HEP Applications with Globus Virtual Workspaces


National Research Council of Canada, Ottawa, Ontario, Canada
University of Victoria, Victoria, British Columbia, Canada

HEPiX Fall 2007, St Louis
Overview

- Motivation
- Virtual Machines on the Grid
- Example Deployment
- Results
The Problem

- In Canada we have computing resources we can’t use. Why?
Virtualization on the Grid

- Virtualization is the solution.
- We can package an application complete with all of its dependencies and move it out to a remote resource.

Scientific
Red Hat
SUSE

Virtual Machine
Real Machine
Virtualization for HEP Apps on the Grid

- Find a virtual machine technology
- Need a middleware
- Movement of Images
- Security
VM: Xen is Useful for HEP

- Xen is a Virtual Machine technology that offers negligible performance penalties unlike more familiar VM systems like VMware.
- Xen uses a technique called “paravirtualization” to allow most instructions to run at their native speed.
  - The penalty is that you must run a modified OS kernel
  - Xen included in Linux Kernel mainline as of 2.6.23.

Figure 3: Relative performance of native Linux (L), Xen/Linux (X), VMware workstation 3.2 (V) and User-Mode Linux (U).
Before Globus Virtual Workspaces

- We first tried developing our own in house solution for GridX1.

- Set of simple Perl scripts to boot VMs on demand.

- Not well integrated with middleware, non-standard interface.

- Rewrite for every cluster.
Security

- Are you giving root away on your clusters?
  - root on domU != root on dom0 (not including recent Xen bugs).

- Sandboxing
  - Globus Virtual Workspaces helps. VMs are booted on BEHALF of users.
  - Different networking sandbox strategies available.
  - We experimented successfully with each worknode NATing its virtual workernodes.

- Authentication
  - Can you verify the source of your image?
Image Signing

First Steps

• We need to verify that the images come from people we trust.
  – Signatures using grid certificates.
  – For VM we run a hash algorithm (sha1) on the image and sign the hash.
• The group allowed to execute VMs doesn’t have to be the same as the group allowed to build them.

Example:

```
$ openssl x509 -in ~/.globus/usercert.pem -pubkey -noout > pubkey.pem
$ openssl dgst -sha1 -sign ~/.globus/userkey.pem -out vm_image.sha1 vm_image.img
$ openssl dgst -sha1 -verify pubkey.pem -signature vm_image.sha1 vm_image.img
```
Experiences

• Test Deployment

• Building Images

• Results
Test Deployments

Goal
• Deploy an example HEP application using Globus Virtual Workspaces.

Configuration
• Deployed Globus Virtual Workspaces on two separate clusters.
  – Scientific Linux(SL) 5.0, Intel machines at the University of Victoria
  – SuSe 10.2, Opteron machines at the National Research Council in Ottawa
• Application is the ATLAS Distribution Kit 13.0.10
  – Selected because it was familiar to us.
Where do we get the VMs?

- Getting the additional flexibility of VM now burdens us with building them.

- Building virtual machines can be a hurdle.
  - If it isn’t easy people won’t do it.

- Several possible approaches.
  - Give users the tools to easily build their own images.
  - Provide users with pre-built images which they can customize.
Building Virtual Machines

- There are many new tools for building images.

SL 5.0 now includes the RedHat Tool ‘virt-manager’ for the creation of Virtual Machines
Other Sources of Images

- Projects like the CERN OS Farm endeavor to create images on the fly at users request.
- Experiments could release pre-certified VM complete with installed application.
Results

• Jet simulation and reconstruction performed using the ATLAS 13.0.10 kit shipped inside a SL 4.5 image to a remote SL 5.0 cluster. Image booted on SuSe cluster (SuSe still needs work).

• Result Verified using ATLAS Run Time Test (RTT).

• More work required to study image portability across common distributions.

• Support from Workspaces developers is excellent. I recommend that you try it out and help make sure that Workspaces ends up suitable for your needs.
Areas of Future Work

• OS kernel of guest image must be present at site.
  – Addressed with addition of pygrub.

• Mechanism for authenticating images.
  – Sign with grid certificates?

• Automatic local image caching.

• Better integration with LRMS (PBS, torque, Maui etc.)

• Integration with Grid Metascheduler
Conclusion

- VMs could allow Canadian HEP access to resources it couldn’t have accessed before.

- Globus Virtual Workspace is in the early stages of providing a mechanism to deploy VMs using existing using GT4.

- Security mechanisms for VMs needs more research.
Question to HEPiX

• How much does booting someone else's VM on your cluster scare you?
Acknowledgements

Globus Virtual Workspaces Developers:

Kate Keahey
Tim Freeman