

Before Cloud

Suppose you want to host a website, these are the following things that you would need to do:



Buy a stack of servers.



Monitoring and Maintain servers.

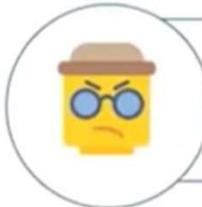


High traffic? More servers.

Potential Problem



If you consider costs then this setup is expensive.

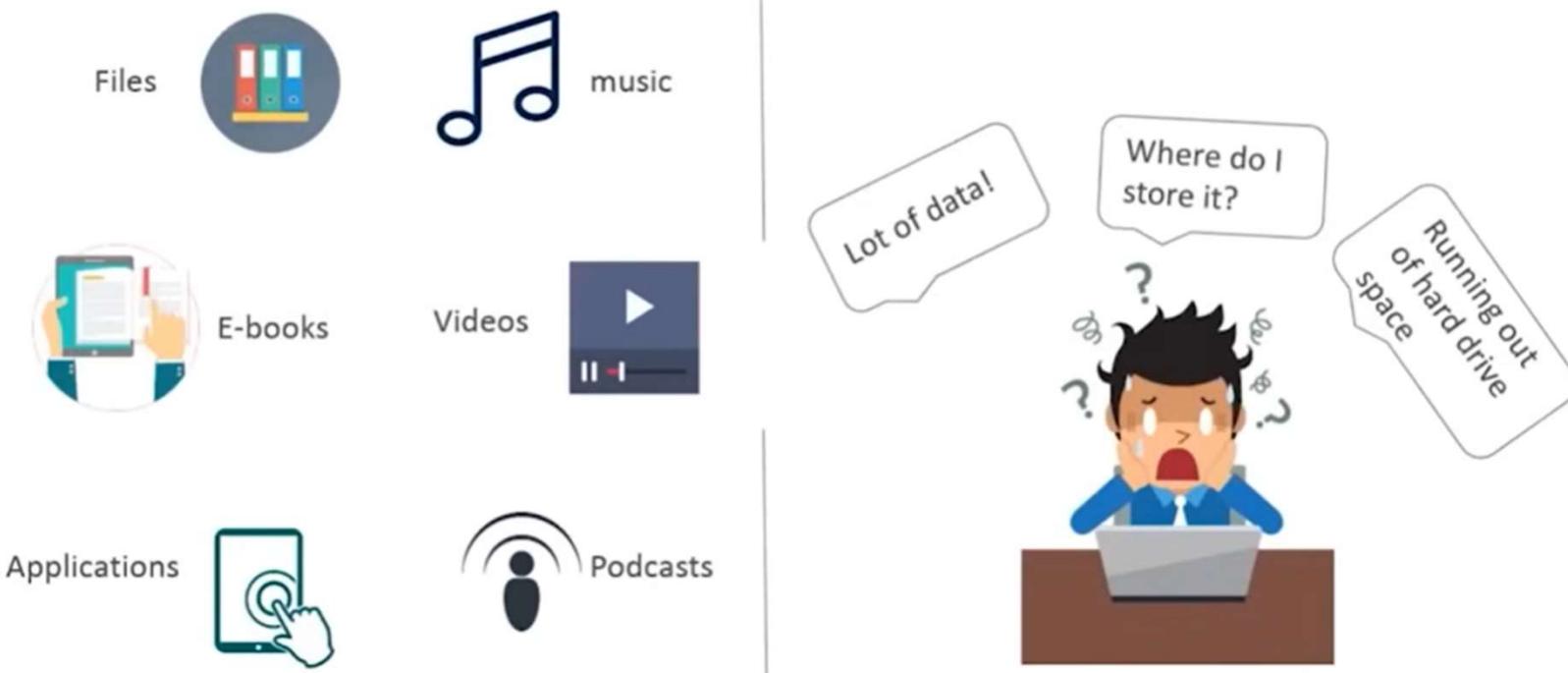


Troubleshooting problems can be tedious and may conflict with your business goals.



Since the traffic is varying, your servers will be idle most of the time.

Why Cloud?



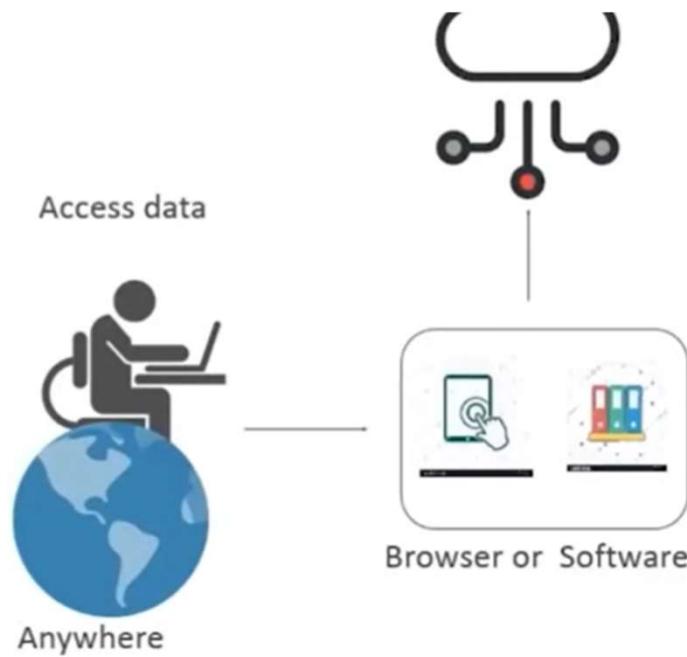
How cloud can support?



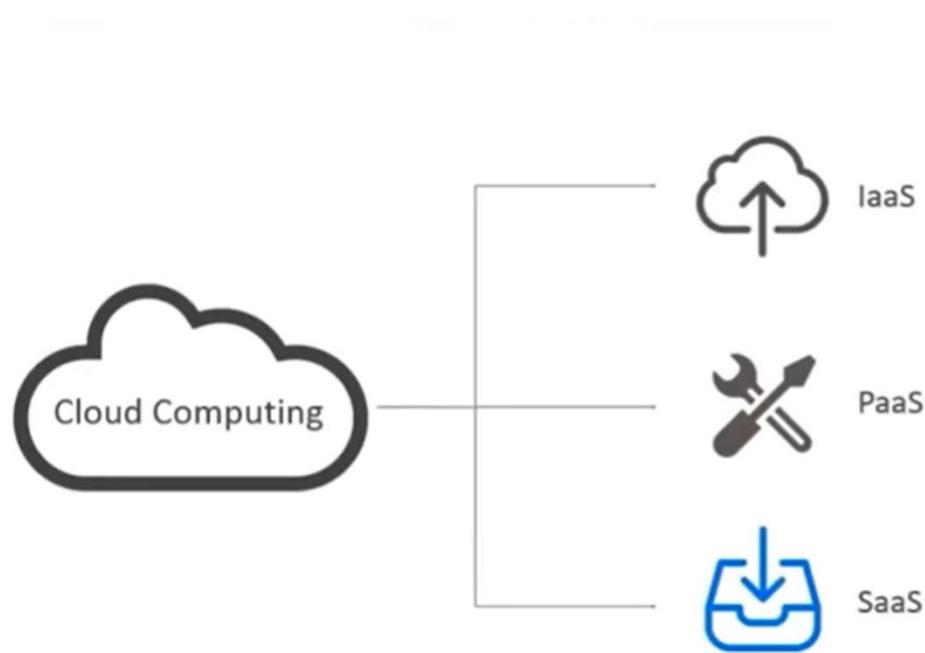
What cloud provide?

Cloud computing is:

- Storing data/applications on remote servers
- Processing data/applications from servers
- Accessing data/applications via Internet



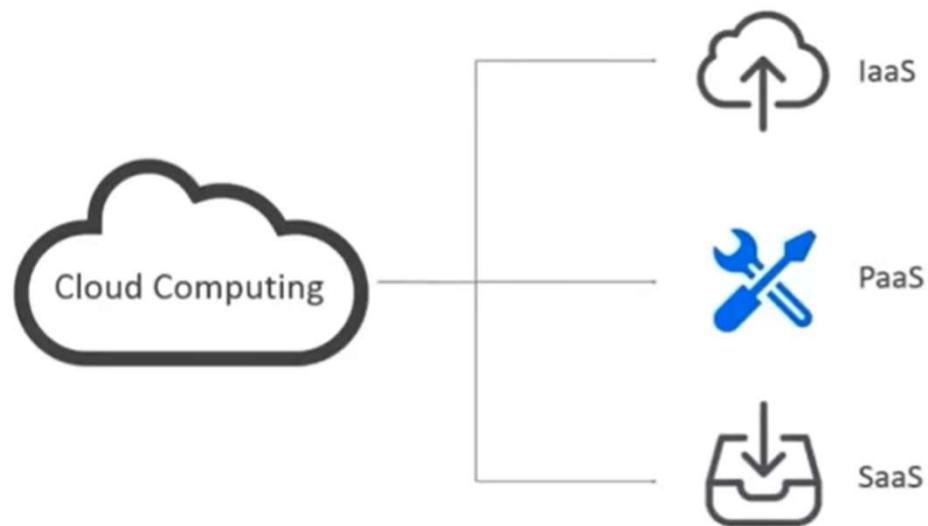
Service Models : SaaS



Software as a Service

- Cloud Provider leases applications or softwares which are owned by them to its client
- Example: salesforce.com provides the CRM(Customer Relation Manager) on a cloud infrastructure to its client and charges them for it, but the software is owned by the salesforce company only

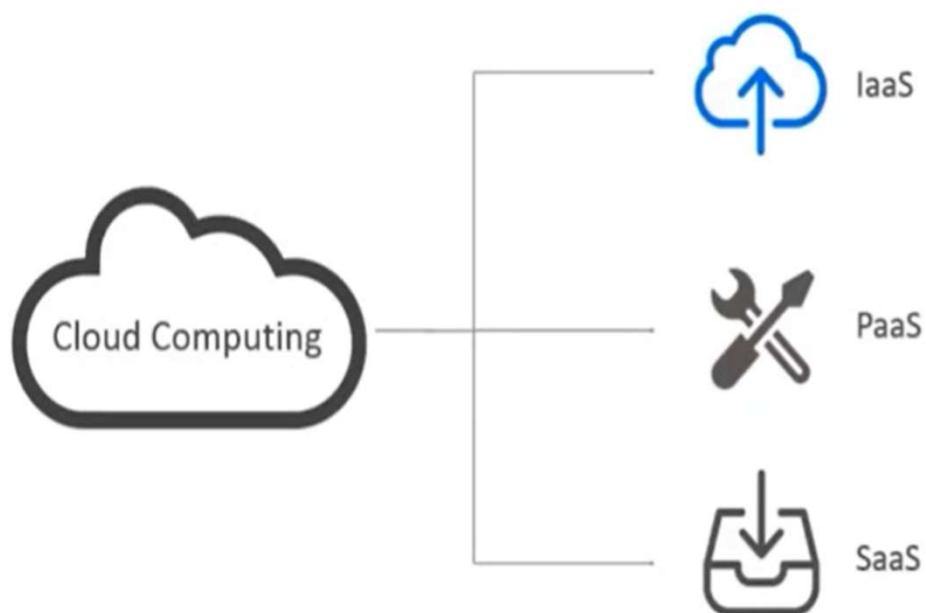
Service Models : PaaS



Platform as a Service

- No control over the underlying architecture including OS, storage, servers etc.
- The Cloud Provider gives the ability to the customer to deploy customer created apps using programming languages, tools etc that are provided by the Cloud Provider.

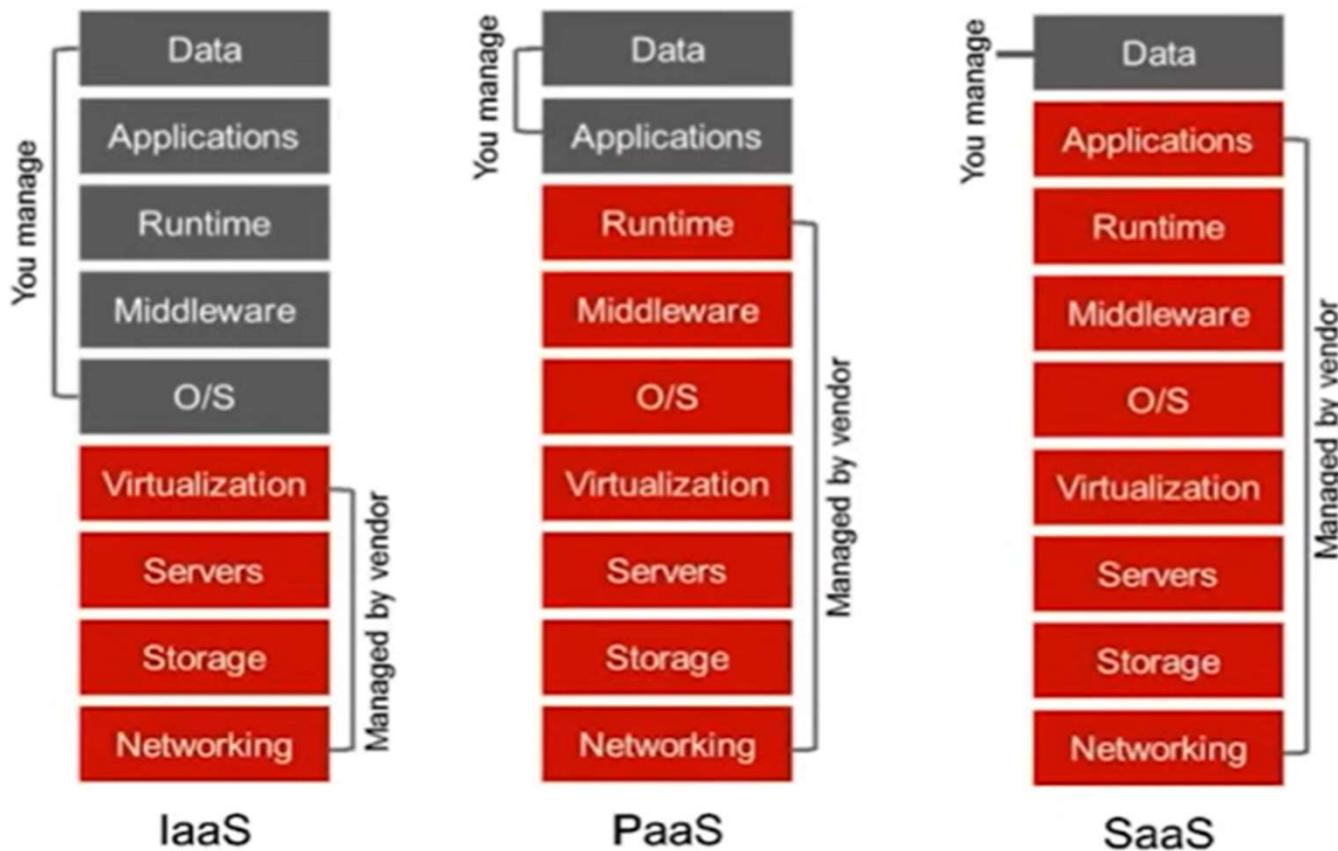
Service Models : IaaS



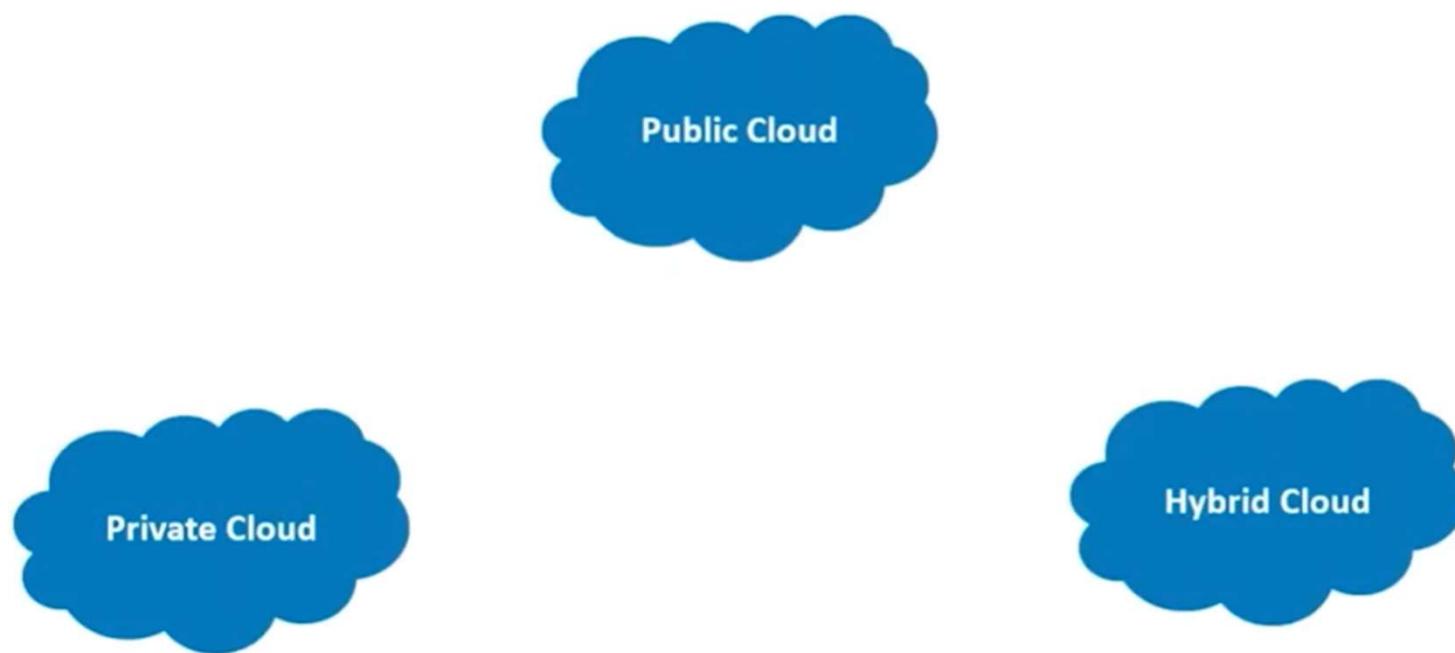
Infrastructure as a Service

- Provides virtualized computing resources over the Internet
- No worries about the underlying physical machine.
- Abstract the user from the physical machine

Service Model Architecture



Cloud Deployment Models



Public



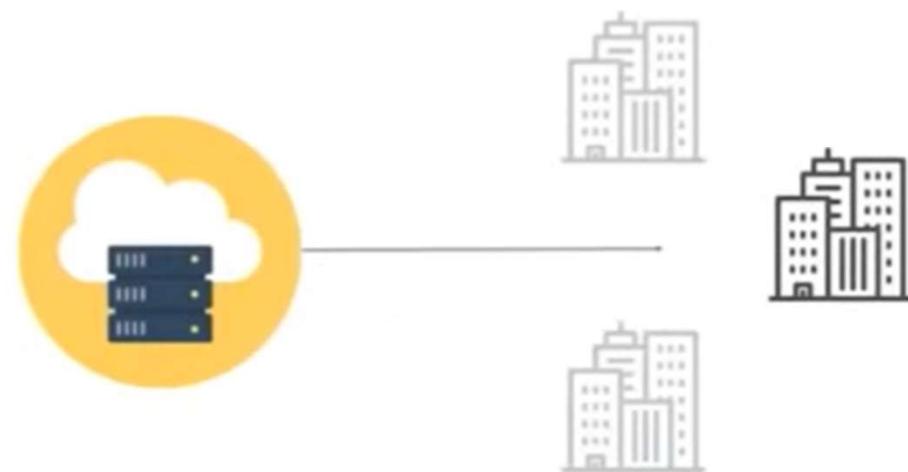
- A service provider makes resources, such as applications and storage, available to the general public over the Internet.
- Easy and inexpensive set-up because hardware, application and bandwidth costs are covered by the provider.
- No wasted resources because you pay for what you use



Private



- Offers hosted services to a limited number of people behind firewall, so it minimizes the security concerns.
- Private cloud gives companies direct control over their data.



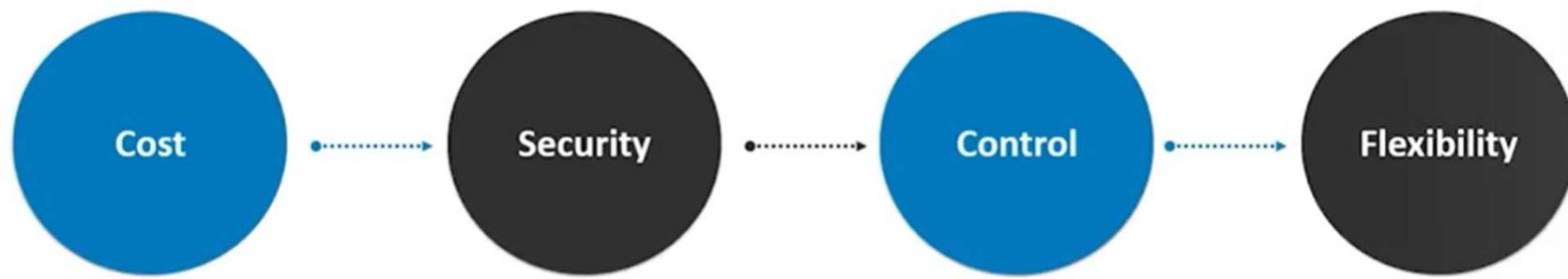
Hybrid



- A cloud computing environment which uses a mix of on-premises, private cloud and third-party, public cloud services
- It helps you leverage the best of both worlds



On Premises VS On Cloud -- Comparison

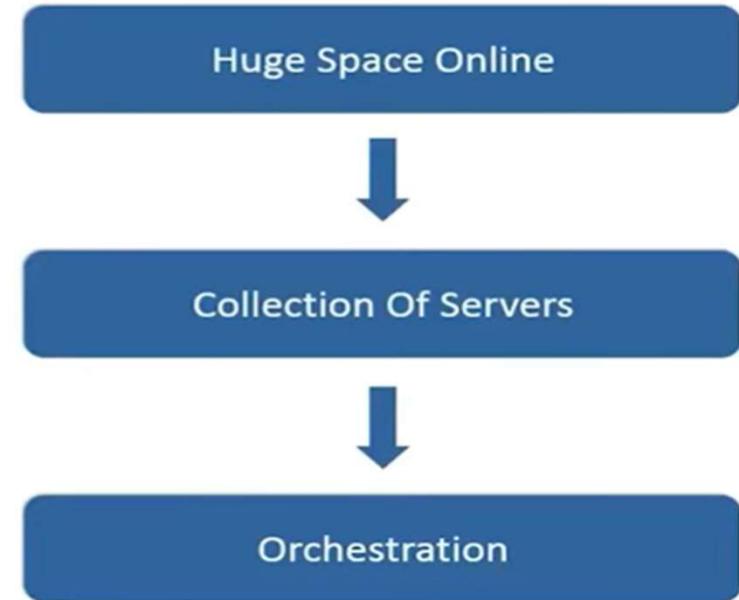
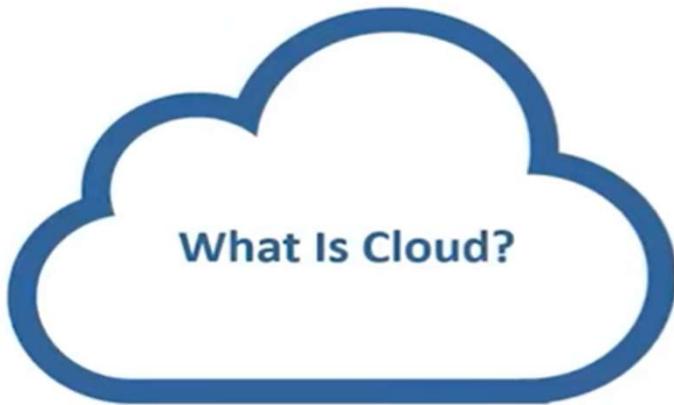


- Write Your own piece of code on your own server
- Provide your own storage
- You will be fully responsible for Security
- You need to take care of availability
- You are supposed to maintain and monitor all applications
- You will have complete control over your data and access

Colo Deployment :Colocation deployment, also known as "colo" deployment, is when a business rents space in a data center to store their computing hardware and servers

- The time to deploy a new workload in an On Prem or Colo deployment is typically at least a week:
- The purchase needs to be approved and the server ordered
- The server needs to be delivered
- It needs to be physically racked up and cabled
- The server team need to install and configure the OS, patches, any standard software and the applications
- The network team need to configure switches, routers and firewalls
- The storage team need to configure the storage system and SAN switches
- The lead time is quicker when provisioning virtual machines but it still takes time for the different teams to manually complete their tasks

On Cloud



Database , Storage , Computation, Compliance, Security , analytics and many other services

YOU JUST FOCUS ON YOUR BUZINESS NEEDS

On-Premise Vs Cloud Computing



Security

- Owner Manages Security
- Quality boils down to owner experience



Security

- Vendor Dependence
- Owners are expected to give up control

Few Issues – On Cloud



Virtualization

Virtualization is a technique, which allows to share single physical instance of an application or resource among multiple organizations or tenants (customers). It does so by **assigning a logical name** to a physical resource and providing a **pointer to that physical resource** on demand.

Virtualization – What it is?

- Creating a virtual machine over existing operating system and hardware is referred as Hardware Virtualization. Virtual Machines provide an environment that is logically separated from the underlying hardware.
- The machine on which the virtual machine is created is known as **host machine** and **virtual machine** is referred as a **guest machine**. This virtual machine is managed by a software or firmware, which is known as **hypervisor**.

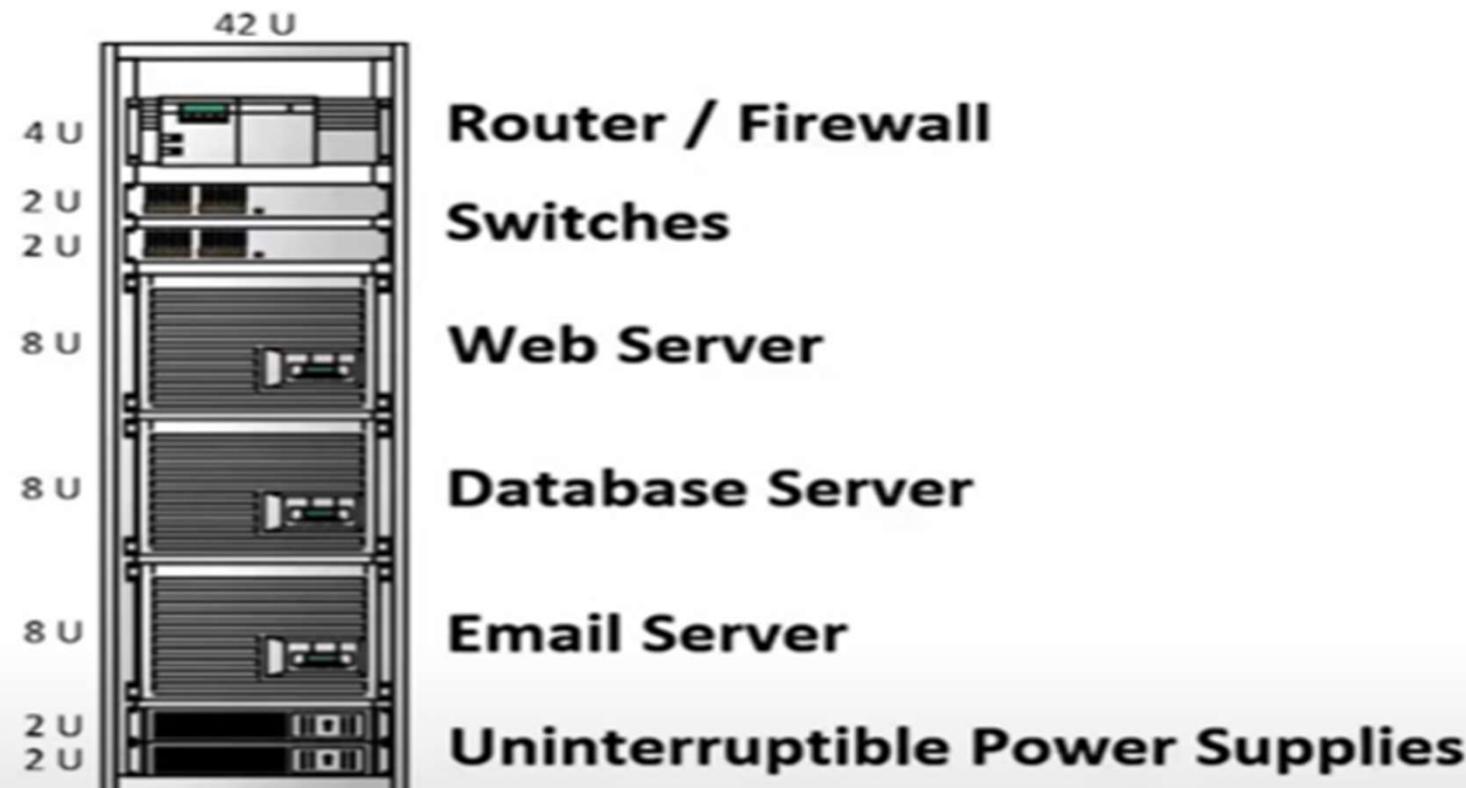
Virtualization .. Cont....

Virtualization is the underlying core technology of cloud computing

Advantages

- • Using virtualization, the physical infrastructure owned by the service provider is shared among many users, increasing the resource utilization.
- • Virtualization provides efficient resource utilization and increased return on investment (ROI).
- • Ultimately, it results in low capital expenditures (CapEx) and operational expenditures (OpEx).
- • Promotes the green IT by reducing energy wastage.
- • Dynamic data center • Improves disaster recovery

Traditional Approach



Traditional Approach

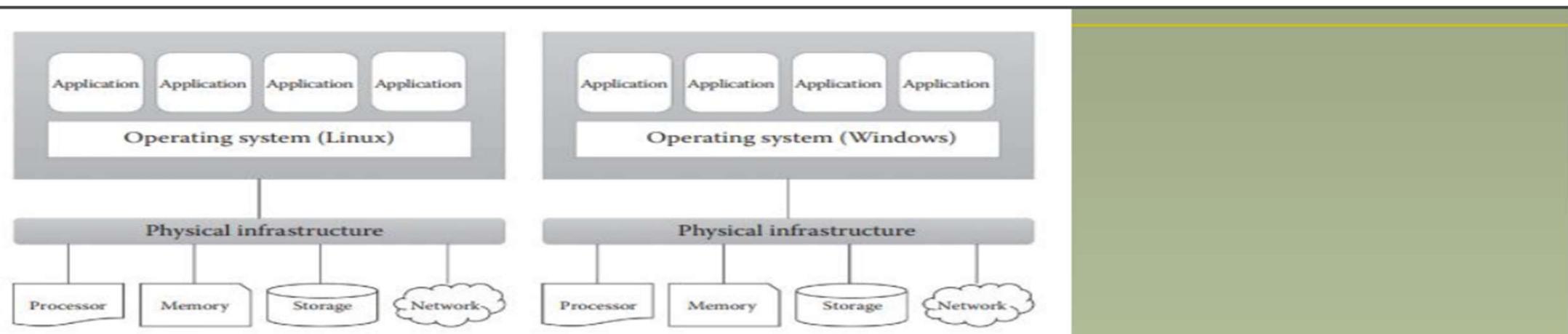


FIGURE 7.1
Before virtualization.

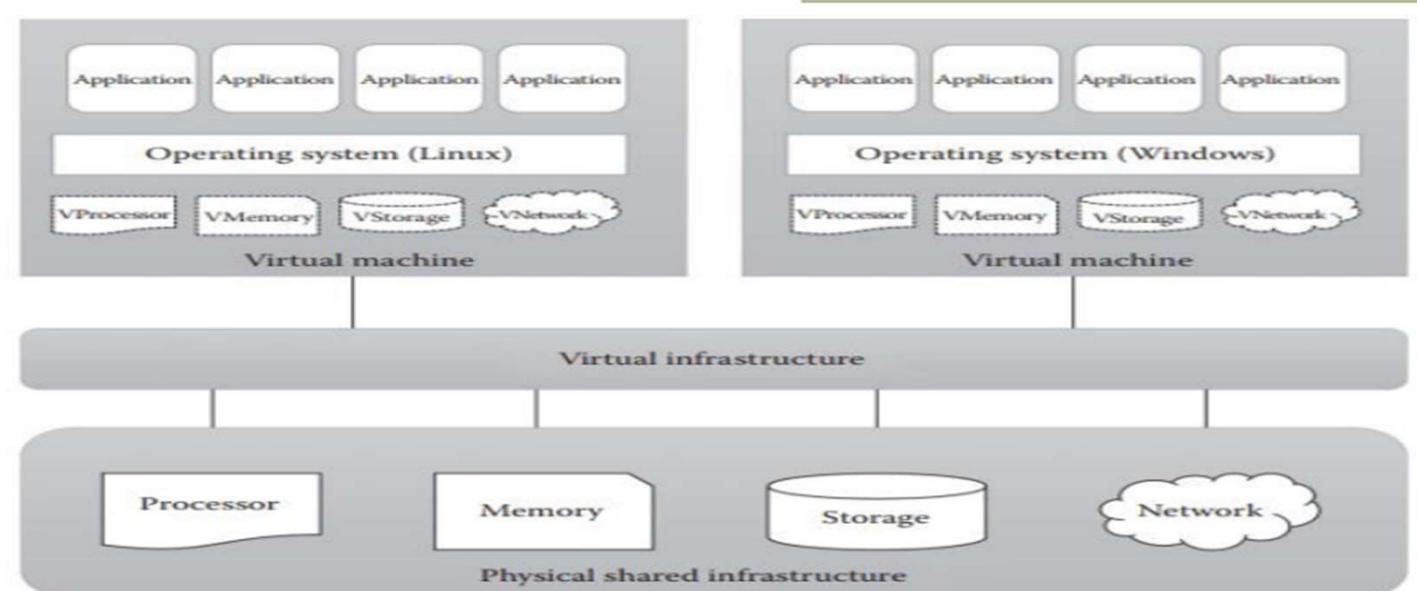
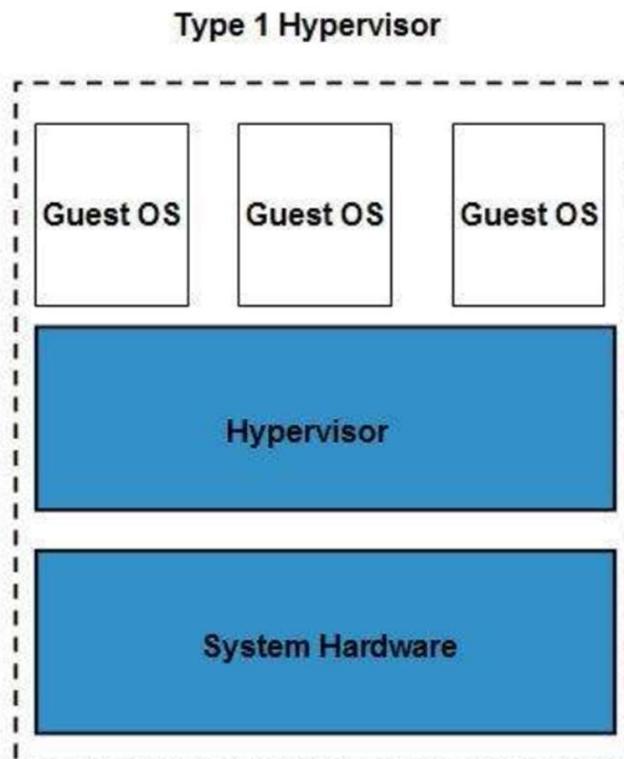


FIGURE 7.2
After virtualization.

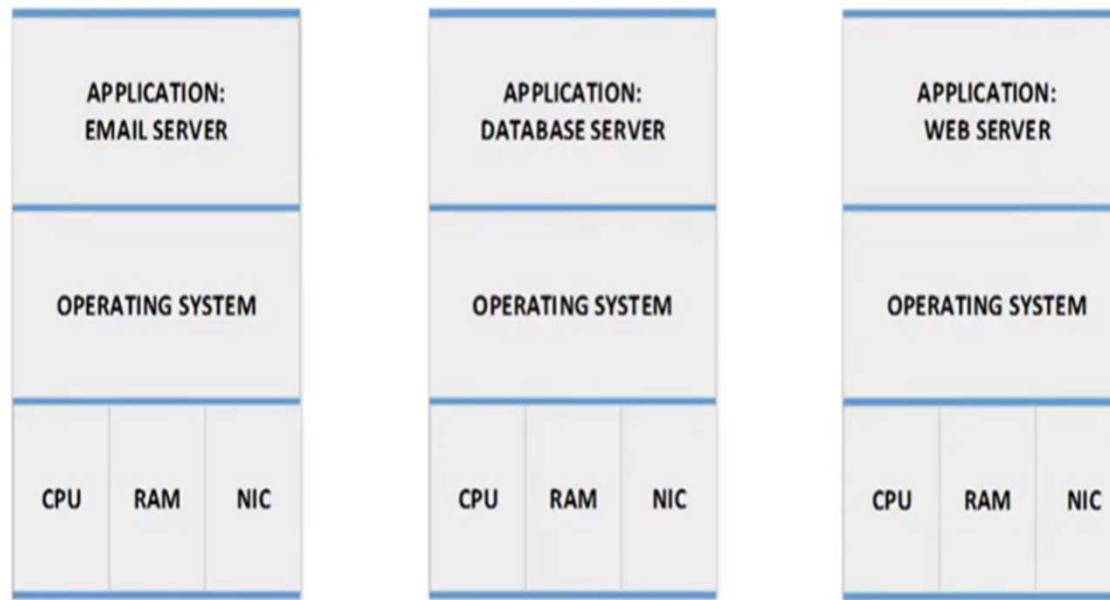
Hypervisor

The **hypervisor** is a firmware or low-level program that acts as a Virtual Machine Manager. There are two types of hypervisor:

Type 1 hypervisor executes on bare system. LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, VirtualLogic VLX are examples of Type 1 hypervisor. The following diagram shows the Type 1 hypervisor.



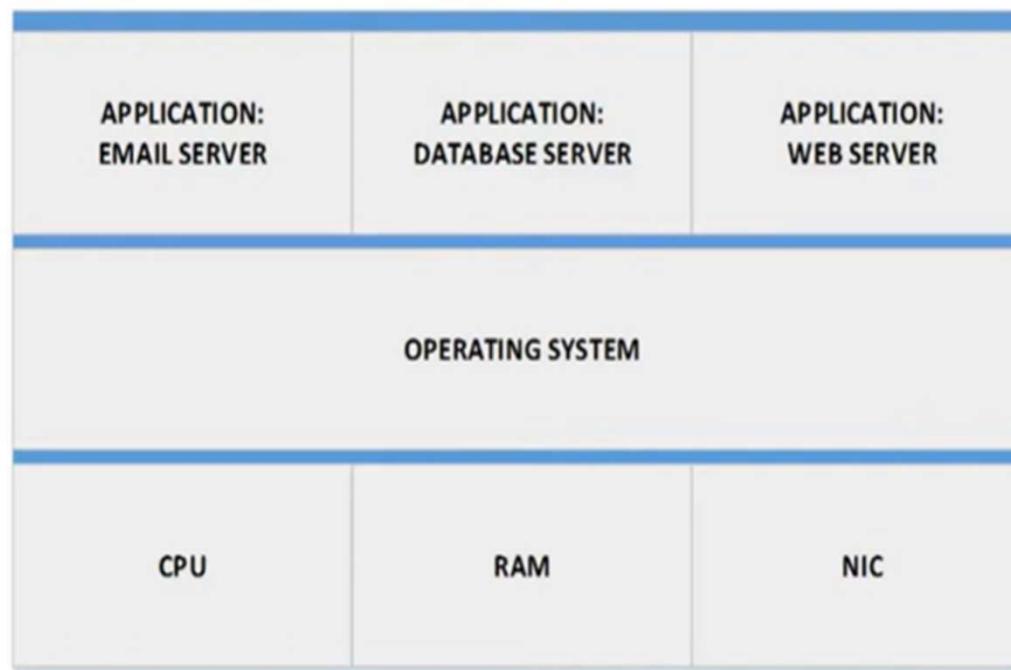
Before Virtualization



- Server utilization (CPU, RAM, NIC etc.) around 15%
- I had to pay for each separate server, and they're all using power, space and cooling

Some common BAD Practice

UPnP / PPT Pre-processor

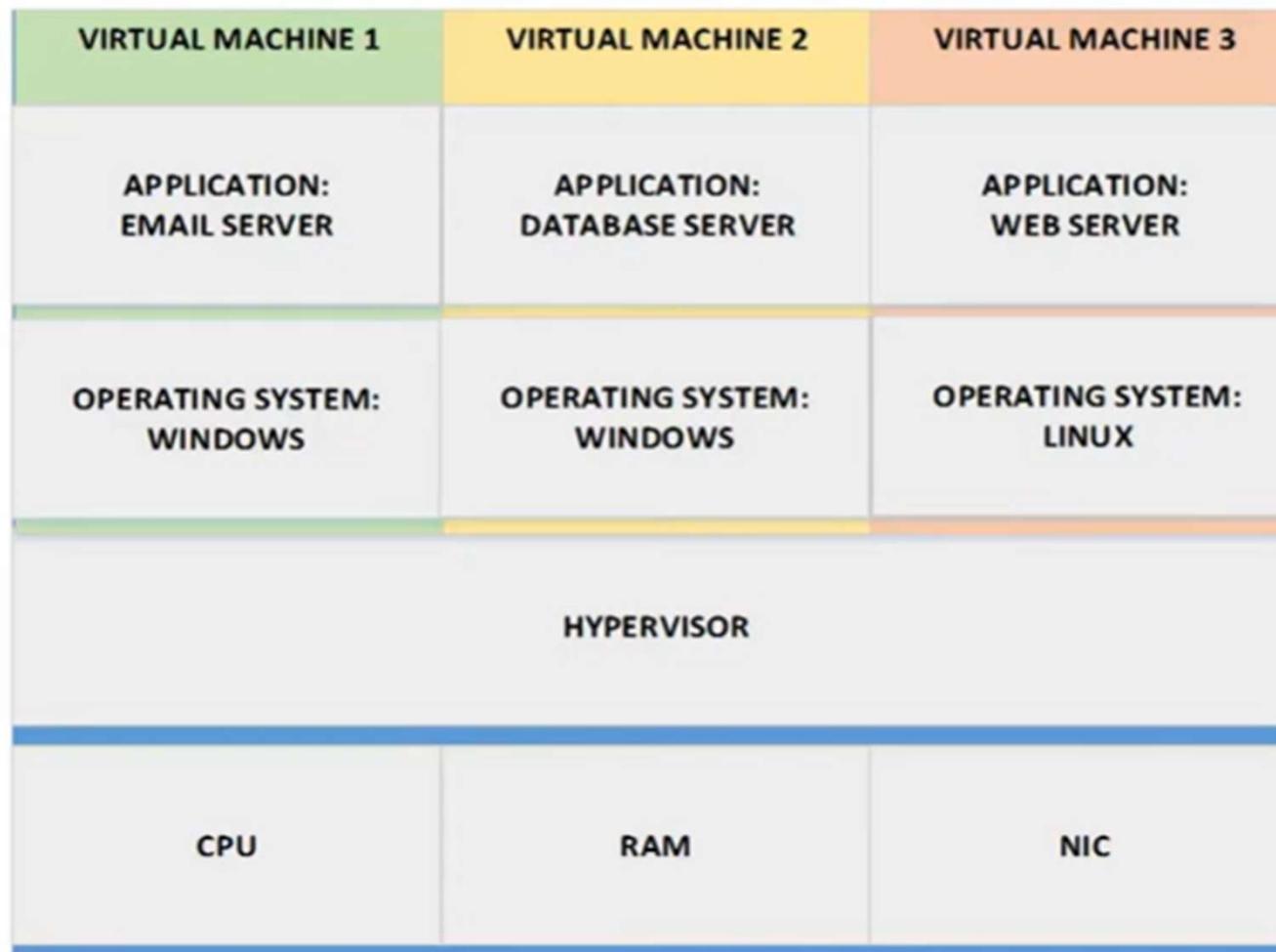


- Putting multiple applications on the same server would improve utilization

Hypervisor –Native/Bare-Metal-type 1

- Type 1 Hypervisors run directly on the system hardware
 - VMware ESXi (part of the vSphere suite)
 - Microsoft Hyper-V
 - Red Hat KVM
 - Oracle VM Server
 - Citrix XenServer

Server Virtualization



Implementation

The screenshot shows the vSphere Web Client interface. On the left, the inventory tree displays two hosts: '10.2.1.11' and '10.2.1.12'. Host '10.2.1.11' contains several virtual machines: 'Open Filer 1', 'Nostalgia2', 'XP-1', 'XP-2', and 'Nostalgia1'. Host '10.2.1.12' contains 'Open Filer 2'. A red circle highlights the host names in the inventory tree. Another red circle highlights the title bar '10.2.1.11 VMware ESXi, 5.0.0, 469512 | Evaluation (60 days remaining)'. A cursor arrow points towards the inventory tree.

Configuration Issues

- ESXi Shell for the host has been enabled
- SSH for the host has been enabled

General

Manufacturer:	VMware, Inc.
Model:	VMware Virtual Platform
CPU Cores:	4 CPUs x 3.206 GHz
Processor Type:	AMD Phenom(tm) II X6 1090T Processor
License:	Evaluation Mode
Processor Sockets:	2
Cores per Socket:	2
Logical Processors:	4
Hyperthreading:	Inactive
Number of NICs:	8
State:	Connected
Virtual Machines and Templates:	4
vMotion Enabled:	Yes
VMware EVC Mode:	Disabled
vSphere HA State:	N/A
Host Configured for FT:	No
Active Tasks:	
Host Profile:	
Image Profile:	ESXi-5.0.0-469512-standard

Resources

Storage	Status	Drive Type
ESXi-1-local-disk	Normal	Non-SSD
NFS-2	Normal	Unknown
OF1-iscsi-1	Normal	Non-SSD
OF2-iscsi-2	Normal	Non-SSD

Network	Type	State
Engineering	Standard port group	Green checkmark
Sales	Standard port group	Green checkmark
VM Network	Standard port group	Green checkmark

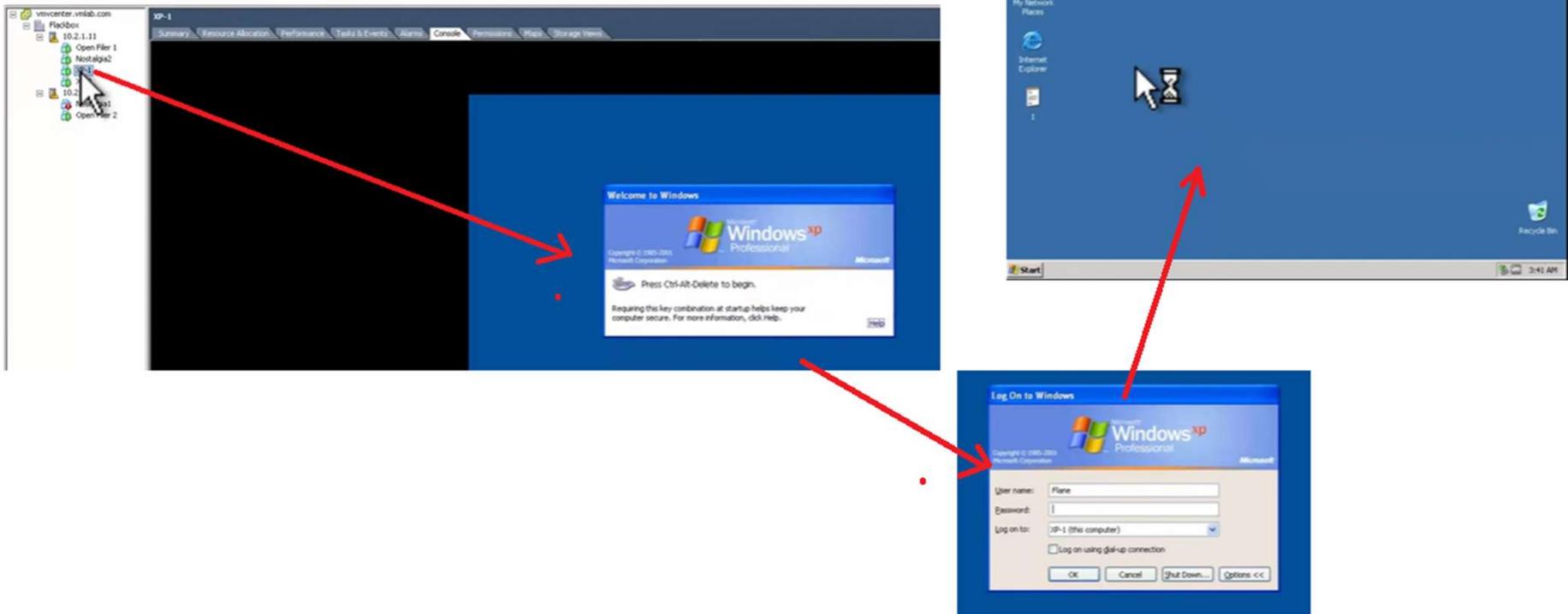
Fault Tolerance

Fault Tolerance Version:	2.0.1-3.0.0-3.0.0
Total Primary VMs:	--

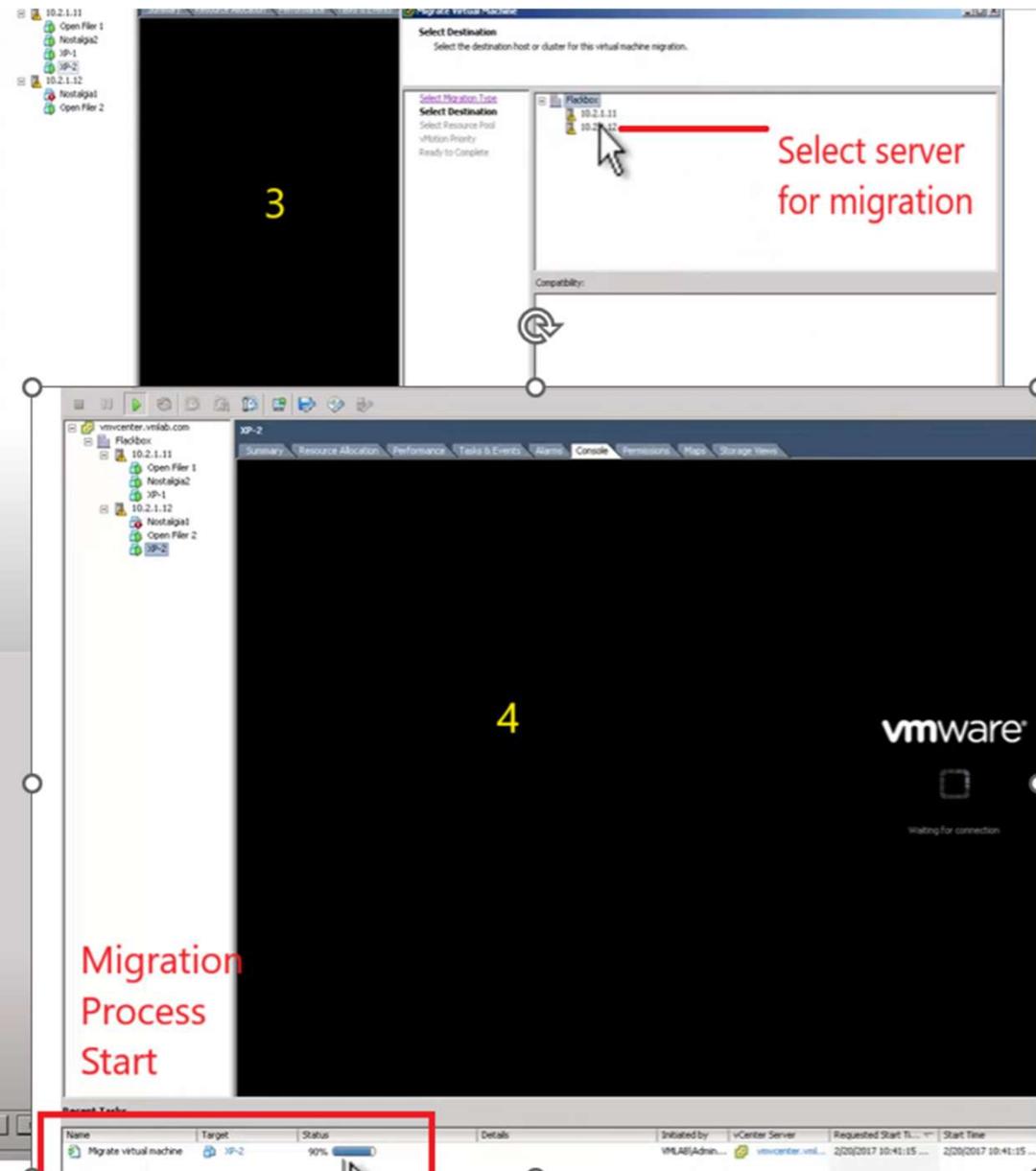
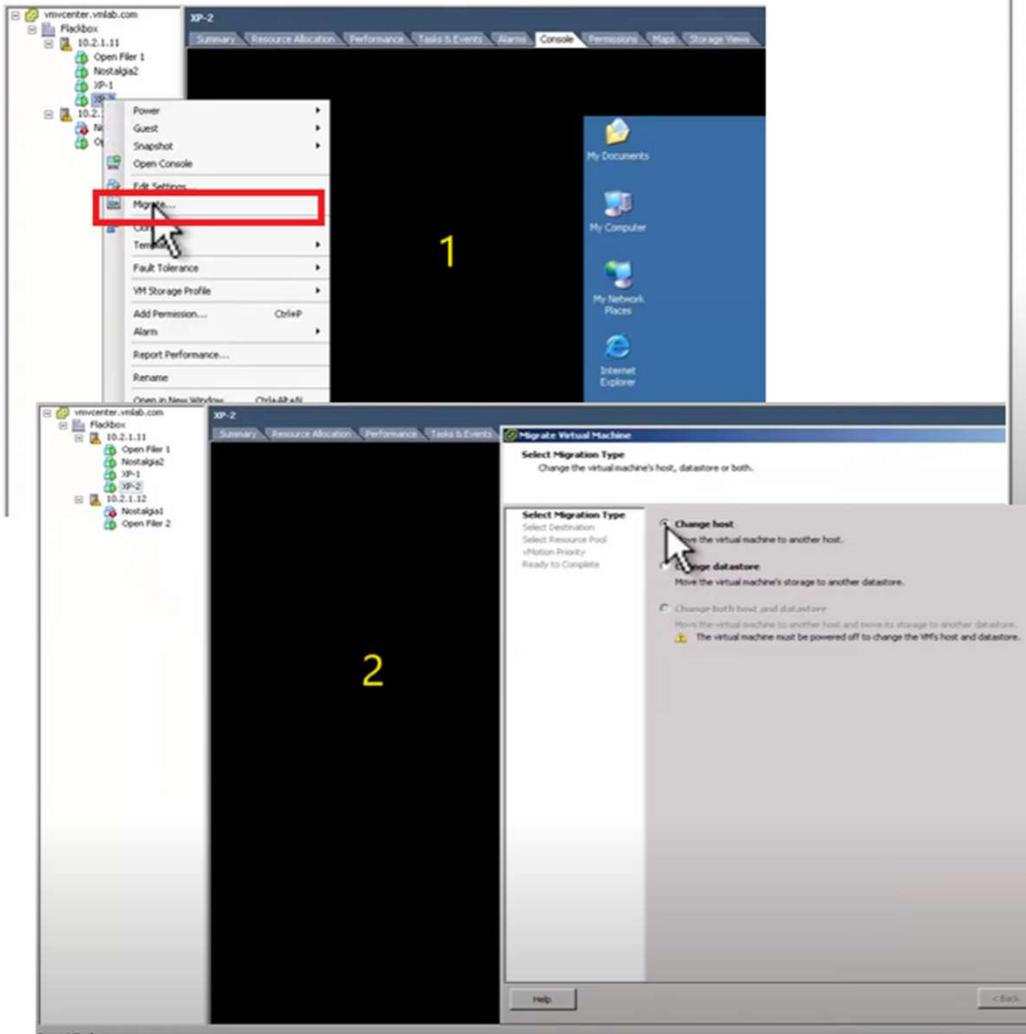
2 ESXi Servers
running
multiple
VM
snapshots

10.11 machine
And 10.12
machine will
have separate
Secondary
memory

XP-1 Instance



Migration



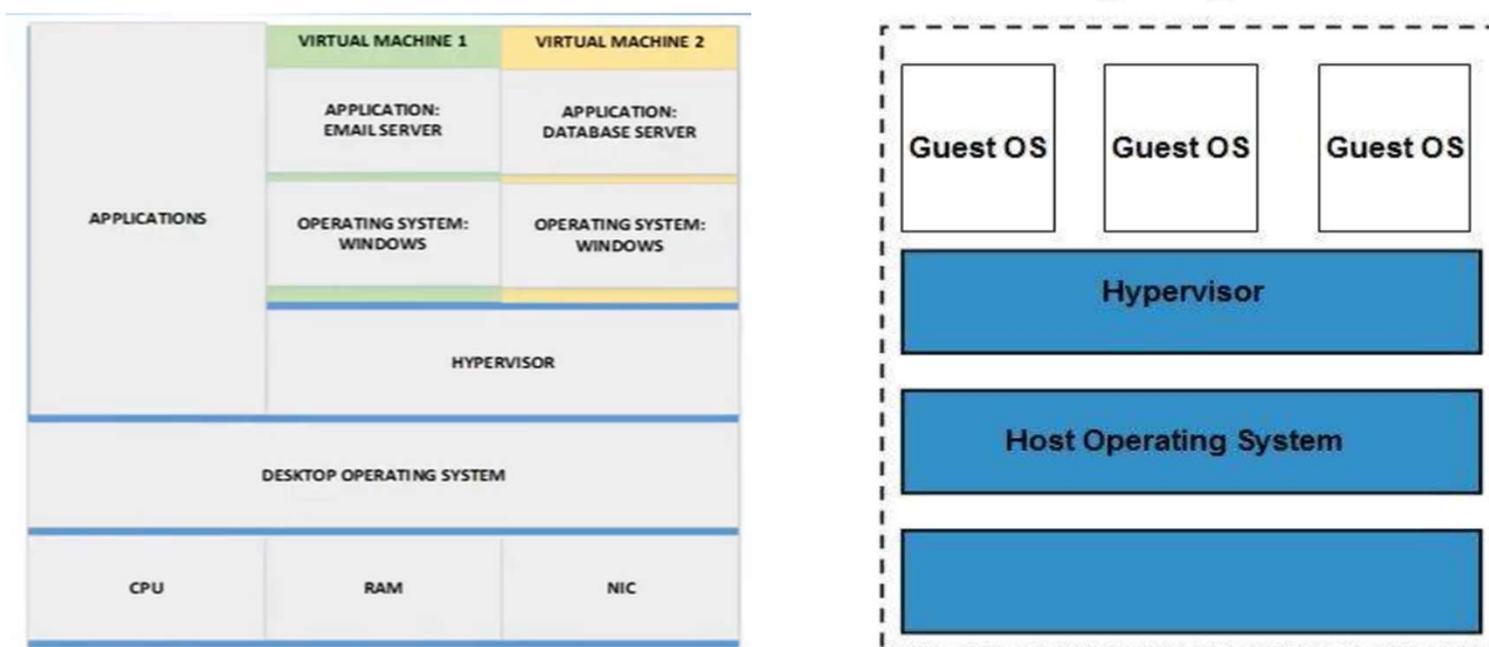
Hypervisor – Hosted Hypervisor- type 2

- Type 2 Hypervisors run on top of a host operating system
 - VMware Workstation, Player and Fusion
 - VirtualBox
 - QEMU
 - Parallels

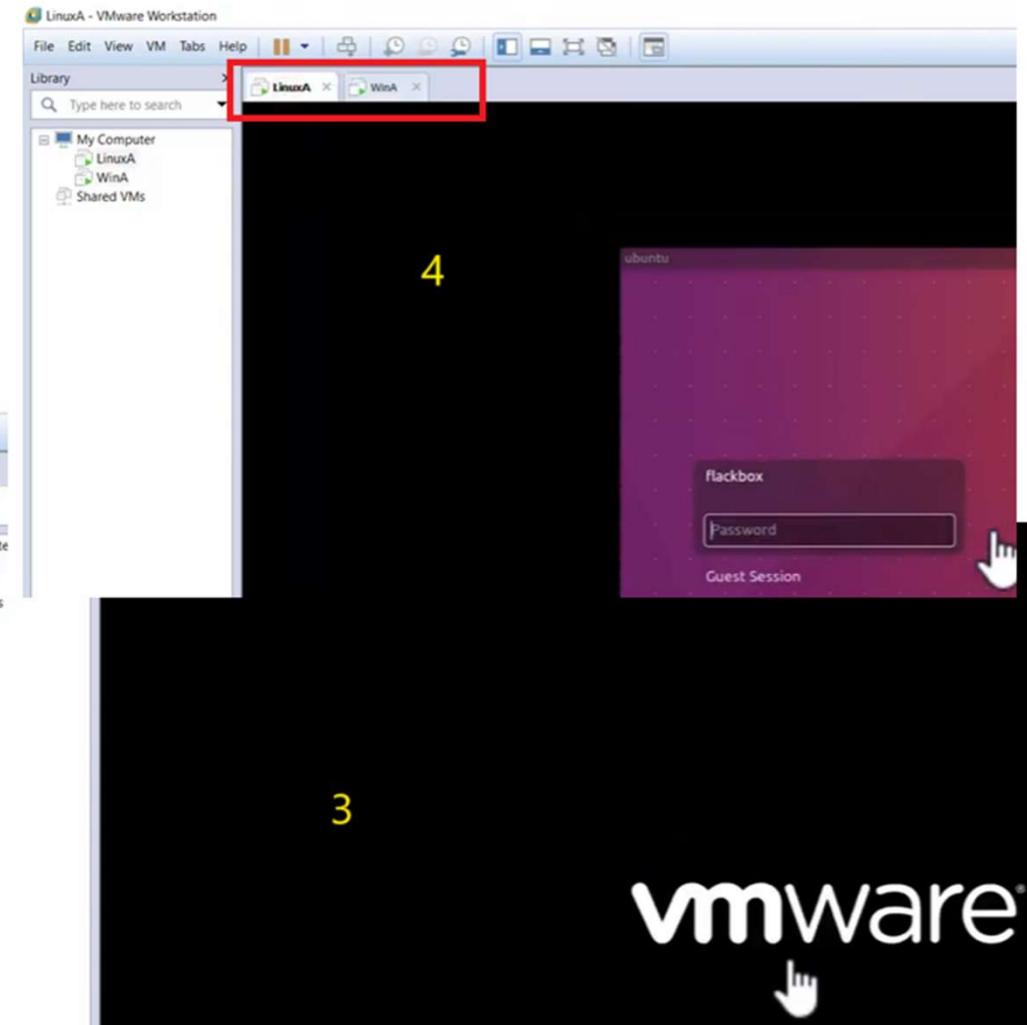
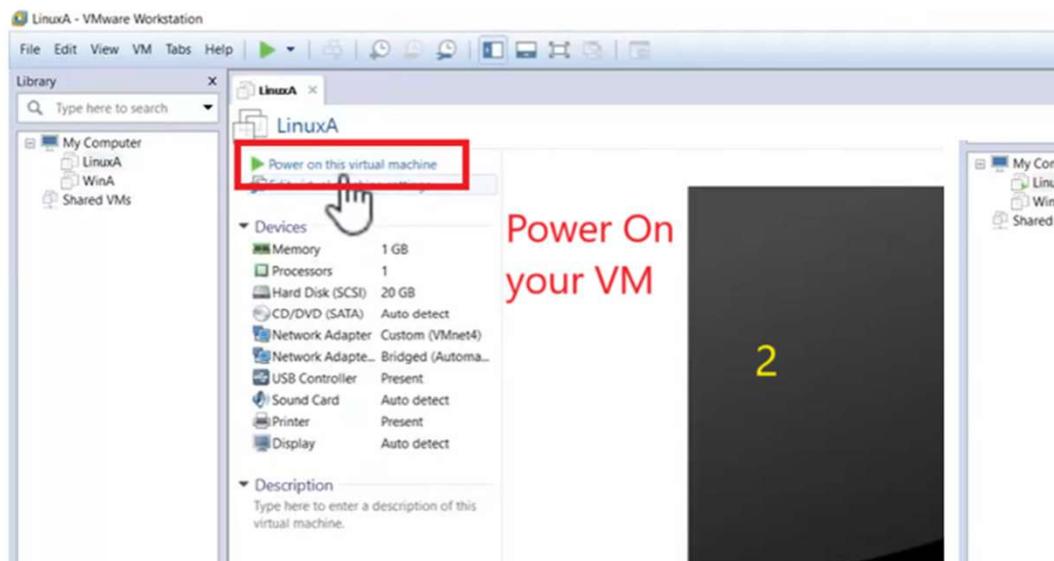
Hypervisor

The **type1 hypervisor** does not have any host operating system because they are installed on a bare system.

Type 2 hypervisor is a software interface that emulates the devices with which a system normally interacts. Containers, KVM, Microsoft Hyper V, VMWare Fusion, Virtual Server 2005 R2, Windows Virtual PC and **VMWare workstation 6.0** are examples of Type 2 hypervisor. The following diagram shows the Type 2 hypervisor.



Hypervision-Type2 Implementation



Summary

Summary of Hypervisors

Hypervisor	Vendor	Type	License
Xen	University of Cambridge Computer Laboratory	Type 1	GNU GPLv2
VMWare ESXi	VMware, Inc.	Type 1	Proprietary
Hyper-V	Microsoft	Type 1	Proprietary
KVM	Open virtualization alliance	Type 2	GNU general public license
VMWare workstation	VMware, Inc.	Type 2	Shareware
Oracle Virtualbox	Oracle Corporation	Type 2	GNU general public license version 2

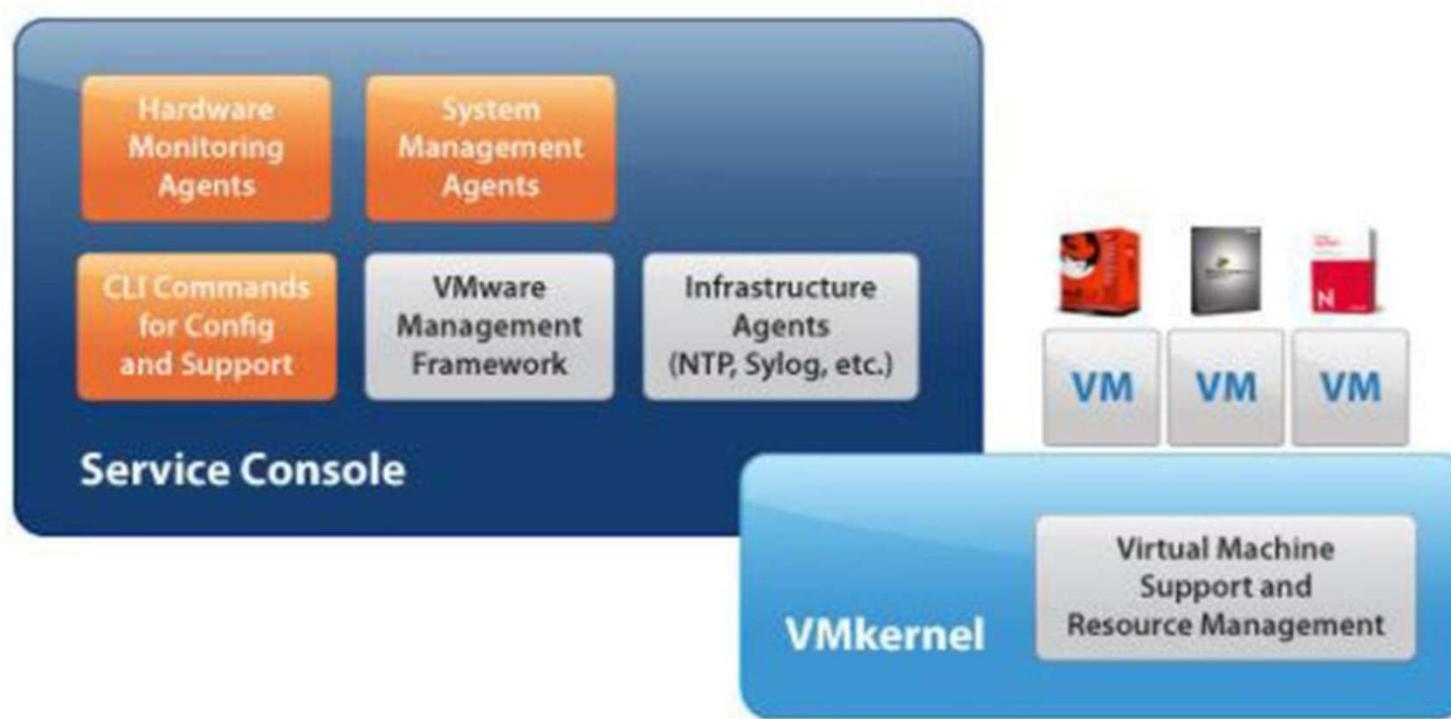
ESX & ESXi

ESX (Elastic Sky X) is the VMware's enterprise server virtualization platform.

In ESX, VMkernel is the virtualization kernel which is managed by a console operating system which is also called as Service console.

Which is linux based and its main purpose is to provide a Management interface for the host and lot of management agents and other third party software agents are installed on the service console to provide the functionaliests like hardware management and monitoring of ESX hypervisor.

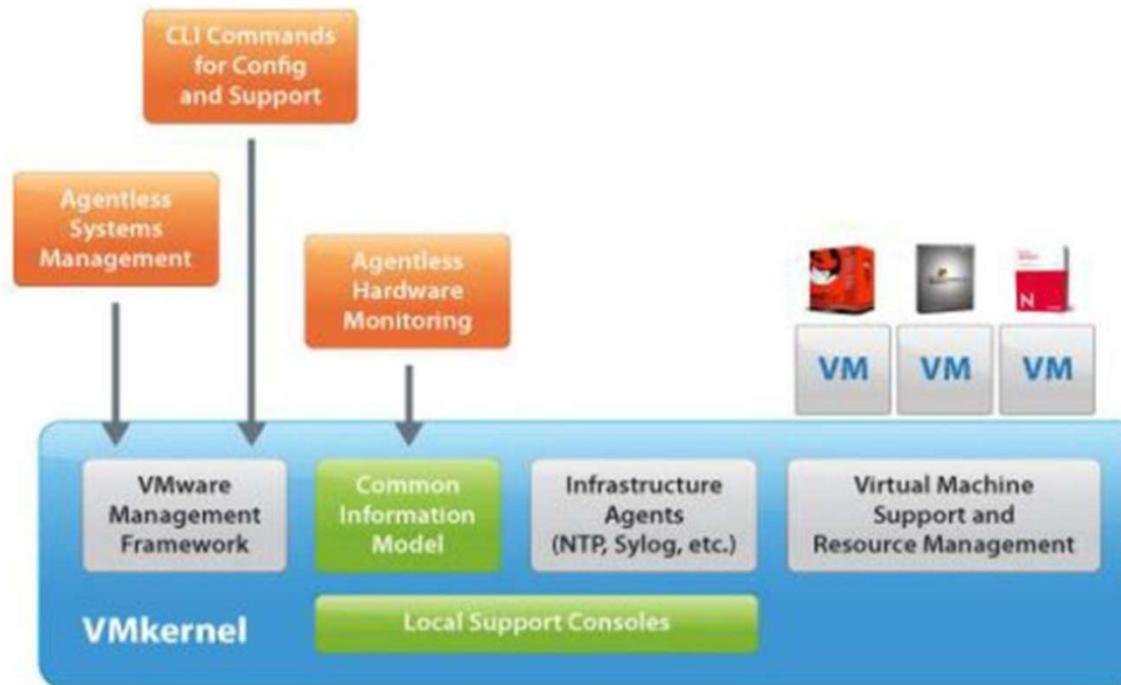
ESX



ESXi

ESXi (Elastic sky X Integrated) is also the VMware's enterprise server virtualization platform. In ESXi, Service console is removed. All the VMware related agents and third party agents such as management and monitoring agents can also run directly on the VMkernel. ESXi uses Direct Console User Interface (DCUI) instead of a service console to perform management of ESXi server. ESXi installation will happen very quickly as compared to ESX installation.

ESXi



Hardware Virtualization Types

Hardware virtualization is further subdivided into the following types:

- ✓ Full Virtualization – In it, the complete simulation of the actual hardware takes place to allow software to run an unmodified guest OS.
- ✓ Para Virtualization – In this type of virtualization, software unmodified runs in modified OS as a separate system.
- ✓ Partial Virtualization – In this type of hardware virtualization, the software may need modification to run.

Full Virtualization

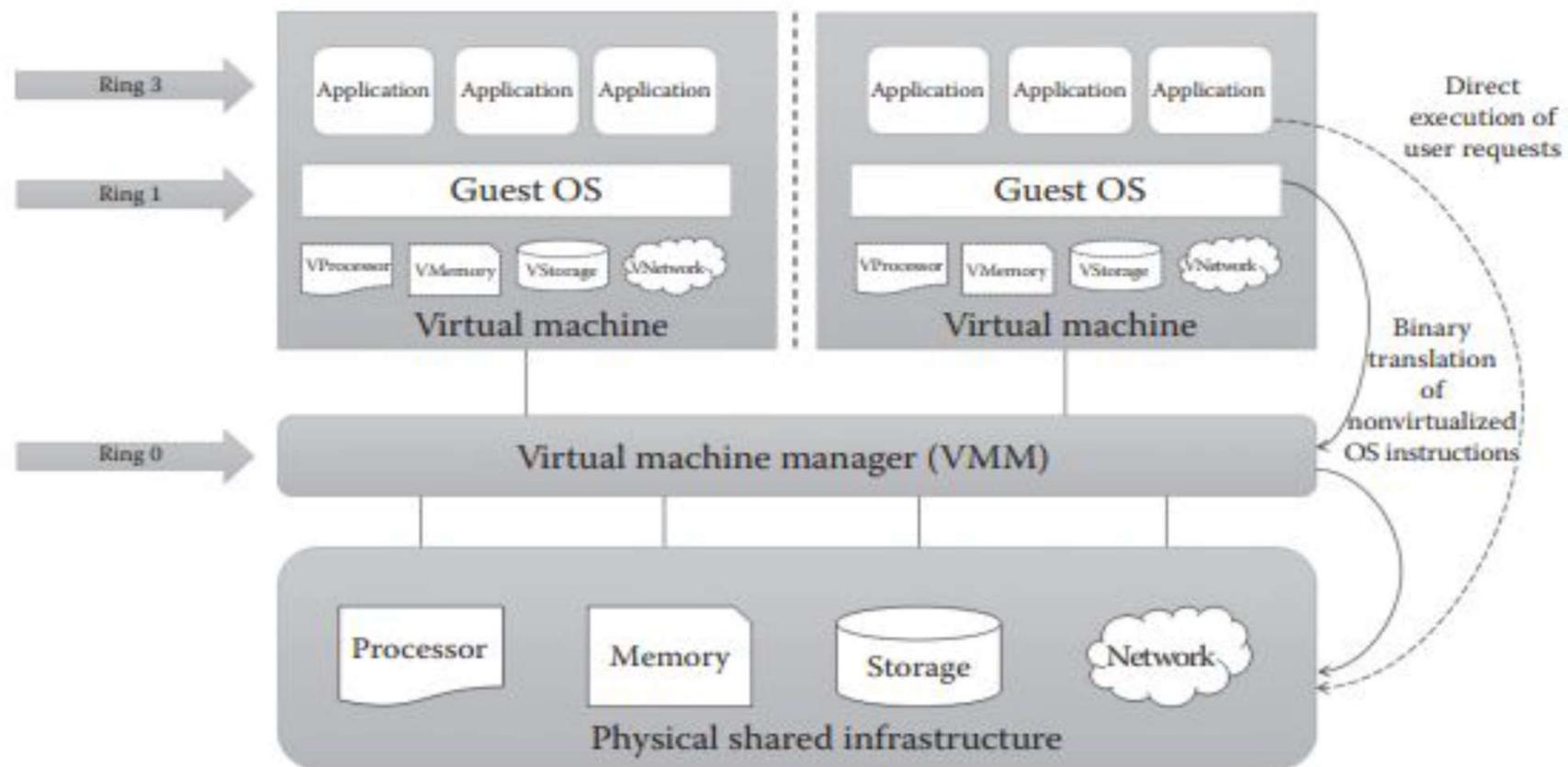


FIGURE 7.10
Full virtualization.

Para Virtualization

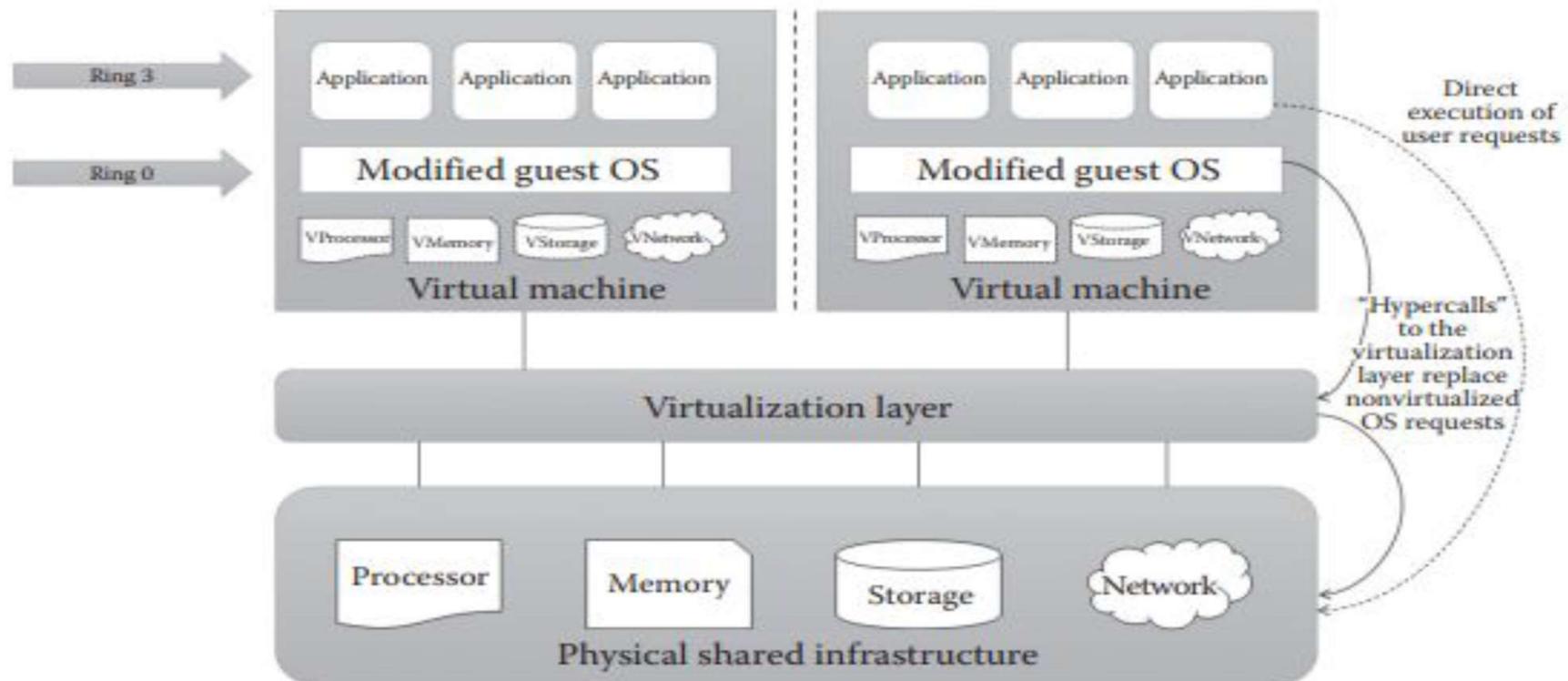
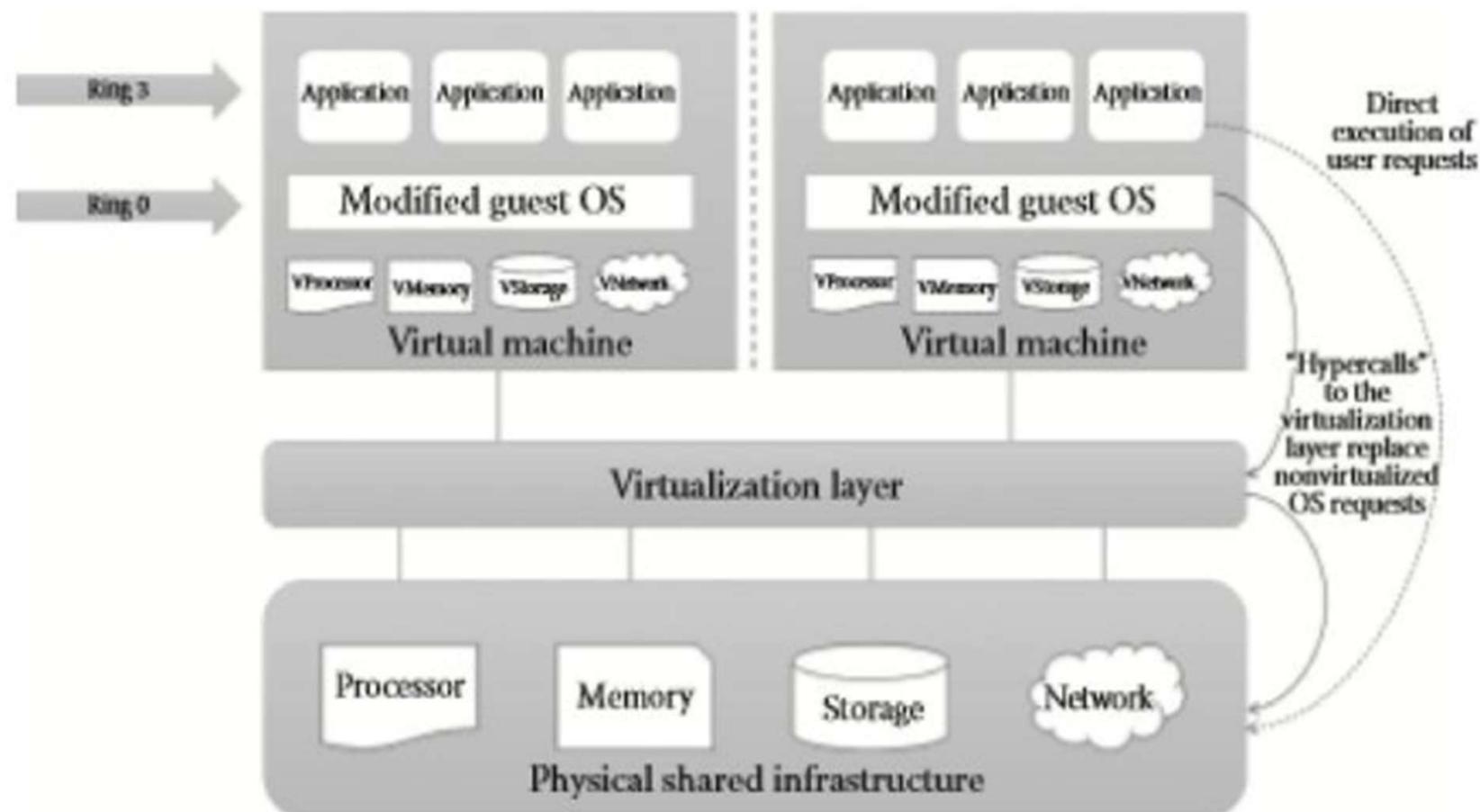


FIGURE 7.11
Paravirtualization.

Partial Virtualization

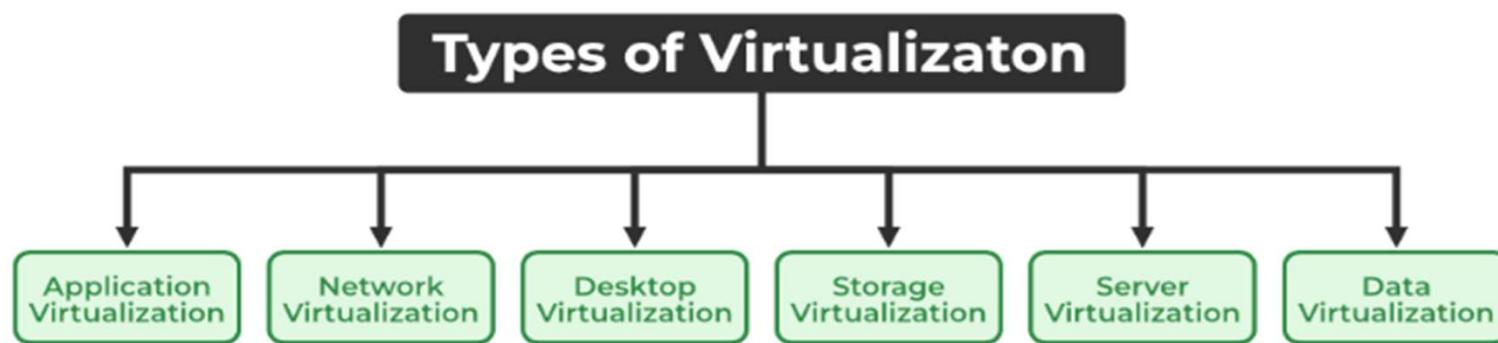


Summary

Summary of the Different Approaches to Virtualization

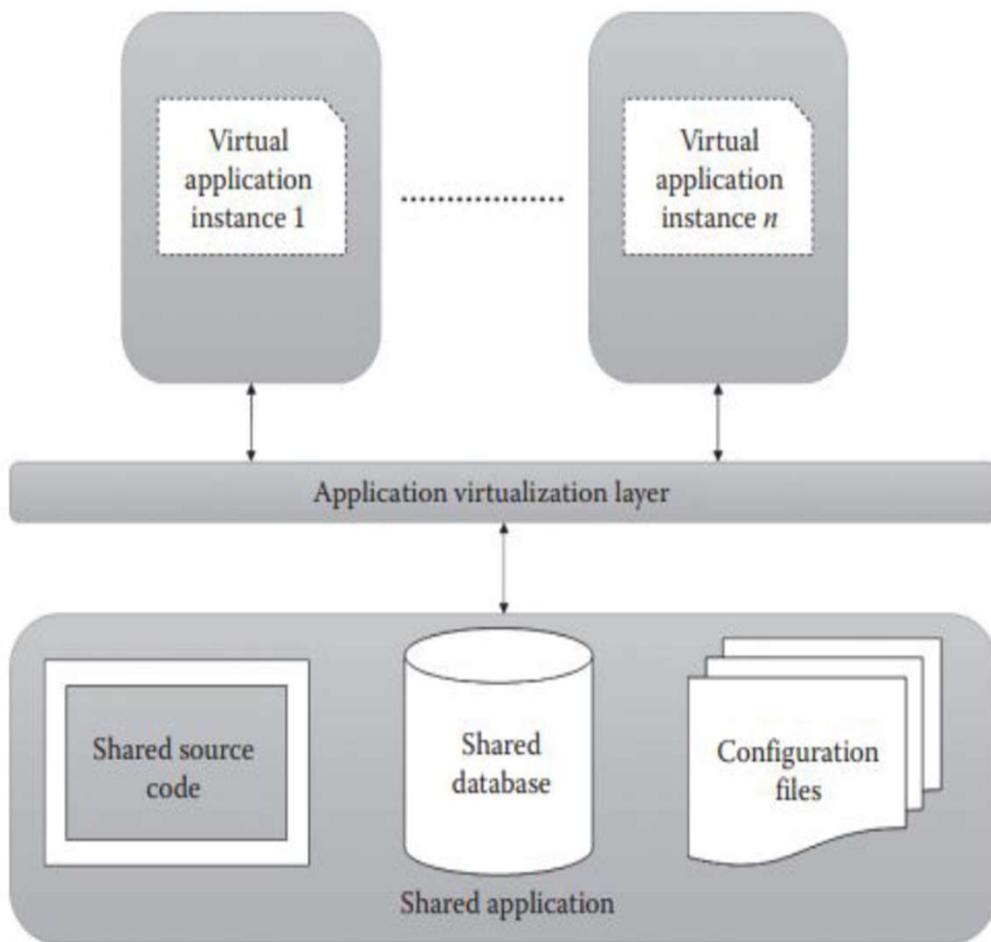
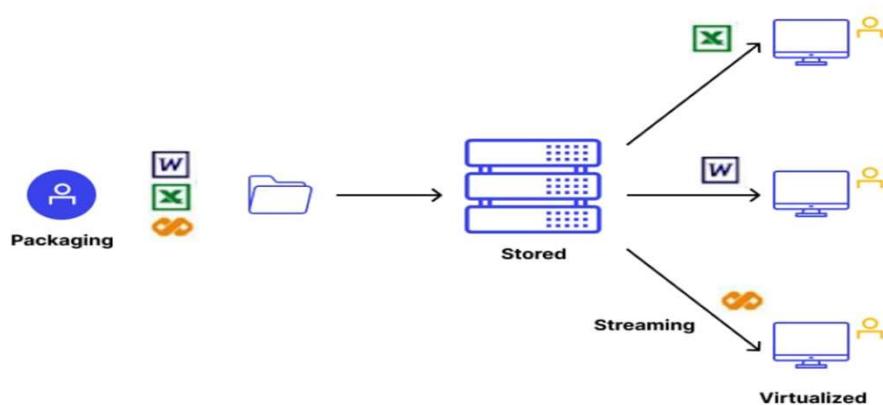
	Full Virtualization	Paravirtualization	Hardware-Assisted Virtualization
Technique	Binary translation and direct execution	Hypercalls	OS requests trap to VMM without binary translation or paravirtualization
Guest OS modification	No	Yes	No
Compatibility	Excellent compatibility	Poor compatibility	Excellent compatibility
Is guest OS hypervisor independent?	Yes	No	Yes
Performance	Good	Better in certain cases	Fair
Position of VMM and privilege level	Ring 0 Root privilege	Below ring 0	Below ring 0 Root privilege
Position of guest OS and privilege level	Ring 1 Nonroot privilege	Ring 0 Root privilege	Ring 0 Nonroot privilege
Popular vendor(s)	VMware ESX	Xen	Microsoft, Virtual Iron, and XenSource

Software Virtualization Types



Software/Application Virtualization

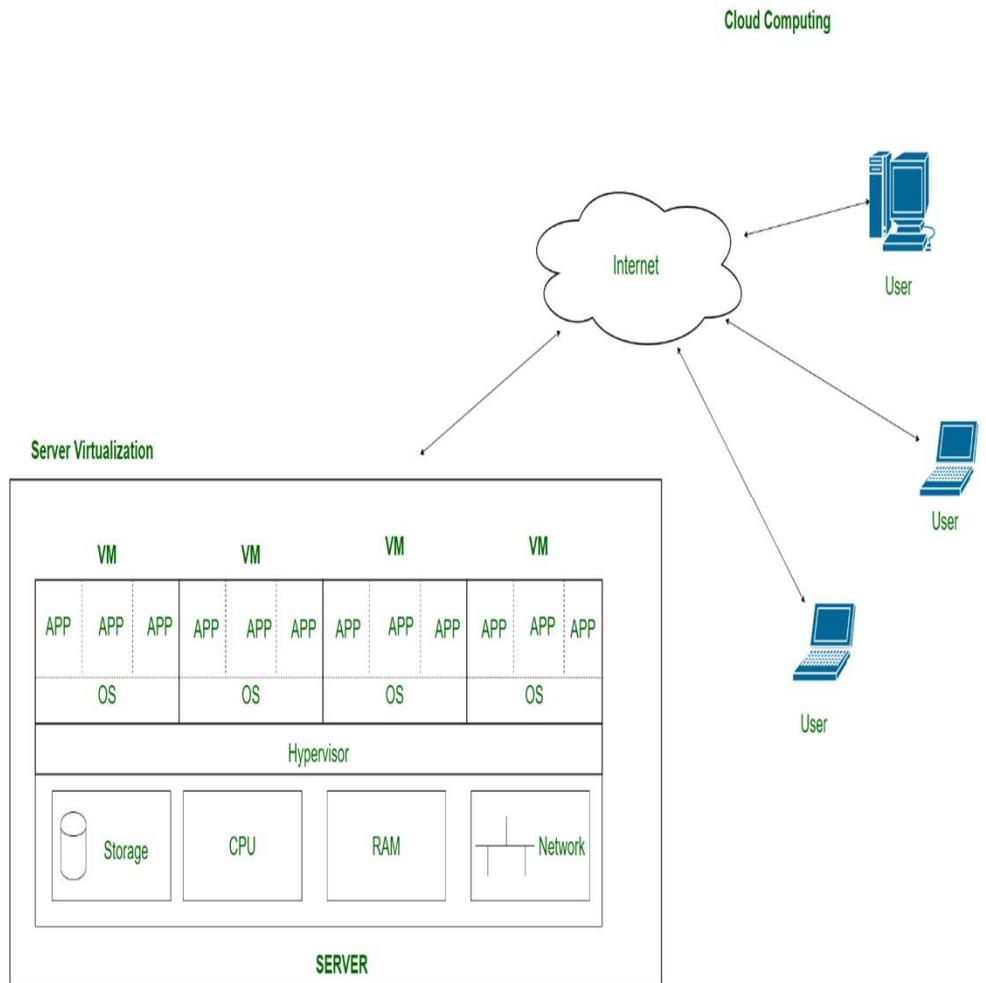
It lessens the time needed for keeping up tasks, like system upgrades and backups. In other words, if an application is deployed on a virtual server, then only essential parts need to be refreshed which leads to fewer changes on the customer's side compared with physical devices. This lets companies keep their business operations running effortlessly while having negligible downtime because of maintenance activities associated with systems



Server Virtualization

Now Consider situation, You are using Mac OS on your machine but particular application for your project can be operated only on Windows. You can either buy new machine running windows or create virtual environment in which windows can be installed and used. Second option is better because of less cost and easy implementation. This scenario is called [Virtualization](#).

In it, virtual CPU, RAM, NIC and other resources are provided to OS which it needed to run. This resources is virtually provided and controlled by an application called Hypervisor.

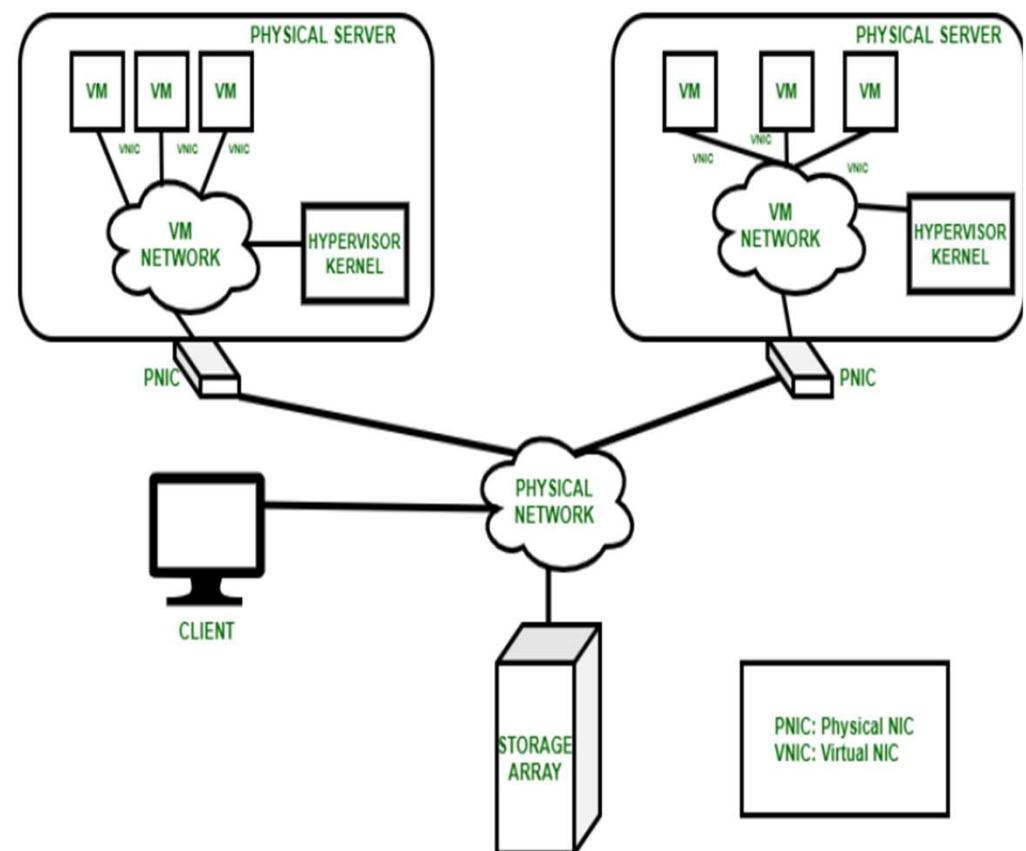


Network Virtualization

It grants organisations the capacity to create separate, logically detached networks on top of their pre-existing infrastructure, allowing for simpler implementation of assorted applications and services.

By virtualising their network configuration, businesses can take advantage of increased scalability, security and dependability while also reducing IT overhead costs.

It essentially abstracts all underlying physical hardware elements and produces multiple “virtual” networking layers which can function independently of one another – granting companies an extra measure of control over their structure as well as greater flexibility when it comes to deployment options!



Network Virtualization

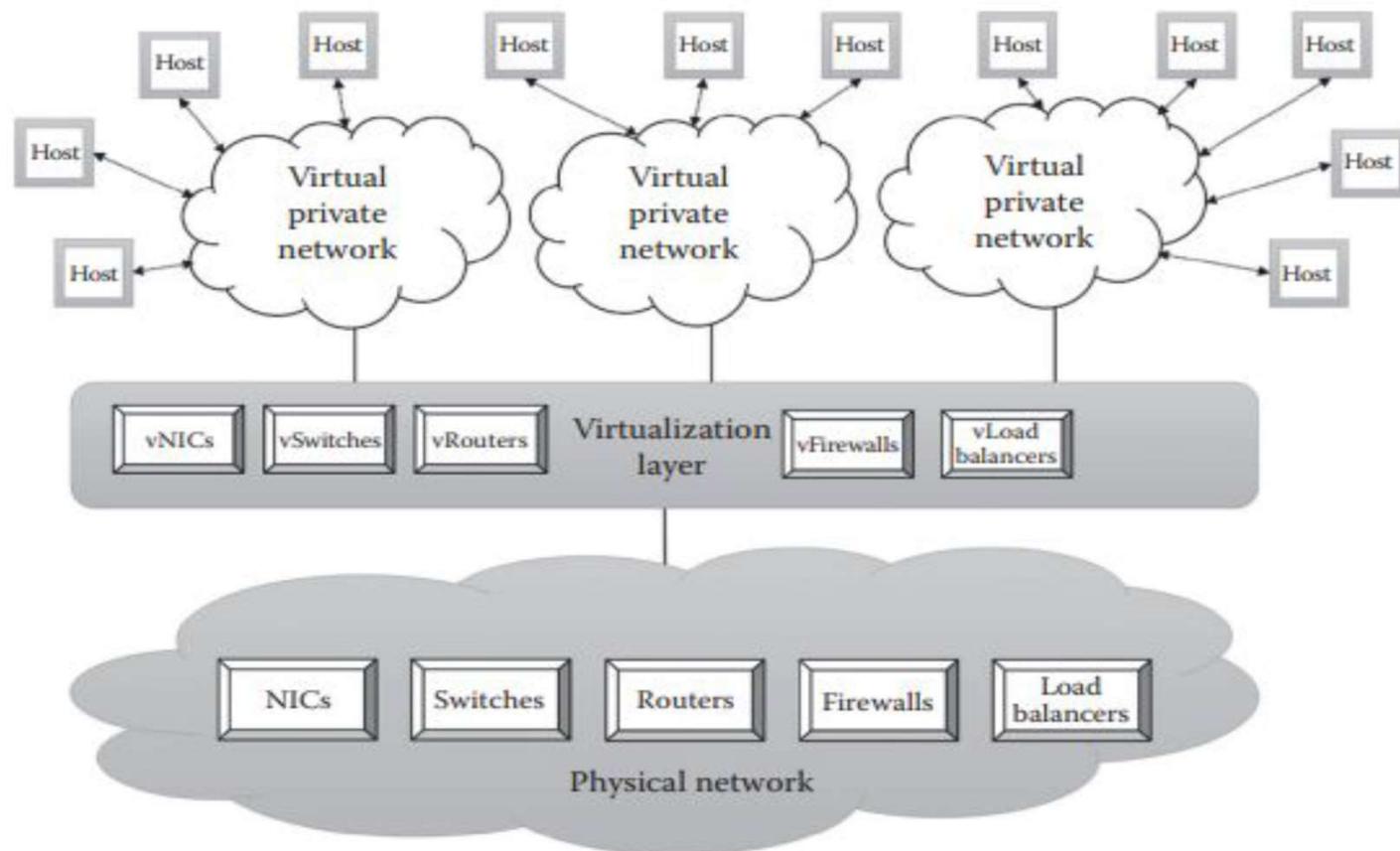


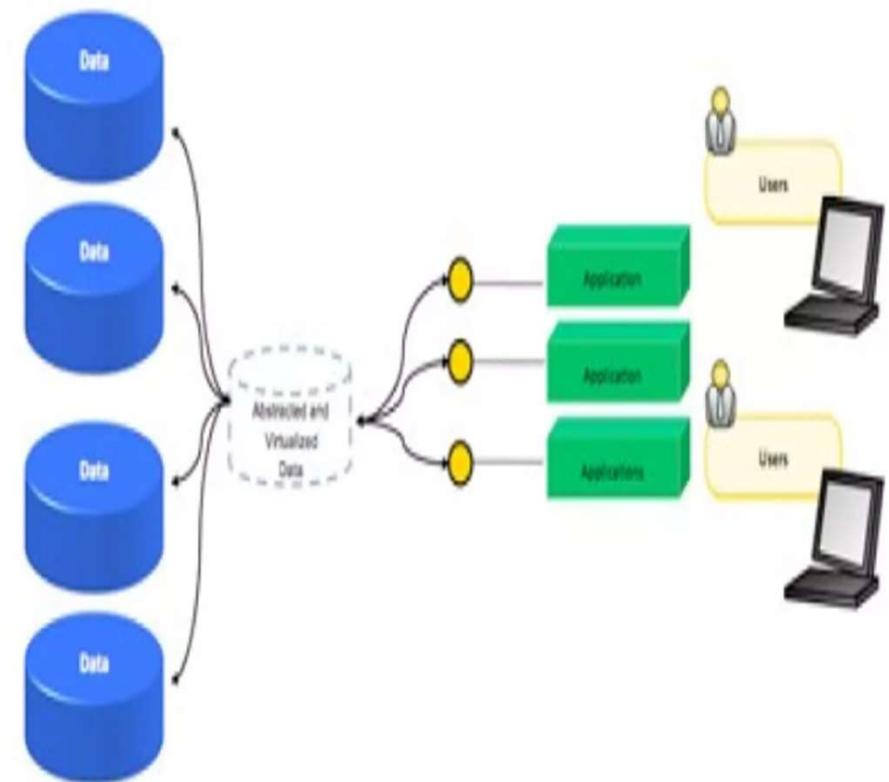
FIGURE 7.6

Network virtualization.

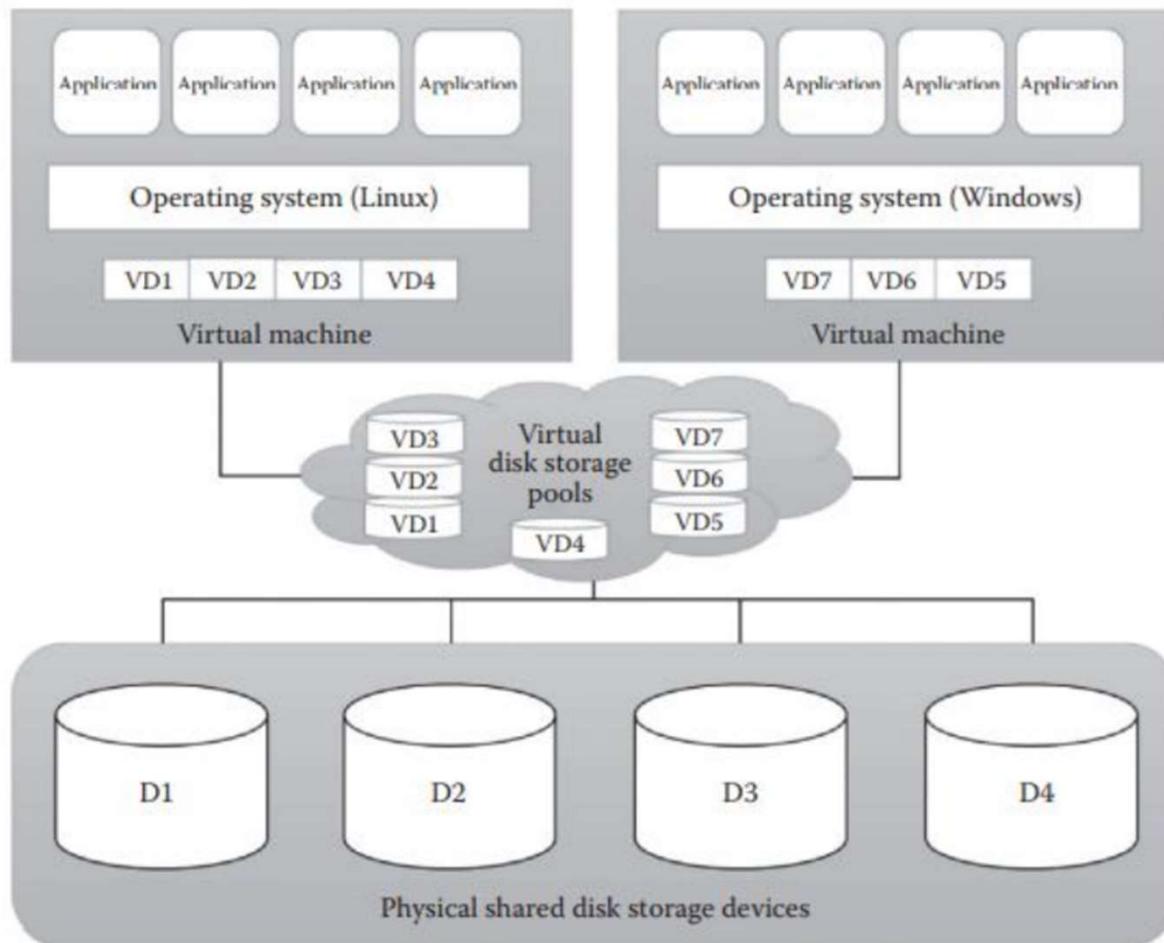
Storage Virtualization

Storage virtualization is one of the most widely used types of virtualization in cloud computing. This technology allows organizations to access, manage and store data across different servers without needing physical storage – allowing them to share resources between their networks as if they were all located together. With this tech, teams can gain access to information from any location regardless of where it's physically stored – which has multiple benefits for businesses wanting to optimize their storage solutions.

- Improved storage management in a heterogeneous IT environment
- Easy updates, better availability
- Reduced downtime
- Better storage utilization
- Automated management



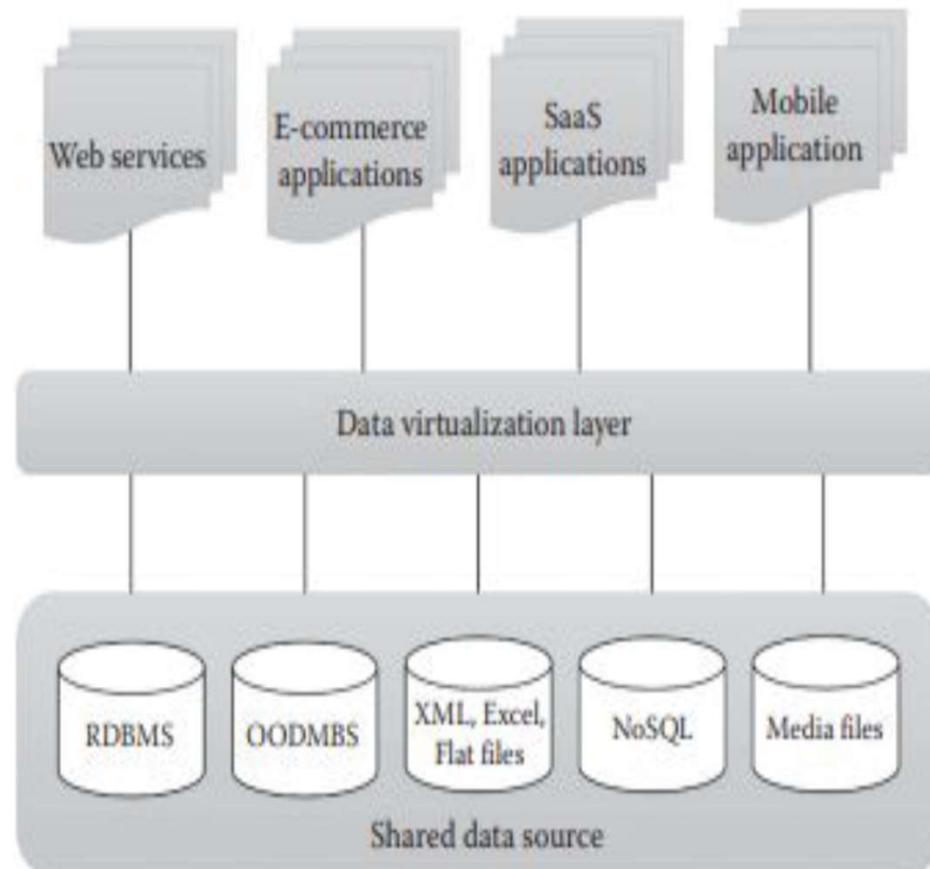
Storage Virtualization



Data Virtualization

Data virtualization is the ability to retrieve the data without knowing its type and the physical location where it is stored.

- It aggregates the heterogeneous data from the different sources to a single logical/virtual volume of data.
- This logical data can be accessed from any applications such as web services, E-commerce applications, web portals, Software as a Service (SaaS) applications, and mobile application.
- Data virtualization hides the type of the data and the location of the data for the application that access it.
- It also ensures the single point access to data by aggregating data from different sources. It is mainly used in data integration, business intelligence, and cloud computing



Desktop Virtualization

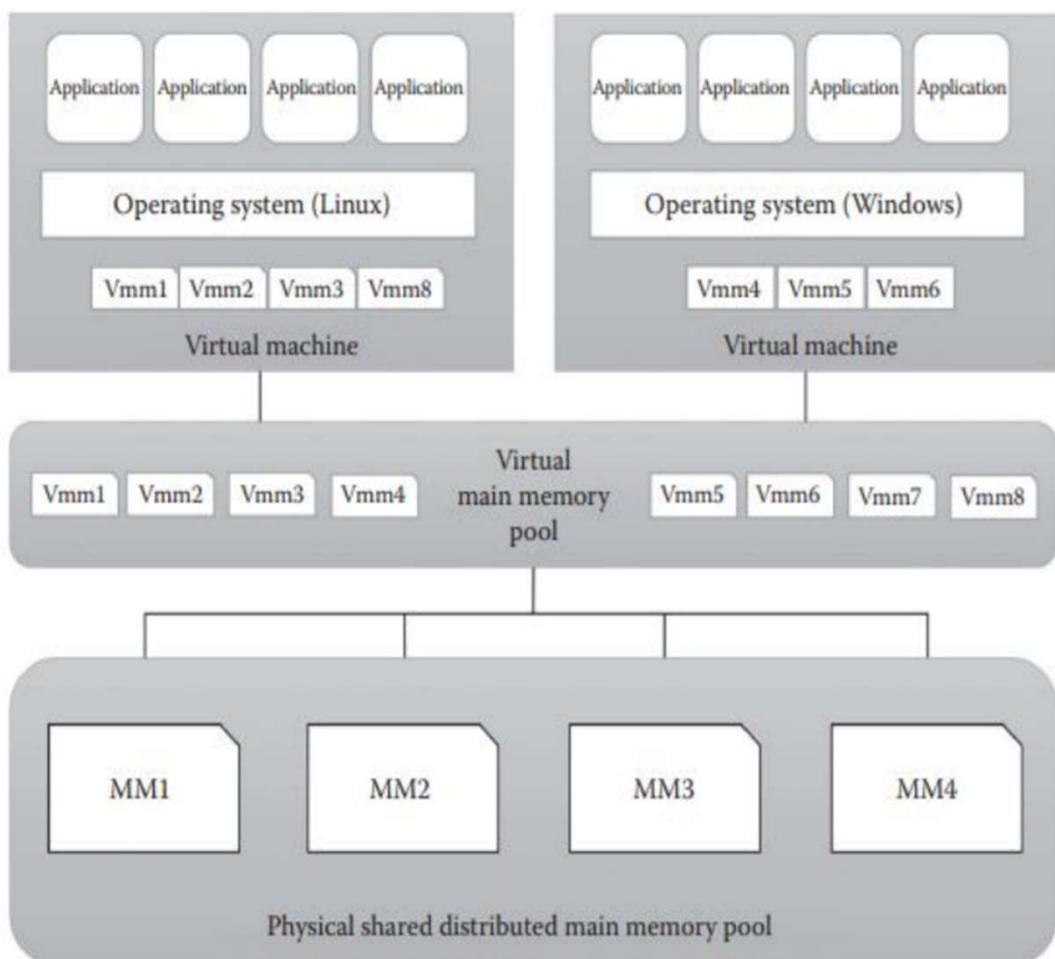
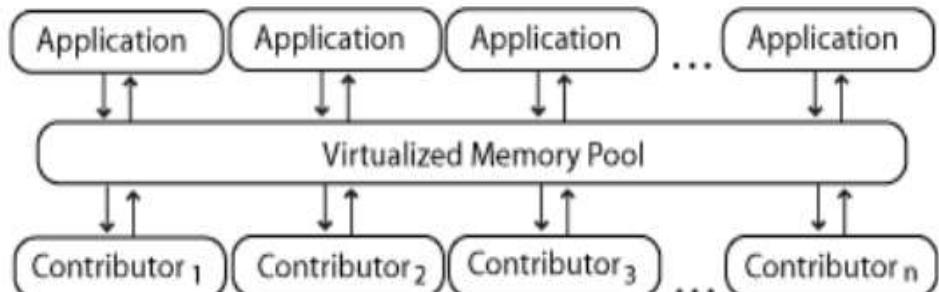
It provides the work convenience and security. As one can access remotely, you are able to work from any location and on any PC. It provides a lot of flexibility for employees to work from home or on the go. It also protects confidential data from being lost or stolen by keeping it safe on central servers.



Memory Virtualization

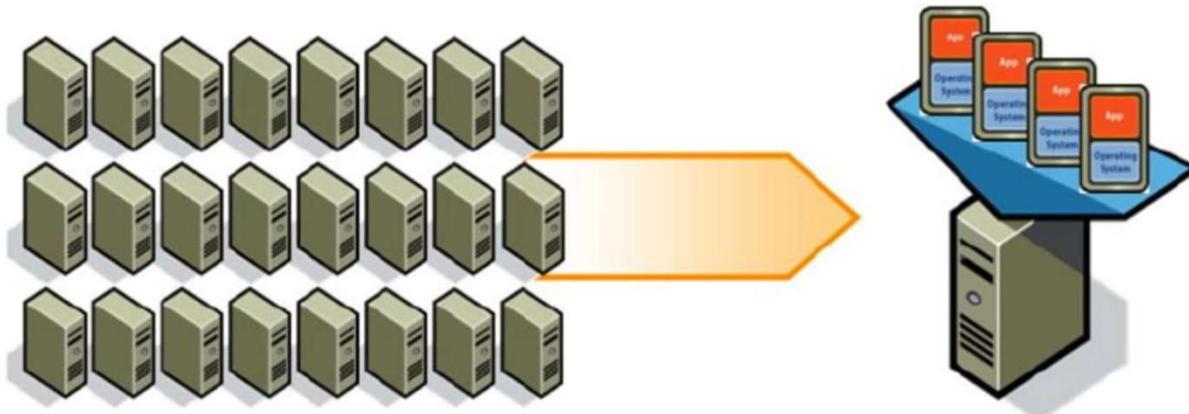
It introduces a way to decouple memory from the server to provide a shared, distributed or networked function. It enhances performance by providing greater memory capacity without any addition to the main memory. That's why a portion of the disk drive serves as an extension of the main memory.

Application-level integration – Applications running on connected computers directly connect to the memory pool through an API or the file system.



Impact of Virtualization

	<u>Before</u>	<u>After</u>
Servers	> 1,000	> 50
Storage	> Direct attach	> Tiered SAN and NAS
Network	> 3000 cables/ports	> 300 cables/ports
Facilities	> 200 racks > 400 power whips	> 10 racks > 20 power whips



Characteristics of Cloud Computing

Characteristics of Cloud Computing



On-Demand Self-Service



Broad Network Access



Resource Pooling



Rapid Elasticity



Measured Service



Resiliency and Availability



Flexibility



Remote Work

1. On-Demand Self-Service

With cloud computing, you can provision computing services, like server time and network storage, automatically. You won't need to interact with the service provider. Cloud customers can access their cloud accounts through a web self-service portal to view their cloud services, monitor their usage, and provision and de-provision services.

2. Broad Network Access

Another essential cloud computing characteristic is broad network access. You can access cloud services over the network and on portable devices like mobile phones, tablets, laptops, and desktop computers. A public cloud uses the internet; a private cloud uses a local area network. Latency and bandwidth both play a major role in cloud computing and broad network access, as they affect the quality of service.

3. Resource Pooling

With resource pooling, multiple customers can share physical resources using a multi-tenant model. This model assigns and reassigns physical and virtual resources based on demand. Multi-tenancy allows customers to share the same applications or infrastructure while maintaining privacy and security. Though customers won't know the exact location of their resources, they may be able to specify the location at a higher level of abstraction, such as a country, state, or data center. Memory, processing, and bandwidth are among the resources that customers can pool.

4. Rapid Elasticity

Cloud services can be elastically provisioned and released, sometimes automatically, so customers can scale quickly based on demand. The capabilities available for provisioning are practically unlimited. Customers can engage with these capabilities at any time in any quantity. Customers can also scale cloud use, capacity, and cost without extra contracts or fees. With rapid elasticity, you won't need to buy computer hardware. Instead, can use the cloud provider's cloud computing resources.

5. Measured Service

In cloud systems, a metering capability optimizes resource usage at a level of abstraction appropriate to the type of service. For example, you can use a measured service for storage, processing, bandwidth, and users. Payment is based on actual consumption by the customer via a pay-for-what-you-use model. Monitoring, controlling, and reporting resource use creates a transparent experience for both consumers and providers of the service.

Other Cloud Computing Characteristics

Resiliency and Availability

Resilience in cloud computing refers to the ability of a service to recover quickly from any disruption. Cloud resiliency is measured by how fast its servers, databases, and networks restart and recover after any damage. To prevent data loss, cloud services create a copy of the stored data. If one server loses data for any reason, the copy version from the other server restores.

Availability is a related key concept in cloud computing. The benefit of cloud services is that you can access them remotely, so there are no geographic restrictions when using cloud resources.

Flexibility

Companies need to scale as their business grows. The cloud provides customers with more freedom to scale as they please without restarting the server. They can also choose from several payment options to avoid overspending on resources they won't need.

Remote Work

Cloud computing helps users work remotely. Remote workers can safely and quickly access corporate data via their devices, including laptops and smartphones. Employees who work remotely can also communicate with each other and perform their jobs effectively using the cloud.

On Demand Services

- ‘A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.’

Rapid Elasticity

- ‘Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.’
- Servers can be quickly provisioned and decommissioned based on current demand.
- Elasticity allows customers to achieve cost savings and is often a core justification for adoption.

Cont..

- If 10 servers are required for a 3 month project, the company can provision them within minutes, pay a small monthly OpEx fee to run them rather than a large upfront CapEx cost, and decommission them at the end of the 3 months.
- If a company experiences seasonal demand, such as an ecommerce store at Christmas, additional front end web servers can be **automatically** provisioned and added to a load balanced server farm, and then automatically decommissioned when the demand dissipates.

NOTE:

Capital expenditures (CapEx) are costs that often yield long-term benefits to a company. CapEx assets often have a useful life of more than one year.

Operating expenses (OpEx) are costs that often have a much shorter-term benefit.