Systems Administration

Introduction to Virtualization

TERNET 2012
Dar es Salaam, Tanzania
April 2012

Objectives

- To revise the core concepts
- To ensure we are using the same terminology

What is it?

- Virtualization -- the abstraction of the a resource from the actual physical instance of that resource.
- What Computing/Network resources can be virtualized?
  - Virtually anything! :)

Anything?

- In the context of this course. We're interested in virtualization along two dimensions:
  - Services ← we have seen some already
  - Hosts

Common virtualization

- Hard drive partitioning
  - A physical drive partitioned into 2 or more logical partitions
- Operating system
  - use of software to allow a piece of hardware to run multiple operating system images

Common virtualization

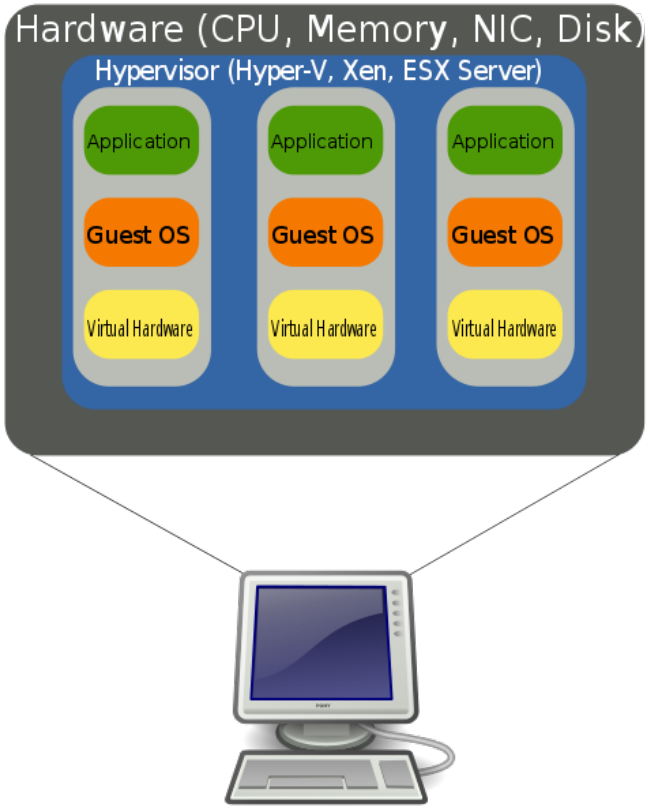
- Network
  - method of combining the available resources in a network by splitting up the available bandwidth into independent channels
- Server
  - masking of server resources (# and IDs of physical servers, cpu, and os) from server users

Resource/Service virtualization

- Examples:
  - Load-balancers
  - DNS Based GLB
  - HTTP(S) Virtual Hosting
  - MX records
  - Virtual Switches
  - Virtual Routers
  - Virtual Firewalls

Host Virtualization

- Examples
  - VMware
  - Virtual-Box
  - KVM (used in class)
  - XEN
  - FreeBSD and Linux Jails
  - Windows Hyper-V
  - Solaris Zones



What problem are we attempting to solve with host virtualization.

- Problem 1 – Idle capacity.
  - Most of the machines in your datacenter are idle most of the time.
  - Capacity you're not using: costs money & reduce return on capital
  - Packing discreet systems into a smaller number of servers provides savings along virtually every dimension.

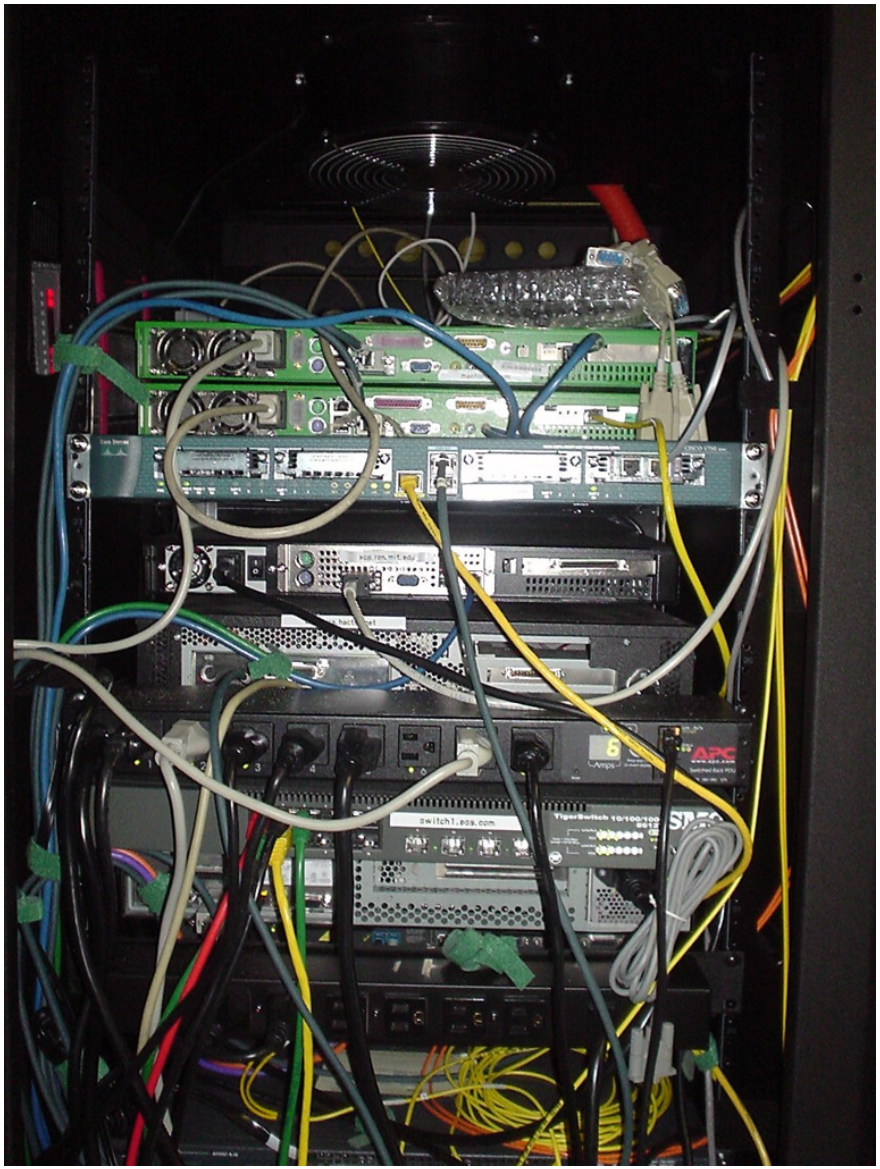
Problems - Continued

- Problem 2 – Provisioning
  - Spinning up a new service involves:
    - Acquiring the hardware
    - Building the server
    - Integration with existing services
  - With virtualization we're aiming to short-circuit that
    - Capacity is a resource
    - Machine instances my be cloned or provisioned from common basic images
    - Resources are purchased in bulk and assigned to applications as necessary.

Problems - Continued

- Problem 3 – Hardware abstraction
  - Operating systems, servers, and applications evolve at different rates.
  - Providing a common set of infrastructure resources means, virtualized systems are portable across servers
  - Hardware failure can more easily be managed.

Examples – Server Virtualization



Provisioning and management

- Is the glue that makes virtualization usable
- In commercial virtualization environments the provisioning/management toolkits represent the bulk of the licensing cost (VMware) and the secret sauce (VMotion, disaster recovery, backup, etc)
- One end of the spectrum:
  - XEN tools – a collection of perl scripts for spinning http://www.xen-tools.org/software/xen-tools/
  - KVM tools - http://www.linux-kvm.org/page/Management_Tools
- The Other:
  - Rightscale - http://www.rightscale.com/products/advantages/managing-systems-not-servers.php

QEMU-KVM

- **Qemu Emulator – the foundation of a number of virtualization products (including VirtualBox)**
- **Emulates the Entire Machine Environment**
  - **BIOS**
  - **CPU(s) SMP-capable**
  - **IDE Controller**
  - **NICs, many types**
  - **Graphics**
  - **USB, Sound, Etc.**
- **qemu-img used to generate Virtual Disks**
  - **supports RAW disks, sparse disks, copy-on-write, and VMDK**

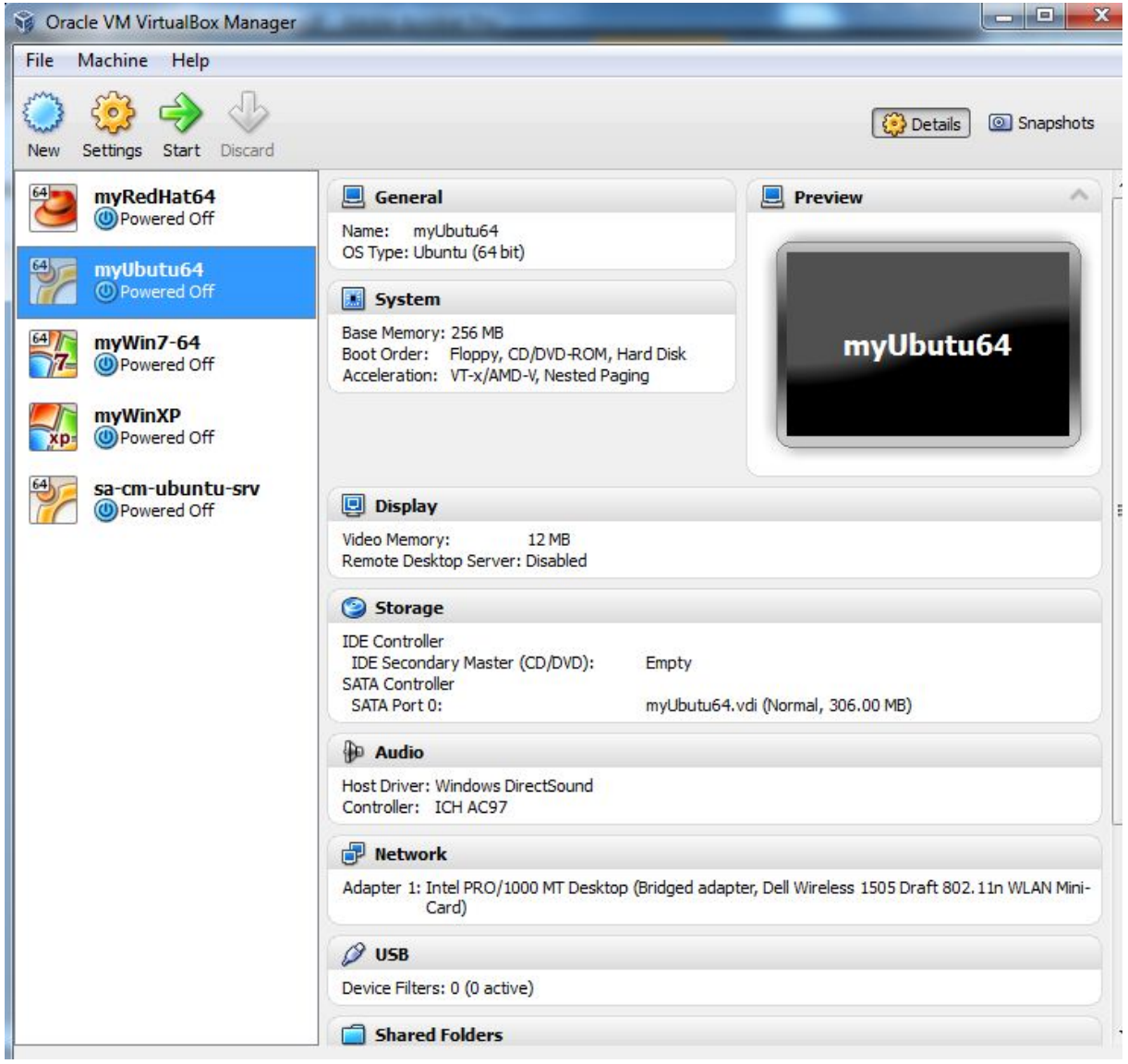
QEMU-KVM

- Why Qemu and not VMWare
  - 1) free open-source software
  - 2) supported by Redhat
  - 3) lots of features
  - 4) lots of support tools in development
- Why **NOT** QEMU-KVM
  - 1) documentation can be missing
  - 2) some features are buggy

2: Oracle VirtualBox

- Uses QEMU emulator
- VirtualBox runs on Windows, Linux, Macintosh, and Solaris hosts -> guest not limited to Windows (NT 4.0, 2000, XP, Server 2003, Vista, Windows 7), DOS/Windows 3.x, Linux, Ubuntu, Solaris and OpenSolaris, OS/2,OpenBSD
- **IQEmu** automated Windows VM creation, application integration
- Free!

Oracle VirtualBox …



Oracle VirtualBox …

- The VM manager shows all guest OS
- VM VB manager is the VM management interface
- Control VMs very easily

1: KVM
http://www.linux-kvm.org

- **KVM (Kernel Virtual Machine)**
- **Kernel modules for accelerating Virtualization**
- **Also provides additional services and I/O functionality**
- **Fully Integrated with current QEMU-KVM Distributions**
- **CPU-Specific, i.e. “kvm_intel” or “kvm_amd”**
- **% cat /proc/cpu | egrep 'vmx|svm'**
- **NOTE: Intel VM Extensions may or may not be enabled in your BIOS by default. Check this before you install a Hypervisor.**

QEMU-KVM

- Can run in a number of display modes:
  - “SDL” VGA Graphics
  - curses” text graphics
  - VNC remote viewing
- Many Network NIC options
  - default is an internal DHCP with NO ICMP support
  - bridged mode support by “virtio” and “tap” interfaces

QEMU Examples

```
qemu -hda /vms/myimg -cdrom /isos/ub10.iso -m 512
qemu -hda img1 -hdb img2 -hdc img3
qemu -hda qemu linux.img \
-net nic,vlan=0 -net tap,vlan=0,ifname=tap0
qemu -hda img.qcow2 -m 512 -daemonize -vnc :5
```

NOTE: the cdrom device can be an ISO file within the Host filesystem, or the Host CDROM drive itself

qemu-img

- qemu-img is the tool used to generate qemu virtual disks
- qcow2 format
  - sparse disk storage
  - copy-on-write (c.o.w.), a.k.a. “snap-shot” support
  - copy-on-write: means freezing a disk image, and using a new file to hold any further writes to that disk. In this way the original disk image is preserved. To roll-back, throw away new file.
- Cabable of reading/converting VirtualBox and VMWare Disks
- Examples:

  qemu-img create myhd.qcow2 6G

  qemu-img convert old.vmdk -O qcow2 newimg

The Qemu Monitor

http://en.wikibooks.org/wiki/QEMU/Monitor

- builtin control console used to jump out of the guest OS and perform operations on the VM
- access with CTRL-ALT-1/CTRL-ALT-2

(Mac uses CTRL-OPT-1/CTRL-OPT-2)

- operations: stop, cont, system_powerdown, change, usb_add, vnc, etc.
- migration: live migration from one site to another

  On site B: % qemu -hda myimg -incoming tcp:0:4444

  On site A: (in monitor) migrate -b tcp:hostB:4444

virsh/virt-manager

- libvirt toolkit – API used to interface with the qemu-kvm (and other vm platforms, xen, etc.)
- provides a uniform interface for controlling VMs
- provides a more consistent management console
- requires user added to groups: kvm, libvirtd
- Examples:

  root/system-level: virsh -c qemu:///system

  user/sessions: virsh -c qemu:///session

  virsh# list -all
- virt-manager: GUI tool for building and controlling VMs

Virtualization - Issues

- “All your eggs in one basket” – a poorly implemented virtual environment can create a large single point of failure
- Virtualization does not magically manufacture additional resources
- High-performance often requires dedicated hardware, ex. 10GB networking, massive Database I/O systems
- Sometimes the virtualized environment does not have all the features of the real one

Virtualization - Summary

- Useful for creating and testing new OS's
- Excellent for creating a dual-head, fully redundant, highly-available set of services with live-migration for failover
- Considerable savings on physical resources: heating, cooling, rack space, etc.
- Copy-On-Write filesystems and Snapshots are useful as for de-duplication and as point-in-time versions of the OS
- Significantly reduces deployment time
- Provides a standard environment for services