UNIT III
Virtualization in Cloud Computing

Course Objectives

- **To study fundamental concepts of cloud computing**
- **To learn various data storage methods on cloud**
- **To understand the implementation of Virtualization in Cloud Computing**
- To learn the application and security on cloud computing
- To study risk management in cloud computing
- To understand the advanced technologies in cloud computing

Course Outcomes

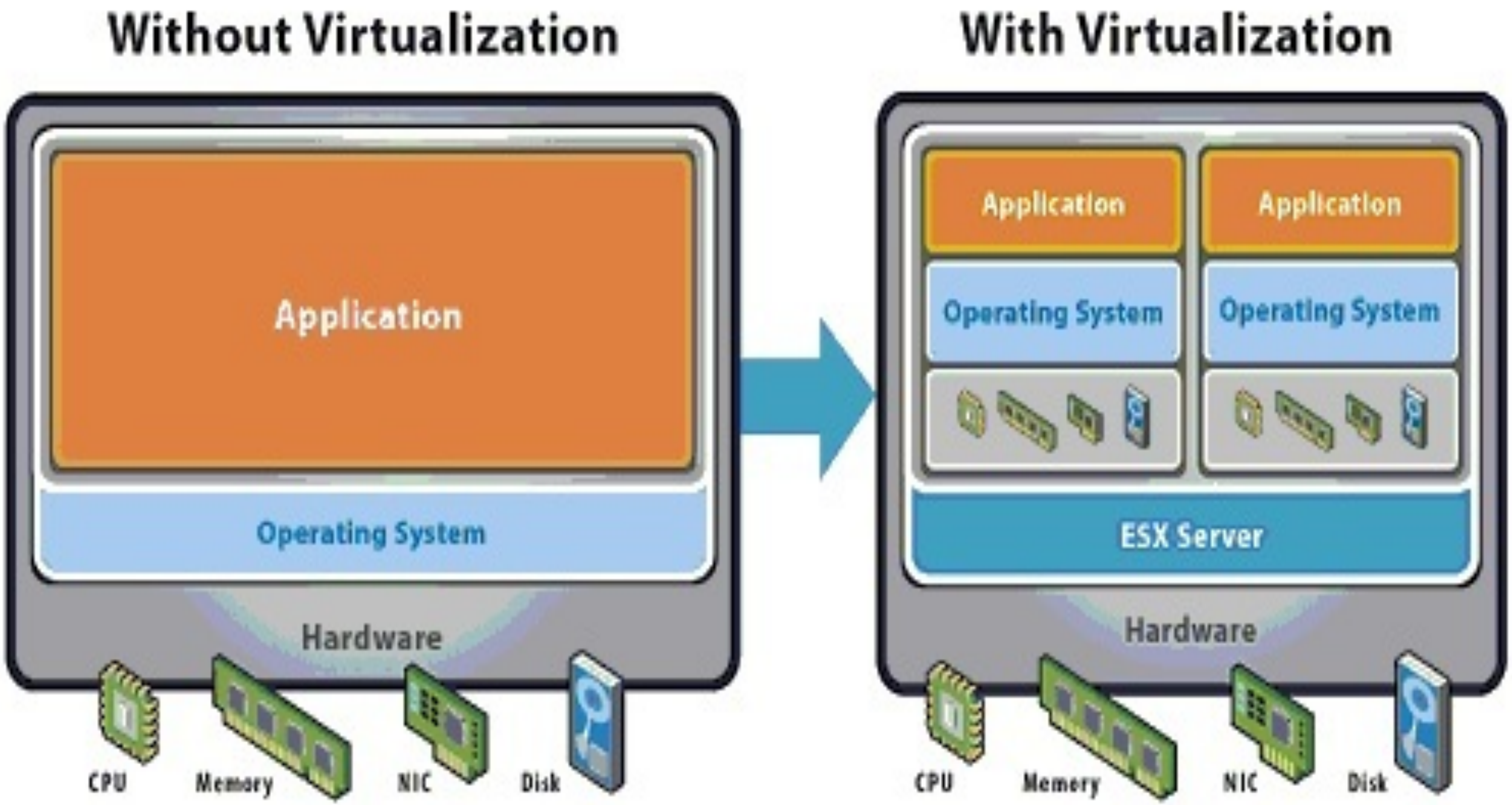
- **CO1: Understand the different Cloud Computing environment**
- **CO2: Use appropriate data storage technique on Cloud, based on Cloud application**
- **CO3: Analyze virtualization technology and install virtualization software**
- CO4: Develop and deploy applications on Cloud
- CO5: Apply security in cloud applications
- CO6: Use advance techniques in Cloud Computing Course Content

What is Virtualization?

- **Virtualization** is the "creation of a virtual (rather than actual) version of something, such as a server, a desktop, a storage device, an operating system or network resources".
- In other words, Virtualization is a technique, which allows to share a single physical instance of a resource or an application among multiple customers and organizations. It does by assigning a logical name to a physical storage and providing a pointer to that physical resource when demanded.

What is the concept behind the Virtualization?

- Creation of a virtual machine over existing operating system and hardware is known as Hardware Virtualization. A Virtual machine provides an environment that is logically separated from the underlying hardware.
- The machine on which the virtual machine is going to create is known as **Host Machine** and that virtual machine is referred as a **Guest Machine.**


Without Virtualization
Application
Operating System
Hardware
CPU
Memory
NIC
Disk
With Virtualization
Application
Application
Operating System
Operating System
ESX Server
Hardware
CPU
Memory
NIC
Disk

Hypervisor

- The hypervisor is also known as Virtual Machine Monitor(VMM).
- It is used to support virtualization of hardware-level on various types of devices such as CPU, memory, disk and network interfaces.
- The hypervisor software is placed exactly between the physical hardware and its operating system
- Hypercalls are provided by the hypervisors for the guest operating system and applications.

Types of hypervisor

- A hypervisor or virtual machine monitor (VMM) is computer software, firmware or hardware that creates and runs virtual machines.
- A computer on which hypervisor runs one or more virtual machines is called a host machine and each virtual machine is called as guest machine.
- The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of guest operating system.

- Multiple instances of a variety of operating systems may share the virtualized hardware resources.
- Hypervisor is the supervisor of the supervisor.
- There are 2 types of hypervisors
    1. Type 1:native or bare metal hypervisor
    2. Type 2: hosted hypervisor

.Type 1:native or bare metal hypervisor

- Type 1 hypervisor is also called as native or bare metal or embedded hypervisor.
- It works directly on the hardware of the host and can monitor operating systems that run above the hypervisor.
- It is totally independent of the operating system.
- The size of the hypervisor is small because its main function is to share and manage hardware resources between various operating system.

.Type 1:native or bare metal hypervisor

- A major advantage of type-1 hypervisor is that any problems occurred in one virtual machine or guest operating system are not going to affect the other guest operating systems on the hypervisor.
- Examples:

✔ Vmware ESXi Server

✔ Microsoft HyperV

✔ Citrix/Xen Server

Type 2: hosted hypervisor

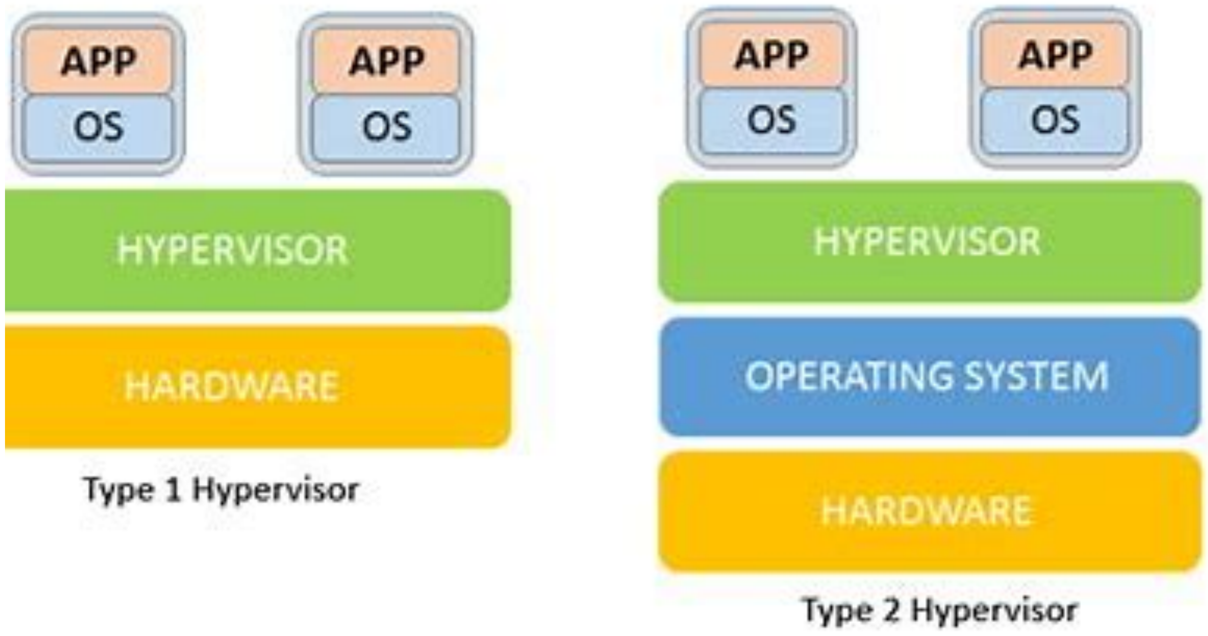
- These hypervisors run on conventional operating system just as another computer program do.
- A guest operating system runs as a process on the host.
- Type -2 hypervisors abstract guest operating systems from the host operating system.
- It is totally depends upon the host operating system for its operations.
- If any problem arise in the base operating system then it affects the entire system
- Examples:

✔ Vmware Workstation

✔ Microsoft Virtual PC

✔ Oracle Virtual Box

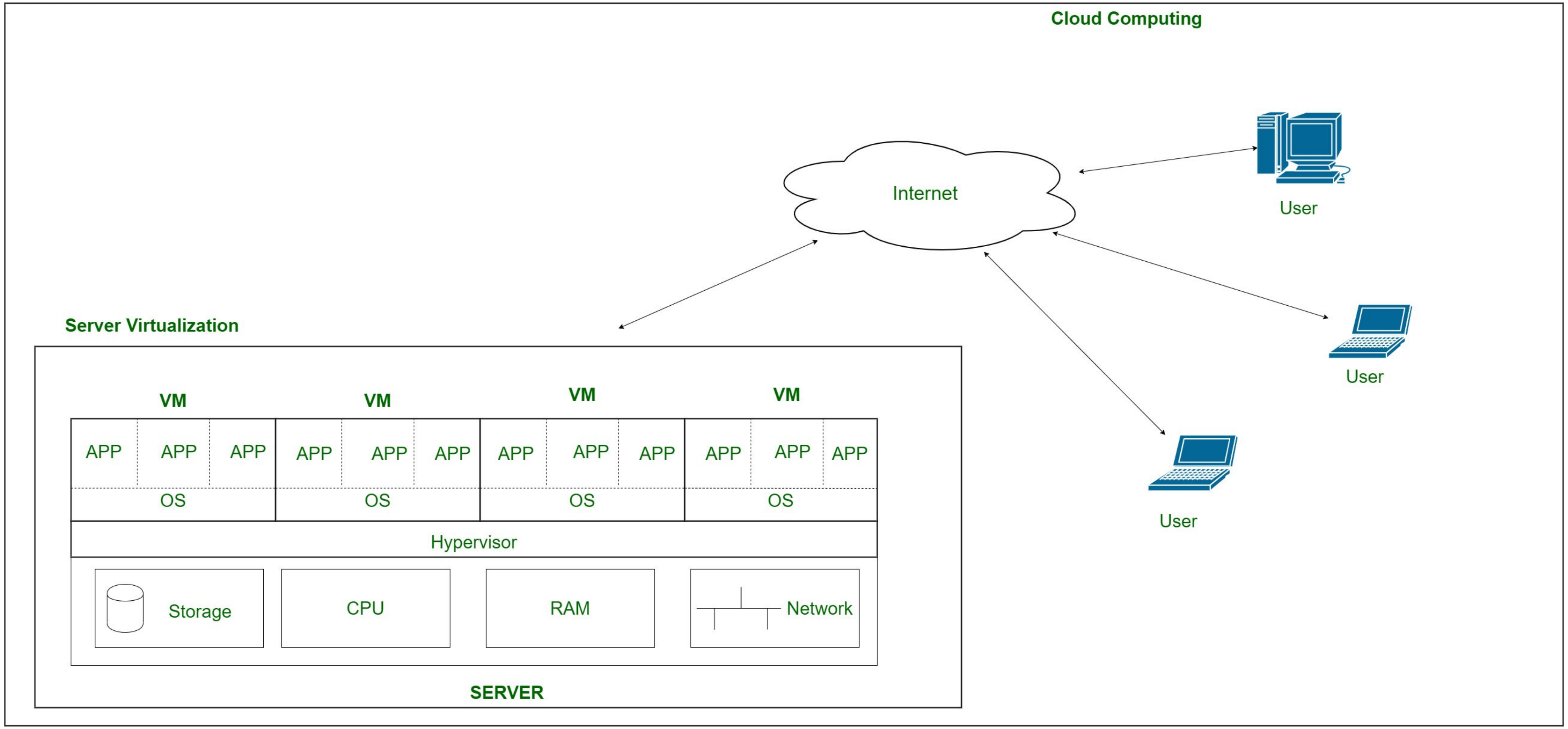
Types of hypervisor



Types of virtualization

- Server virtualization
- Network virtualization
- Storage virtualization
- Desktop virtualization
- Application Virtualization
- Management virtualization

Server Virtualization



- To implement **Server Virtualization**, hypervisor is installed on server which manages and allocates host hardware requirements to each virtual machine.
- This hypervisor sits over server hardware and regulates resources of each VM.
- A user can increase or decrease resources or can delete entire VM as per his/her need.
- This servers with VM created on them is called server virtualization and concept of controlling this VM by users through internet is called **Cloud Computing.**

Types of server virtualization

- Full virtualization
- Paravirtualization
- Os-level virtualization

Binary Translation with full virtualization

- Hardware virtualization can be classified into :
  1. Full virtualization
  2. Host-based virtualization

  - Hypervisor as well as VMM approaches are considered as full virtualization
  - In full virtualization, there is no need to modify the host OS.
  - It depends upon the binary translation for the process of trapping and to virtualized the execution specific, sensitive, non-virtualizable instructions.

- There are number of non-critical and critical instructions in guest OS and their applications.
- In a host based system, both of the host OS as well as the guest OS are used.
- A virtualization software layer is established in between the host OS and the guest OS.

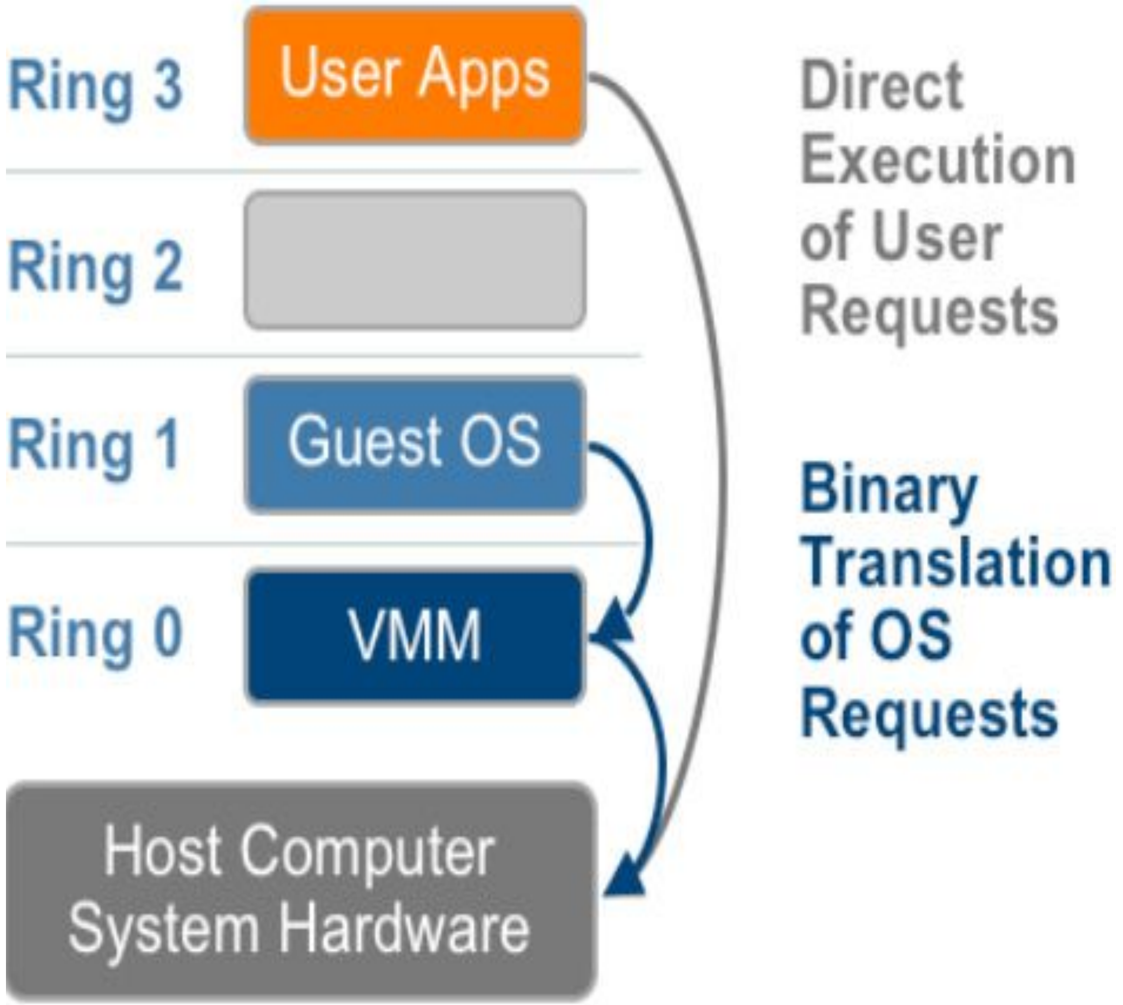
Full Virtualization

- With full virtualization, noncritical instructions run on the hardware directly while critical instructions are discovered and replaced with traps into the VMM to be emulated by software.
- Both the hypervisor and VMM approaches are considered full virtualization. Why are only critical instructions trapped into the VMM? This is because binary translation can incur a large performance overhead.
- Noncritical instructions do not control hardware or threaten the security of the system, but critical instructions do.
- Therefore, running noncritical instructions on hardware not only can promote efficiency, but also can ensure system security.

Binary Translation of Guest OS Requests using a VMM

- This approach was implemented by VMware and many other software companies.
- VMware puts the VMM at Ring 0 and the guest OS at Ring 1.
- The 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. Consequently, the guest OS is unaware that it is being virtualized.

Binary Translation with full virtualization



Host-based Virtualization

- An alternative VM architecture is to install a virtualization layer on top of the host OS.
- This host OS is still responsible for managing the hardware.
- The guest OSes are installed and run on top of the virtualization layer.
- Dedicated applications may run on the VMs. Certainly, some other applications can also run with the host OS directly.
- This host-based architecture has some distinct advantages, as enumerated next.
- First, the user can install this VM architecture without modifying the host OS.
- The virtualizing software can rely on the host OS to provide device drivers and other low-level services. This will simplify the VM design and ease its deployment.

- Second, the host-based approach appeals to many host machine configurations.
- Compared to the hypervisor/VMM architecture, the performance of the host-based architecture may also be low.
- When an application requests hardware access, it involves four layers of mapping which downgrades performance significantly.
- When the ISA of a guest OS is different from the ISA of the underlying hardware, binary translation must be adopted.
- Although the host-based architecture has flexibility, the performance is too low to be useful in practice.

Para-Virtualization with Compiler Support

- Para-virtualization needs to modify the guest operating systems.
- A para-virtualized VM provides special APIs requiring substantial OS modifications in user applications.
- Performance degradation is a critical issue of a virtualized system.
- No one wants to use a VM if it is much slower than using a physical machine.
- The virtualization layer can be inserted at different positions in a machine soft-ware stack.
- However, para-virtualization attempts to reduce the virtualization overhead, and thus improve performance by modifying only the guest OS kernel.

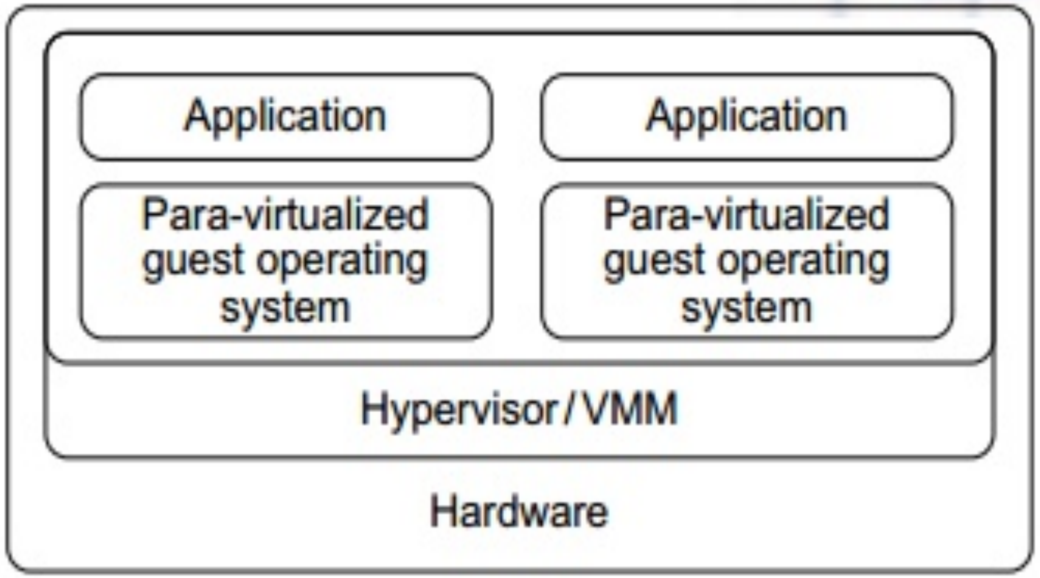


FIGURE 3.7

Para-virtualized VM architecture, which involves modifying the guest OS kernel to replace nonvirtualizable instructions with hypercalls for the hypervisor or the VMM to carry out the virtualization process (See Figure 3.8 for more details.)

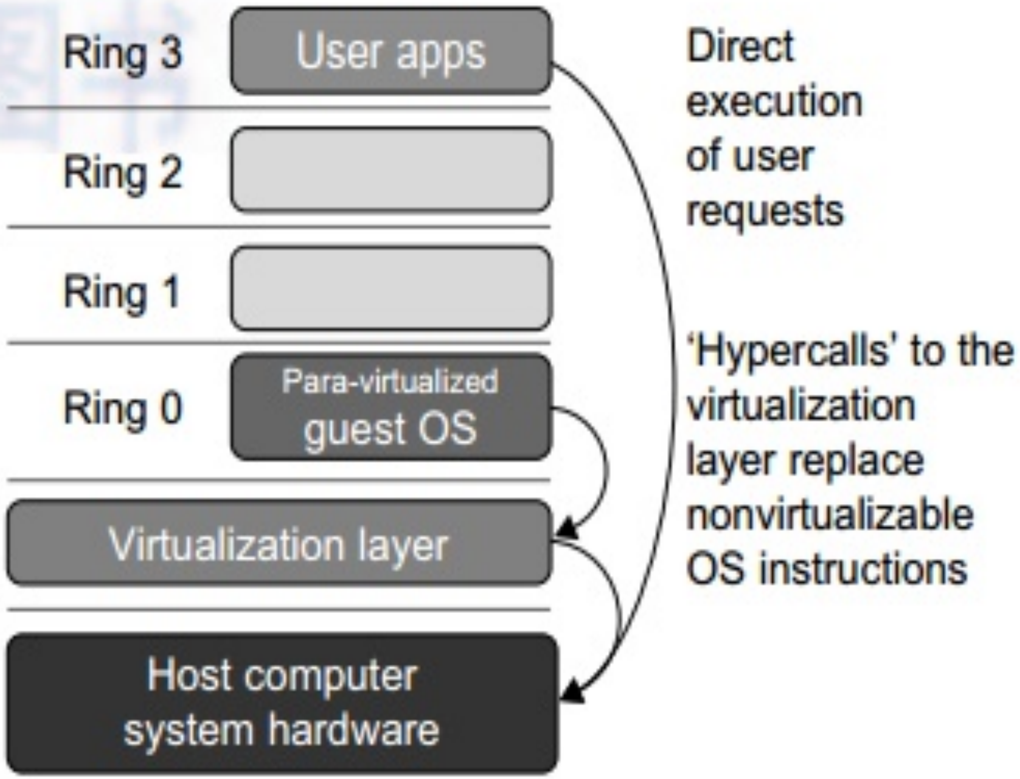


FIGURE 3.8

The use of a para-virtualized guest OS assisted by an intelligent compiler to replace nonvirtualizable OS instructions by hypercalls.

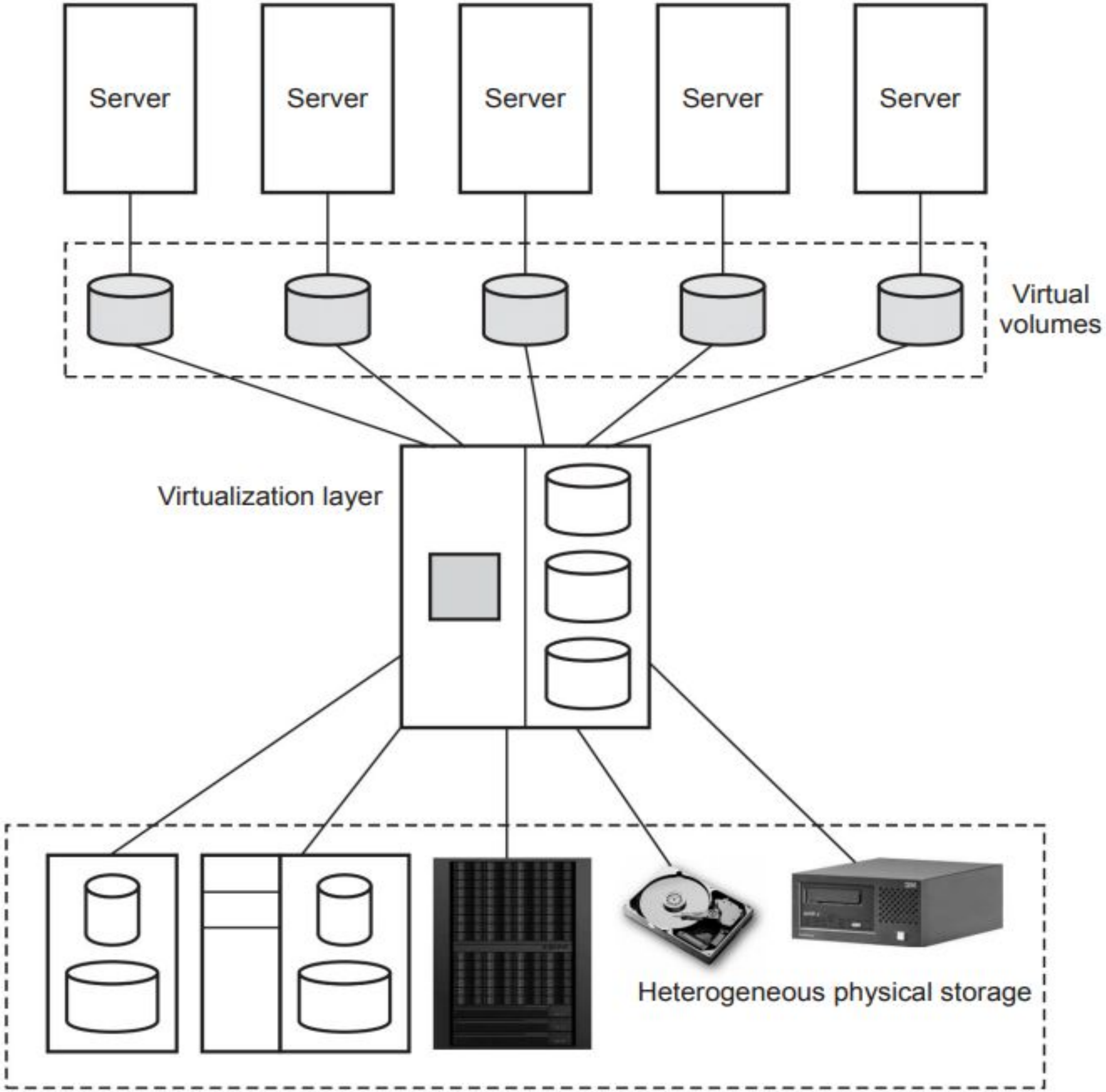
(*Courtesy of VMWare [71]*)

Para-Virtualization with Compiler Support

- Unlike the full virtualization architecture which intercepts and emulates privileged and sensitive instructions at runtime, para-virtualization handles these instructions at compile time.
- The guest OS kernel is modified to replace the privileged and sensitive instructions with hypercalls to the hypervisor or VMM.
- Xen assumes such a para-virtualization architecture.

Storage Virtualization

- Storage virtualization is the pooling of physical storage from multiple storage devices into what appears to be a single storage device -- or pool of available storage capacity. A central console manages the storage.
- The technology relies on software to identify available storage capacity from physical devices and to then aggregate that capacity as a pool of storage that can be used by traditional architecture servers or in a virtual environment by virtual machines (VMs).


Server
Server
Server
Server
Server
Virtual
volumes
Virtualization layer
Heterogeneous physical storage

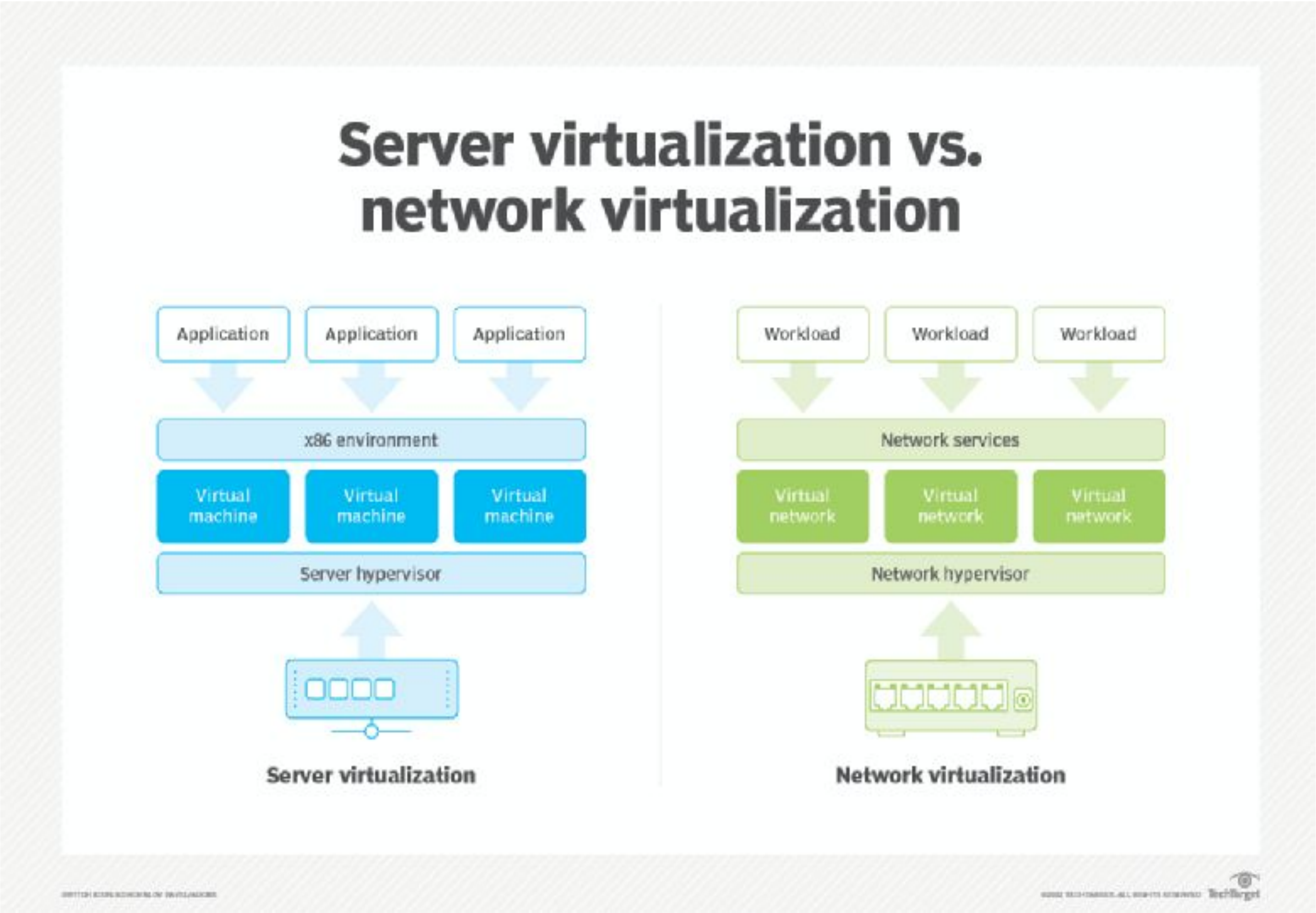
Network Virtualization

- Network virtualization is the process of abstracting and combining physical network resources into virtualized, software-defined networks.
- This enables the creation of multiple virtual networks that can run on a single physical network infrastructure, allowing for more efficient use of resources, flexibility, and scalability.

Network Virtualization

There are two main components of network virtualization:

1. **Virtual Network Functions (VNFs):** These are software-based network functions that replace traditional hardware-based network appliances like routers, firewalls, and load balancers. VNFs run on virtual machines or containers, enabling more flexible and scalable network configurations.
2. **Network Overlay:** This involves creating a virtual layer over the physical network, where virtual networks are constructed and managed independently of the underlying hardware. This allows for the segmentation of traffic, improved security, and easier management.



Server virtualization vs. network virtualization
Application
Application
Application
x86 environment
Virtual machine
Virtual machine
Virtual machine
Server hypervisor
Server virtualization
Workload
Workload
Workload
Network services
Virtual network
Virtual network
Virtual network
Network hypervisor
Network virtualization
TechTarget

OS Virtualization

- Operating system-based Virtualization refers to an operating system feature in which the kernel enables the existence of various isolated user-space instances.
- The installation of virtualization software also refers to Operating system-based virtualization.
- It is installed over a pre-existing operating system and that operating system is called the host operating system.

OS Virtualization

- In this virtualization, a user installs the virtualization software in the operating system of his system like any other program and utilizes this application to operate and generate various virtual machines.
- Here, the virtualization software allows direct access to any of the created virtual machines to the user.

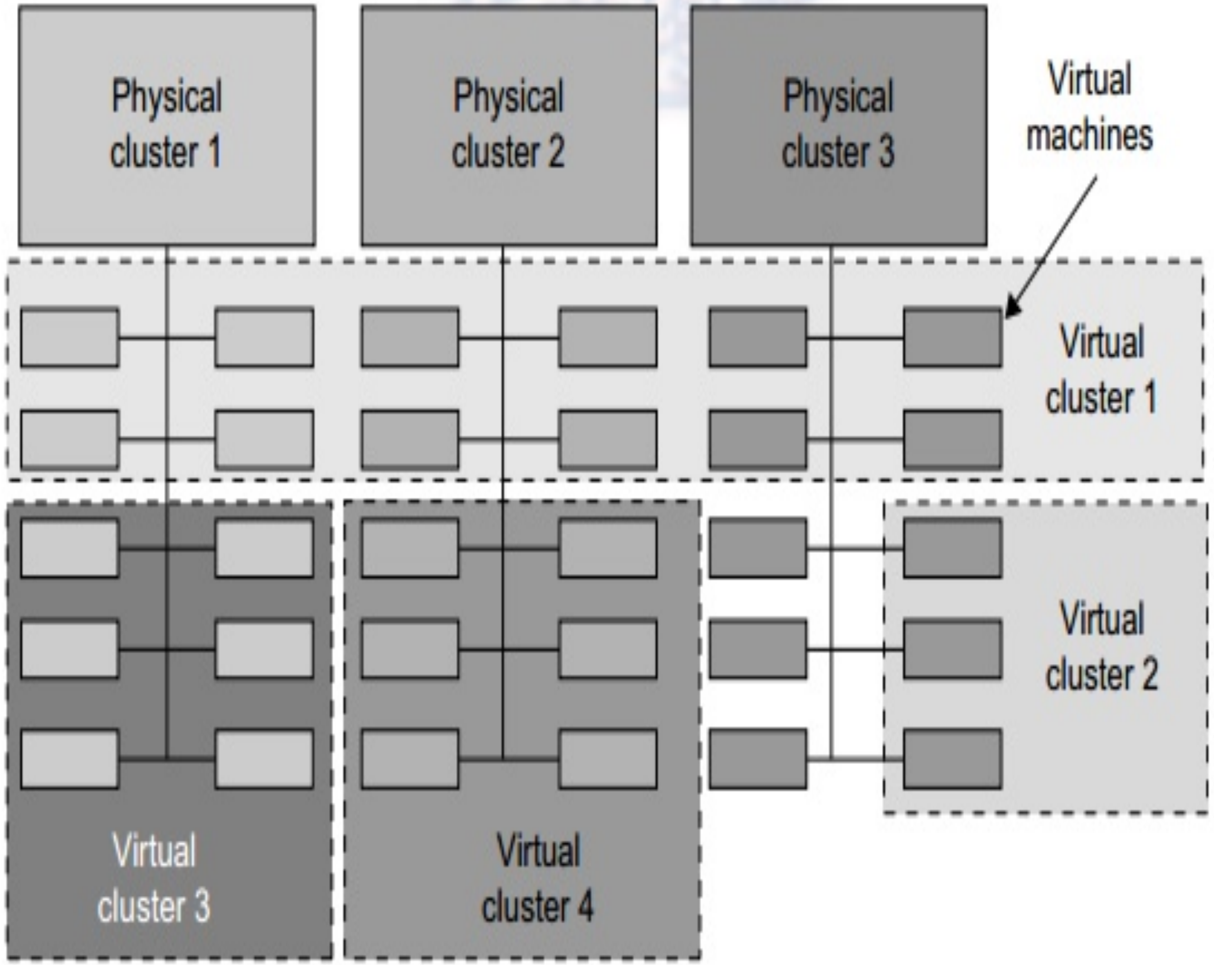
Virtual Clustering

Virtual Clusters and Resource Management

- A physical cluster is a collection of servers (physical machines) interconnected by a physical network such as a LAN.
- In this section, we will study three critical design issues of virtual clusters:
- live migration of VMs,
- memory and file migrations,
- dynamic deployment of virtual clusters.

Physical versus Virtual Clusters

- Virtual clusters are built with VMs installed at distributed servers from one or more physical clusters.
- The VMs in a virtual cluster are interconnected logically by a virtual network across several physical networks.
- Each virtual cluster is formed with physical machines or a VM hosted by multiple physical clusters.
- The virtual cluster boundaries are shown as distinct boundaries.

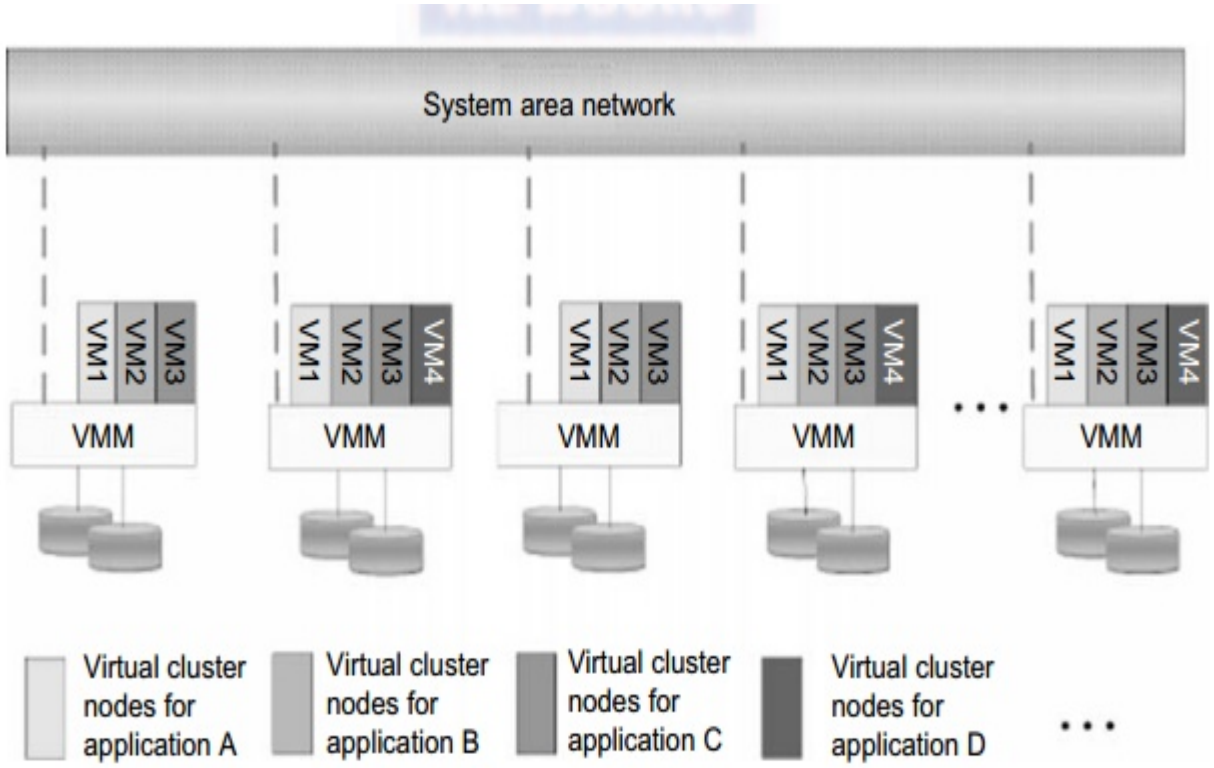

Physical
cluster 1
Physical
cluster 2
Physical
cluster 3
Virtual
machines
Virtual
cluster 1
Virtual
cluster 2
Virtual
cluster 3
Virtual
cluster 4

Properties

- The virtual cluster nodes can be either physical or virtual machines. Multiple VMs running with different OSes can be deployed on the same physical node.
- A VM runs with a guest OS, which is often different from the host OS, that manages the resources in the physical machine, where the VM is implemented.
- The purpose of using VMs is to consolidate multiple functionalities on the same server. This will greatly enhance server utilization and application flexibility.

- VMs can be colonized (replicated) in multiple servers for the purpose of promoting distributed parallelism, fault tolerance, and disaster recovery.
- The size (number of nodes) of a virtual cluster can grow or shrink dynamically, similar to the way an overlay network varies in size in a peer-to-peer (P2P) network.
- The failure of any physical nodes may disable some VMs installed on the failing nodes. But the failure of VMs will not pull down the host system.

Virtual cluster based on application partitioning



Virtualization Application

- Application virtualization helps a user to have remote access to an application from a server.
- The server stores all personal information and other characteristics of the application but can still run on a local workstation through the internet.
- An example of this would be a user who needs to run two different versions of the same software.
- Technologies that use application virtualization are hosted applications and packaged applications.

Pitfalls of virtualization

- Virtualization is a technology that creates virtual representations of various computing resources, allowing more efficient utilization of physical hardware.
- It is widely used in cloud computing to create multiple virtual instances of servers, desktops, storage devices, operating systems, etc. Here are some of the key pros and cons of virtualization:

Pros of Virtualization

- **1. Efficient Hardware Utilization:** Virtualization allows multiple virtual machines (VMs) to run on a single physical server, leading to efficient utilization of hardware resources. This consolidation reduces the need for physical hardware, resulting in cost savings for both users and service providers[1].
- **2. High Availability:** Virtualization provides advanced features that ensure virtual instances are always available. This minimizes downtime and maintains service continuity, which is crucial for businesses[1].
- **3. Disaster Recovery:** Virtualization simplifies disaster recovery by enabling easy data backup, duplication, and real-time data recovery. This ensures minimal data loss and quick restoration in case of a disaster[1].
- **4. Energy Savings:** By reducing the number of physical servers, virtualization helps save energy. This leads to lower power and cooling costs, contributing to overall cost savings[1].
- **5. Quick and Easy Setup:** Setting up virtual servers is much faster and easier compared to physical servers. This reduces the time and effort required for deployment, leading to increased productivity[1].
- **6. Cloud Migration:** Virtualization facilitates easy migration to cloud services. Data from existing servers can be easily transferred to cloud servers, reducing maintenance costs and improving efficiency[1].
- **7. Resource Optimization:** Virtualization allows efficient utilization of physical hardware by running multiple VMs on a single server. This leads to cost savings in terms of hardware, power, cooling, and space[1].

Cons of Virtualization

- **1. High Initial Investment:** While virtualization reduces costs in the long run, the initial setup costs for storage and servers can be higher than traditional setups[1].
- **2. Complexity:** Managing virtualized environments can be complex, especially as the number of VMs increases. This requires skilled staff and proper management tools[1].
- **3. Security Risks:** Virtualization introduces additional layers, which may pose security risks if not properly configured and monitored. Data hosted on shared resources can be vulnerable to unauthorized access[1].
- **4. Learning Curve:** Organizations may need to hire new IT staff with relevant skills or provide training to existing staff, which can increase costs[1].
- **5. Data Vulnerability:** Working on virtual instances on shared resources means that data is hosted on third-party resources, making it vulnerable to attacks and unauthorized access[1].
- **6. Performance Overhead:** Virtualization introduces a slight overhead due to the hypervisor layer. Proper resource allocation and monitoring are crucial to minimize performance impact[2].

Virtualization in grid

- Virtualization in grid computing refers to the process of creating a virtual environment that allows multiple applications and operating systems to run on a distributed network of computers, known as a grid. Here's a brief overview:

- Grid Computing: This involves using a decentralized network of computers to work together on large-scale tasks. Each computer in the grid can be thought of as a node that contributes its processing power, storage, and network resources1.

- Virtualization: In the context of grid computing, virtualization allows for the creation of virtual machines (VMs) or virtual environments on these nodes. This means that a single physical machine can host multiple VMs, each running its own operating system and applications2.

- Benefits: Virtualization in grid computing enhances resource utilization, scalability, and flexibility. It allows for better management of computational resources and can improve the efficiency of large-scale computations2.

- Applications: Grid computing with virtualization is used in various fields such as scientific research, drug development, market analysis, and seismic activity monitoring

Virtualization vs cloud computing

Feature	Virtualization	Cloud Computing
Technology	Creates virtual machines (VMs) on physical hardware	Delivers IT resources like servers, storage, and software over the internet
Location	VMs run on local hardware	Resources are located in remote data centers managed by cloud providers
Control	Users maintain control over VMs and underlying hardware	Users access resources on-demand with limited control over underlying infrastructure
Scalability	Can be scaled up or down by adding or removing physical hardware	Highly scalable – resources can be easily provisioned and released as needed

Feature	Virtualization	Cloud Computing
Cost	Requires upfront investment in hardware and virtualization software	Typically pay-as-you-go model based on resource usage
Management	Requires in-house IT expertise to manage VMs and hardware	Cloud provider handles infrastructure management and maintenance
Use Cases	Server consolidation, application isolation, disaster recovery	Remote work, web applications, data storage, big data analytics
Examples	Dedicated virtual servers (VPS/VDS), virtual desktops (VDI)	Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS)