Infrastructure as a Service [IaaS]

Virtualization

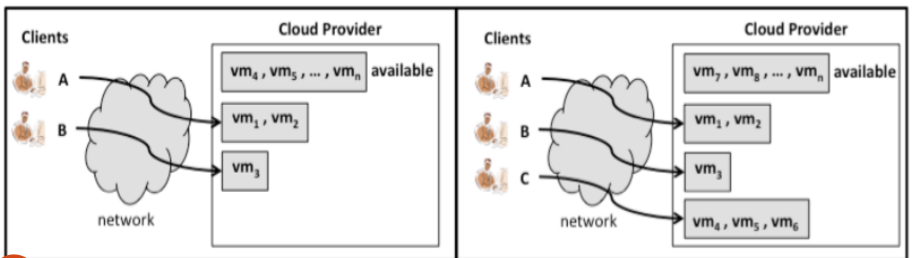
IaaS – Infrastructure as a Service

- What does a subscriber get?
  - Access to virtual computers, network-accessible storage, network infrastructure components such as firewalls, and configuration services.
- How are usage fees calculated?
  - Typically, per CPU hour, data GB stored per hour, network bandwidth consumed, network infrastructure used (e.g., IP addresses) per hour, value-added services used (e.g., monitoring, automatic scaling)

2

IaaS Provider/Subscriber Interaction Dynamics

- The provider has a number of available virtual machines (vm) that it can allocate to clients.
  - Client A has access to vm1 and vm2, Client B has access to vm3 and Client C has access to vm4, vm5 and vm6
  - Provider retains only vm7 through vm_N

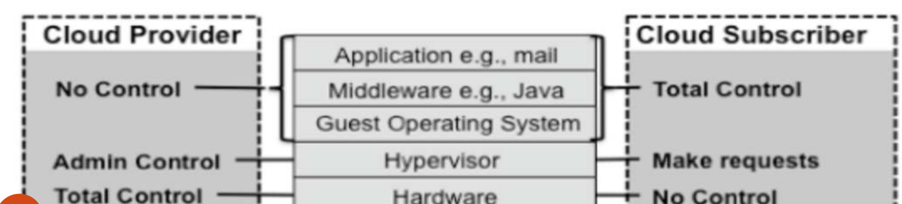


3

Source: LeeBadger, and Tim Grance "NIST DRAFT Cloud Computing Synopsis and Recommendations"

IaaS Component Stack and Scope of Control

- IaaS component stack comprises of hardware, operating system, middleware, and applications layers.
- Operating system layer is split into two layers.
  - Lower (and more privileged) layer is occupied by the Virtual Machine Monitor (VMM), which is also called the Hypervisor
  - Higher layer is occupied by an operating system running within a VM called a guest operating system



4

Source: LeeBadger, and Tim Grance "NIST DRAFT Cloud Computing Synopsis and Recommendations"

IaaS Component Stack and Scope of Control

- In IaaS Cloud provider maintains total control over the physical hardware and administrative control over the hypervisor layer
- Subscriber controls the Guest OS, Middleware and Applications layers.
- Subscriber is free (using the provider's utilities) to load any supported operating system software desired into the VM.
- Subscriber typically maintains complete control over the operation of the guest operating system in each VM.

5

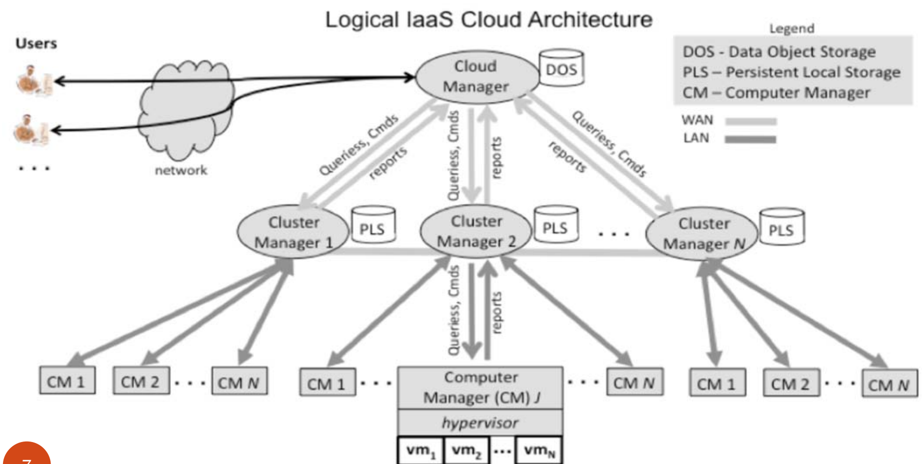
IaaS Component Stack and Scope of Control

- A hypervisor uses the hardware to synthesize one or more Virtual Machines (VMs); each VM is "an efficient, isolated duplicate of a real machine" .
- Subscriber rents access to a VM, the VM appears to the subscriber as actual computer hardware that can be administered (e.g., powered on/off, peripherals configured) via commands sent over a network to the provider.

6

IaaS Cloud Architecture

- Logical view of IaaS cloud structure and operation



Source: LeeBadger, and Tim Grance "NIST DRAFT Cloud Computing Synopsis and Recommendations "

IaaS Cloud Architecture

- Three-level hierarchy of components in IaaS cloud systems
  - *Top level* is responsible for *central control*
  - *Middle level* is responsible for *management of possibly large computer clusters* that may be *geographically distant* from one another
  - *Bottom level* is responsible for *running the host computer systems* on which virtual machines are created.
- Subscriber queries and commands generally flow into the system at the top and are forwarded down through the layers that either answer the queries or execute the commands

IaaS Cloud Architecture

- Cluster Manager can be geographically distributed
- Within a cluster manger computer manger is connected via high speed network.

9

Operation of the Cloud Manager

- Cloud Manager is the public access point to the cloud where subscribers sign up for accounts, manage the resources they rent from the cloud, and access data stored in the cloud.
- Cloud Manager has mechanism for:
  - Authenticating subscribers
  - Generating or validating access credentials that subscriber uses when communicating with VMs.
  - Top-level resource management.
- For a subscriber's request cloud manager determines if the cloud has enough free resources to satisfy the request

10

Data Object Storage (DOS)

- DOS generally stores the subscriber's metadata like user credentials, operating system images.
- DOS service is (usually) single for a cloud.

11

Operation of the Cluster Managers

- Each *Cluster Manager* is responsible for the operation of a collection of computers that are connected via high speed local area networks
- *Cluster Manager* receives resource allocation commands and queries from the *Cloud Manager*, and calculates whether part or all of a command can be satisfied using the resources of the computers in the cluster.
- *Cluster Manager* queries the *Computer Managers* for the computers in the cluster to determine resource availability, and returns messages to the *Cloud Manager*

12

Operation of the Cluster Managers

- Directed by the Cloud Manager, a Cluster Manager then instructs the Computer Managers to perform resource allocation, and reconfigures the virtual network infrastructure to give the subscriber uniform access.
- Each Cluster Manager is connected to Persistent Local Storage (PLS)
- PLS provide persistent disk-like storage to Virtual Machine

13

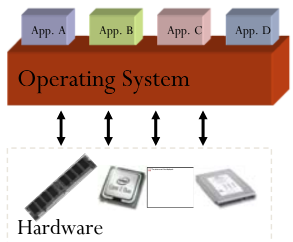
Operation of the Computer Managers

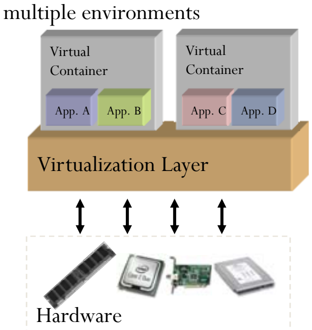
- At the lowest level in the hierarchy computer manger runs on each computer system and uses the concept of virtualization to provide Virtual Machines to subscribers
- Computer Manger maintains status information including how many virtual machines are running and how many can still be started
- Computer Manager uses the command interface of its hypervisor to start, stop, suspend, and reconfigure virtual machines

14

Virtualization

- Virtualization is a broad term (virtual memory, storage, network, etc)
- Focus: **Platform virtualization**
- Virtualization basically allows one computer to do the job of multiple computers, by sharing the resources of a single hardware across multiple environments





'Non-virtualized' system
A single OS controls all hardware platform resources

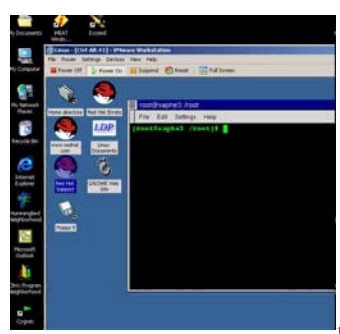
Virtualized system
It makes it possible to run multiple Virtual Containers on a single physical platform

15

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Virtualization

- Virtualization is way to run **multiple operating systems** and **user applications** on the same hardware
  - E.g., run both Windows and Linux on the same laptop
- How is it different from **dual-boot**?
  - Both OSes run **simultaneously**
- The OSes are completely **isolated** from each other



16

Virtualization Requirements

- A Model of Third Generation Machines
  - Machine states: S = (E, M, P, R), *executable storage E, processor mode M, program counter P, and relocation-bounds register R*
  - Instructions classification
    - Privileged instructions
    - Control sensitive instructions
    - Behavior sensitive instructions
- Properties for a Virtual Machine Monitor
  - Equivalence
  - Resource control
  - Efficiency

17

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Hypervisor or Virtual Machine Monitor

Research Paper :Popek and Goldberg, "Formal requirements for virtualizable third generation architectures", CACM 1974 (http://portal.acm.org/citation.cfm?doid=361011.361073).

A **hypervisor** or **virtual machine monitor** runs the guest OS directly on the CPU. (This only works if the guest OS uses the same instruction set as the host OS.) Since the guest OS is running in user mode, privileged instructions must be intercepted or replaced. This further imposes restrictions on the instruction set for the CPU, as observed in a now-famous paper by Popek and Goldberg identify three goals for a virtual machine architecture:

- Equivalence: The VM should be indistinguishable from the underlying hardware.
- Resource control: The VM should be in complete control of any virtualized resources.
- Efficiency: Most VM instructions should be executed directly on the underlying CPU without involving the hypervisor.

18

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Hypervisor or Virtual Machine Monitor

Popek and Goldberg describe (and give a formal proof of) the requirements for the CPU's instruction set to allow these properties. The main idea here is to classify instructions into

- **privileged** instructions, which cause a trap if executed in user mode, and
- **sensitive** instructions, which change the underlying resources (e.g. doing I/O or changing the page tables) or observe information that indicates the current privilege level (thus exposing the fact that the guest OS is not running on the bare hardware).
- The former class of sensitive instructions are called **control sensitive** and the latter **behavior sensitive** in the paper, but the distinction is not particularly important.

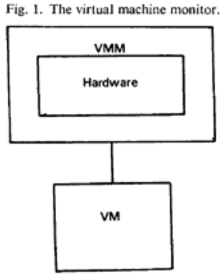
What Popek and Goldberg show is that we can only run a virtual machine with all three desired properties if the sensitive instructions are a subset of the privileged instructions. If this is the case, then we can run most instructions directly, and any sensitive instructions trap to the hypervisor which can then emulate them (hopefully without much slowdown).

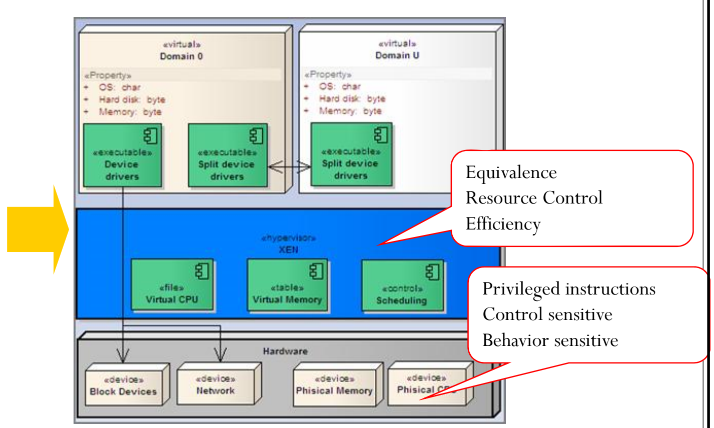
19

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

VMM and VM

Fig. 1. The virtual machine monitor.





- For any conventional third generation computer, a VMM may be constructed if the set of sensitive instructions for that computer is a subset of the set of privileged instructions
- A conventional third generation computer is recursively virtualizable if it is virtualizable and a VMM without any timing dependencies can be constructed for it.

20

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

The evolution of virtualization

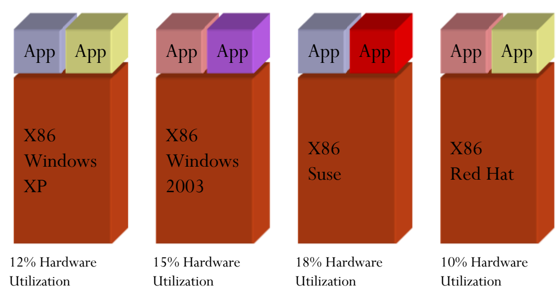
How did it start?

- Server virtualization has existed for several decades
  - IBM pioneered more than 4 decades ago with the capability to "multitask"
- The inception was in specialized, proprietary, high-end server and mainframe systems
- By 1980/90 servers virtualization adoption initiated a reduction
  - Inexpensive x86 hardware platforms
  - Windows/Linux adopted as server OSs

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Computing Infrastructure – 2000

- 1 machine → 1 OS → several applications
- Applications can affect each other
- Big disadvantage: machine utilization is very low, most of the times it is below than 25%



23

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Virtualization…

x86 server deployments introduced new IT challenges:

- Low server infrastructure utilization (10-18%)
- Increasing physical infrastructure costs (facilities, power, cooling, etc)
- Increasing IT management costs (configuration, deployment, updates, etc)
- Insufficient failover and disaster protection

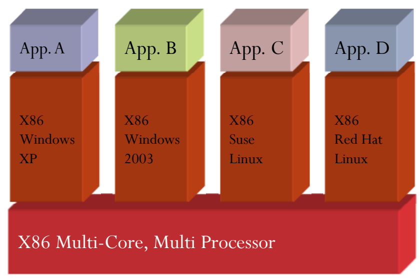
The solution for all these problems was to virtualize x86 platforms

24

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Computing Infrastructure - Virtualization

- It matches the benefits of high hardware utilization with running several operating systems (applications) in separated virtualized environments
  - Each application runs in its own operating system
  - Each operating system does not know it is sharing the underlying hardware with others



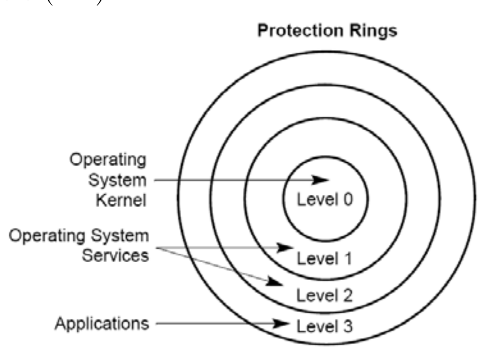
70% Hardware Utilization

25

rce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

x86 modes: Privilege Levels

- x86 processor's segment-protection mechanism recognizes 4 privilege levels (0-high, 3-low level)
- Recognizes the following three types of privilege levels:
  - Current privilege level (CPL)
  - Descriptor privilege level (DPL)
  - Requested privilege level (RPL)



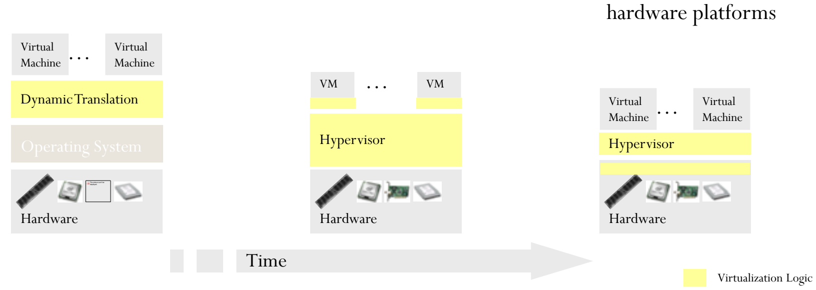
26

rce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Approaches to Server Virtualization

Evolution of Software solutions

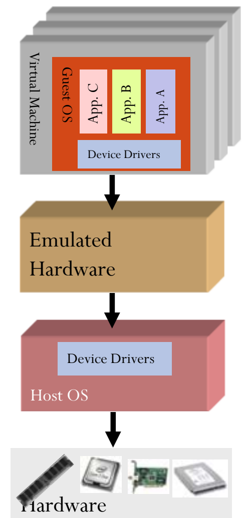
- 1st Generation: Full virtualization (Binary rewriting)
  - Software Based
  - VMware and Microsoft
- 2nd Generation: Para-virtualization
  - Cooperative virtualization
  - Modified guest
  - VMware, Xen
- 3rd Generation: Silicon-based (Hardware-assisted) virtualization
  - Unmodified guest
  - VMware and Xen on virtualization-aware hardware platforms



rce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Full Virtualization

- 1st Generation offering of x86/x64 server virtualization
- Dynamic binary translation
  - The emulation layer talks to an operating system which talks to the computer hardware
  - The guest OS doesn't see that it is used in an emulated environment
- All of the hardware is emulated including the CPU
- Two popular open source emulators are QEMU and Bochs



29

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Full Virtualization - Advantages

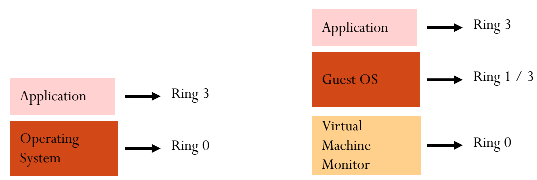
- Emulation layer
  - Isolates VMs from the host OS and from each other
  - Controls individual VM access to system resources, preventing an unstable VM from impacting system performance
- Total VM portability
  - By emulating a consistent set of system hardware, VMs have the ability to transparently move between hosts with dissimilar hardware without any problems
    - It is possible to run an operating system that was developed for another architecture on your own architecture
    - A VM running on a Dell server can be relocated to a Hewlett-Packard server

30

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Full Virtualization - Drawbacks

- Hardware emulation comes with a performance price
- In traditional x86 architectures, OS kernels expect to run privileged code in Ring 0
  - However, because Ring 0 is controlled by the host OS, VMs are forced to execute at Ring 1/3, which requires the VMM to trap and emulate instructions
- Due to these performance limitations, para-virtualization and hardware-assisted virtualization were developed



Traditional x86 Architecture   Full Virtualization

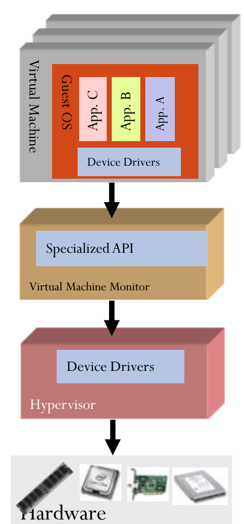
31

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Server virtualization approaches

Para-Virtualization

- The Guest OS is modified and thus run kernel-level operations at Ring 1 (or 3)
  - the guest is fully aware of how to process privileged instructions
  - thus, privileged instruction translation by the VMM is no longer necessary
  - The guest operating system uses a specialized API to talk to the VMM and, in this way, execute the privileged instructions
- The VMM is responsible for handling the virtualization requests and putting them to the hardware



32

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

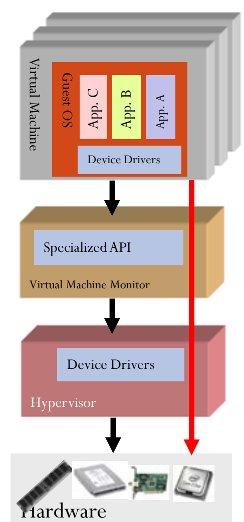
Para-Virtualization

- VM guest operating systems are para-virtualized using two different approaches:
- – Recompiling the OS kernel
  - Para-virtualization drivers and APIs must reside in the guest operating system kernel
  - You do need a modified operating system that includes this specific API, requiring a compiling operating systems to be virtualization aware
    - – Some vendors (such as Novell) have embraced para-virtualization and have provided para-virtualized OS builds, while other vendors (such as Microsoft) have not
- – Installing para-virtualized drivers
  - In some operating systems it is not possible to use complete para-virtualization, as it requires a specialized version of the operating system
  - To ensure good performance in such environments, para-virtualization can be applied for individual devices
  - For example, the instructions generated by network boards or graphical interface cards can be modified before they leave the virtualized machine by using para-virtualized drivers

rce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Hardware-assisted virtualization

- Guest OS runs at ring 0
- The VMM uses processor extensions (such as Intel®-VT or AMD-V) to intercept and emulate privileged operations in the guest
- Hardware-assisted virtualization removes many of the problems that make writing a VMM a challenge
- The VMM runs in a more privileged ring than 0, a virtual -1 ring is created



rce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Hardware-assisted virtualization

- Pros
  - It allows to run unmodified OSs (so legacy OS can be run without problems)
- Cons
  - Speed and Flexibility
    - An unmodified OS does not know it is running in a virtualized environment and so, it can't take advantage of any of the virtualization features
      - It can be resolved using para-virtualization partially

35

Source: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Approaches to desktop virtualization

36

Extending the concept of virtualization for desktops

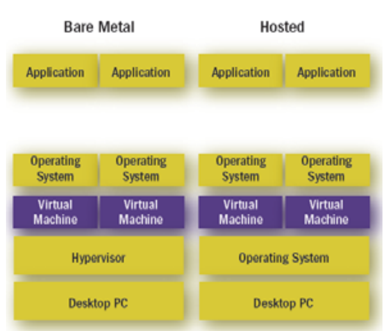
- Servers
  - Hosted virtualization - mainframes
  - VMMs / Bare Metal hypervisors
  - OS virtualization
- Desktops
  - Desktop virtualization
  - Server-side workspace virtualization
  - Client-side workspace virtualization
- Application virtualization
  - Application isolation
  - Application streaming

37

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Desktop Virtualization

- A VMM or hypervisor running on a physical desktop
- Examples include:
  - Microsoft Virtual PC
  - Parallels Desktop for Mac
  - VMware Fusion
  - WINE.
- Use cases include:
  - Emulating Windows games on the Macintosh,
  - Testing code inside VMs
  - Underpinning client-side workspace virtualization
- Desktop hypervisors and VMMs don't necessarily scale to meet enterprise needs; that's why most of the providers have server products as well

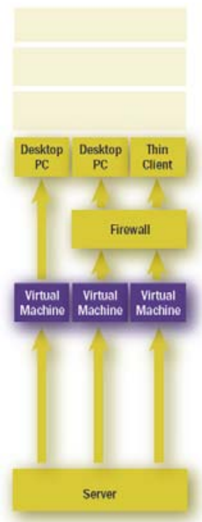


38

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Server-side workspace virtualization

- A workspace (desktop operating system with custom configuration) running inside a virtual machine hosted on a server
- Examples include:
  - VMware VDI
- Use cases include:
  - Centrally managed desktop infrastructure
  - Security enforcement and lockdown
- A pool of virtual workspaces resides on the server. Remote users log into them from any networked device via Microsoft's Remote Desktop Protocol (RDP)
- Users can customize their virtual workspace to their heart's content, while operators enjoy the relatively straightforward task of managing desktop configuration on one central server
- Connection brokers arbitrate between a pool of virtual workspaces residing on a central server
- The biggest problem with server-hosted workspace virtualization is that it's a bandwidth hog. Performance is constrained by the performance of your network

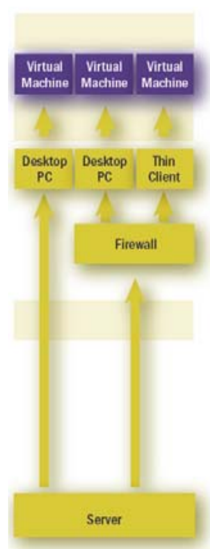


39

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Client-side workspace virtualization

- A workspace (desktop operating system with custom configuration) running inside a virtual machine hosted on a desktop
- Use cases include:
  - Secure remote access
  - Protection of sensitive data for defense, healthcare industries
  - Personal computer running corporate desktops remotely
- A virtual workspace is served out to execute on the client device
- Centralizes management
- Its big advantage over other models is the security and isolation of data and logic on the client
- It's the right model for organizations that need to ensure the security of environments served to remote users
  - Defense contractors
  - Healthcare providers



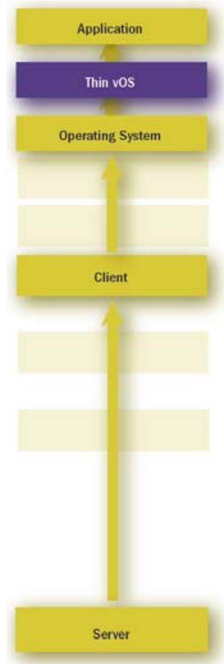
40

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Application Isolation

- An application packaged with its own virtual copies of the operating system resources it might otherwise need to change (registries, file systems, libraries)
- Examples include:
  - Thinstall
  - Trigence
- Use cases include:
  - Preventing DLL hell
  - Sandboxing desktop applications for secure execution

- Applications use a virtual registry (Thinstall) and file system embedded in the package with the application
  - These extra tools insulate applications from changes to and incompatibility with the underlying desktop operating system
- Mostly in Windows, although Linux and Solaris as well
- Drawback: increased footprint of the application package and the correspondingly greater memory requirements



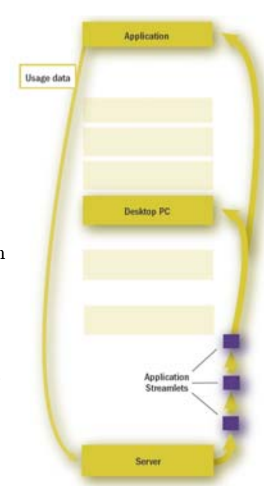
41

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Desktop virtualization approaches

Application Streaming

- Just-in-time delivery of a server-hosted application to the desktop, such that the desktop application can execute before the entire file has been downloaded from the server
- Examples include: AppStream
- Use cases include:
  - Managing the number of instances of running applications, in the case of license constraints
- Superset of Application Isolation, including a delivery method and an execution mode
  - You stream the application code to the desktop, where it runs in isolation
- No full PC environment, just the application, so you have to provide a workspace
  - Requires to maintain the client-side operating system and ensuring compatibility. This may be why application streaming, which has been around for a long time , has not really lived up to its early hype.



42

urce: www.dc.uba.ar/events/eci/2008/courses/n2/Virtualization-Introduction.ppt

Uses of virtualization

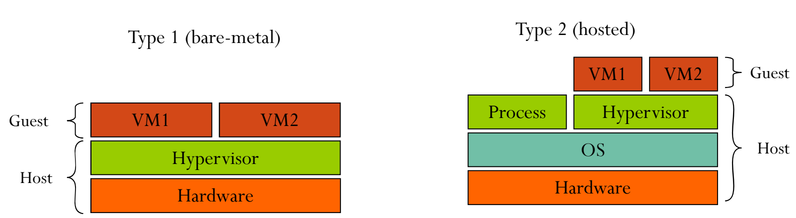
- Server consolidation
  - Run a **web server** and a **mail server** on the **same physical server**
- Easier development
  - Develop critical **operating system components** (file system, disk driver) without affecting **computer stability**
- QA
  - Testing a network product (e.g., a firewall) may require **tens of computers**
  - Try testing thoroughly a product at each pre-release milestone… and have a straight face when your boss shows you the **electricity bill**
- Cloud computing
  - The modern buzz-word
  - Amazon sells computing power
  - You pay for e.g., 2 CPU cores for 3 hours plus 10GB of network traffic

43

Source: http://webcourse.cs.technion.ac.il/234123/Spring2010/ho/WCFiles/virtualization.ppt

Two types of hypervisors

- Definitions
  - **Hypervisor** (or **VMM** – Virtual Machine Monitor) is a software layer that allows several **virtual machines** to **run** on a **physical machine**
  - The physical OS and hardware are called the **Host**
  - The virtual machine OS and applications are called the **Guest**



VMware ESX, Microsoft Hyper-V, Xen

VMware Workstation, Microsoft Virtual PC, Sun VirtualBox, QEMU, KVM

44

Source: http://webcourse.cs.technion.ac.il/234123/Spring2010/ho/WCFiles/virtualization.ppt

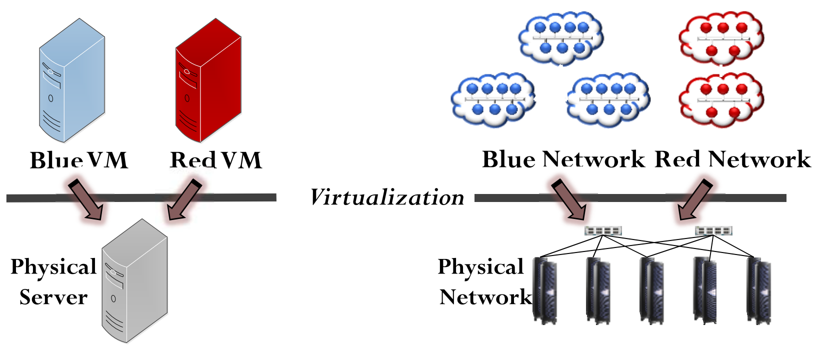
Bare-metal or hosted?

- **Bare-metal**
  - Has complete **control over hardware**
  - Doesn't have to **"fight" an OS**
- **Hosted**
  - Avoid **code duplication**: need not code a **process scheduler**, **memory management** system – the **OS already does** that
  - Can run native **processes alongside** VMs
  - Familiar environment – **how much CPU** and **memory** does a VM take? Use **top**! How big is the **virtual disk**? **ls –l**
  - Easy management
- A combination
  - Mostly hosted, but some parts are inside the OS kernel for performance reasons
  - E.g., **KVM**

45

Source: http://webcourse.cs.technion.ac.il/234123/Spring2010/ho/WCFiles/virtualization.ppt

Network Virtualization



Server Virtualization

- Run multiple virtual servers on a physical server
- Each VM has illusion it is running as a physical server

Network Virtualization

- Run multiple virtual networks on a physical network
- Each virtual network has illusion it is running as a physical network

46

Source: http://www.microsoft.com/en-us/download/details.aspx?id=34782

Network Virtualization Benefits

To Workload Owners
- Seamless migration to the cloud
- Move n-tier topology to the cloud
- Preserve policies, VM settings, IP addresses

To Enterprises
- Private Cloud datacenter consolidation and efficiencies
- Extension of datacenter into hybrid cloud
- Incremental integration of acquired company network infrastructure

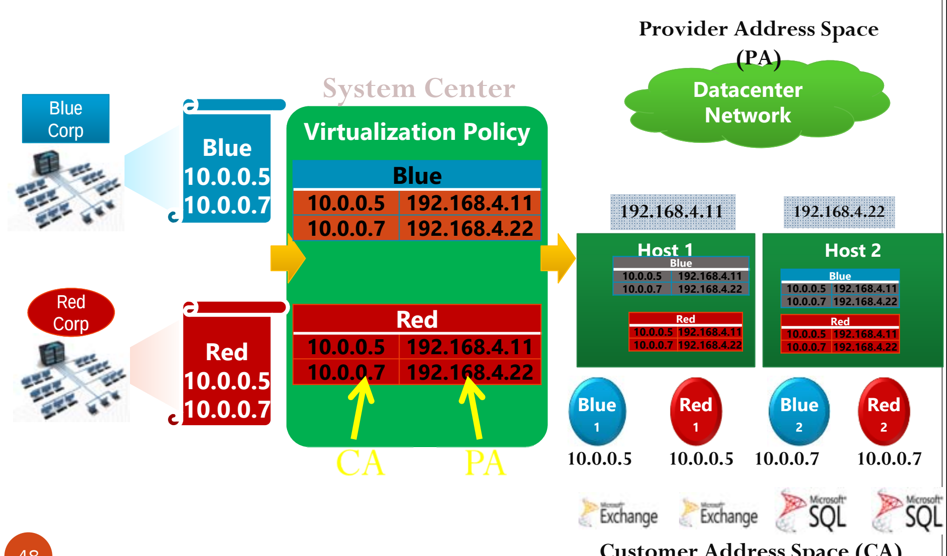
To Hosters
- Bring Your own IP
- Bring Your network topology
- Scalable multi-tenancy

To Private/Public Cloud Datacenter Admins
- Flexible VM placement without reconfiguration
- Decoupling of server and network admin roles increases agility

47

Source: http://www.microsoft.com/en-us/download/details.aspx?id=34782

Virtualize Customer Addresses

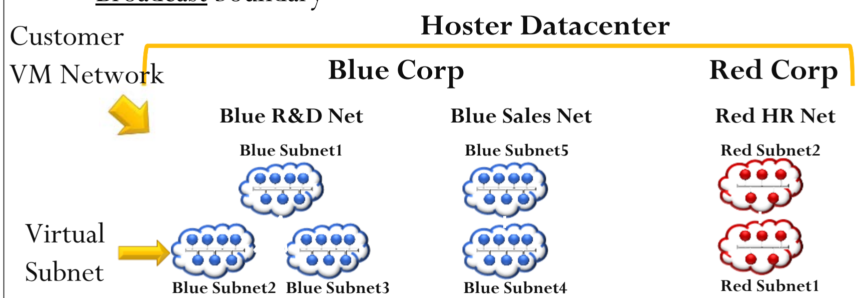


48

Source: http://www.microsoft.com/en-us/download/details.aspx?id=34782

Network Virtualization Concepts

- Customer VM Network
  - One or more virtual subnets forming an isolation boundary
  - A customer may have multiple Customer VM Networks
    e.g. Blue R&D and Blue Sales are isolated from each other
- Virtual Subnet
  Broadcast boundary



49

Source: http://www.microsoft.com/en-us/download/details.aspx?id=34782

Storage Virtualization

50

Storage Virtualization *is the next frontier in Storage Advances that* ***aims to provide a layer of*** abstraction to reduce complexity.

Storage Networking Industry Association (SNIA) defines Storage Virtualization as:

1. The act of abstracting, hiding, or isolating the internal functions of a storage (sub) system or service from applications, host computers, or general network resources, for the purpose of enabling application and network-independent management of storage or data.

2. The application of virtualization to storage services or devices for the purpose of aggregating functions or devices, hiding complexity, or adding new capabilities to lower level storage resources.

51

Storage Virtualization

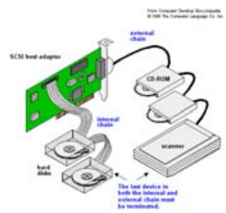
Server side storage innovations… a combination of **storage devices**, **storage interfaces** and **storage software** innovations have helped enterprises cope with exponential growth of data storage requirement !

Storage devices have evolved from tapes to hard drives to RAID hard drives increasing capacity and resiliency.

52

<u>Storage interface innovations</u> have evolved from SCSI to ISCI, Fiber Channel (FCP) and Infini-Band to inter **connect devices and transport the data faster.**

SCSI



ISCSI

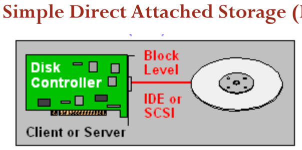
FCP

Infiniband

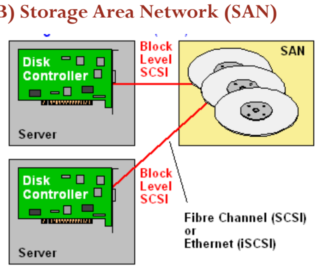
53

<u>Storage Software</u> from simple back-up and restore to advanced <u>storage networks</u> and <u>storage management</u> software functions.

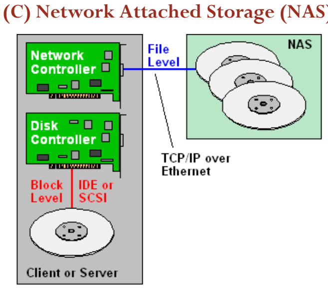
(A) Simple Direct Attached Storage (DAS)



(B) Storage Area Network (SAN)



(C) Network Attached Storage (NAS)

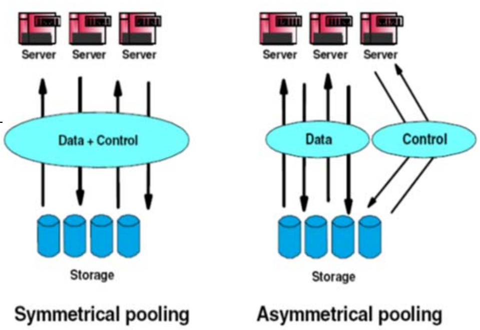


54

How storage is virtualized at the enterprise level

Currently Networks are virtualized using Metadata or Storage Volume Controllers. There are two types of network virtualization…

Metadata or Storage Volume Controllers (SVC) are placed (in-band) or in the path of data flow.



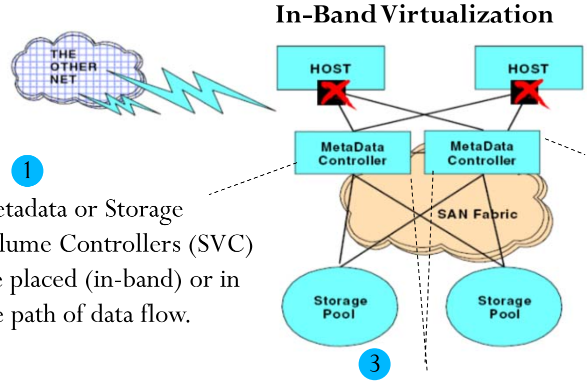
Metadata or Storage Volume Controllers are placed (out of band) outside the path of data flow.

55

urce: IBM Redbook Page 8 "http://www.redbooks.ibm.com/redbooks/pdfs/sg246210.pdf"

How storage is virtualized at the enterprise level

In-Band Virtualization



1 Metadata or Storage Volume Controllers (SVC) are placed (in-band) or in the path of data flow.

2 SVC controls who can get access to the storage device controls, how storage can be accessed, how storage is allocated, etc.

3 SVC are managed through Storage Management Software.

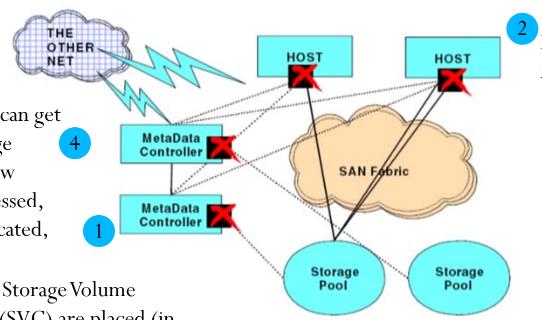
4 Key Challenge is the potential IO bottlenecks

56

Source: http://www.redbooks.ibm.com/redbooks/pdfs/sg246210.pdf

How storage is virtualized at the enterprise level

Out-of-Band Network Virtualization



Host sends Metadata to SVC

SVC controls who can get access to the storage device controls, how storage can be accessed, how storage is allocated, etc.

Metadata or Storage Volume Controllers (SVC) are placed (in-band) or in the path of data flow.

Storage Pool sends Metadata to SVC

57

Source: http://www.redbooks.ibm.com/redbooks/pdfs/sg246210.pdf

Definition

- A SAN (Storage Area Network) is a network designed to transfer data from servers to targets, and it is alternative to a directly attached target architecture, or to a DAS architecture, where the storage is connected to the servers on general purpose networks
- Additional definitions of a SAN imply that the SAN should also be highly performing, and should be such to enable storage devices to communicate with one another and with computer systems

58

Storage Area Network(SAN)

59

SAN

- Multiple technology can be used when building a SAN; traditionally the dominant technology is Fiber Channel, but IP based solutions are also quite popular for specific applications
- The concept of SAN is also independent from the devices that are attached to it. Can be disks, tapes, RAIDs, file servers, or other

60

Basic Building Blocks of SAN

SAN interconnections

- Different technologies can be used to interconnect the network nodes, extending the Disk interface outside the server
- Fiber Channel is a dedicated channel based high performance and highly available network based on Fiber Channel Protocols
- iSCSI is SCSI protocol carried over an IP network. In this case the network infrastructure can be shared with other applications
- SCSI network is an extension of the internal SCSI bus, used for short distances due to its parallel architecture

Initiator and Target in FC SANs

- Fiber Channel Node: can be the source or the destination of information
- If the node is an Initiator (source), it is usually connected to the network via an HBA (Host Bus Adaptor), which is the physical connection interface, and can be based either on electrical or (more often) optical technology
- If the node is a target (destination), it can be a JBOD (Just a Bunch of Disks), a RAID (Redundant Array of Independent Disks), or a Storage array

63

Hard Drives

- The basic storage element is an Hard Drive. They are made into complex devices composed of platters, heads, cylinders and tracks
- The Logical Block Addressing (LBA) addresses the sector within the disk. Modern drives have 512 byte sectors
- File systems arrange files into sectors so that they can be stored and retrieved
- The File system usually deals with clusters of blocks and uses a FAT (File Allocation Table) to map a file to the sectors

64

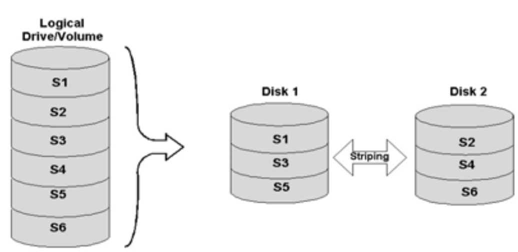
JBODs and RAIDs

- While a Jbod is a group of disks packaged in an enclosure and connected via a FC loop, a RAID is a more sophisticated device, that may improve performance and/or reliability of the storage device
- RAID is improving performances reading/writing information from a set of disks at the same time, and reliability adding parity and/or mirroring information on multiple disks of the array
- RAID can be performed in HW via a controller embedded in the enclosure or software on the host

65

RAID 0 or striping

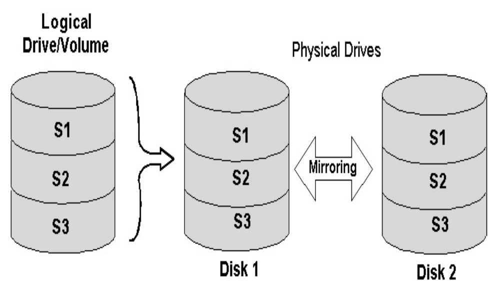
- Data are split onto different disks for performance increase: performances depend on information unit size vs stripe size
- No redundancy added
- Cost is limited (no additional hardware)



66

RAID 1 - mirroring

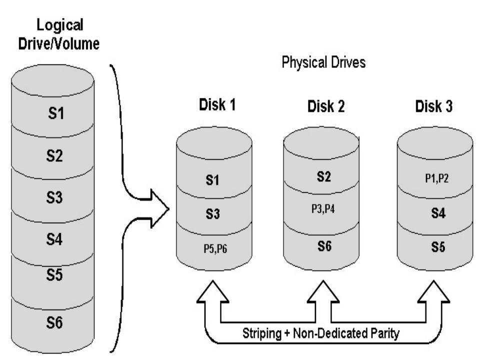
- Data are replicated on multiple disks for redundancy
- Performance may be impacted if copy is done serially
- Increase of cost proportional to the amount of redundancy
- More complex algorithm to manage multiple copies



67

RAID 3

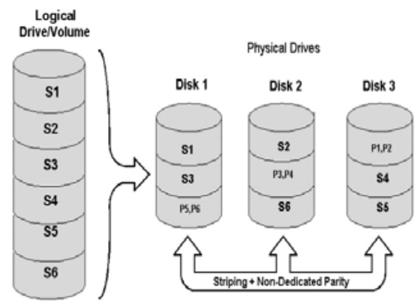
- Data protection via ECC (Error Correction Control code)
- Good redundancy
- Performance not changed for reads but lower for writes since the ECC need to be calculated
- Cost is only 1 extra disk for the entire logical array



68

RAID 5

- Data protection with ECC, but parity is spread on the array
- Good redundancy
- Same speed reads, slower writes
- One disk per array of added cost



69

Storage Architecture

70

RAID array vs DAS vs NAS vs SAN (I)

- A RAID array is an enclosure containing a set of disks and a RAID controller providing in hardware the features of a RAID 0 – 5 and usually some caching engine
- A DAS (Direct Attached Storage) is an architecture for which the storage is "privately" attached to the servers: cannot be shared, it is hard to scale, expensive and complex to manage. 80% of the market it is still DAS

71

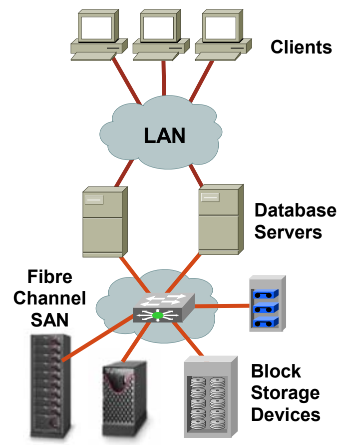
RAID array vs DAS vs NAS vs SAN (II)

- NAS (Netwrok Attached Storage) is an architecture for which the storage is attached to the servers via a multi-purpose network, and it is accessed at a file level via protocols like CIFS or NFS
- The network is usually an IP network
- TCP can be tuned to optimize storage transport

72

SAN architecture

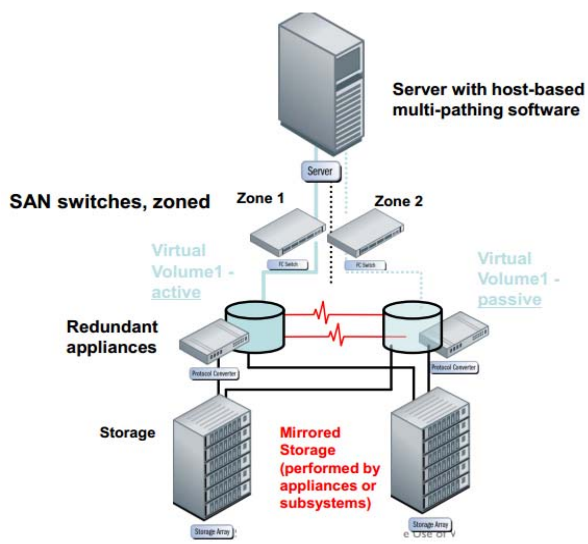
- Storage is accessed at block level not at file level
- Very high performances
- Storage is shared
- Good management tools
- Interoperability issues



Storage Area Network (SAN)

73

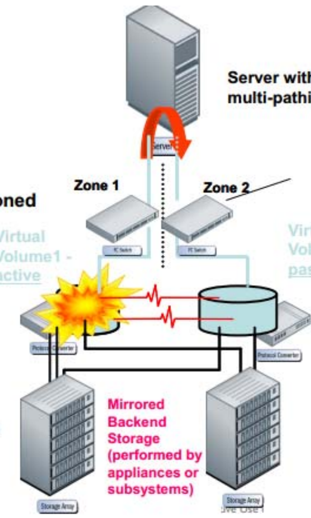
Achieving High Availability



74

Source: RobPeglar, "Storage Virtualization-II Effective use of virtualizaition" Xiotech Corporation

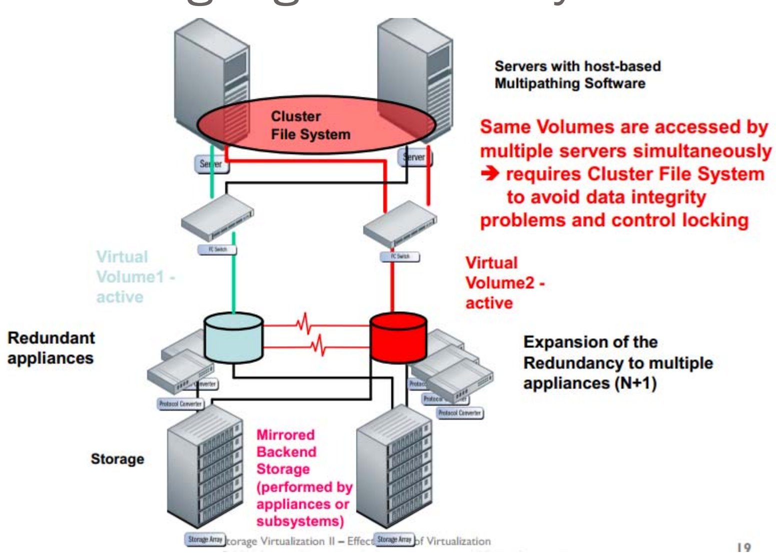
Achieving High Availability



75

Source: RobPeglar, "Storage Virtualization-II Effective use of virtualizaition" Xiotech Corporation

Achieving High Availability



76

Source: RobPeglar, "Storage Virtualization-II Effective use of virtualizaition" Xiotech Corporation

Achieving High Availability

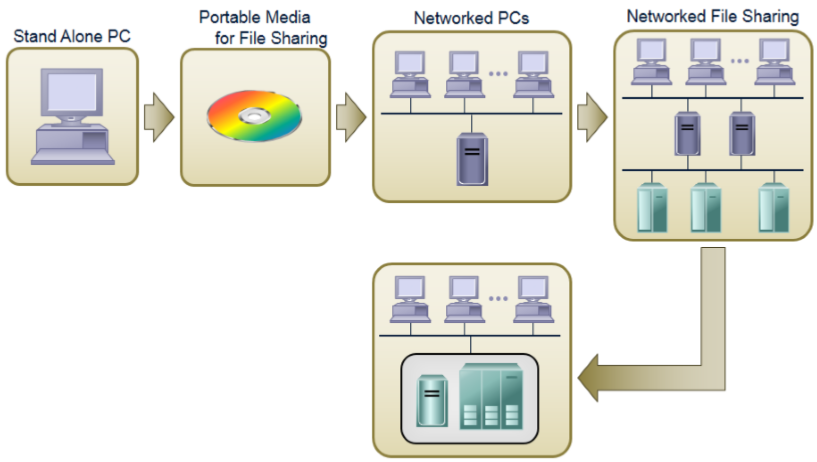
- To support high-availability configurations, virtual storage management must be distributed across two or more switches

77

Network Attached Storage(NAS)

78

NAS Evolution

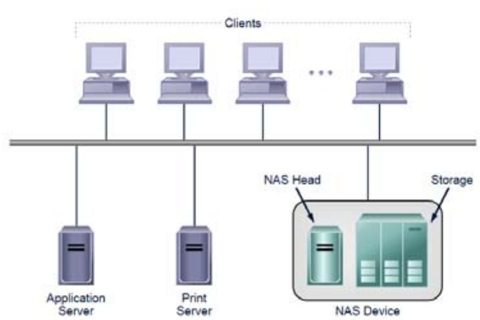


79

Source: EMC² Corporation, "Network Attached Storage"

What is NAS?

- NAS is shared storage on a network
- A NAS server is a storage device that consists of a high performance file server and attached to a LAN
- NAS head (as illustrated) could be remote from its storage (gateway) or contained within the same cabinet as its storage

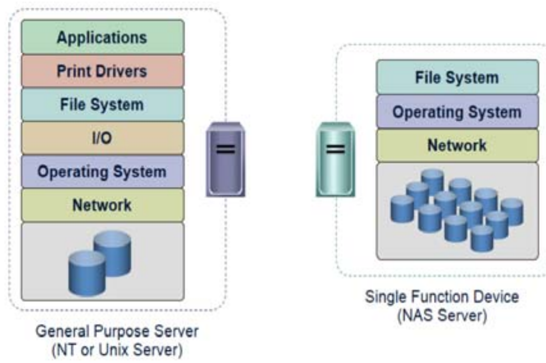


80

Source: EMC² Corporation, "Network Attached Storage"

General purpose Servers vs NAS

- A single function NAS device provides:
  - Real-time OS dedicated to file serving
  - Open standard protocols
  - Built-in native clustering for high availability

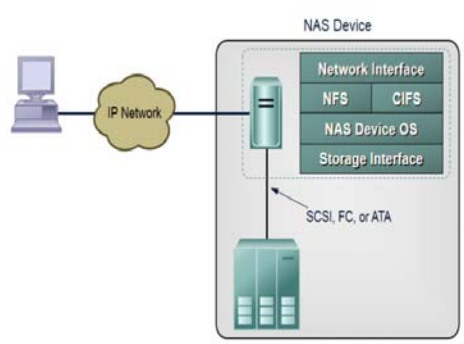


81

Source: EMC² Corporation, "Network Attached Storage"

NAS Device Components

- A NAS device is made up of the following components:
  - Network Interface via one or more Network Interface Cards (NICs)
    - Examples: Gigabit Ethernet (1000 Mb/s), Fast Ethernet (10Mb/s), ATM, and FDDI.
  - Network File Systems (NFS) and Common Internet File Systems (CIFS) protocols



82

Source: EMC² Corporation, "Network Attached Storage"

NAS Device Components

- Proprietary, optimized Windows, UNIX, or LINUX based OS. Examples:
  - DART - Data Access in Real Time (EMC)
  - Data ONTAP (Network Appliance)
- Industry standard storage protocols to connect to and manage physical disk storage resources.
  - Examples: Serial ATA (SATA), SCSI, or Fibre Channel

83