1 Software as a Service (SaaS)
2 Platform as a Service (PaaS)
CLOUD COMPUTING
3 Infrastructure as a Service (IaaS)
1 On-demand Self-service
2 Resource Pooling
3 Broad Network Access
4 Rapid Elasticity
5 Measured Services

PMCA506L: Cloud Computing

Module 3 : Virtual Machines

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Courtesy : *Ming Lian , Dogules E Comer & Other Sources of Internet*

Dr. R. K. Nadesh

Virtualization

- Virtualization is a broad term that refers to the abstraction of resources across many aspects of computing

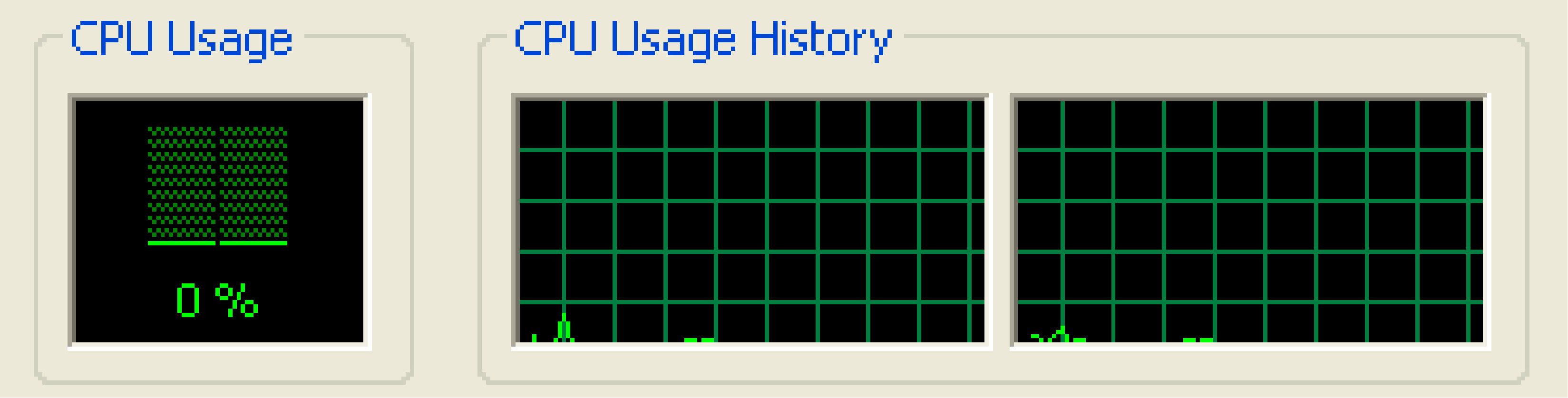
- For our purposes - One physical machine to support multiple virtual machines that run in parallel.

Problem Assessment – Why Virtualization?

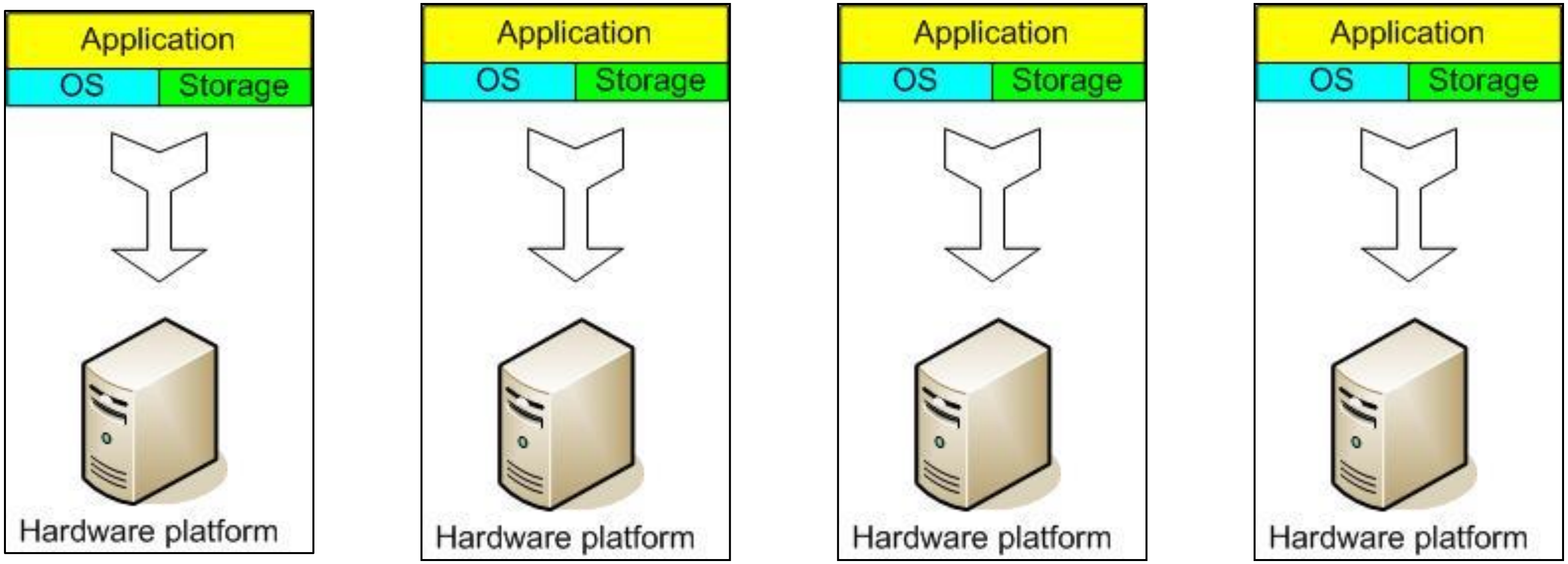
- Too many servers for too little work
- Aging hardware reaching end of usable life
- High infrastructure requirements
- Limited flexibility in shared environments

Problem Assessment

Low utilization metrics in servers across the organization…

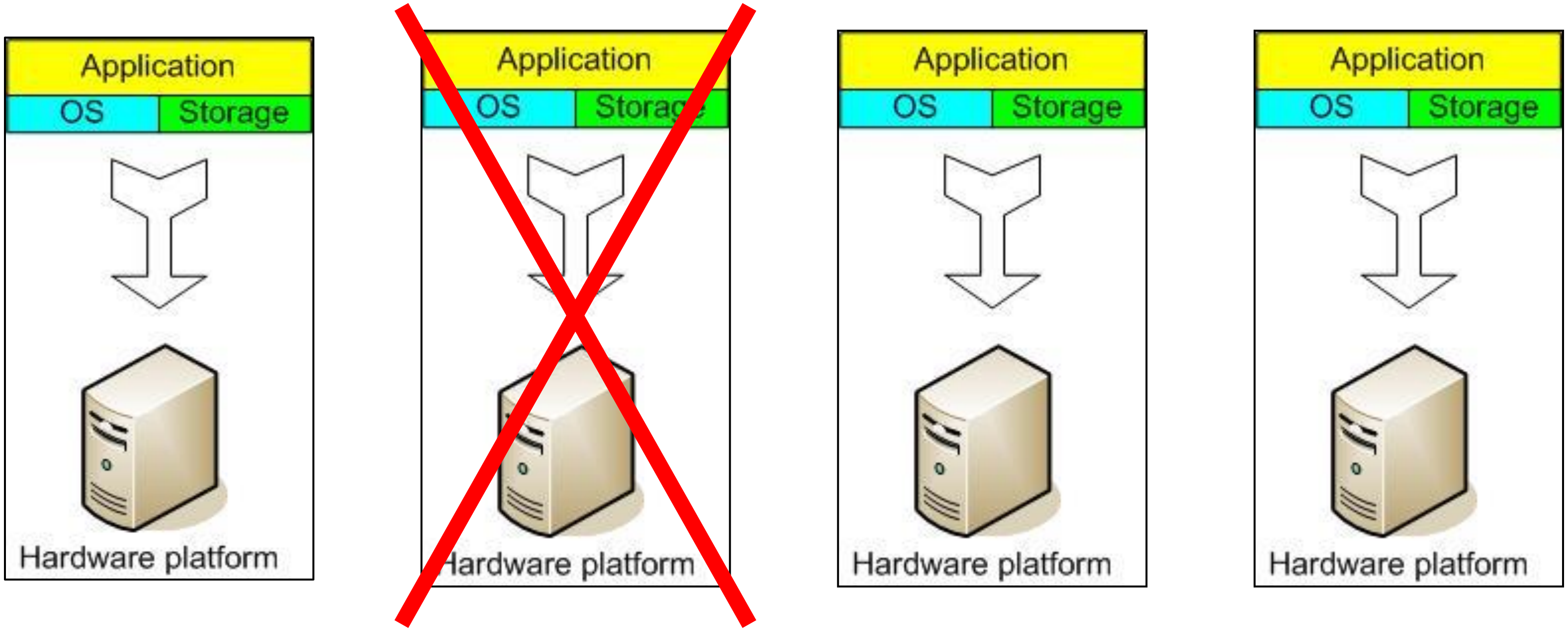


The Traditional Server Concept



Web Server	App Server	DB Server	EMail
Windows	Linux	Linux	Windows
IIS	Glassfish	MySQL	Exchange

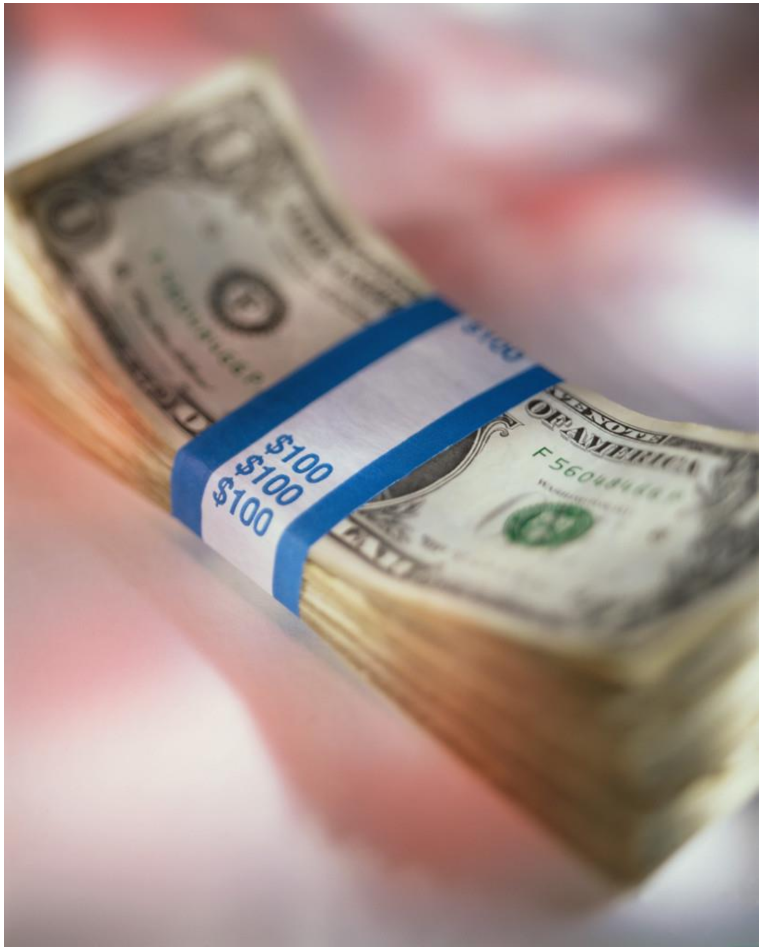
And if something goes wrong …



Web Server	App Server	DB Server	EMail
Windows	DOWN!	Linux	Windows
IIS		MySQL	Exchange

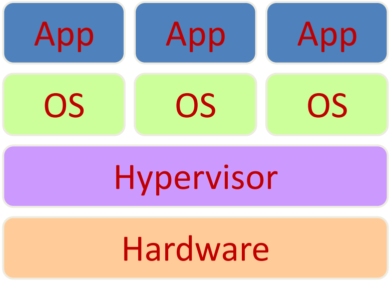
Problem Assessment

- High costs and infrastructure needs
  - Maintenance
  - Leases
  - Networking
  - Floor space
  - Cooling
  - Power
  - Disaster Recovery



Virtualization

- Virtual workspaces:
  - An abstraction of an execution environment that can be made dynamically available to authorized clients by using well-defined protocols,
  - Resource quota (e.g. CPU, memory share),
  - Software configuration (e.g. O/S, provided services).
- Implement on Virtual Machines (VMs):
  - Abstraction of a physical host machine,
  - Hypervisor intercepts and emulates instructions from VMs, and allows management of VMs,
  - VMWare, Xen, etc.
- Provide infrastructure API:
  - Plug-ins to hardware/support structures

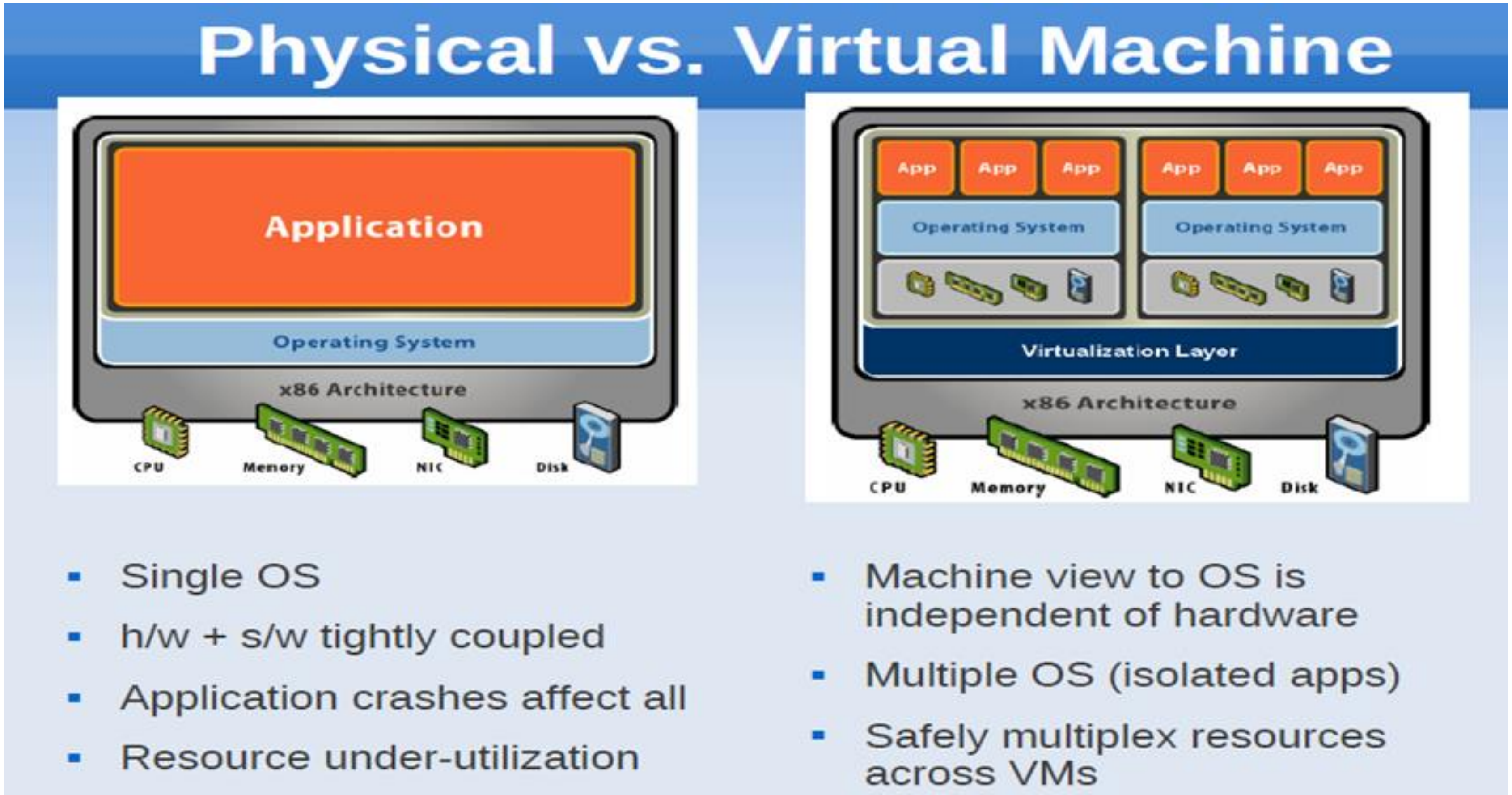


Virtualized Stack

Core Technology

- The Hypervisor
  - A computing layer which allows multiple operating systems to run on a host computer at the same time.
  - Originally developed in the 1970s as part of the IBM S/360
  - Many modern day variants from different developers

Comparison



Physical vs. Virtual Machine

- Single OS
- h/w + s/w tightly coupled
- Application crashes affect all
- Resource under-utilization

- Machine view to OS is independent of hardware
- Multiple OS (isolated apps)
- Safely multiplex resources across VMs

Uses of Virtualization

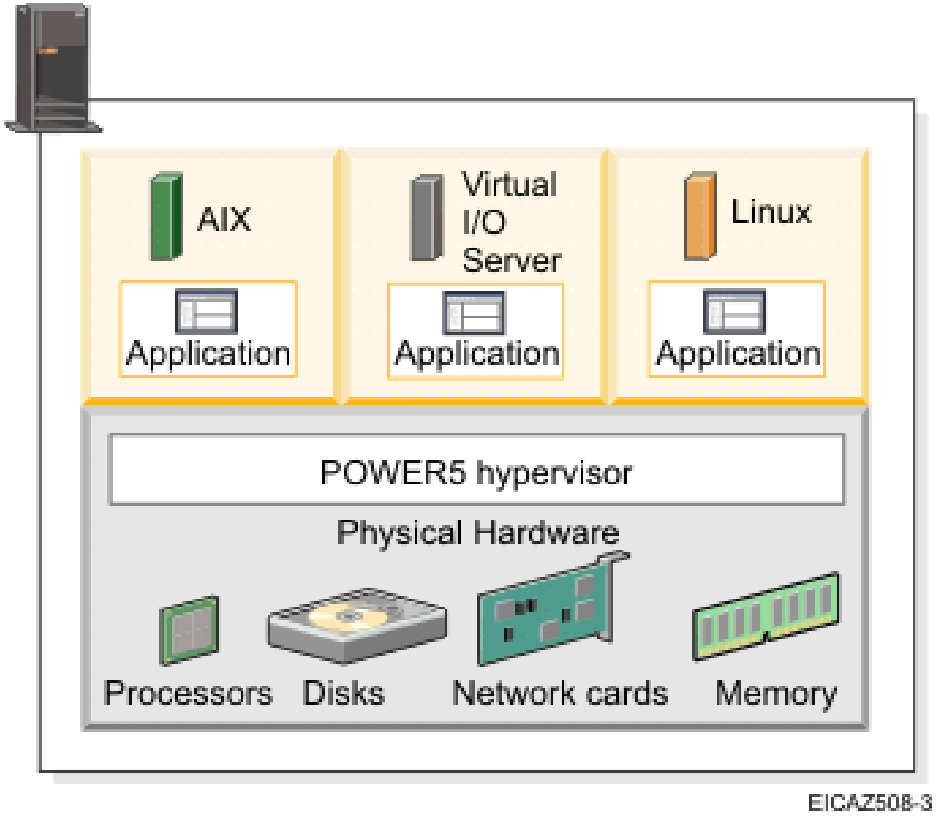
- Server consolidation
  - Run a **web server** and a **mail server** on the **same physical server**
- Easier development
  - Develop critical **operating system components** (file system, disk driver) without affecting **computer stability**
- QA
  - Testing a network product (e.g., a firewall) may 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

Virtualization Scenarios

- Hardware Virtualization
- Software Virtualization
  - Full Virtualization
  - Para-Virtualization

Hardware Virtualization (example)

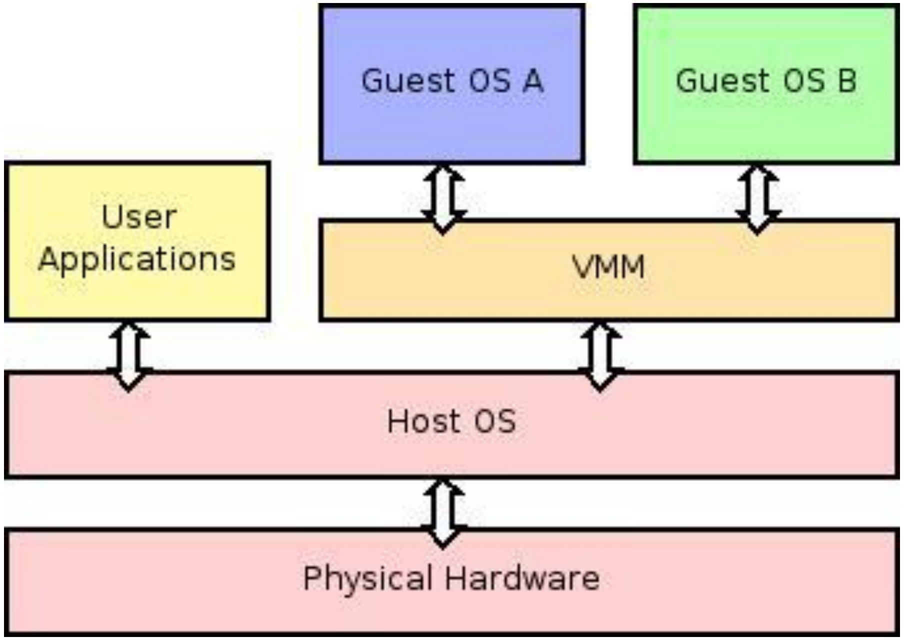
- IBM pSeries Servers



http://publib.boulder.ibm.com/infocenter/eserver/v1r2/topic/eicaz/eicaz508.gif

Software Virtualization (example)

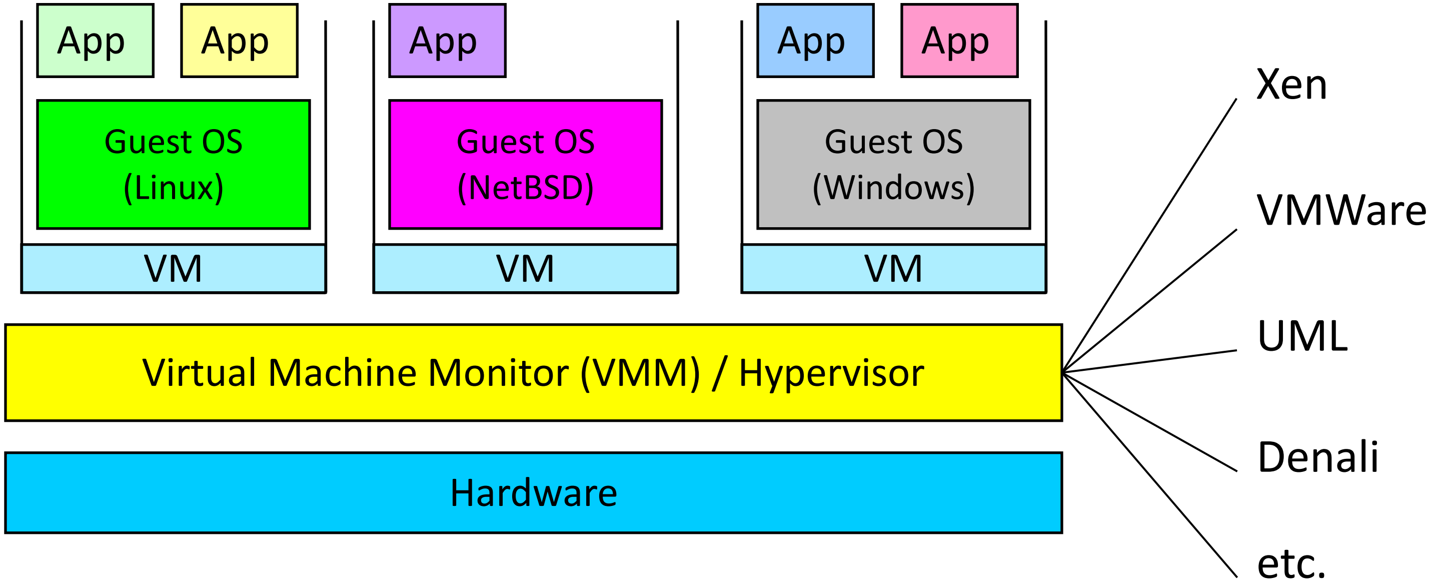
- VMware Server (GSX)



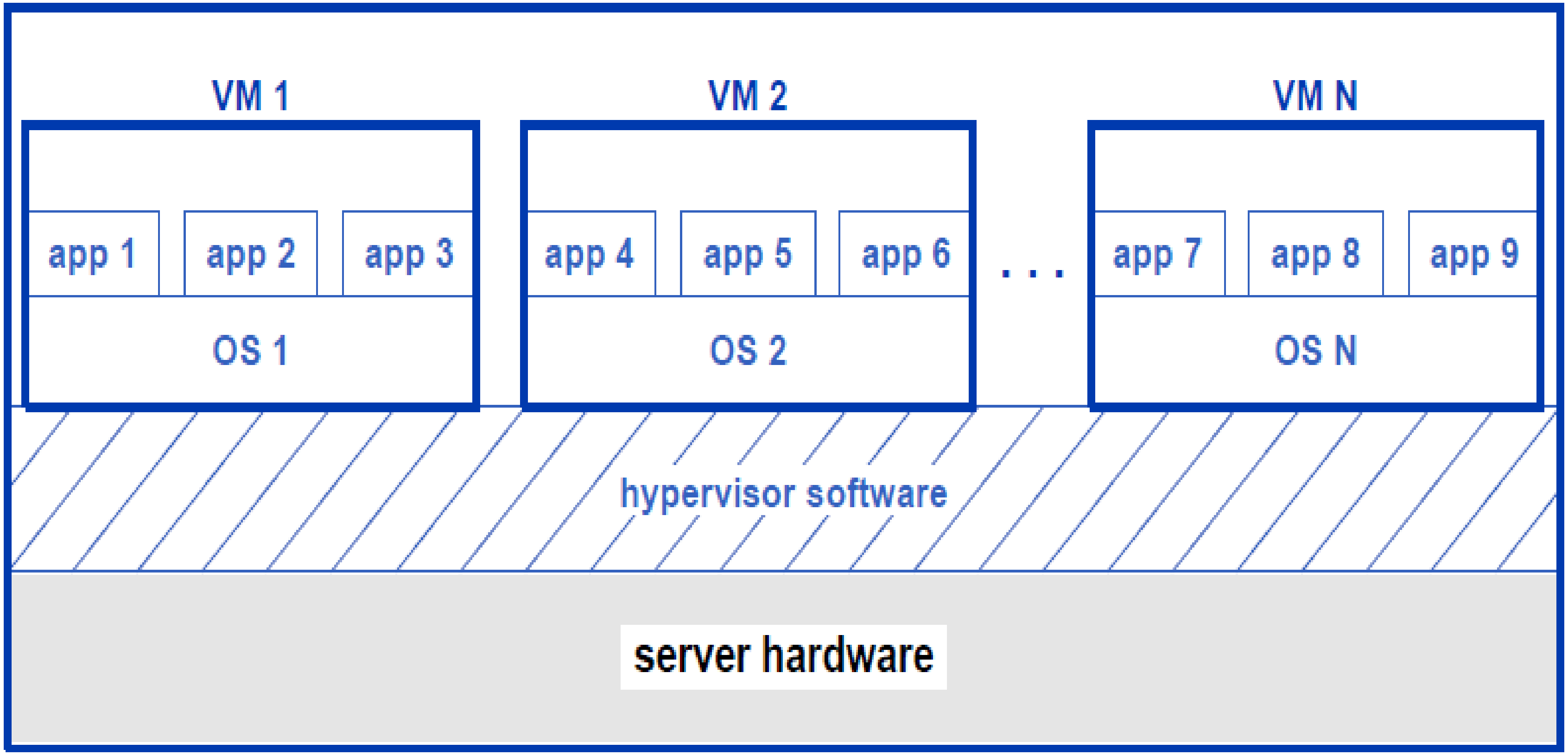
http://openlab-mu-internal.web.cern.ch/openlab-mu-internal/openlab-II_Projects/Platform_Competence_Centre/Virtualization/Virtualization.asp

Virtual Machines

- VM technology allows multiple virtual machines to run on a single physical machine.



Conceptual Organization Of VM Systems

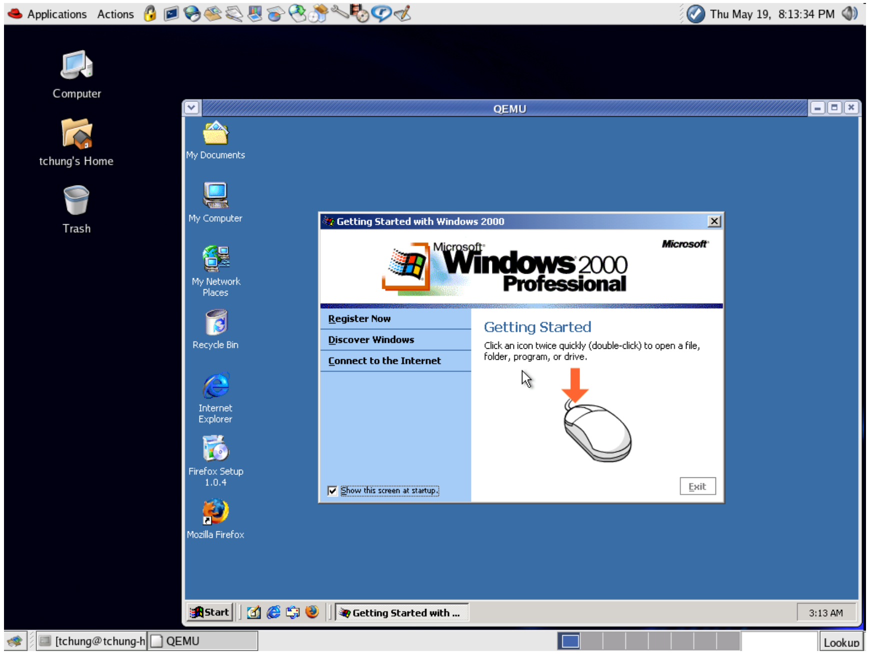


Advantages of virtual machines

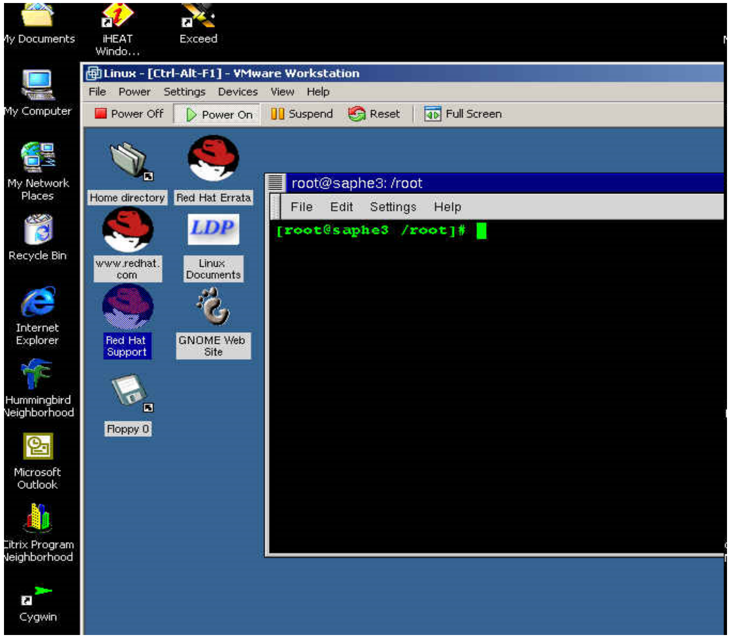
- Run operating systems where the physical hardware is unavailable,
- Easier to create new machines, backup machines, etc.,
- Software testing using “clean” installs of operating systems and software,
- Emulate more machines than are physically available,
- Timeshare lightly loaded systems on one host,
- Debug problems (suspend and resume the problem machine),
- Easy migration of virtual machines (shutdown needed or not).
- Run legacy systems!

How Virtualization Different from Dual Boot ?

- 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

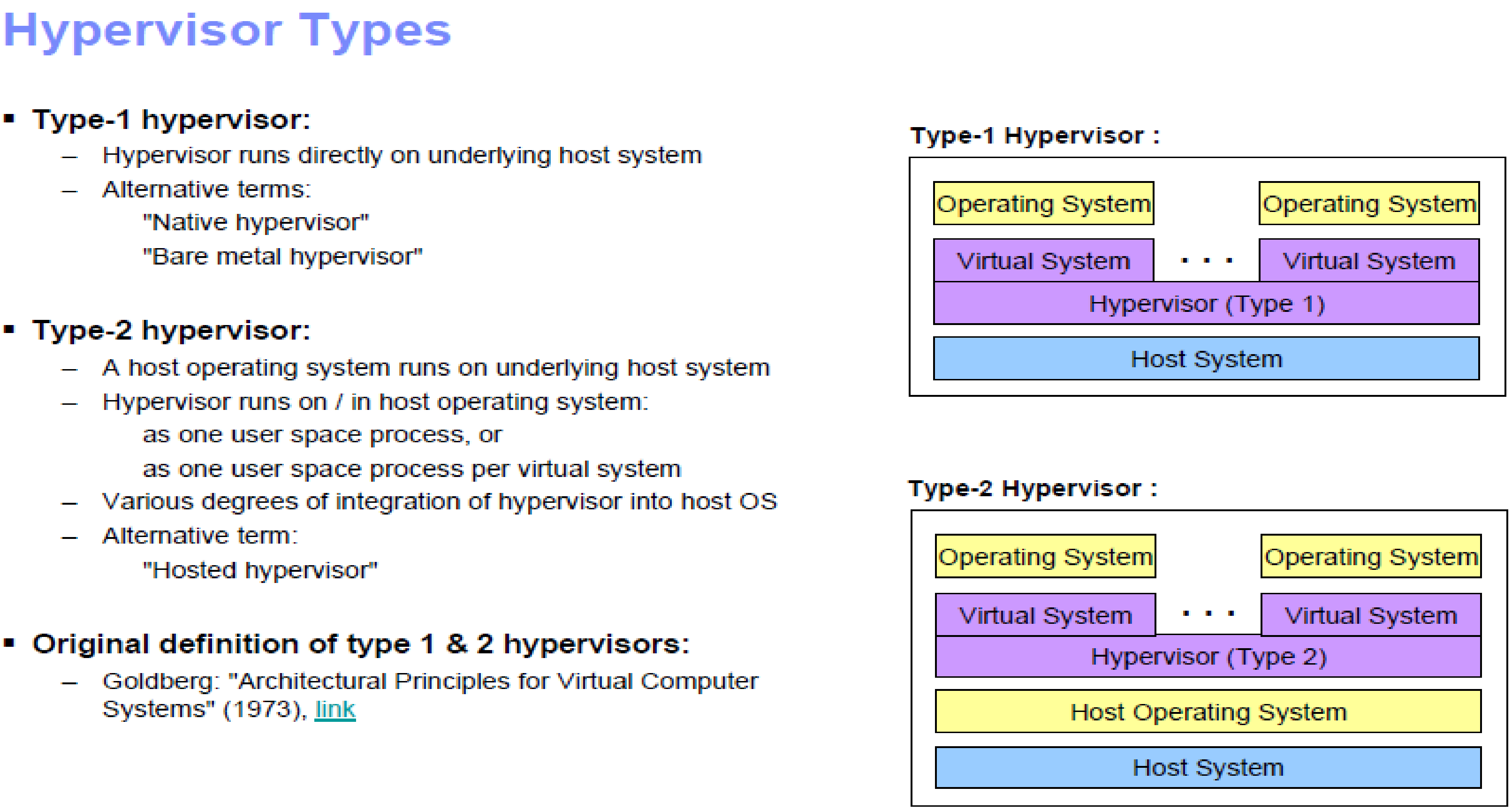


Click an icon twice quickly (double-click) to open a file, folder, program, or drive.



```
[root@saphe3 /root]#
```

HyperVisor (Virtual Machine Monitor)



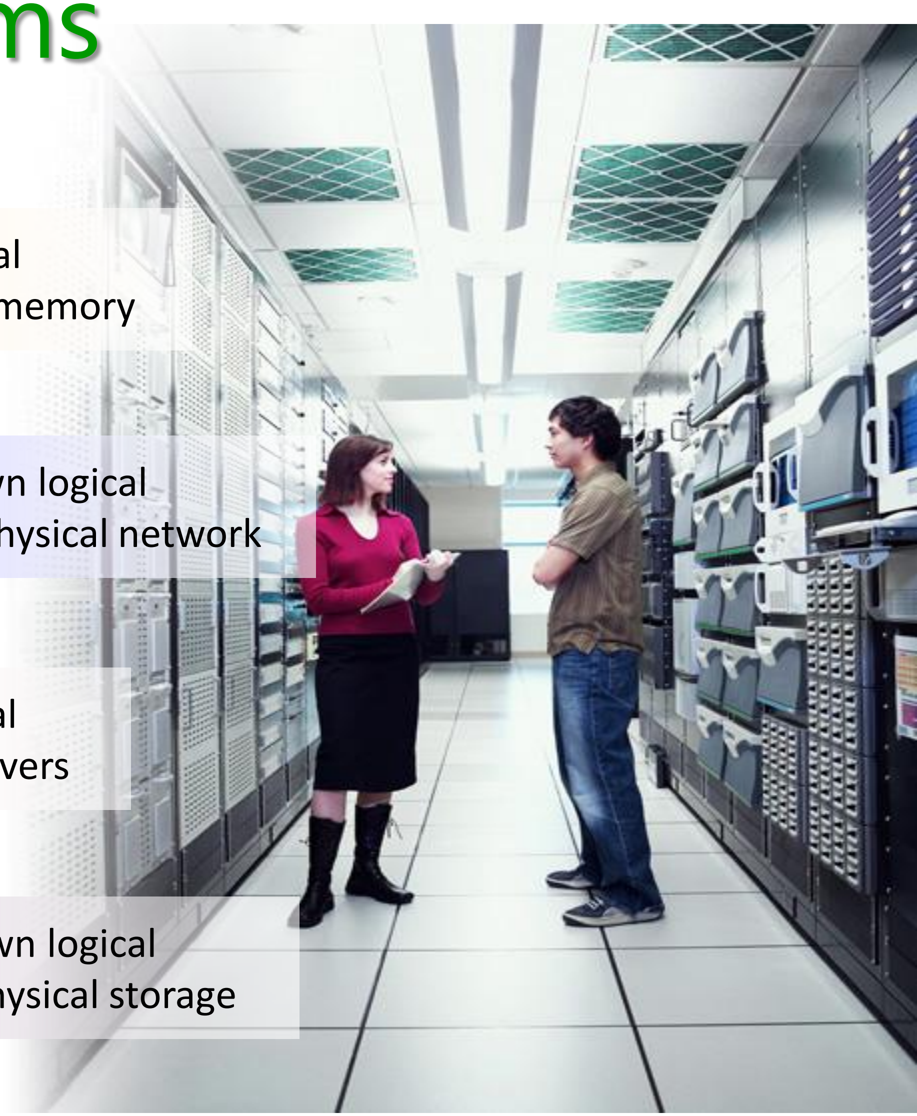
Hypervisor Types

- **Type-1 hypervisor:**
  - Hypervisor runs directly on underlying host system
  - Alternative terms:
    "Native hypervisor"
    "Bare metal hypervisor"
- **Type-2 hypervisor:**
  - A host operating system runs on underlying host system
  - Hypervisor runs on / in host operating system:
    as one user space process, or
    as one user space process per virtual system
  - Various degrees of integration of hypervisor into host OS
  - Alternative term:
    "Hosted hypervisor"
- **Original definition of type 1 & 2 hypervisors:**
  - Goldberg: "Architectural Principles for Virtual Computer Systems" (1973), link

Type-1 Hypervisor :

Type-2 Hypervisor :

Virtualization Comes in Many Forms



Virtual Memory

Each application sees its own logical **memory,** independent of physical memory

Virtual Networks

Each application sees its own logical **network,** independent of physical network

Virtual Servers

Each application sees its own logical **server,** independent of physical servers

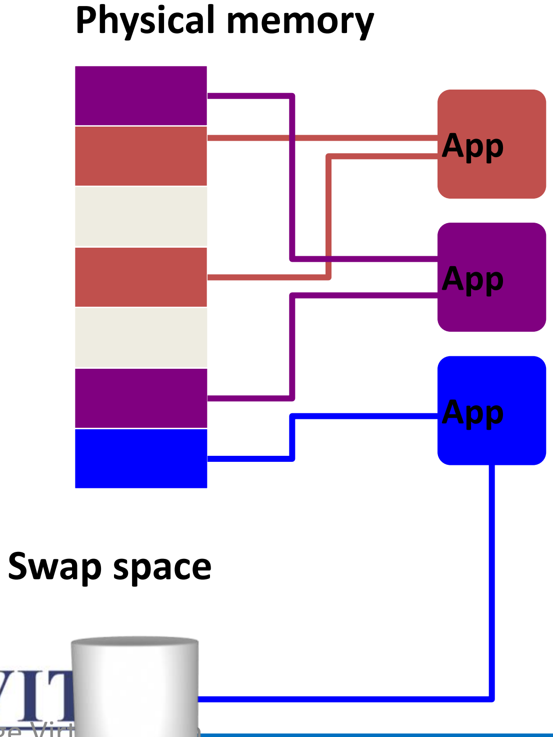
Virtual Storage

Each application sees its own logical **storage,** independent of physical storage

Memory Virtualization

Virtual Memory

Each application sees its own logical **memory,** independent of physical memory



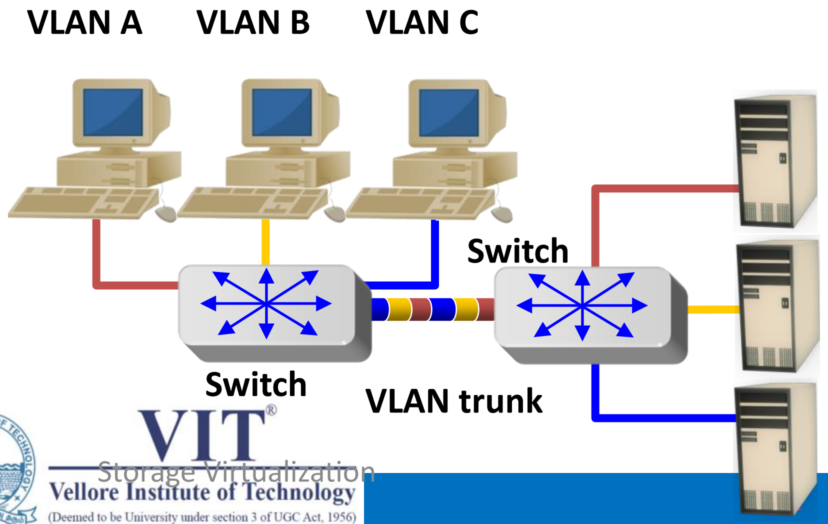
Benefits of Virtual Memory

- Remove physical-memory limits
- Run multiple applications at once

Network Virtualization

Virtual Networks

Each application sees its own logical **network,** independent of physical network

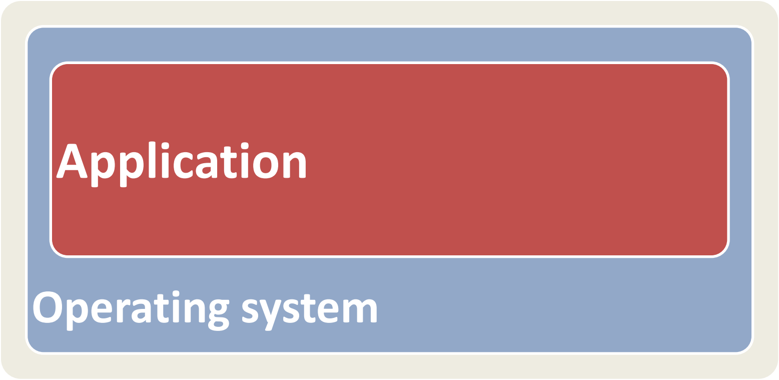


Benefits of Virtual Networks

- Common network links with access-control properties of separate links
- Manage logical networks instead of physical networks
- **Virtual SANs** provide similar benefits for storage-area networks

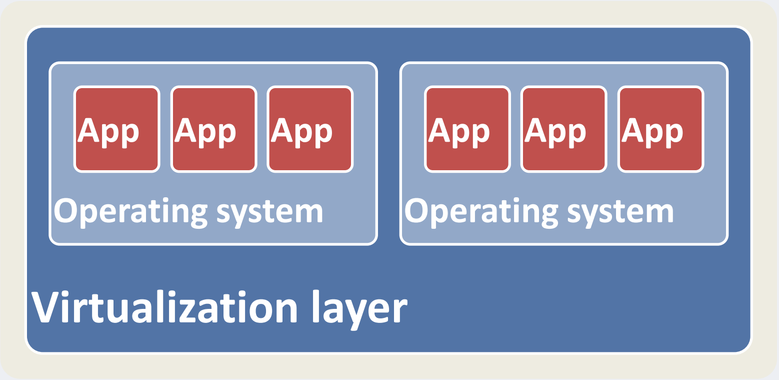
Server Virtualization

Before Server Virtualization:



- Single operating system image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources

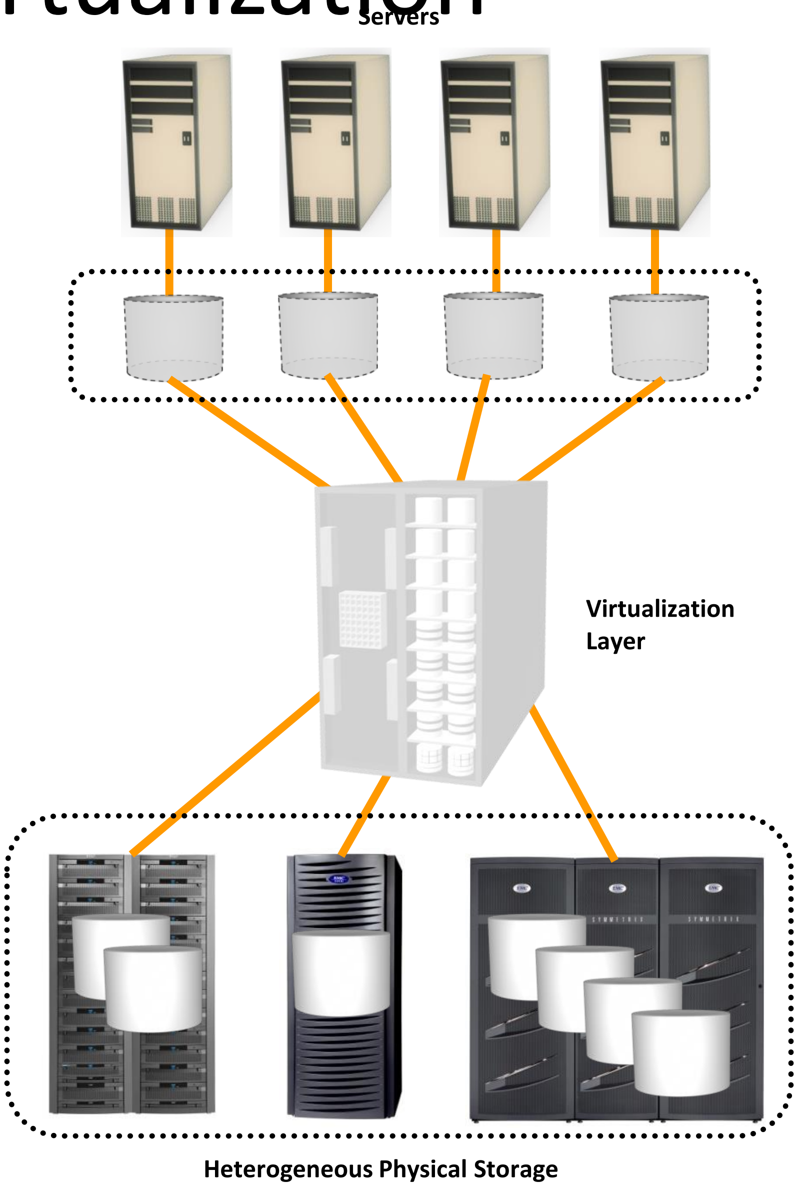
After Server Virtualization:



- Virtual Machines (VMs) break dependencies between operating system and hardware
- Manage operating system and application as single unit by encapsulating them into VMs
- Strong fault and security isolation
- Hardware-independent

Storage Virtualization

- Process of presenting a logical view of physical storage resources to hosts
- Logical storage appears and behaves as physical storage directly connected to host
- Examples of storage virtualization are:
  - Host-based volume management
  - LUN creation
  - Tape virtualization
- Benefits of storage virtualization:
  - Increased storage utilization
  - Adding or deleting storage without affecting application's availability
  - Non-disruptive data migration



Efficient Execution And Processor Privilege Levels

- When a user launches an app, the operating system loads the code for the app into the computer’s memory.

- The operating system then instructs the processor to start executing the code.

- Execution proceeds at the hardware rate because the processor executes code for the app directly without going “through” the operating system.

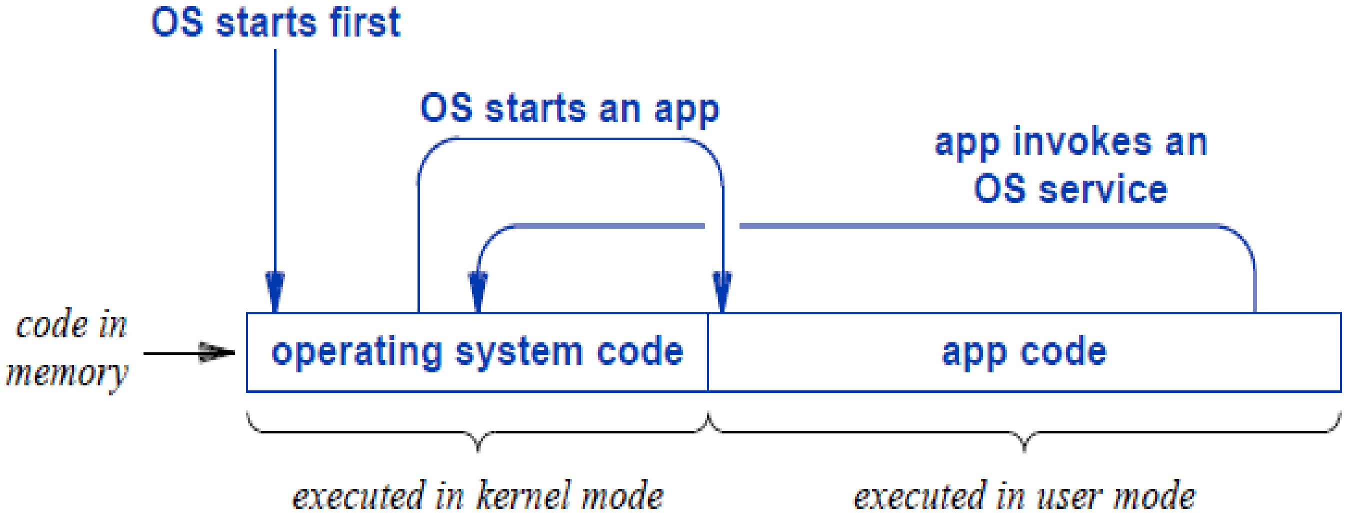
- An applicationcannotbeallowedtoexecuteallpossibleinstructionsorthecomputerwouldbevulnerabletohackerswhomightstealinformationorusethecom puter in a crime.

- To prevent such problems, the processor hardware used in a conventional computer has two ***privilege levels* or *modes of operation*.**

***Kernel mode*, User Mode**

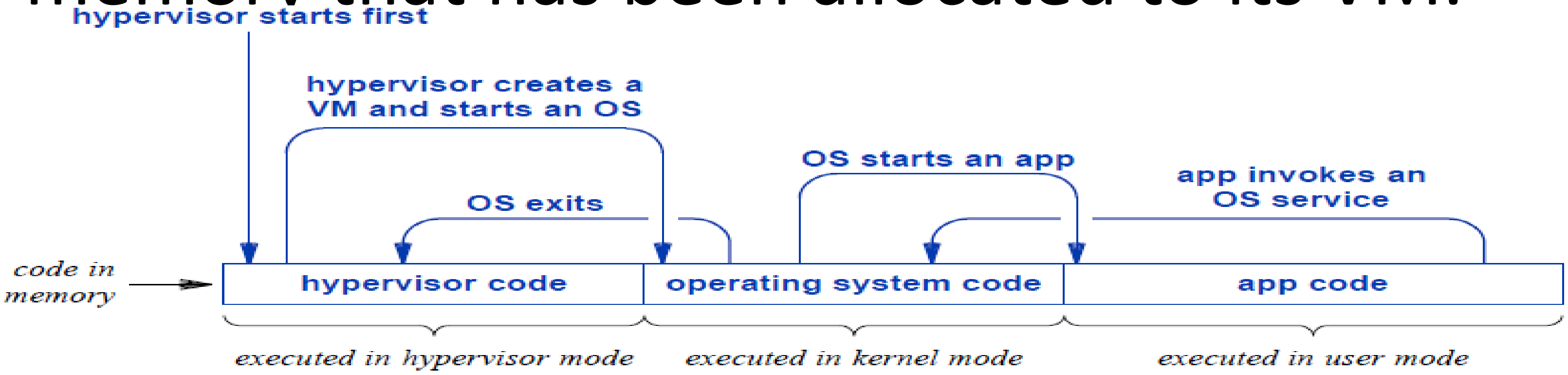
Illustration of operating system and app code in memory.

- The processor executes each at the same high speed, changing mode when transitioning from one to the other.



Extending Privilege To A Hypervisor

- Three levels of privilege: one for the hypervisor, a second for an operating system, and a third for apps.
- Only the hypervisor can create a VM and allocate memory to the VM.
- The operating system is restricted to the memory that has been allocated to its VM.



Hypervisor

- The hypervisor sits directly between the physical hardware and its OS.
- The hypervisor provides ***hyper calls*** for the guest OSes and applications.
- *Microkernel architecture* like the Microsoft Hyper-V
- *Monolithic hypervisor architecture* like the VMware ESX for server virtualization.

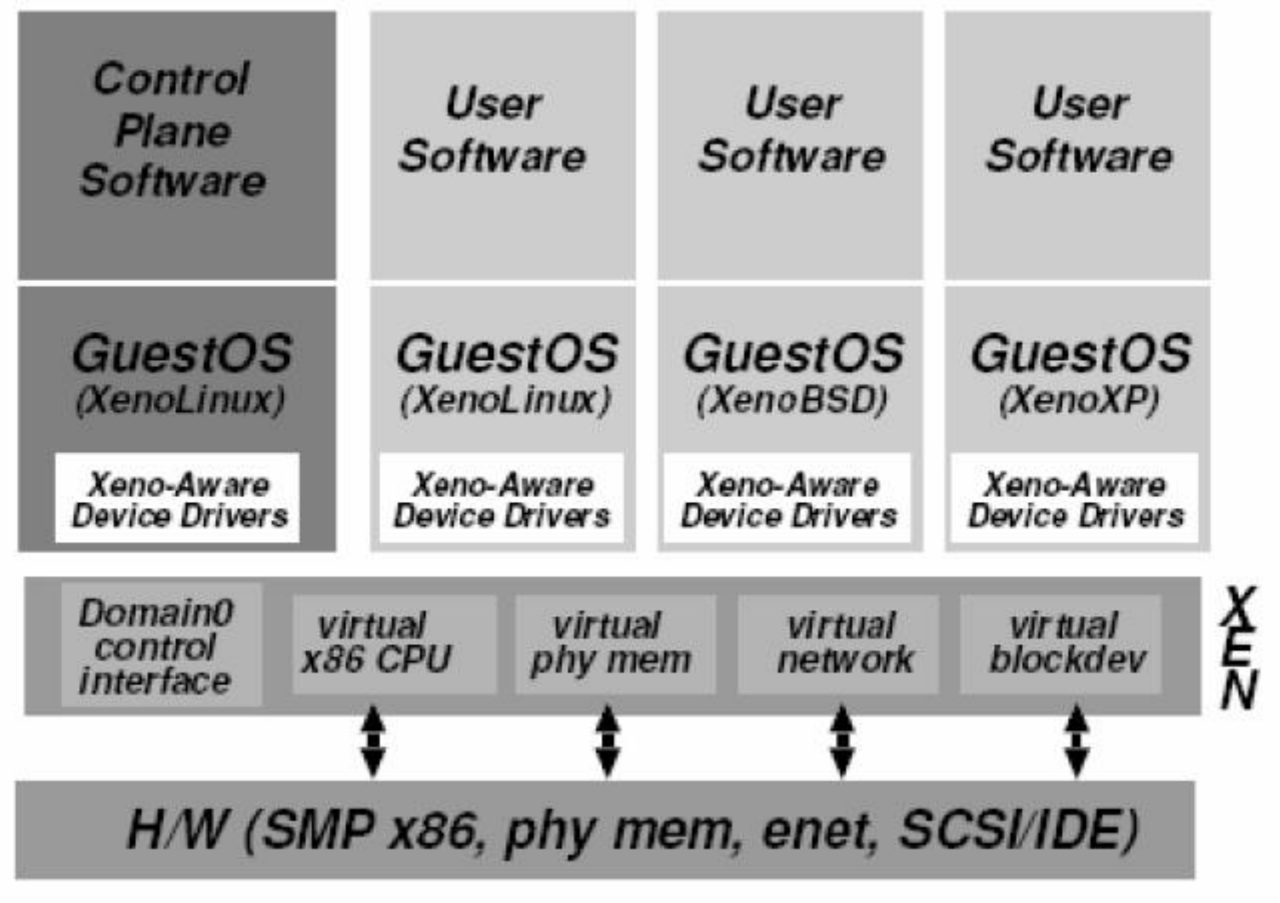
Hypervisor

- A micro-kernel hypervisor includes only the basic and unchanging functions .

  (such as physical memory management and processor scheduling).

- The device drivers and other changeable components are outside the hypervisor.
- A monolithic hypervisor implements all the aforementioned functions, including those of the device drivers.

Xen Architecture(Micro Kernel)



Xen(Open Source)

- The core components of a Xen system are the hypervisor, kernel, and applications.

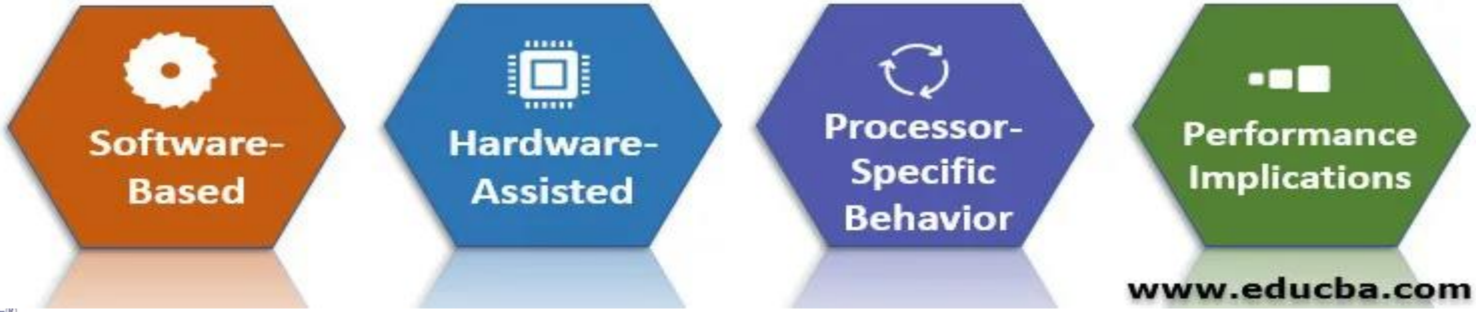
- Many guest OSes can run on top of the hypervisor. However, not all guest OSes are created equal, and one in particular controls the others.

- The guest OS, which has control ability, is called Domain 0, and the others are called Domain U. Domain 0 is a privileged guest OS of Xen.

CPU Virtualization

- CPU Virtualization emphasizes running programs and instructions through a virtual machine, giving the feeling of working on a physical workstation.

- All the operations are handled by an emulator that controls software to run according to it.



Software-Based CPU Virtualization

- CPU Virtualization is software-based where with the help of it, application code gets executed on the processor and the privileged code gets translated first, and that translated code gets executed directly on the processor.

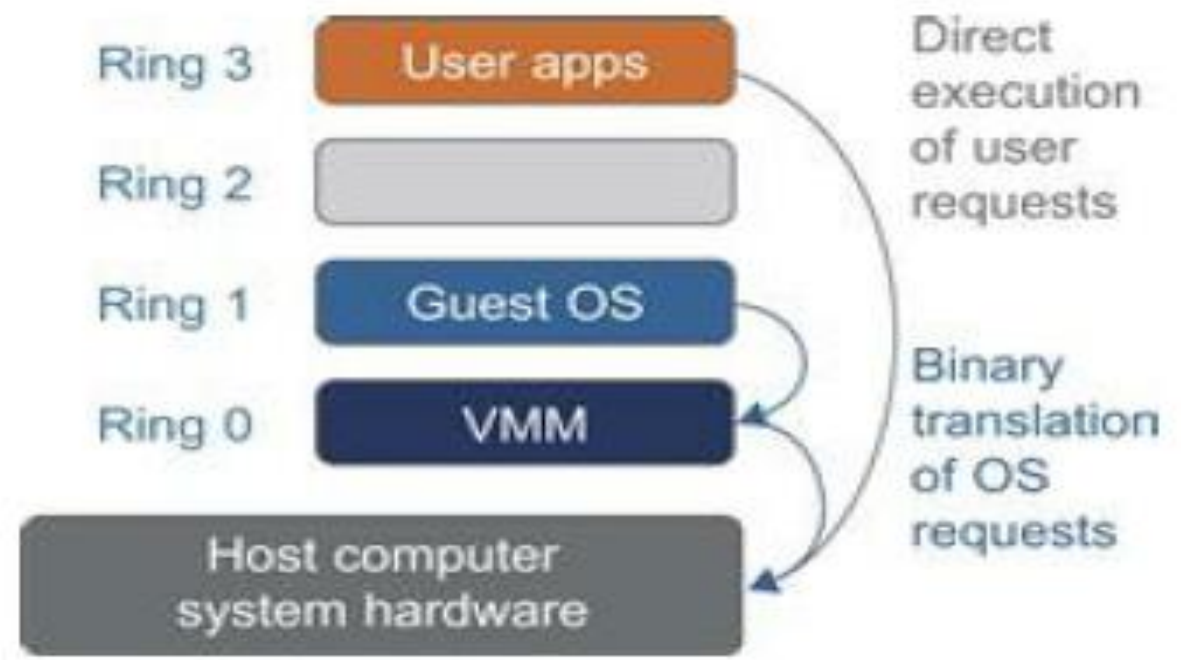
- This translation is purely known as Binary Translation (BT).

- The code that gets translated is very large in size and also slow at the same time on execution.

- The guest programs that are based on privileged coding runs very smooth and fast.

- The code programs or the applications that are based on privileged code components that are significant such as system calls, run at a slower rate in the virtual environment.

CPU VIRTUALIZATION


Ring 3
User apps
Ring 2
Ring 1
Guest OS
Ring 0
VMM
Host computer system hardware
Direct execution of user requests
Binary translation of OS requests

VMM scans the instruction stream and identifies the privileged, control- and behavior sensitive instructions.

When these instructions are identified, they are trapped into the VMM, which emulates the behavior of these instructions.

The method used in this emulation is called *binary translation*.

Therefore, full virtualization combines binary translation and direct execution.

The guest OS is completely decoupled from the underlying hardware.

Hardware-Assisted CPU Virtualization

- The guest user uses a different version of code and mode of execution known as a guest mode. The guest code mainly runs on guest mode.

- The best part in hardware-assisted CPU Virtualization is that there is no requirement for translation while using it for hardware assistance.

- For this, the system calls runs faster than expected.

- Workloads that require the updation of page tables get a chance of exiting from guest mode to root mode that eventually slows down the program’s performance and efficiency.

Memory Virtualization

- In a traditional execution environment, the operating system maintains mappings of *virtual memory* to *machine memory* using page tables, which is a one-stage mapping from virtual memory to machine memory.

- All modern x86 CPUs include a *memory management unit (MMU)* and a *Translation Lookaside Buffer (TLB)* to optimize virtual memory performance.

- Virtual memory virtualization involves sharing the physical system memory in RAM and dynamically allocating it to the *physical memory* of the VMs.

Virtual Memory Virtualization

- Two-stage mapping process should be maintained by the guest OS and the VMM, respectively:

- Virtual memory to physical memory

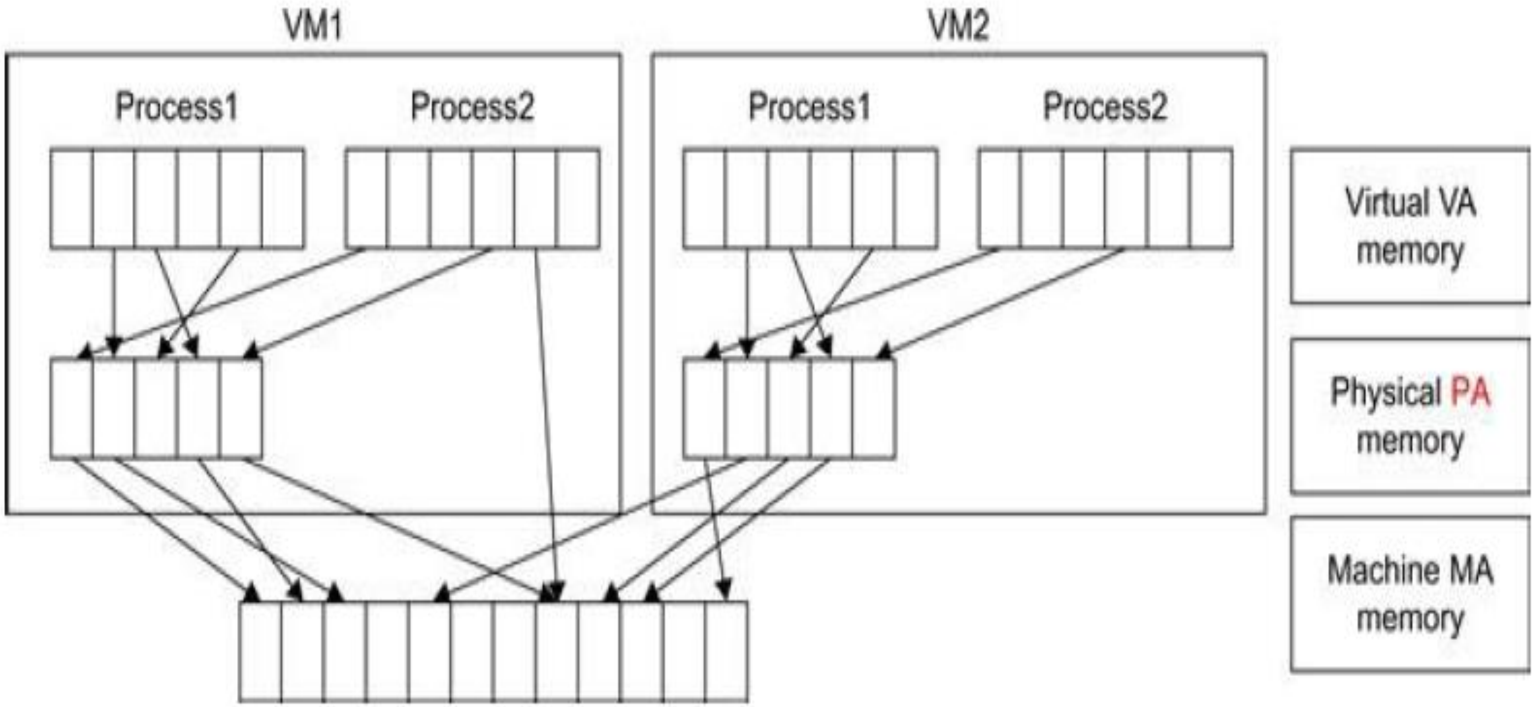
  and

  physical memory to machine memory.

Virtual Memory Virtualization

- MMU virtualization should be supported, which is transparent to the guest OS.
- The guest OS continues to control the mapping of virtual addresses to the physical memory addresses of VMs. But the guest OS cannot directly access the actual machine memory.
- The VMM is responsible for mapping the guest physical memory to the actual machine memory

Two-level memory mapping procedure



Shadow Page Table

- Each page table of the guest OSes has a separate page table in the VMM corresponding to it, the VMM page table is called the shadow page table.

- VMware uses shadow page tables to perform virtual-memory-to-machine-memory address translation.

- Processors use TLB hardware to map the virtual memory directly to the machine memory to avoid the two levels of translation on every access.

Levels Of Trust And I/O Devices

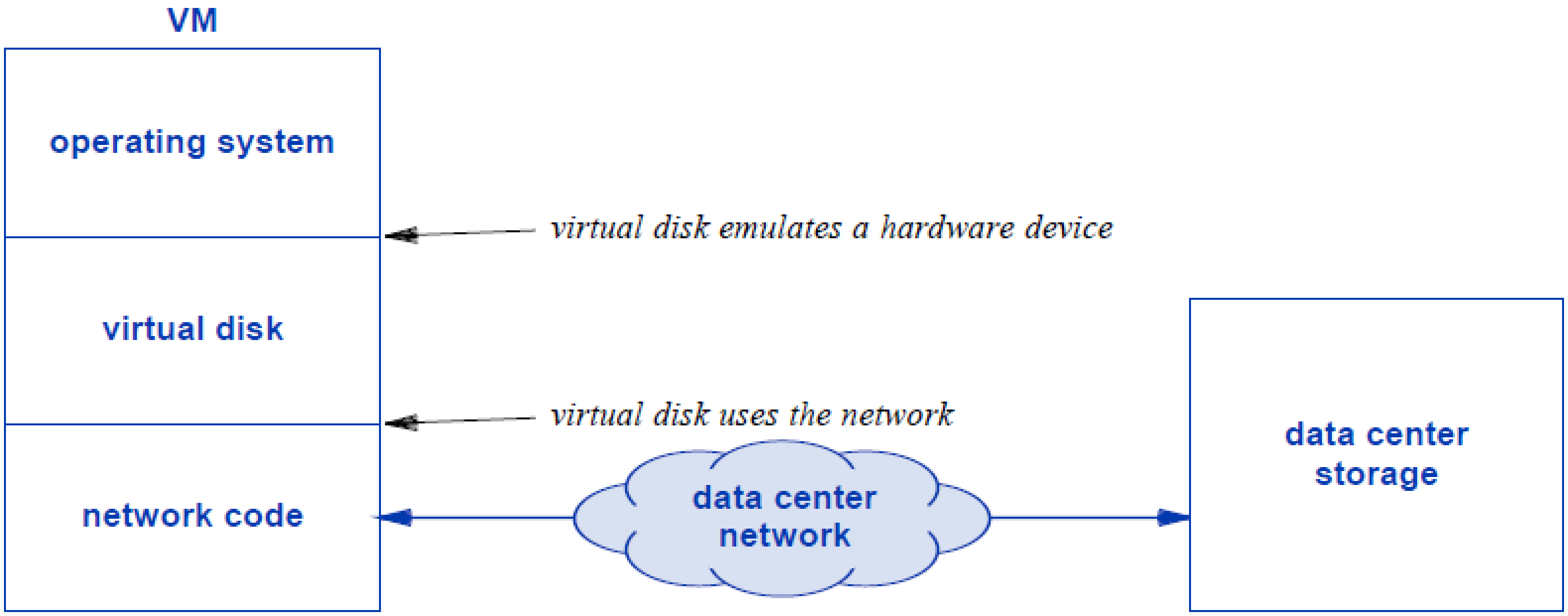
How an operating system manages I/O devices (e.g., a screen, keyboard, disk, and network interface) on a conventional computer?

- Operating system uses a hardware mechanism known as a *bus* to communicate with I/O devices.
- The first step consists of sending a series of requests across the bus to form a list of all I/O devices that are present.
- The operating system must include *device driver* software for each device.
- The operating system uses the device driver code for a given device to control the device hardware and handle all communication with the device.

Virtual I/O Devices

- Hypervisor creates a set of *virtual I/O devices* for the VM to use.
- A virtual I/O device is implemented by software.
- When an operating system on a VM attempts to use the bus to access an I/O device, the access violates privilege, which means the hypervisor is invoked.
- The hypervisor runs the appropriate virtual device, software, and then arranges to send the response to the operating system as if a physical device responded.

An Example Virtual Device



VM As A Digital Object

- A hypervisor must keep a record of the VM
- Memory that have been allocated to the VM
- The virtual I/O devices that have been created for the VM(including disk space that has been allocated in the data center storage facility).
- Current status of the VM
- VM can be turned into a digital object. That is, the entire VM can be transformed into a set of bytes
- Imagine, for example, that they are placed in a special file.
- ***All the pieces of a VM can be collected together into a digital object.***

VM Migration

- The movement of VMs from one resource to another, such as from one physical host to another physical host, or data store to data store, is known as VM migration.

- There are two types of VM migration:

✓ **Cold**

✓ **Hot (Live)**

Cold-Hot (Live)

- Cold migration occurs when the VM is shut down.

- Live migration is the process of moving a running virtual machine without stopping the OS and other applications from source host to destination host.

Cold Migration

- Cold migration is the migration of powered off or suspended virtual machines between hosts across clusters, data centers, and vCenter Server instances.

- By using cold migration, you can also move associated disks from one datastore to another.

- If you attempt to migrate a powered off virtual machine that is configured with a 64-bit operating system to a host that does not support 64-bit operating systems, vCenter Server generates a warning.

Data Store

- **Datastores** in VMware vSphere are storage containers for files.
- They could be located on a local server hard drive or across the network on a SAN.
- Datastores hide the specifics of each storage device and provide a uniform model for storing virtual machine files.
- Datastores are used to hold virtual machine files, templates, and ISO images.
- They can be formatted with **VMFS (Virtual Machine File System**, a clustered file system from VMware)

VMware vSphere

VMware vSphere is a software suite that includes components like ESXi, vCenter Server, vSphere Client, vCenter Orchestrator, vSphere Update Manager, etc.

vSphere components provide virtualization, management, resource optimization and many other features useful for a virtual environment

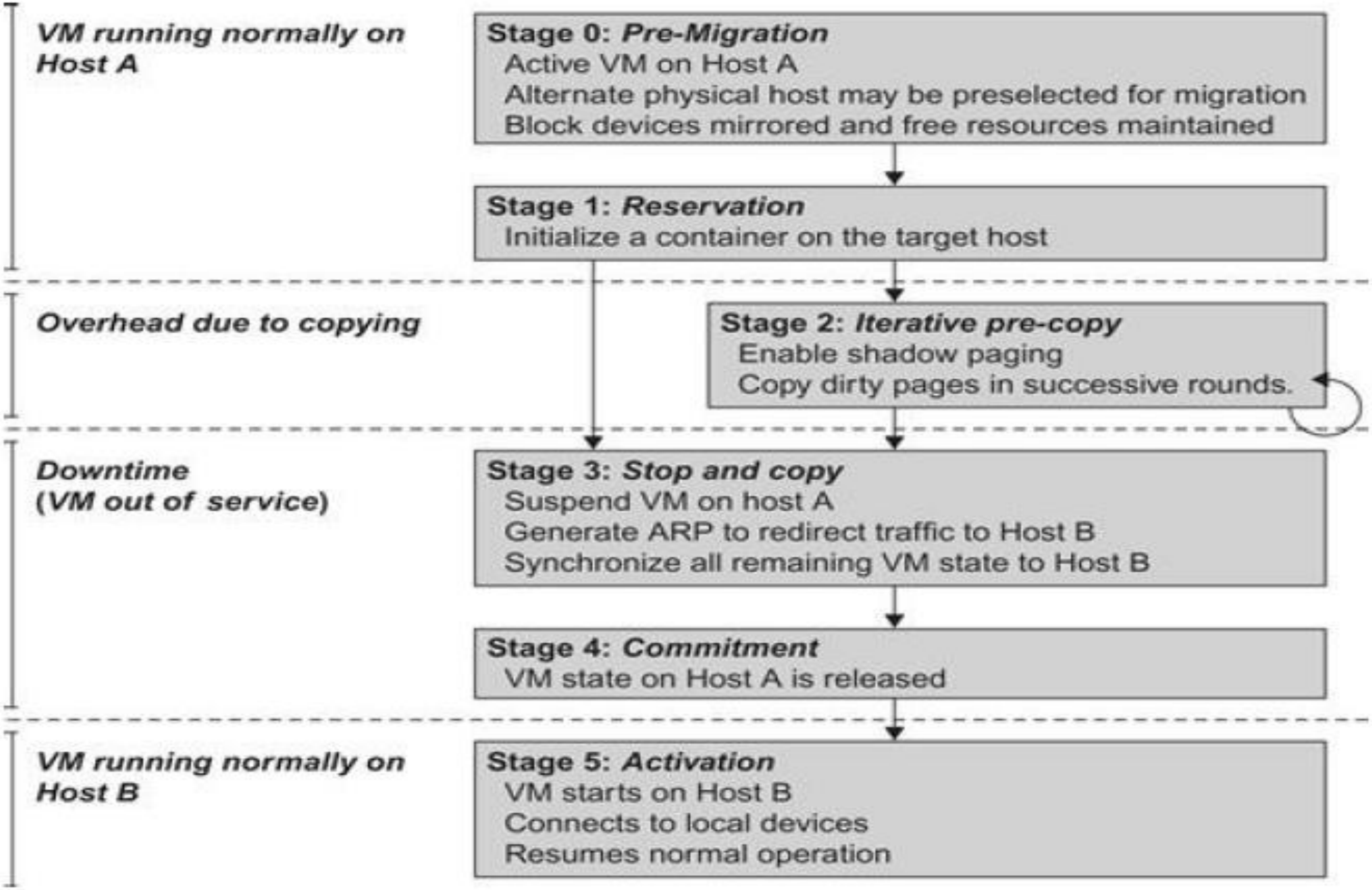
- https://www.vmware.com/products/vsphere.html

LIVE VM migration

- **Live migration** refers to the process of moving a **running** virtual machine or application between different physical machines without disconnecting the client or application.

- Memory, storage, and network connectivity of the virtual machine are transferred from the original guest machine to the destination.

Live migration process of a VM from one host to another



VM running normally on Host A

Stage 0: ***Pre-Migration***
Active VM on Host A
Alternate physical host may be preselected for migration
Block devices mirrored and free resources maintained

Stage 1: ***Reservation***
Initialize a container on the target host

Overhead due to copying

Stage 2: ***Iterative pre-copy***
Enable shadow paging
Copy dirty pages in successive rounds.

Downtime
(VM out of service)

Stage 3: ***Stop and copy***
Suspend VM on host A
Generate ARP to redirect traffic to Host B
Synchronize all remaining VM state to Host B

Stage 4: ***Commitment***
VM state on Host A is released

VM running normally on Host B

Stage 5: ***Activation***
VM starts on Host B
Connects to local devices
Resumes normal operation

VMware vSphere vMotion

- VMware vSphere vMotion is a zero downtime live migration of workloads from one server to another.
- This capability is possible across vSwitches, Clusters, and even Clouds (depending of the vSphere edition that you have).
- During the workload migration, the application is still running and users continue to have access to the systems they need.