FermiGrid Virtualization and Xen

Steven Timm

Feb 28, 2008

Fermilab Computing Techniques Seminar

Outline

- Virtual machines, brief history
- Virtualization in x86 hardware space
- Paravirtualized machines
- Hardware virtualized machines
- Implementations
- Local applications

What is a Virtual Machine

- Capable of virtualizing a whole set of resources, including processor(s), memory, storage, and peripherals
- Three properties of interest (Popek and Goldberg, 1974):
  - Equivalence—program run on the VM should exhibit behavior identical to running on the equivalent machine directly
  - Control—Virtual Machine Manager must be in control of the virtualized resources
  - Efficiency—Majority of instructions must be executed without intervention of the Virtual Machine Manager.

Virtual Machines--A Brief History

- IBM first released VM/370 for System 370 mainframe in 1972 after earlier prototypes on S-360.
- VM continues to present day, can support TSE, OS, AIX, Linux, and other instances of VM.
- Most commonly used client OS was CMS, lightweight single-user operating system.
- Term “hypervisor” first coined by IBM to describe the function of software that managed many virtual machines
- First example of full virtualization—complete simulation of the underlying hardware.

Virtualization in x86 hardware space

- X86 virtualization originally thought to be difficult—have to account for 17 unprivileged instructions that are sensitive to machine state. Two ways to do it:
- Paravirtualization—Modified device drivers in the kernel (Early VMware and Xen)
- Full virtualization—Intel VT-x technology, AMD-V “Pacifica” HVM mode (Later VMware and Xen)

Xen Hardware and OS support

- Paravirtualized:
  - Works on most Intel or AMD based hardware from 2003 and later. Requires Physical Address Extensions which some laptops don't have.
  - Supports most newer Linux (SUSE>=10, RHEL>=4, Fedora>=6, Ubuntu, Debian Etch)
- HVM
  - Requires Intel VT-x or AMD Pacifica extensions, most machines 2005 and later, and BIOS support on the motherboard. Hardware compatibility list available.
  - In addition to above, can run Windows XP, OpenBSD, Solaris x86, and legacy linuxes.

Xen Capabilities

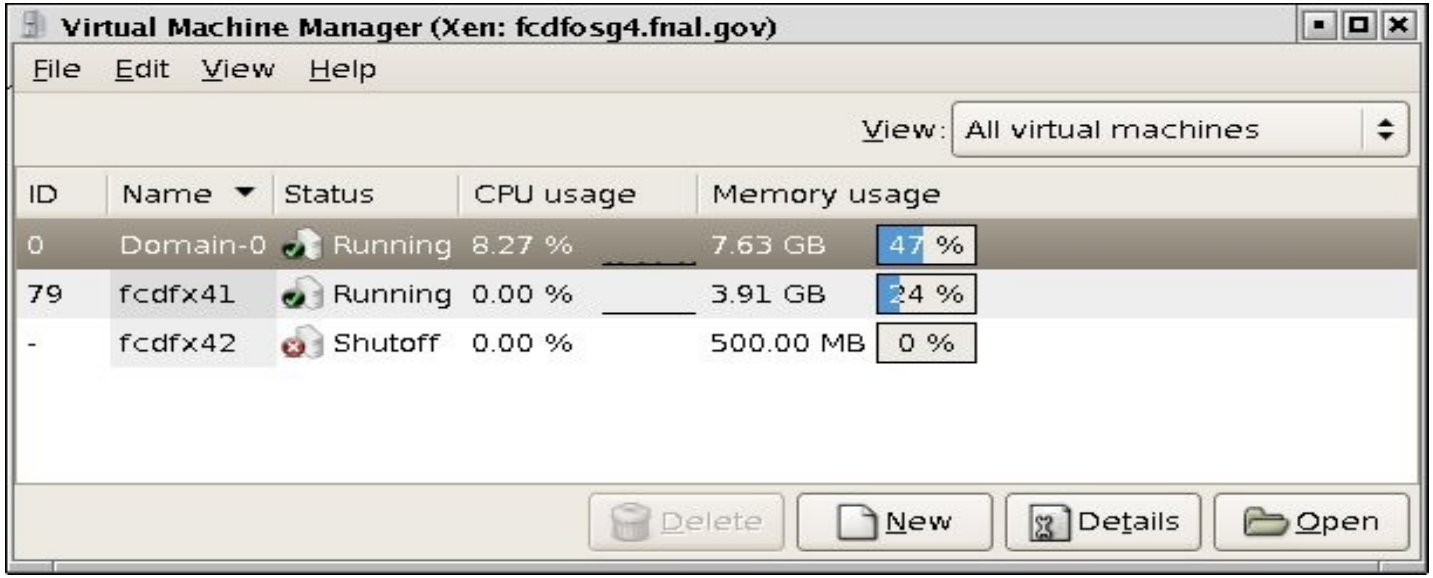
- Base OS of the machine is called “dom0”, runs the hypervisor
- Xen guests are referred to as “domU”, however many of them there are.
- Live migration of guest domains from one dom0 to the other.
- I/O and CPU throttling—machines can be allocated a percentage of total I/O and percentages of CPU usage. XenSource claims that this makes them denial-of-service proof.
- We expect this feature will be used by “VO Box / Edge Services” of LCG/OSG respectively.
- FermiGrid hasn’t tried to use either of these two features yet.
- Reboots faster! Xen instance can reboot in 5-10 seconds as opposed to 2-4 minutes for Dell PowerEdge 2950.

Where to get Xen from

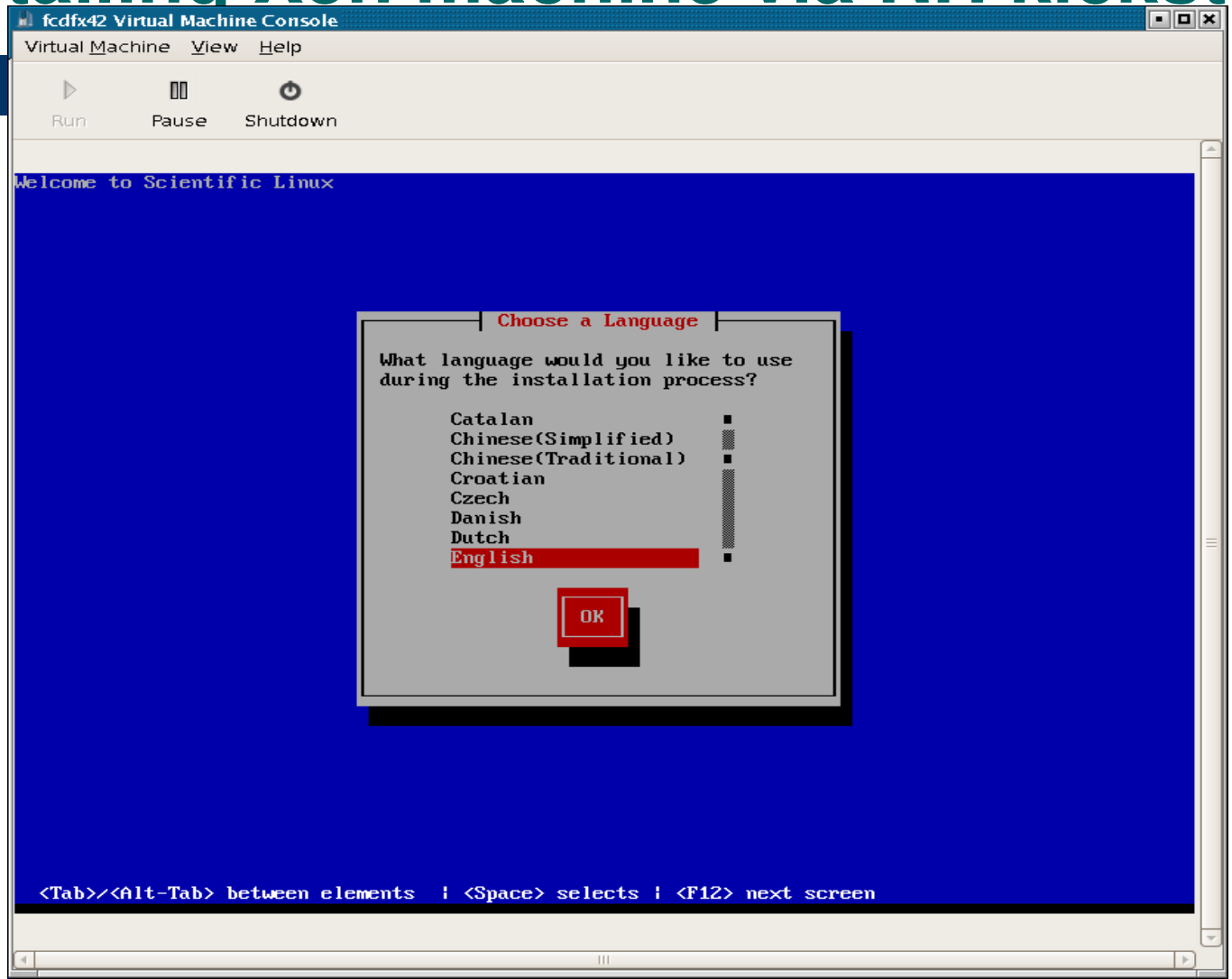
- Xen included in Fedora >=6, Red Hat (Centos, Scientific Linux) >=5.1, SuSE (SLES and OpenSuSE 10.x), Ubuntu.
- Source tarball and instructions of how to build it into the kernel are on www.xen.org, also i386-flavor rpms.

Xen Provisioning

- Virt-manager (part of Redhat/SL)



Installing Xen machine via RH kickstart

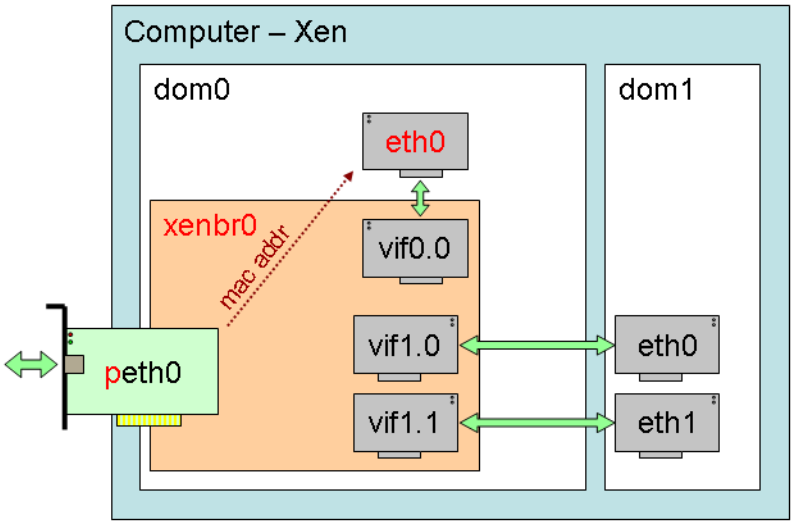


Provisioning the hard way

- Install a normal SL4 or SL5 machine
- Get the Xen binary tarball from xen.org
  - Make install
  - This gives a Xen-modified kernel (needed both for host and PVE guests)
- Adjust grub and reboot your machine with Xen kernel
- Take a known good OS install (could be one from a different machine that you want to migrate)
- Rsync it into the partition of what is going to be your root install for the Xen machine.
- Modify the network files appropriately.

Xen networking

- Two major ways to get network access from Xen instances to the outside world:
- Bridging (shown at left)
- NAT
- All FermiGrid setups use bridging.



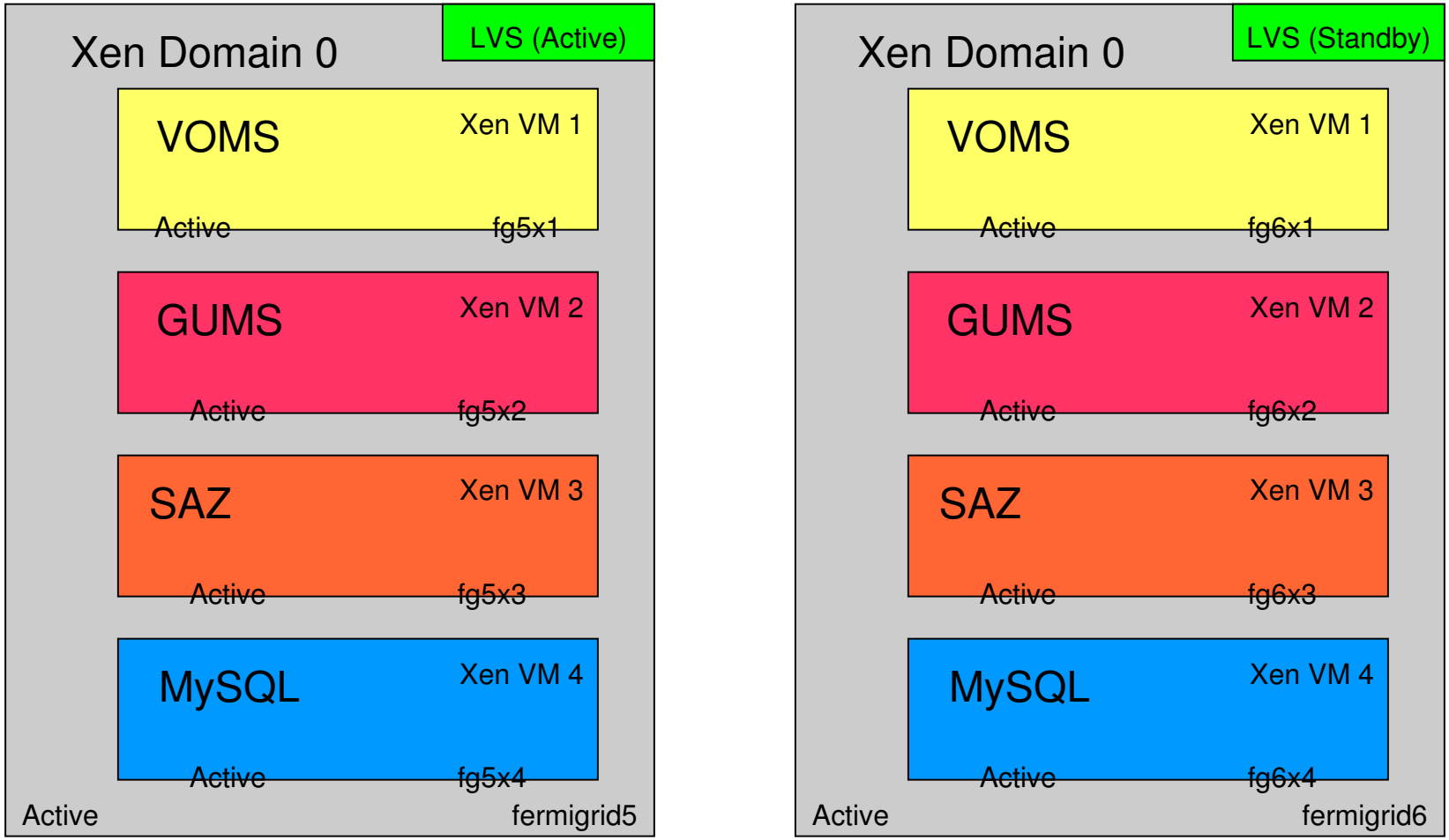
Xen and Citrix/Xensource.com

- Xensource.com was founded to market the hypervisor
- Taken over by Citrix in 2007
- Their goal mainly to market turnkey Xen appliances, special-purpose hypervisors
- We are currently in dialogue with them and will probably download evaluation version of their product.
- First impression—the value they propose to provide isn't worth the price they want ($25000).

Uses of Xen at Fermilab

- Development and Integration instances (USCMS for OSG-ITB, FermiGrid for OSG-ITB and Gratia development machines.)
- FermiGrid High Availability (see next slide)
- Soon to come, on individual cluster gatekeepers too, (GP Grid Cluster, CDF Grid Cluster 3).

Current FermiGrid High Availability



FermiGrid HA, Hardware and OS

- Currently 2 physical machines
  - Dell 2950
  - Dual core, dual CPU, 3GHz
  - 16GB RAM
  - Dual Gigabit ethernet NICs
  - 150GB RAID 1
  - Redundant Power Supplies
  - Base OS is SLF 5.0
- 4 Xen hosts apiece
  - fermigrid5 hosts fg5x1, fg5x2, fg5x3, fg5x4
  - fermigrid6 hosts fg6x1, fg6x2, fg6x3, fg6x4

Future uses for Xen in FermiGrid

- In next couple weeks we will move the LVS server inside of a Xen instance as well.
- High availability Xen instances for Squid, MyProxy, ReSS Information Gatherer
- High availability globus gatekeepers, Web Service containers, condor_schedd's. (These require a shared file system and would have to be active-passive).

FermiGrid Xen experience

- Why virtualize at all—
  - Services (VOMS, GUMS, SAZ) are designed to run on their own machine in their own tomcat instance
  - Some don't use much memory or CPU, a full server would be a waste.
- Why use paravirtualized Xen
  - Testing has showed performance is within a couple percent of native machine performance
  - It was free and it worked.
  - Early test hardware and versions of Xen didn't support HVM at the time.

FermiGrid Xen experience cont'd.

- On FermiGrid-HA, host OS is x86_64 and guests are i386.
- Only supported on Xen 3.1.0 and greater.
- TUV ships 5.1 with something called Xen 3.0.3 which has most but not all of the Xen 3.1.0 features back-ported.
- Unfortunately—not the 32 bit guest on 64 bit hosts.
- So we are using Xen straight from Xen.org
- Xen 3.1 supplied binary tarballs, for Xen 3.2 we will have to build from source (unless TUV gets their act together in time.)

Conclusions

- Xen instances, in combination with Linux Virtual Server and MySQL replication, allow us to run more services on less hardware, with improved reliability, less cost, and improved throughput!

Helpful Web Sites on Xen

- Open source Xen project www.xen.org
- Xen wiki http://wiki.xensource.com/xenwiki/
  - In particular the Networking section of the wiki.
- These slides, in DOCDB