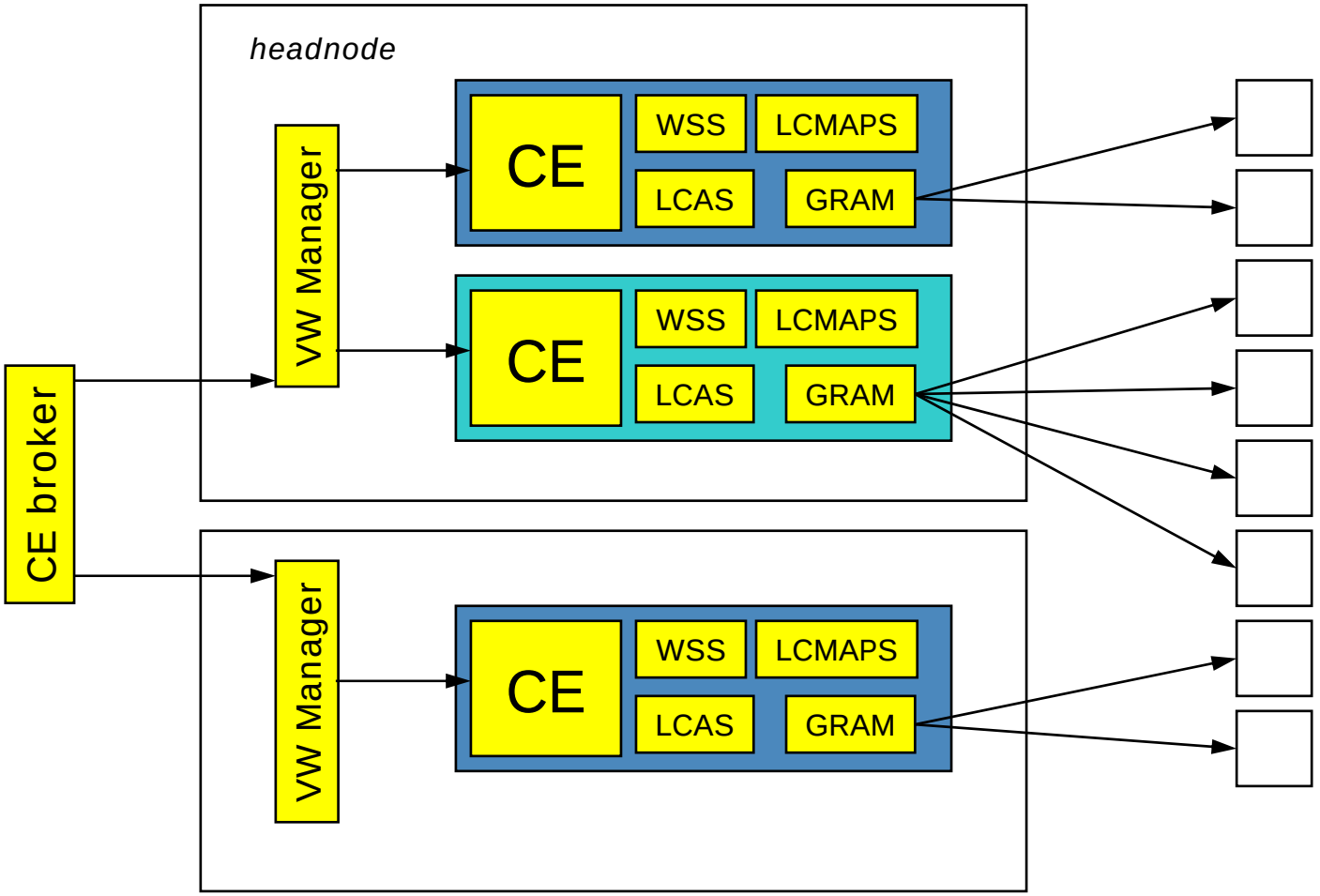
Deployment Concerns

- Distribution/Installation
  - Para-virtualization (Xen) requires kernel modifications
    - Yes, but … everything else stays the same
    - Work in progress on making Xen part of Linux kernel
      - False information of its conclusion seen recently!
    - Support from many Linux distributors: Fedora, Debian, SUSE, Gentoo, Mandrake, etc.
- Privilege level(Xen)
  - Domain0 is a privileged domain, not a good environment for sharing.
  - If Xen configuration is going to be permanent using DomainU is recommended
    - Performance impact needs to be considered

The Xen of Enforcement

- CPU
  - Schedulers: BVT, FBVT, Round Robin, Atropos/SEDF
  - May be selected at boot time; BVT is default
  - Borrowed Virtual Time (BVT)
    - Fair share of CPU based on weights assigned to the domains
    - Work-conserving
  - Simple Earliest Deadline First (SEDF)
    - Reserves absolute shares of CPU for domains
- Memory
  - Memory size specified in a configuration file
  - Can be readjusted from domain0
- Disk
  - Export partitions
  - Logical Volume Manager (LVM) allows to grow and shrink the disk size
- Networking
  - Standard Linux deployment tools: Domain0 can do traffic shaping for user domains.

Workspace as a CE Environment



Pros and Cons

- Problems that VMs solve for us
  - Environment management
    - Configuration management
      - Running two different versions of CE software side-by-side
    - Enforcement and isolation
  - Graceful load management
    - Renegotiating the resource allocation
    - Live migration across nodes
- Problems that VMs don't solve for us
  - Job management: jobs within an environment still need to be managed
    - Job throttling
    - Job persistence, restart, etc.
  - The cost of perfect enforcement
    - Each CE will run a copy of similar services leading to potential inefficiencies
    - There has been some work in sharing e.g. shared libraries between VMs, but is inconclusive right now

Other things VMs are not helpful for...

- Digestion
- Walking your dog
- Laundry
- Most sports
- Painting walls
- Knitting
- Grocery shopping (unless you shop via peapod)

Meanwhile, in a galaxy far, far away...

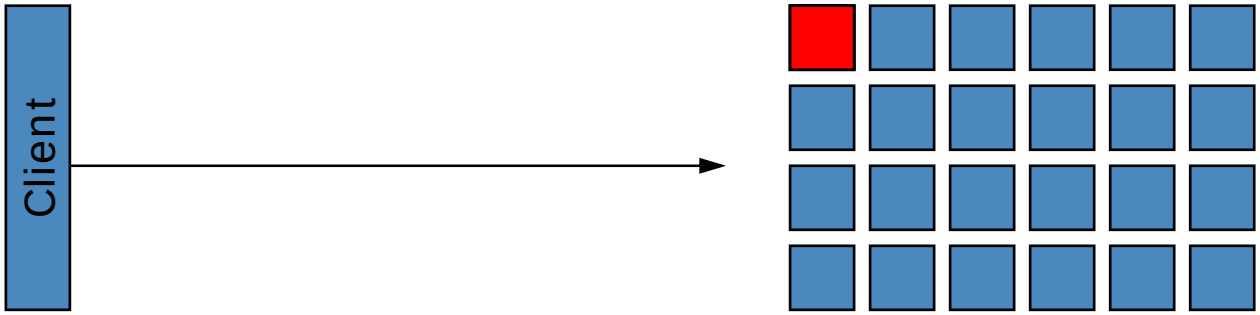
- Similar ideas in US projects
  - Edge Services
    - effort led by Frank Wuertherwein in the context of OSG/CMS
  - Management of submit nodes
    - work with Rob Gardner and Mike Wilde in the context of Grid3/Atlas
- Requirements:
  - Install and manage a complex configuration
    - Easy upgrades based on pre-configured images, consistent configuration across sites, version management, etc.
  - Control of resources
    - Guaranteed dedicated use of resources
  - Flexible load balancing
    - Widening the interface to a cluster based on need

Edge Services

- Edge Services: Services executing on the edge of a private/public network boundary
  - Typical configuration of today's resources
    - Resources within a site are available only on a private network
    - Site can be accessed through a limited number of public addresses
  - Examples: CE, SE, GK, and others
- Edge Services will be deployed in VM-based workspaces
  - Role-based deployment
  - Initially no advance reservations, no load balancing
  - A proof-of-concept activity
- Draft document available:
  - http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=167

Submit Node Management

- Similar to the Edge Service activity with particular emphasis on:
  - Configuration management
    - Configure once, copy and deploy many times
  - Load balancing: widening the submit bottleneck to clusters based on need



Conclusions

- Workspaces solve environment management problems
  - Configuration management
    - Configure once, copy and deploy many times
    - Upgrading service versions
    - Running conflicting or incompatible services side-by-side
  - Enforcement
    - Guaranteed resources with respect to other users
- There seems to be a confluence of ideas
  - Similar ideas in three different contexts
  - Coincidence?
- Workspaces do not solve job management problems
  - Which leads into subjects we'll talk about later…