Server and Storage Virtualization

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides and audio/video recordings of this class lecture are at:

http://www.cse.wustl.edu/~jain/cse570-15/

1. Why Virtualize?
2. Server Virtualization Concepts
3. Storage Virtualization
4. Open Virtualization Format (OVF)

Note: Network Virtualization will be discussed in subsequent lectures

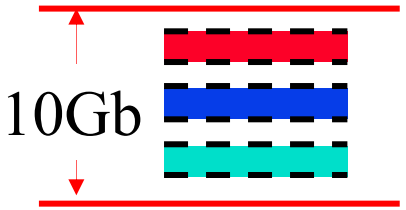
Virtualization

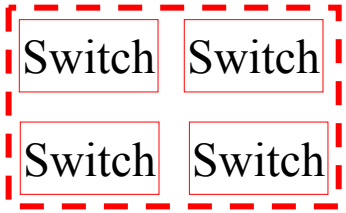
"*Virtualization means that Applications can use a resource without any concern for where it resides, what the technical interface is, how it has been implemented, which platform it uses, and how much of it is available.*"

-Rick F. Van der Lans
in Data Virtualization for Business Intelligence Systems

5 Reasons to Virtualize

1. Sharing: Break up a large resource
   Large Capacity or high-speed
   E.g., Servers
2. Isolation: Protection from other tenants
   E.g., Virtual Private Network
3. Aggregating: Combine many resources in to one, e.g., storage
4. Dynamics: Fast allocation, Change/Mobility, load balancing, e.g., virtual machines
5. Ease of Management ⇒ Easy distribution, deployment, testing





Advantages of Virtualization

- Minimize hardware costs (CapEx)
  Multiple virtual servers on one physical hardware
- Easily move VMs to other data centers
  - Provide disaster recovery. Hardware maintenance.
  - Follow the sun (active users) or follow the moon (cheap power)
- Consolidate idle workloads. Usage is bursty and asynchronous.
  Increase device utilization
- Conserve power
  Free up unused physical resources
- Easier automation (Lower OpEx)
  Simplified provisioning/administration of hardware and software
- Scalability and Flexibility: Multiple operating systems

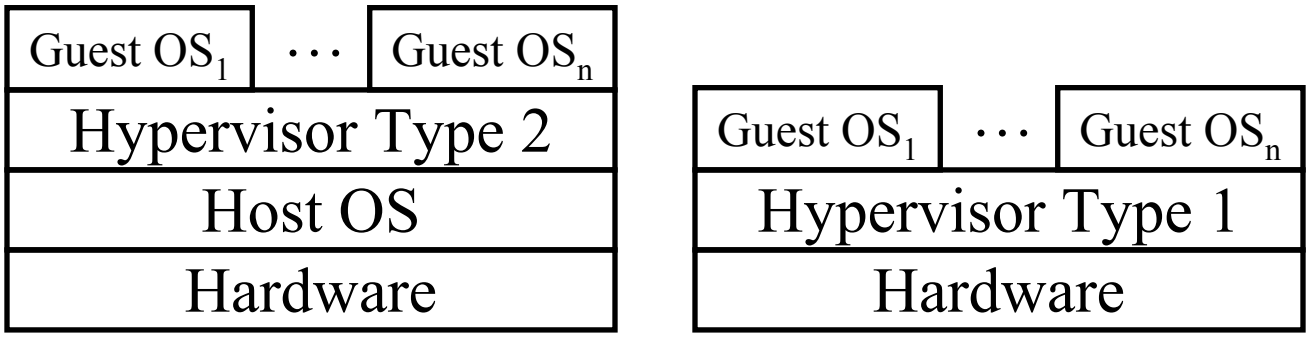
Ref: http://en.wikipedia.org/wiki/Platform_virtualization
Ref: K. Hess, A. Newman, "Practical Virtualization Solutions: Virtualization from the Trenches," Prentice Hall, 2009, ISBN:0137142978

Virtualization in Computing

- **Storage**:
  - Virtual Memory ⇒ L1, L2, L3, ... ⇒ Recursive
  - Virtual CDs, Virtual Disks (RAID), Cloud storage
- **Computing**:
  - Virtual Desktop ⇒ Virtual Server ⇒ Virtual Datacenter

    Thin Client ⇒ VMs ⇒ Cloud
- **Networking**: Plumbing of computing
  - Virtual Channels, Virtual LANs, Virtual Private Networks

Server Virtualization Concepts

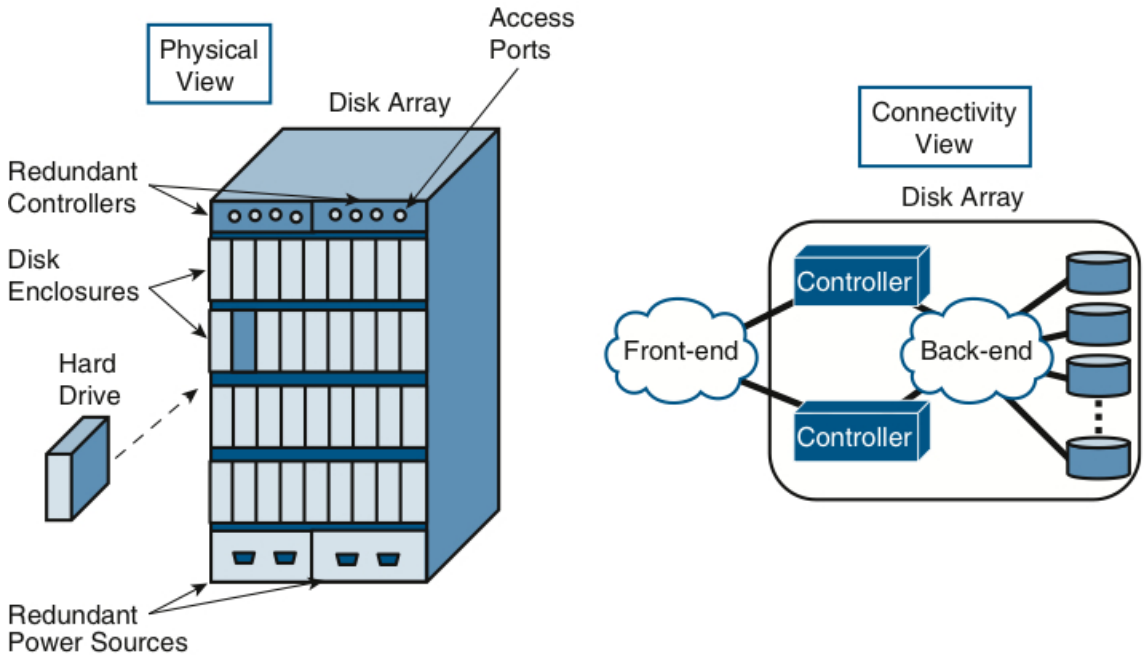


- Host OS: Runs on the bare metal
- Guest OS: Runs on the host OS,
  e.g., Windows XP Mode on Win 7
- Hypervisor: Software to support multiple virtual machines
  - Type 1: Runs on bare metal, e.g., Xen, VMware ESXi
  - Type 2: Runs on a host OS, e.g., MS Virtual PC
  - Type 0: Both 1 and 2, e.g., Linux KVM

Ref: http://en.wikipedia.org/wiki/Hypervisor

Disk Arrays

- In data centers, all disks are external to the server
  ⇒Data accessible by other servers in case of a server failure
- JBODs (Just a bunch of disks): Difficult to manage
- Disk Arrays: An easy to manage pool of disks with redundancy



Ref: G. Santana, "Data Center Virtualization Fundamentals," Cisco Press, 2014, ISBN:1587143240

Data Access Methods

Three ways for applications to access data:

- **Block Access**: A fixed number of bytes (block-size), e.g., 1 sector, 4 sectors, 16 sectors
- **File Access**: A set of bytes with name, creation date, and other meta data.
  - May or may not be contiguous.
  - A file system, such as, FAT-32 (File Allocation Table) or NTFS (New Technology File System) defines how the meta-data is stored and files are organized.
  - File systems vary with the operating systems.
- **Record Access**: Used for highly structured data in databases. Each record has a particular format and set of fields. Accessed using Structured Query Language (SQL), Open DataBase Connectivity (ODBC), Java DataBase Connectivity (JDBC)
- Storage systems provide block access. A logical volume manager in the OS provides other "virtual" views, e.g., file or record

What is Storage Virtualization?

- Restating Rick F. Van der Lans: *Storage virtualization means that Applications can use storage without any concern for where it resides, what the technical interface is, how it has been implemented, which platform it uses, and how much of it is available*
- Distance: Remote storage devices appear local
- Size: Multiple smaller volume appear as a single large volume
- Spread: Data is spread over multiple physical disks to improve reliability and performance
- File System: Windows, Linux, and UNIX all use the same storage device
- Virtual Interface: A SCSI disk connected to a computer with no SCSI interface
- Advantages: High availability, Disaster recovery, improved performance, sharing (better CapEx)

Benefits of Storage Virtualization

- Much larger distances
- Greater performance
- Increased disk utilization
- Higher availability with multiple access path
- Higher availability due to redundant storage
- Disaster recovery capability
- Continuous on-line back
- Easier testing
- Increased scalability
- Allows thin provisioning (Appears as if there is more disk than physical)

Open Virtualization Format (OVF)

- Standalone software can be distributed as a virtual machine image, called, virtual appliance
- Independent of hypervisor or processor architecture
- OVF is the standard format for virtual appliances
  Standardized by DTMF (Distributed Management Task Force)
  Now ISO/IEC standard.
- OVF package consists of several files in a directory.
  An XML file with extension .ovf or a compliant format, e.g., .vmdk in the directory contains all the meta data required to run the package, e.g., hardware requirements, descriptions, security certificates, etc.
- VMware, Microsoft, Oracle, Citrix, IBM and many others support OVF

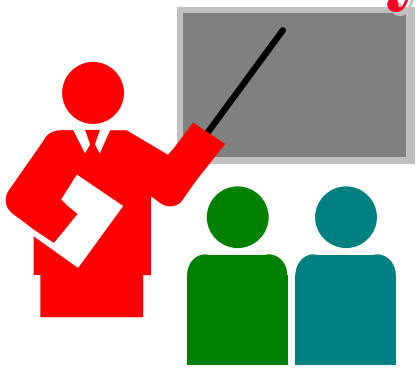
Ref: http://en.wikipedia.org/wiki/Virtual_appliance
Ref: http://en.wikipedia.org/wiki/Open_Virtualization_Format

OVF (Cont)

- OVF V1.1.0 supports single VM packages and packages containing multiple VMs constituting a multitier service
- Other popular format are Microsoft's Virtual Hard Disk (VHD), and VMware's Virtual Machine Disk (VMDK)

Ref: http://en.wikipedia.org/wiki/VHD_%28file_format%29
Ref: http://en.wikipedia.org/wiki/VMDK

Summary



1. Virtualization allows computation to be done anywhere anytime on any infrastructure ⇒Easy and efficient resource scheduling and management
2. Servers, storage, and network all need to be virtualized
3. Hypervisors of type 1 run on bare metal. Type 2 require a host OS.
4. OVF is the standard format for virtual images

Acronyms

API	Application Programming Interface
CapEx	Capital Expenditure
CD	Compact Disk
DTMF	Distributed Management Task Force
DVD	Digital Video Disk
ESX	VMware Product Name
FAT	File Allocation Table
FreeBSD	Free Berkeley System Distribution
IEC	International Electro technical Commission
ISO	International Standards Organization
JBOD	Just a bunch of disks
JVM	Java Virtual Machine
KVM	Kernel Virtual Machine
MS	Microsoft
NTFS	New Technology File System

Acronyms (Cont)

ODBC	Open Database Connectivity
OpEx	Operational Expenses
OS	Operating System
OVF	Open Virtualization Format
PC	Personal Computer
RAID	Redundant Array of Independent Disks
SCSI	Small Computer Systems Interface
SMP	Symmetric Multiprocessing
SQL	Structured Query Language
UML	User-Mode Linux
USB	Universal Serial Bus
VHD	Virtual Hard Disk
VM	Virtual Machine
VMDK	Virtual machine disk
XML	eXtensible Markup Language

Reading List

- G. Santana, "Data Center Virtualization Fundamentals," Cisco Press, 2014, ISBN:1587143240K. Hess, A. Newman, "Practical Virtualization Solutions: Virtualization from the Trenches," Prentice Hall, 2009, ISBN:0137142978 (Safari Book)
- L. C. Miller, "Server Virtualization for Dummies," Wiley, 2012, Oracle Special Edition, http://www.oracle.com/oms/hardware/extremeperformance/assets/ept-eb-dummies-server-1641465.pdf
- C. Scheffy, "Virtualization for Dummies," Wiley 2007, AMD Special Edition, http://www.amd.com/us/Documents/Virt_for_Dummies.pdf
- B. Golden, "Virtualization for Dummies," Wiley, 2011, HP special edition, https://ssl.www8.hp.com/de/de/pdf/virtuallisation_tcm_144_1147500.pdf

Wikipedia Links

- http://en.wikipedia.org/wiki/Desktop_virtualization
- http://en.wikipedia.org/wiki/Hardware-assisted_virtualization
- http://en.wikipedia.org/wiki/Hardware_emulation
- http://en.wikipedia.org/wiki/Hardware_virtualization
- http://en.wikipedia.org/wiki/Hypervisor
- http://en.wikipedia.org/wiki/Open_Virtualization_Format
- http://en.wikipedia.org/wiki/Platform_virtualization
- http://en.wikipedia.org/wiki/VHD_%28file_format%29
- http://en.wikipedia.org/wiki/VMDK
- http://en.wikipedia.org/wiki/Operating_system-level_virtualization
- http://en.wikipedia.org/wiki/Virtual_appliance
- http://en.wikipedia.org/wiki/Virtual_machine
- http://en.wikipedia.org/wiki/Virtualbox
- http://en.wikipedia.org/wiki/Virtualization
- http://en.wikipedia.org/wiki/Windows_Virtual_PC
- http://en.wikipedia.org/wiki/Xen

Related Web Sites

- Open Virtualization Format (OVF), http://www.dmtf.org/standards/ovf (OVF standard documents)