

Virtualization Users Workshop Report

Owen Synge
D-Cache Desy

Outline

- Virtualization Basics.
 - What products.
- Implications.
 - What does/could this mean for us in HEP.
- Virtualization Workshop
 - Who was thier?
 - What happened at our meeting?
 - Notable comments?

Virtualization Basics

- Unix runs with one OS per host
 - Is this the right model for grids?
- Virtualization products exist
 - Xen, Vmware, Solaris containers, UML, Vserver
- Most favor Solaris and Xen solutions
- Available and production grade solutions exist now.
 - Some sites already using Virtualization extensively.

Implications

- Little performance overhead with Xen.
- Virtualization for availability
 - Not my area (See Thomas Finnern's talk)
 - Migrating/pausing services
- Virtualization for security
 - Isolation of hardware access.
- More operating systems to manage.
 - More than one concurrent operating system flavor is now possible

Workshop: Who was present?

- Virtualization Users
 - It was called the Virtualization user workshop
- Solution developers
 - Globus, Metacentre, Universitaet Karlsruhe
- LCG Grid integrators
- OSG/LCG Grid Administrators

What was presented?

- Using Virtualization for deployment
 - Grid Ireland : Trinity College Dublin
- Using Virtualization for testing
 - Dcache : Desy , gLite certification: CERN
- Using Virtualization on worker node
 - PBS and Xen : MetaCentre
 - Torque and Xen : Universitaet Karlsruhe
 - GridView : Trinity collage Dublin
- Globus Workspaces
 - WSRF control of worker nodes

What was demonstrated

- GridView for managing Xen images
 - Web interface for managing images
- Cerns
 - Cluster of virtual machine management application
- Desys
 - xen-image-manager
 - Simple image management application

What was talked about?

- Benefits of virtualization in a Grid.
- Worker node models and their use cases
- How HEP Grid could be best served?

Benefits of Virtualization

- Greater job throughput (Performance)
 - Running in 32/64 bit mode correctly.
- Greater Job isolation (Security)
 - Memory/OS is not shared.
- Backfilling (Utilization)
 - Pausing jobs while parallel jobs run.
- Management (Administrator)
 - Pausing, and migration of jobs would help.

Benefits of Virtualization 2

- Hardware OS, user OS isolation.
 - Admin runs OS they want, user get OS they want
- Security
 - Eases state management
 - Wiping nodes
 - Root on VM cannot escape VM
 - Daemon jobs and proxy stealing impossible
- Dynamic OS selection
 - Could be different OS per Job

Models for worker node 1

- One persistent VM with single image.
 - Benefit
 - Job isolation from base image
 - No user access to DOM0 hence DOM0 is consistent
 - Eases hardware abstraction.
 - Cost
 - Performance (Slight)
 - Maintenance (Two OS's)
 - Not integrated as a solution for large sites.

Models for worker node 2

- multiple persistent VM, multiple images.
 - Benefits
 - Currently in production at 3 sites
 - Can support “Back filling”
 - Memory is not shared between VM's.
 - Penalties
 - Lots of images, lots of integration
 - Memory is not shared between VM's.
- Model 2 has Cost/Benefits of 1

Models for worker node 3

- Non persistent Virtual machine.
 - Benefits
 - Greatly enhanced security
 - Jobs demonized and job cleaning automatic
 - Minimal modification to batch system required
 - Costs
 - VM boot time for jobs and image installation
 - 45 seconds typically using rsync
- Model 3 has all Cost/Benefits of 1 & 2

Models for worker node 4

- Non persistent VM's with a library of Images.
 - Benefits
 - Great flexibility
 - Experiments don't have to coordinate (eg SL3/SL4)
 - Costs
 - Additional integration for Batch system and VM
 - Information systems need changes
 - JDL needs changes
 - Not all batch queues will be integrated quickly
- Model 4 has all costs benefits of 1,2 & 3

Models for worker node 5

- User defined Images on VM's
 - Benefit
 - No need for anyone to agree OS
 - Cost
 - New infrastructure needed (Picking OS image)
 - Need to define images (For sites to generate)
 - User supplied raw image is potentially dangerous
- Model 5 has all Cost/Benefits of 1,2,3 & 4

Models for Worker node Summary

- Model 4 seems to cover all current HEP use cases, as VO's are relatively large.
 - Dynamic image generation not seen as needed right now
- Security greatly enhanced with non persistent VM's
- Model 5 might be needed in long term.
 - Will wait and see user demand.
- Expect to support at least 3 Worker node Operating systems eg (i386/x64/sl4/sl5)

Work required and expected before Virtualization on WN

- For Short term use
 - VM integration for most Batch schedulers
 - Automated deployment
 - Recipes for deployment
 - Good kernel support in OS layer (RHEL5)
- For long term
 - Labeling and defining images for dynamic images. (eg ALI/rpath)

Root and the VM

- Should we give end users “root” privileges on non persistent VM's?
 - Leads to poor deployment practices
 - Dangers with rfio, dcap and nfs
 - Advantages in users will be able to setup things like at their site.
 - Conclusion
 - Resist users getting root unless we need them too and don't tell them about it being possible

Sites bringing virtualisation to the Grid worker node

- Following Model 2
 - Metacenter
 - Universitaet Karlsruhe
- Following Model 5
 - Globus

Workshop Conclusions

- Virtualization on the grid is an inevitable development in the next 5 years.
 - Dues to benefits for admins and users.
- Virtualization provides increased job isolation and security.
 - Memory and hardware resource sharing can be enforced
- Heterogeneous Operating system support on same hardware is major advantage.