GridX1

Deploying HEP Applications Using Xen and Globus Virtual Workspaces

I. Gable, A. Agarwal, A. Charbonneau, R. Desmarais, R. Enge, D. Grundy, A. Norton, D. Penfold-Brown, R. Seuster, R.J. Sobie, D. C. Vanderster

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

CHEP 2007, Victoria, BC

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.

GridX1

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.
- “Evaluation of Virtual Machines for HEP Grids”, Proceedings of CHEP 2006, Mumbai India.

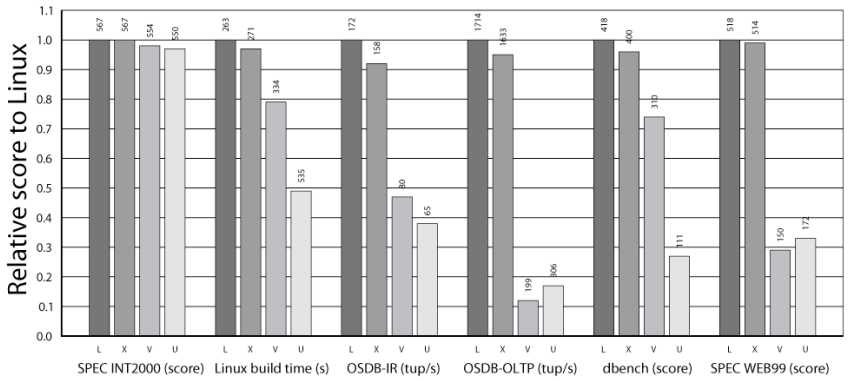
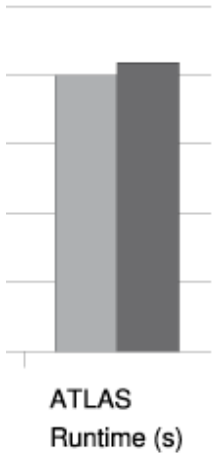


Figure 3: Relative performance of native Linux (L), XenoLinux (X), VMware workstation 3.2 (V) and User-Mode Linux (U).



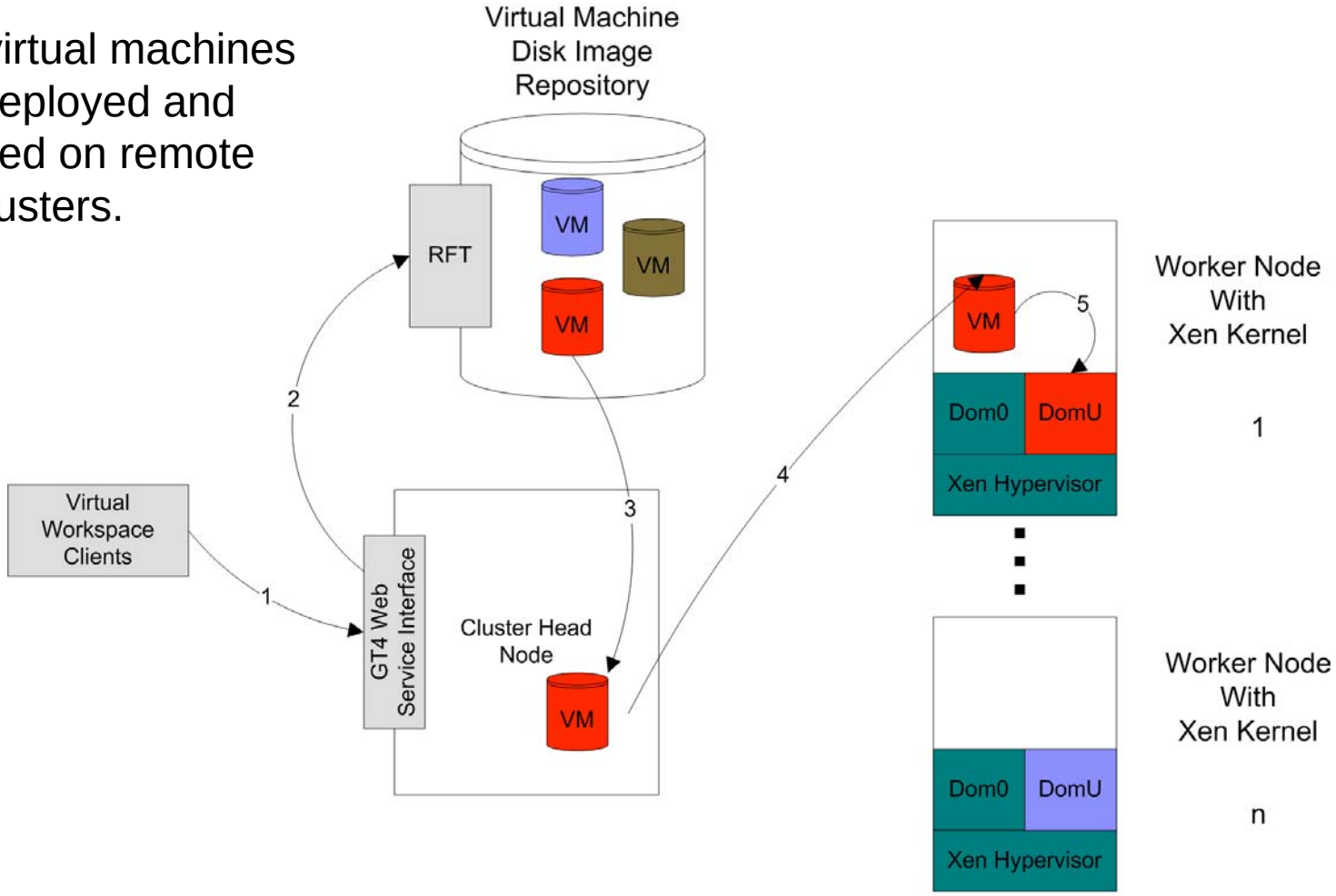
Middleware: Globus Virtual Workspaces

- We first tried developing our own in house solution
  - GridX1 GT2 based grid.
  - Set of simple Perl scripts to boot VMs on demand.
  - Not well integrated with middleware, non-standard interface.
  - Rewrite for every cluster.
- Globus Virtual Workspaces
  - Globus Project from Mathematics and Computer Science Division of Argonne National Laboratory.
  - Uses the Globus Toolkit Version 4 to present a **Web Services** Interface for the deployment and management of VMs on remote clusters.
  - Runs like any other Globus 4 Service.
  - In early stages of development, technology preview release available.
  - We are among a small number of users of this package and are contributing bug reports and feature request back to the developers.

GridX1

Movement of Images

Allow virtual machines to be deployed and managed on remote GT4 clusters.



Some Limitations

- OS kernel of guest image must be present at site.
  - Addressed with addition of pygrub
- No mechanism for authenticating images.
  - Sign with grid certificates?
- No automatic local image caching.

Security

- Are you giving root away on your clusters?
  - root on domU != root on dom0.
- 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?

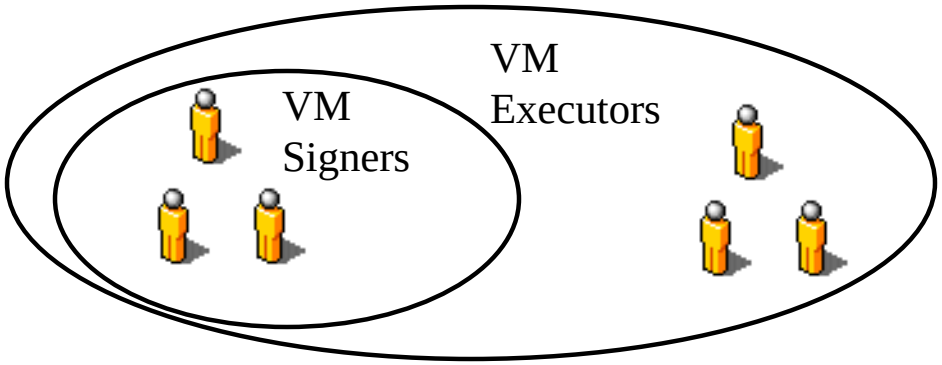
Image Signing

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.

Quite Simple:

```
$ openssl x509 -in ~/.globus/usercert.pem -pubkey -noout > pubkey.pem
$ openssl dgst -sha1  -sign ~/.ssh/userkey.pem -out vm_image.sha1 vm_image.img
$ openssl dgst -sha1  -verify pub.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, i686 machines at the University of Victoria
  - SuSe 10.0 i686 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?

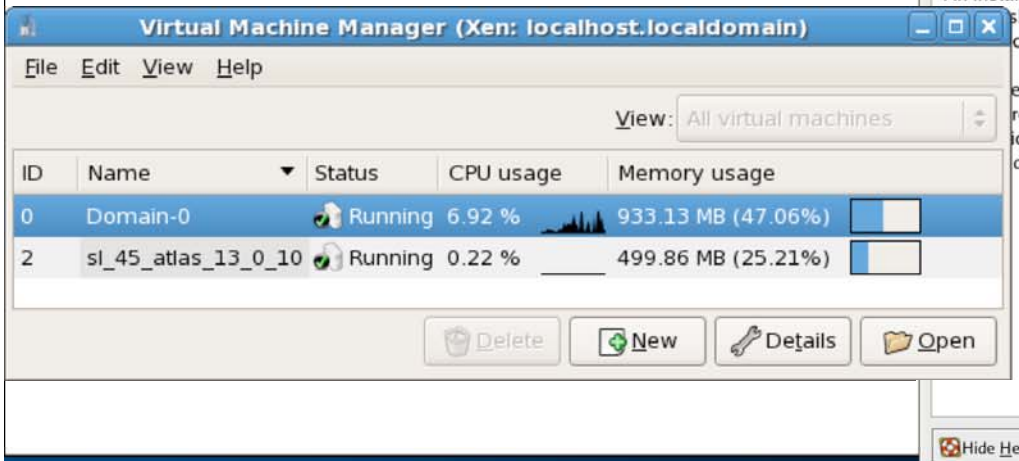
- Getting the additional flexibility of VM now burdens us with building them.
- Building virtual machines can be a hurdle.
  - If it isn't easy people won't do it.
- Several possible approaches.
  - Give users the tools to easily build their own images.
  - Provide users with pre-built images which they can customize.

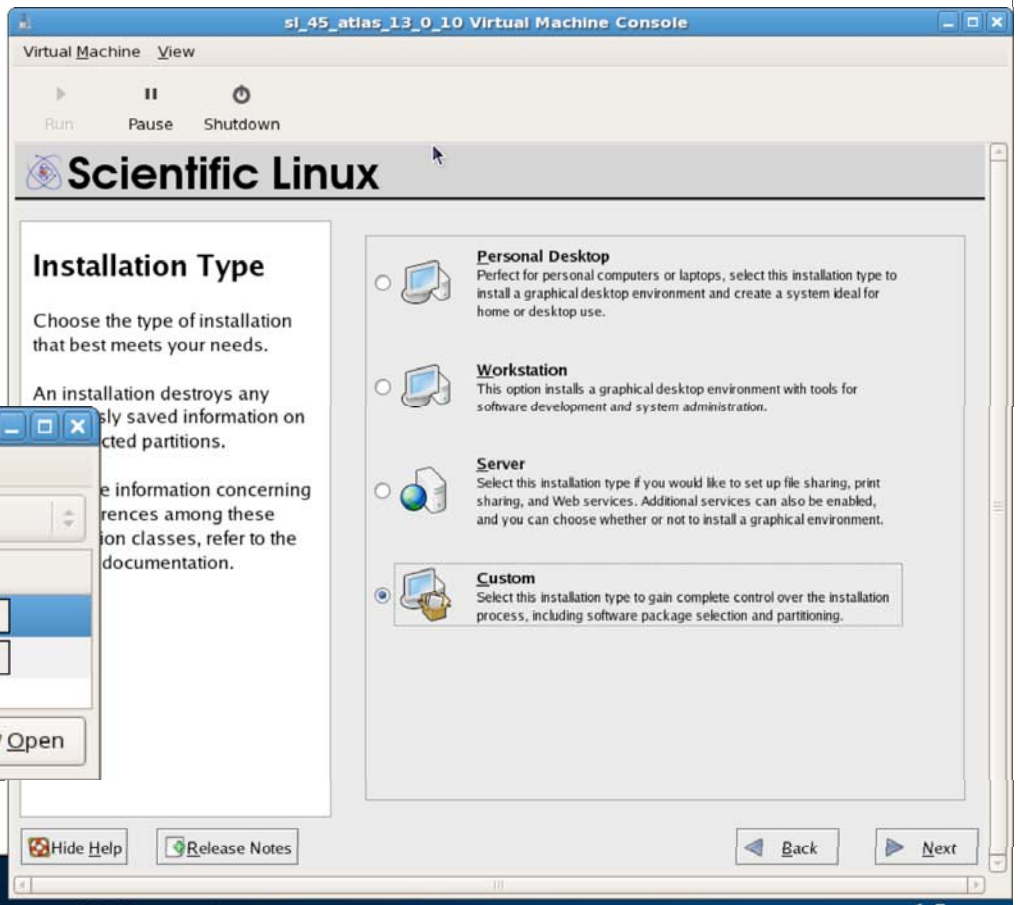
GridX1

Building Virtual Machines

- There are many new tools for building images.

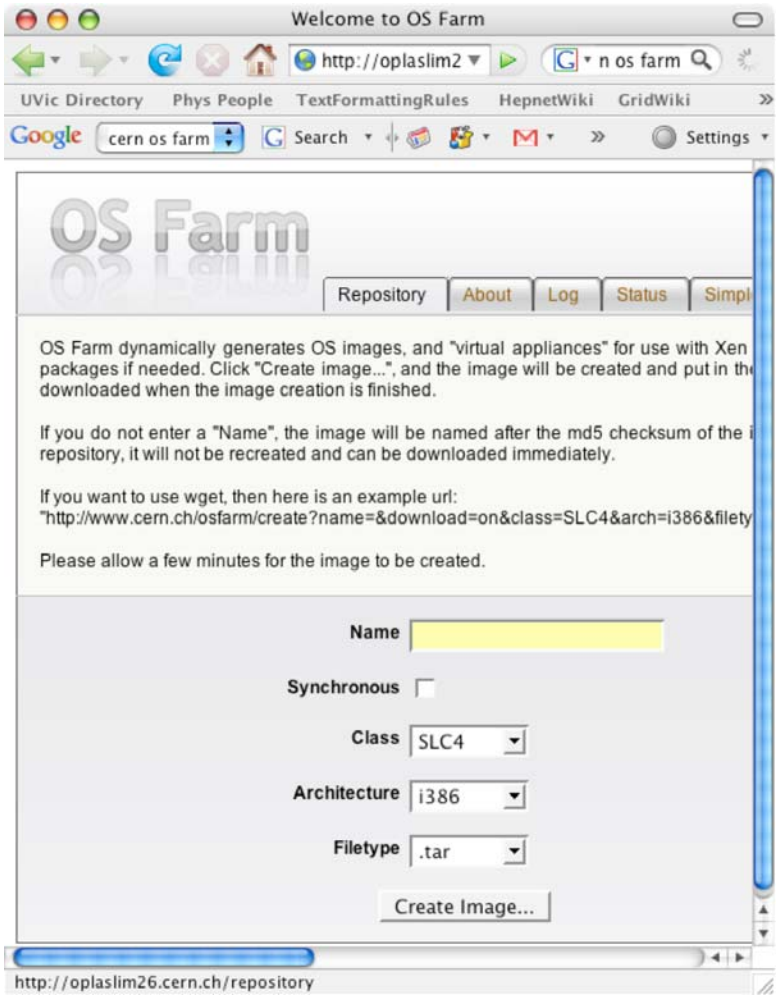
SL 5.0 now includes the RedHat Tool ‘virt-manager’ for the creation of Virtual Machines



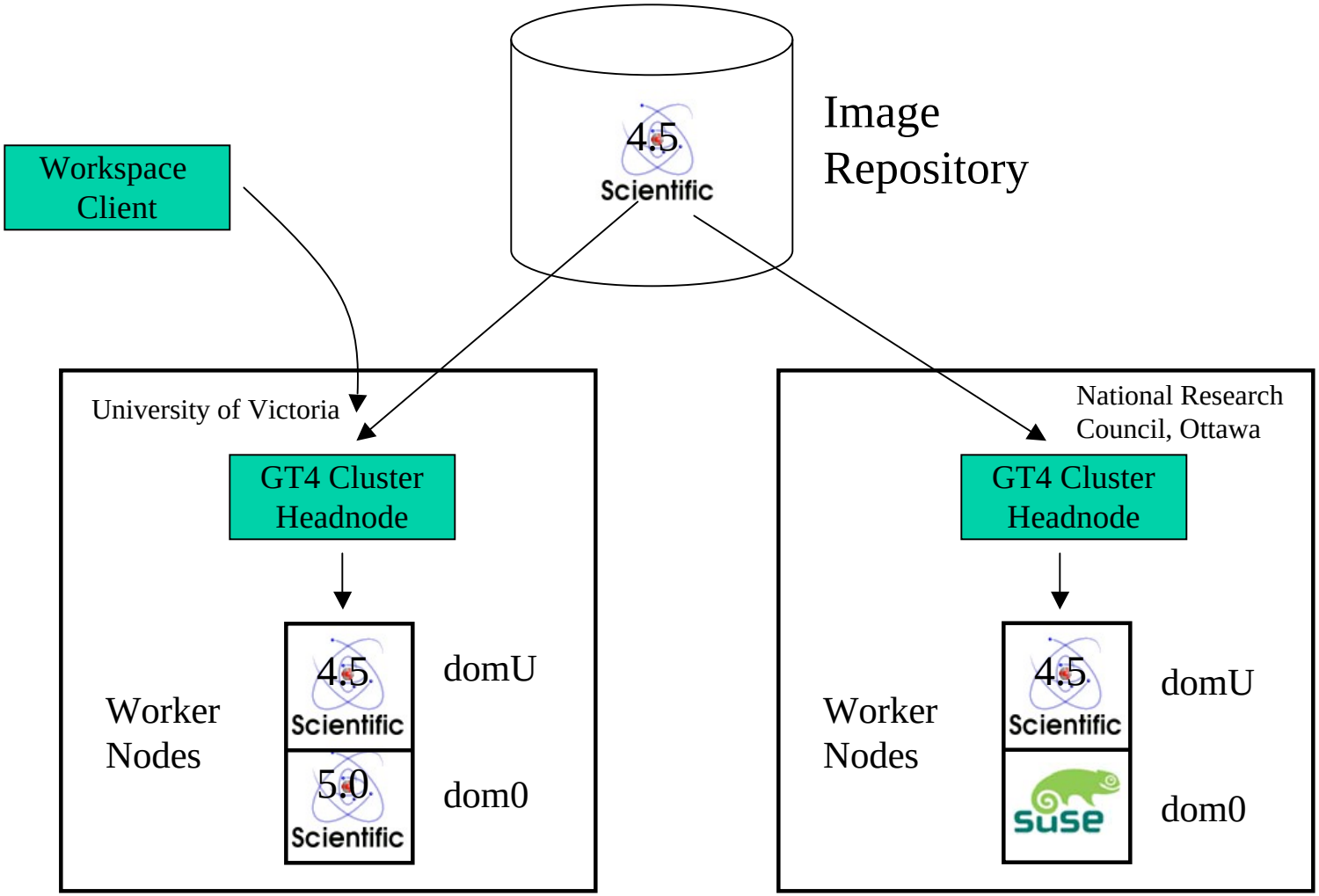


Other Sources of Images

- 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.



Test Deployment



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.
- Result Verified using ATLAS Run Time Test (RTT).
- As Globus Virtual Workspaces is in the early stages of development some stability problems were encountered as expected.

Conclusion

- VMs could allow HEP access to resources it couldn't have accessed before.
- Globus Virtual Workspace is in the early stages of providing a mechanism deploy VMs using existing Grid Technologies.
- Security mechanisms for VMs needs more research.

Future Work

- LRMS integration.
- Image signing.
- Local Image caching.

Acknowledgements

Globus Virtual Workspaces Developers:

Kate Keahey

Tim Freeman