

Virtualization in Cloud Computing



Virtualization Basics

A virtual machine is a software computer that, like a physical computer, runs an operating system and applications.

Each virtual machine contains its own virtual, or software-based, hardware, including a virtual CPU, memory, hard disk, and network interface card.

Virtualization is a technique of how to separate a service from the underlying physical delivery of that service. It is the process of creating a virtual version of something like computer hardware.

With the help of Virtualization, multiple operating systems and applications can run on same machine and its same hardware at the same time, increasing the utilization and flexibility of hardware.

One of the main cost effective, hardware reducing, and energy saving techniques used by cloud providers is virtualization

Virtualization Basics

Virtualization allows to share a single physical instance of a resource or an application among multiple customers and organizations at one time.

It does this by assigning a logical name to a physical storage and providing a pointer to that physical resource on demand.

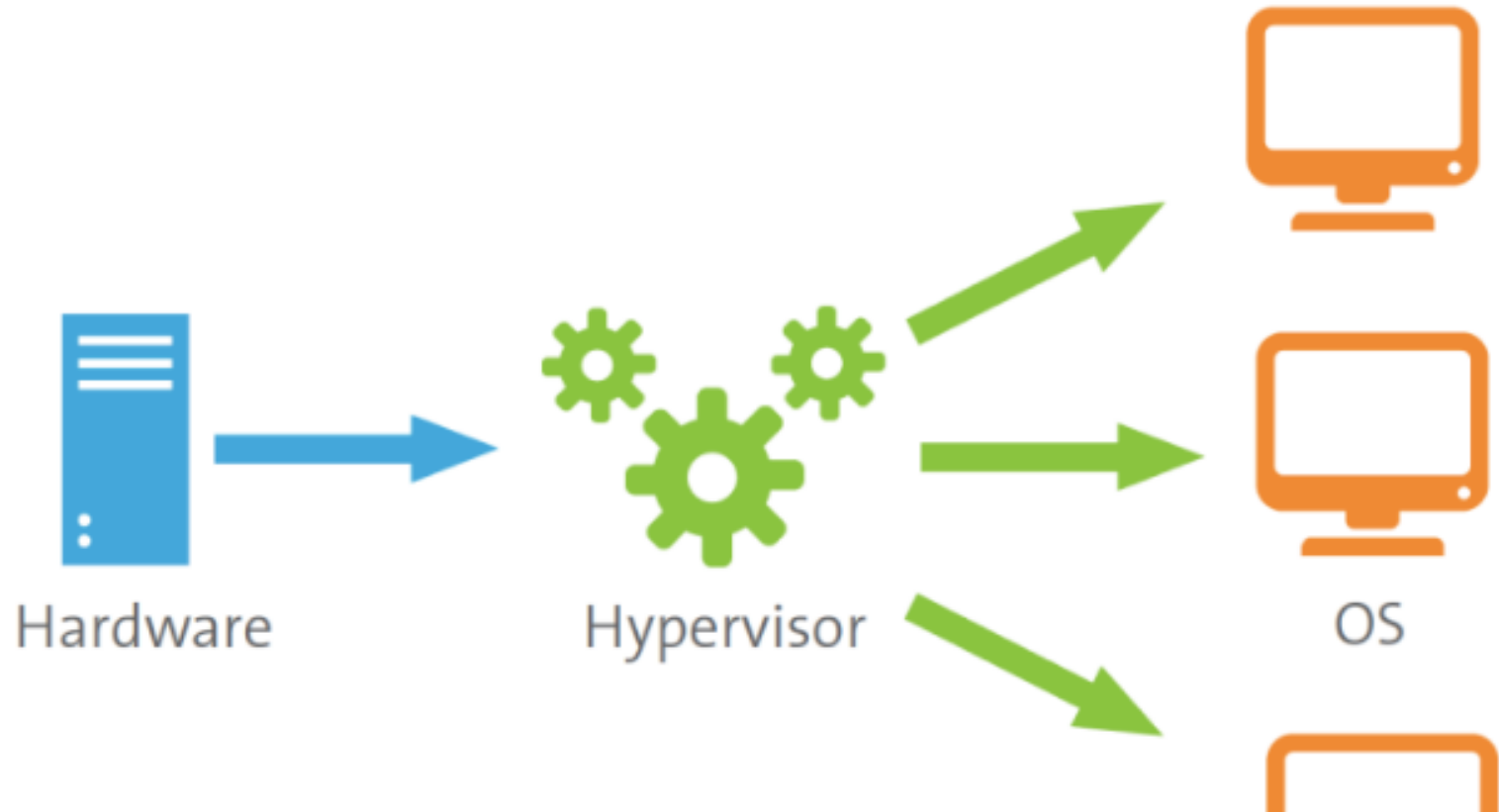
The term virtualization is often synonymous with hardware virtualization, which plays a fundamental role in efficiently delivering Infrastructure-as-a-Service (IaaS) solutions for cloud computing.

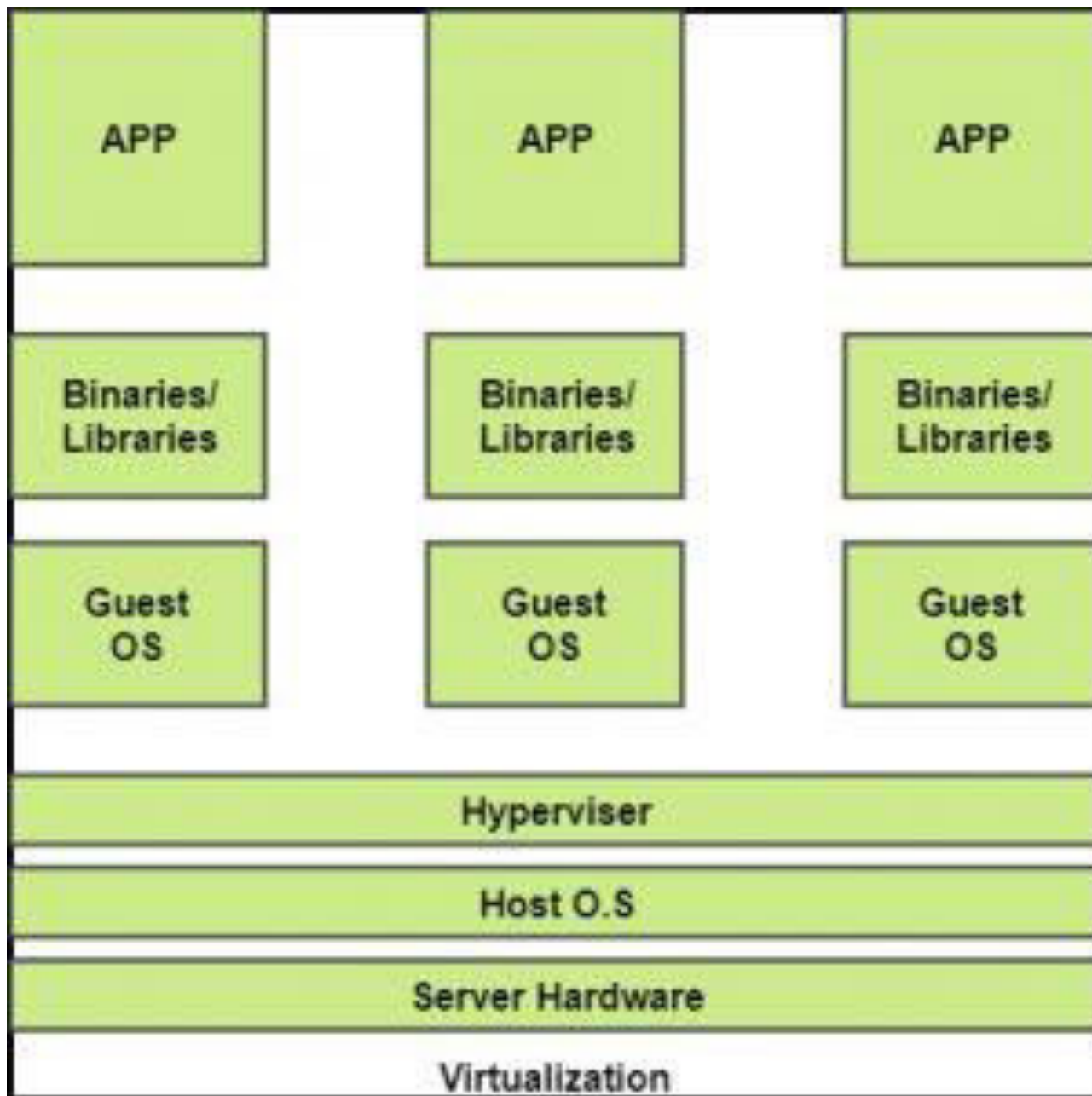
Moreover, virtualization technologies provide a virtual environment for not only executing applications but also for storage, memory, and networking.

The machine on which the virtual machine is going to be built is known as Host Machine and that virtual machine is referred as a Guest Machine.

What is a hypervisor?

A **hypervisor**, also known as a virtual machine monitor or VMM, is software that creates and runs virtual machines (VMs). A hypervisor allows one host computer to support multiple guest VMs by virtually sharing its resources, such as memory and processing.





BENEFITS OF VIRTUALIZATION

1. More flexible and efficient allocation of resources.
2. Enhance development productivity.
3. It lowers the cost of IT infrastructure.
4. Remote access and rapid scalability.
5. High availability and disaster recovery.
6. Pay peruse of the IT infrastructure on demand.
7. Enables running multiple operating systems.

Types of Virtualization

**Application
Virtualization**

**Network
Virtualization**

**Desktop
Virtualization**

**Storage
Virtualization**

**Server
Virtualization**

**Data
Virtualization**

Application Virtualization

Application virtualization helps a user to have remote access of 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.

Application virtualization software allows users to access and use an application from a separate computer than the one on which the application is installed.

Using application virtualization software, IT admins can set up remote applications on a server and deliver the apps to an end user's computer.

For the user, the experience of the virtualized app is the same as using the installed app on a physical machine.

Network Virtualization

The ability to run multiple virtual networks with each has a separate control and data plan.

It co-exists together on top of one physical network. It can be managed by individual parties that potentially confidential to each other.

Network virtualization provides a facility to create and provision virtual networks—logical switches, routers, firewalls, load balancer, Virtual Private Network (VPN), and workload security within days or even in weeks.

Network Virtualization is a process of logically grouping physical networks and making them operate as single or multiple independent networks called Virtual Networks.

Desktop Virtualization

Desktop virtualization allows the users' OS to be remotely stored on a server in the data centre.

It allows the user to access their desktop virtually, from any location by a different machine.

Users who want specific operating systems other than Windows Server will need to have a virtual desktop.

Main benefits of desktop virtualization are user mobility, portability, easy management of software installation, updates, and patches.

Desktop virtualization is technology that lets users simulate a workstation load to access a desktop from a connected device. It separates the desktop environment and its applications from the physical client device used to access it.

Storage Virtualization

Storage virtualization is an array of servers that are managed by a virtual storage system.

The servers aren't aware of exactly where their data is stored, and instead function more like worker bees in a hive. It makes managing storage from multiple sources to be managed and utilized as a single repository.

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 -- that is managed from a central console.

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

Server Virtualization

Server virtualization is the process of dividing a physical server into multiple unique and isolated virtual servers by means of a software application. Each virtual server can run its own operating systems independently.

This is a kind of virtualization in which masking of server resources takes place.

Here, the central-server(physical server) is divided into multiple different virtual servers by changing the identity number, processors. So, each system can operate its own operating systems in isolate manner.

Where each sub-server knows the identity of the central server. It causes an increase in the performance and reduces the operating cost by the deployment of main server resources into a sub-server resource. It's beneficial in virtual migration, reduce energy consumption, reduce infrastructural cost, etc

Data Virtualization

This is the kind of virtualization in which the data is collected from various sources and managed that at a single place without knowing more about the technical information like how data is collected, stored & formatted then arranged that data logically so that its virtual view can be accessed by its interested people and stakeholders, and users through the various cloud services remotely.

Many big giant companies are providing their services like Oracle, IBM, Cdata, etc.

It can be used to performing various kind of tasks such as:

Data-integration

Business-integration

Service-oriented architecture data-services

Searching organizational data

Components of Modern Virtualization

Three main components of modern virtualization come together to create the virtual machines and associated infrastructure that power the modern world.

Virtual Compute

When we talk about virtualized compute resources, we mean the CPU – the “brain” of the computer.

In a virtualized system, the Hypervisor software creates virtual CPUs in software, and presents them to its guests for their use.

The guests only see their own virtual CPUs, and the Hypervisor passes their requests on to the physical CPUs

Components of Modern Virtualization

The benefits of **Virtual Compute** are immense.

Because the Hypervisor can create more virtual CPUs (often creating many more than the number of physical CPUs), the physical CPUs spend far less time in an idle state, making it possible to extract much more performance from them.

In a non-virtualized architecture, many CPU cycles (and therefore electricity and other resources) are wasted because most applications do not run a modern physical CPU at full tilt all the time.

Components of Modern Virtualization

Virtual Storage

In the beginning, there were disk drives.

Later, there was RAID (Redundant Array of Independent Disks). RAID allowed us to combine multiple disks into one, gaining additional speed and reliability.

Later still, the RAID arrays that used to live in each server moved into specialized, larger storage arrays which could provide virtual disks to any number of servers over a specialized high-speed network, known as a SAN (Storage Area Network).

All of these methods are simply different ways to virtualize storage.

Over the years, the trend has been to abstract away individual storage devices, consolidate them, and distribute work across them. The ways we do this have gotten better, faster, and become networked – but the basic concept is the same

Quick Question



RAID, or “Redundant Arrays of Independent Disks” is a technique which makes use of a combination of multiple disks instead of using a single disk for increased performance, data redundancy or both

Components of Modern Virtualization

Virtual Networking

Network virtualization uses software to abstract away the physical network from the virtual environment.

The Hypervisor may provide its guest systems with many virtual network devices (switches, network interfaces, etc.) that have little to no correspondence with the physical network connecting the physical servers.

Within the Hypervisor's configuration, there is a mapping of virtual to physical network resources, but this is invisible to the guests.

If two virtual computers in the same environment need to talk to each other over the network, that traffic may never touch the physical network at all if they are on the same host system.

<https://www.kelsercorp.com/blog/the-7-types-of-virtualization>

IMPLEMENTATION LEVELS OF VIRTUALIZATION

Talking of the Implementation levels of virtualization in cloud computing, there are a total of five levels that are commonly used.

- **Instruction Set Architecture Level (ISA)**
- **Hardware Abstraction Level (HAL)**
- **Operating System Level**
- **Library Level**
- **Application Level**

Instruction Set Architecture Level (ISA)

ISA virtualization can work through ISA emulation. This is used to run many legacy codes that were written for a different configuration of hardware.

These codes run on any virtual machine using the ISA. With this, a binary code that originally needed some additional layers to run is now capable of running on the x86 machines.

It can also be tweaked to run on the x64 machine.

For the basic emulation, an interpreter is needed, which interprets the source code and then converts it into a hardware format that can be read.

This then allows processing. This is one of the five implementation levels of virtualization in cloud computing.

Hardware Abstraction Level (HAL)

True to its name HAL lets the virtualization perform at the level of the hardware. This makes use of a hypervisor which is used for functioning.

At this level, the virtual machine is formed, and this manages the hardware using the process of virtualization. It allows the virtualization of each of the hardware components, which could be the input-output device, the memory, the processor, etc.

Multiple users will not be able to use the same hardware and also use multiple virtualization instances at the very same time. This is mostly used in the cloud-based infrastructure.

Virtualization at the HAL exploits the similarity in architectures of the guest and host platforms to cut down the interpretation latency. Virtualization technique helps map the virtual resources to physical resources and use the native hardware for computations in the VM.

Operating System Level

At the level of the operating system, the virtualization model is capable of creating a layer that is abstract between the operating system and the application.

This is an isolated container that is on the operating system and the physical server, which makes use of the software and hardware. Each of these then functions in the form of a server.

When there are several users, and no one wants to share the hardware, then this is where the virtualization level is used. Every user will get his virtual environment using a virtual hardware resource that is dedicated. In this way, there is no question of any conflict.

Operating system virtualization (OS virtualization) is a server virtualization technology that involves tailoring a standard operating system so that it can run different applications handled by multiple users on a single computer at a time

Library Level

The operating system is cumbersome, and this is when the applications make use of the API that is from the libraries at a user level.

These APIs are documented well, and this is why the library virtualization level is preferred in these scenarios.

API hooks make it possible as it controls the link of communication from the application to the system.

Quick Question



API stands for Application Programming Interface. In the context of APIs, the word Application refers to any software with a distinct function. Interface can be thought of as a contract of service between two applications. This contract defines how the two communicate with each other using requests and responses.

API calls and hooks

Application programming interfaces (APIs) are a way for one program to interact with another. API calls are **the medium by which they interact**. An API call, or API request, is a message sent to a server asking an API to provide a service or information

API hooking is a technique by which we can instrument and modify the behavior and flow of API calls.

<https://resources.infosecinstitute.com/topic/api-hooking/>

[https://www.cloudflare.com/learning/security/api/what-is-api-call/#:~:text=Application%20programming%20interfaces%20\(APIs\)%20are,provide%20a%20service%20or%20information.](https://www.cloudflare.com/learning/security/api/what-is-api-call/#:~:text=Application%20programming%20interfaces%20(APIs)%20are,provide%20a%20service%20or%20information.)

Application Level

The application-level virtualization is used when there is a desire to virtualize only one application and is the last of the implementation levels of virtualization in cloud computing. One does not need to virtualize the entire environment of the platform.

This is generally used when you run virtual machines that use high-level languages. The application will sit above the virtualization layer, which in turn sits on the application program.

It lets the high-level language programs compiled to be used in the application level of the virtual machine run seamlessly.

Application virtualization is a process that deceives a standard app into believing that it interfaces directly with an operating system's capacities when, in fact, it does not. This ruse requires a virtualization layer inserted between the app and the OS.

VIRTUALIZATION STRUCTURES/TOOLS AND MECHANISMS

In general, there are three typical classes of VM architecture.

Before virtualization, the operating system manages the hardware.

After virtualization, a virtualization layer is inserted between the hardware and the operating system. In such a case, the virtualization layer is responsible for converting portions of the real hardware into virtual hardware.

Therefore, different operating systems such as Linux and Windows can run on the same physical machine, simultaneously.

Depending on the position of the virtualization layer, there are several classes of VM architectures, namely the hypervisor architecture, para-virtualization, and host-based virtualization.

Hypervisor and Xen Architecture

The hypervisor supports hardware-level virtualization on bare metal devices like CPU, memory, disk and network interfaces.

The hypervisor software sits directly between the physical hardware and its OS. This virtualization layer is referred to as either the VMM or the hypervisor. The hypervisor provides hypercalls for the guest OSes and applications.

TYPE-1 Hypervisor:

The hypervisor runs directly on the underlying host system. It is also known as a “Native Hypervisor” or “Bare metal hypervisor”. It does not require any base server operating system. It has direct access to hardware resources.

TYPE-2 Hypervisor:

A Host operating system runs on the underlying host system. It is also known as ‘Hosted Hypervisor’. Such kind of hypervisors doesn’t run directly over the underlying hardware rather they run as an application in a Host system(physical machine)

Hypervisor and Xen Architecture

The Xen Architecture: Xen is an open source hypervisor program developed by Cambridge University. Xen is a micro-kernel hypervisor, which separates the policy from the mechanism.

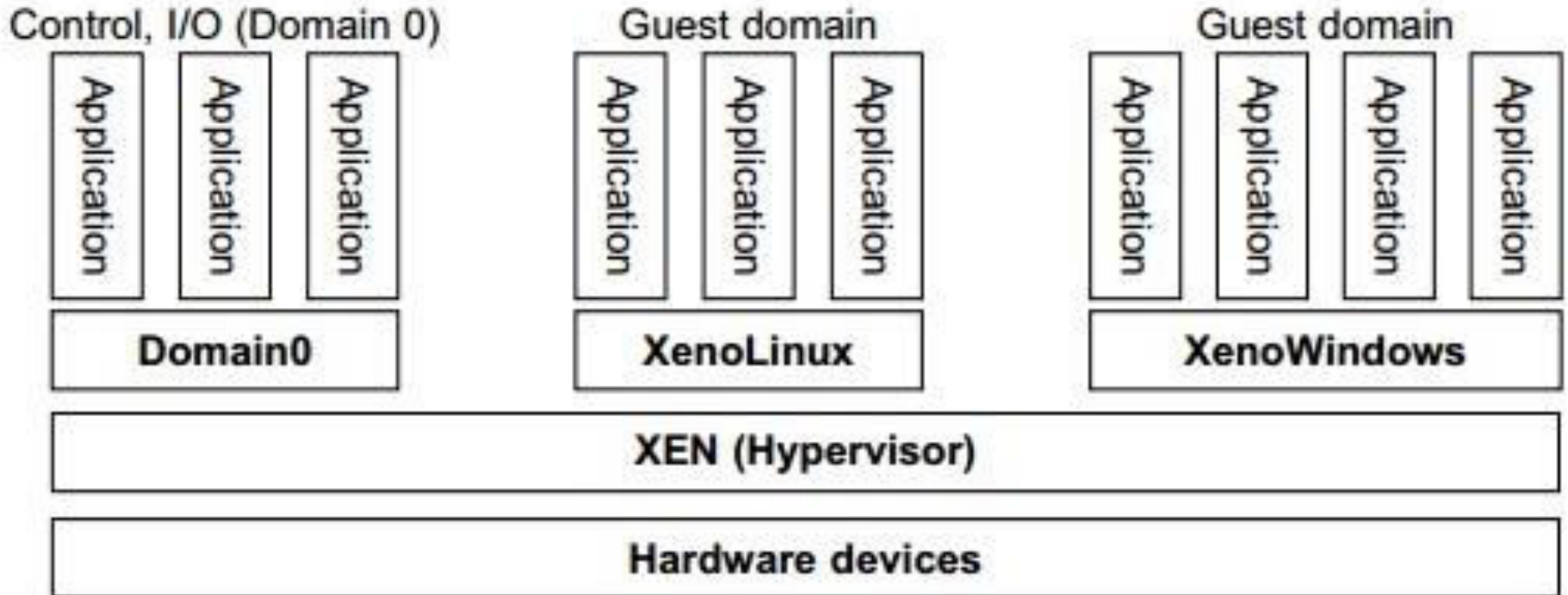
The Xen hypervisor implements all the mechanisms, leaving the policy to be handled by Domain 0.

Xen does not include any device drivers natively. It just provides a mechanism by which a guest OS can have direct access to the physical devices.

As a result, the size of the Xen hypervisor is kept rather small. Xen provides a virtual environment located between the hardware and the OS.

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

Xen Architecture



Xen Architecture

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. It is first loaded when Xen boots without any file system drivers being available.

Domain 0 is designed to access hardware directly and manage devices. Therefore, one of the responsibilities of Domain 0 is to allocate and map hardware resources for the guest domains (the Domain U domains).

Xen Architecture

Pros:

- a) Enables the system to develop lighter and flexible hypervisor that delivers their functionalities in an optimized manner.
- b) Xen supports balancing of large workload efficiently that capture CPU, Memory, disk input-output and network input-output of data. It offers two modes to handle this workload: Performance enhancement, and For handling data density.
- c) It also comes equipped with a special storage feature that we call Citrix storage link. Which allows a system administrator to uses the features of arrays from Giant companies- Hp, Netapp, Dell Equal logic etc.
- d) It also supports multiple processor, live migration one machine to another, physical server to virtual machine or virtual server to virtual machine conversion tools, centralized multiserver management, real time performance monitoring over window and linux.

Xen Architecture

Cons:

- a) Xen is more reliable over linux rather than on window.
- b) Xen relies on 3rd-party component to manage the resources like drivers, storage, backup, recovery & fault tolerance.
- c) Xen deployment could be a burden some on your Linux kernel system as time passes.
- d) Xen sometimes may cause increase in load on your resources by high input-output rate and may cause starvation of other VM's.

Binary Translation with Full Virtualization

Depending on implementation technologies, hardware virtualization can be classified into two categories: full virtualization and host-based virtualization.

Full virtualization does not need to modify the host OS. It relies on binary translation to trap and to virtualize the execution of certain sensitive, nonvirtualizable instructions.

The guest OSes and their applications consist of noncritical and critical instructions.

In a host-based system, both a host OS and a guest OS are used.

A virtualization software layer is built between the host OS and guest OS

Binary Translation with Full Virtualization

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.

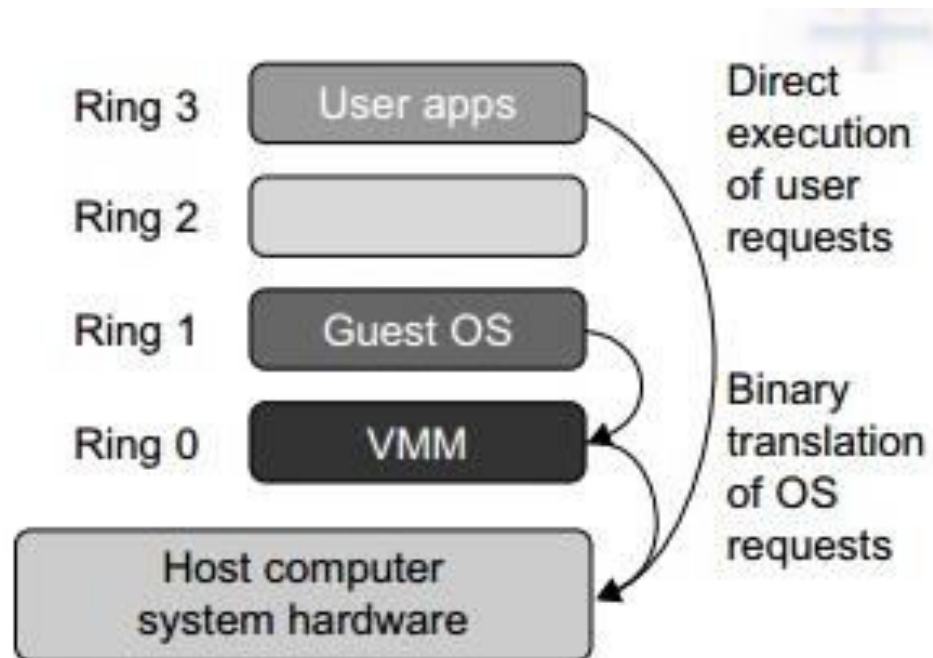
Therefore, running noncritical instructions on hardware not only can promote efficiency, but also can ensure system security.

Binary Translation 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



Binary Translation with Full Virtualization

Host-Based Virtualization

This host-based architecture has some distinct advantages:

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

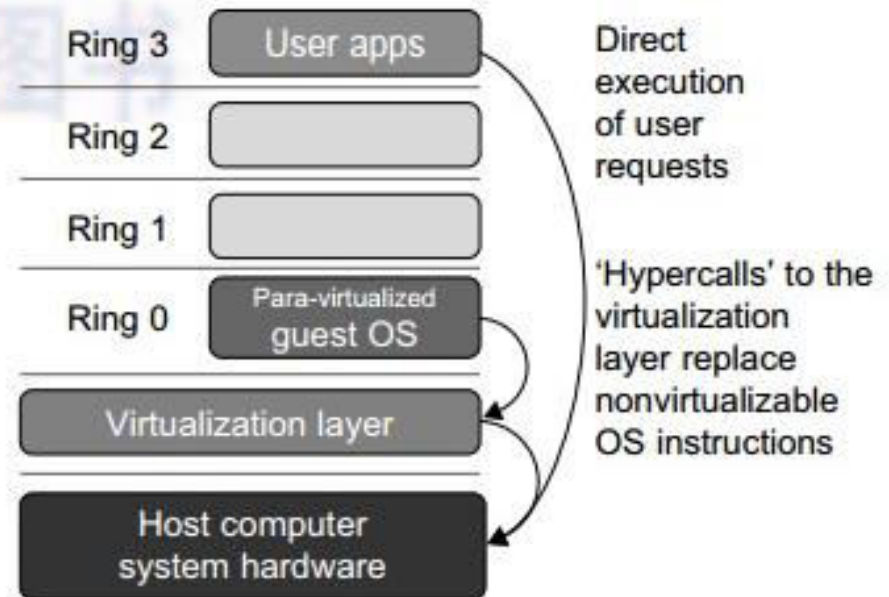
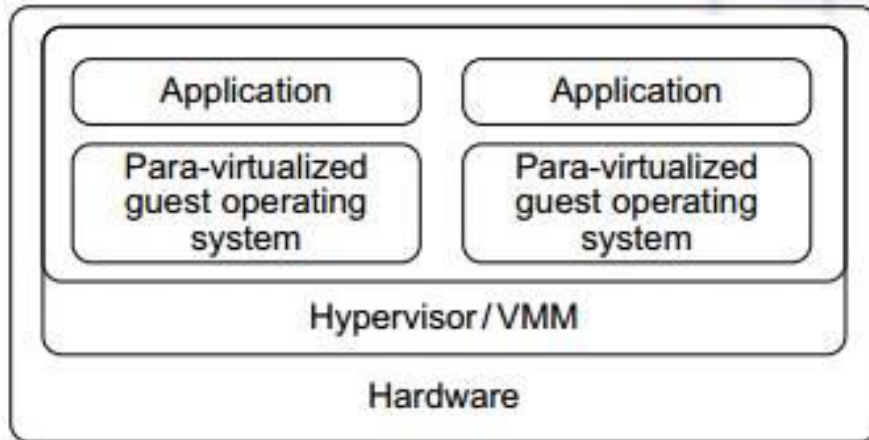
Para-virtualization needs to modify the guest operating systems. A para-virtualized VM provides special APIs requiring substantial OS modifications in user applications.

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 software stack. However, para-virtualization attempts to reduce the virtualization overhead, and thus improve performance by modifying only the guest OS kernel.

The guest operating systems are para-virtualized. They are assisted by an intelligent compiler to replace the nonvirtualizable OS instructions by hypercalls

Para-Virtualization Architecture



VIRTUALIZATION OF CPU, MEMORY, AND I/O DEVICES

To support virtualization, processors such as the x86 employ a special running mode and instructions, known as hardware-assisted virtualization.

In this way, the VMM and guest OS run in different modes and all sensitive instructions of the guest OS and its applications are trapped in the VMM.

To save processor states, mode switching is completed by hardware.

For the x86 architecture, Intel and AMD have proprietary technologies for hardware-assisted virtualization.

CPU Virtualization

A VM is a duplicate of an existing computer system in which a majority of the VM instructions are executed on the host processor in native mode.

Thus, unprivileged instructions of VMs run directly on the host machine for higher efficiency.

The critical instructions are divided into three categories: privileged instructions, control-sensitive instructions, and behavior-sensitive instructions.

Privileged instructions execute in a privileged mode and will be trapped if executed outside this mode.

Control-sensitive instructions attempt to change the configuration of resources used.

Behavior-sensitive instructions have different behaviors depending on the configuration of resources, including the load and store operations over the virtual memory.

CPU Virtualization

A CPU architecture is virtualizable if it supports the ability to run the VM's privileged and unprivileged instructions in the CPU's user mode while the VMM runs in supervisor mode.

When the privileged instructions including control- and behavior-sensitive instructions of a VM are executed, they are trapped in the VMM.

In this case, the VMM acts as a unified mediator for hardware access from different VMs to guarantee the correctness and stability of the whole system.

However, not all CPU architectures are virtualizable. RISC CPU architectures can be naturally virtualized because all control- and behavior-sensitive instructions are privileged instructions.

On the contrary, x86 CPU architectures are not primarily designed to support virtualization.

Memory Virtualization

Virtual memory virtualization is similar to the virtual memory support provided by modern operating systems.

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.

However, in a virtual execution environment, virtual memory virtualization involves sharing the physical system memory in RAM and dynamically allocating it to the physical memory of the VMs.

Memory Virtualization

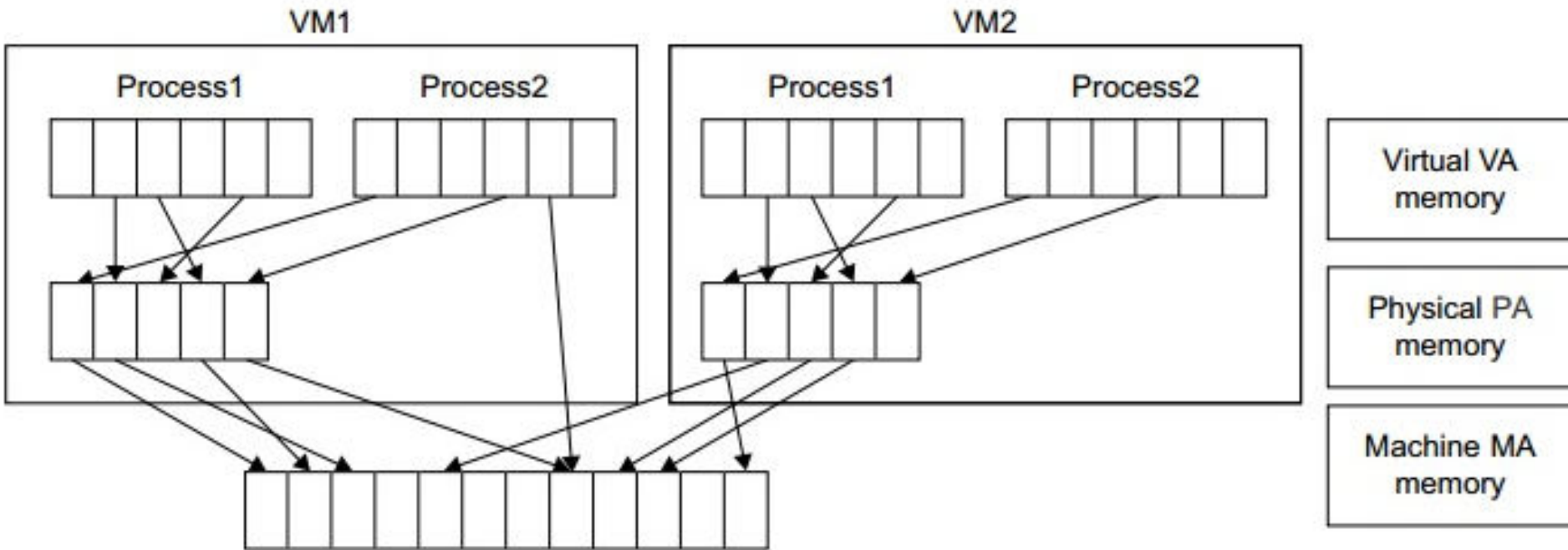
A two-stage mapping process should be maintained by the guest OS and the VMM, respectively: virtual memory to physical memory and physical memory to machine memory.

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



I/O Virtualization

I/O virtualization involves managing the routing of I/O requests between virtual devices and the shared physical hardware.

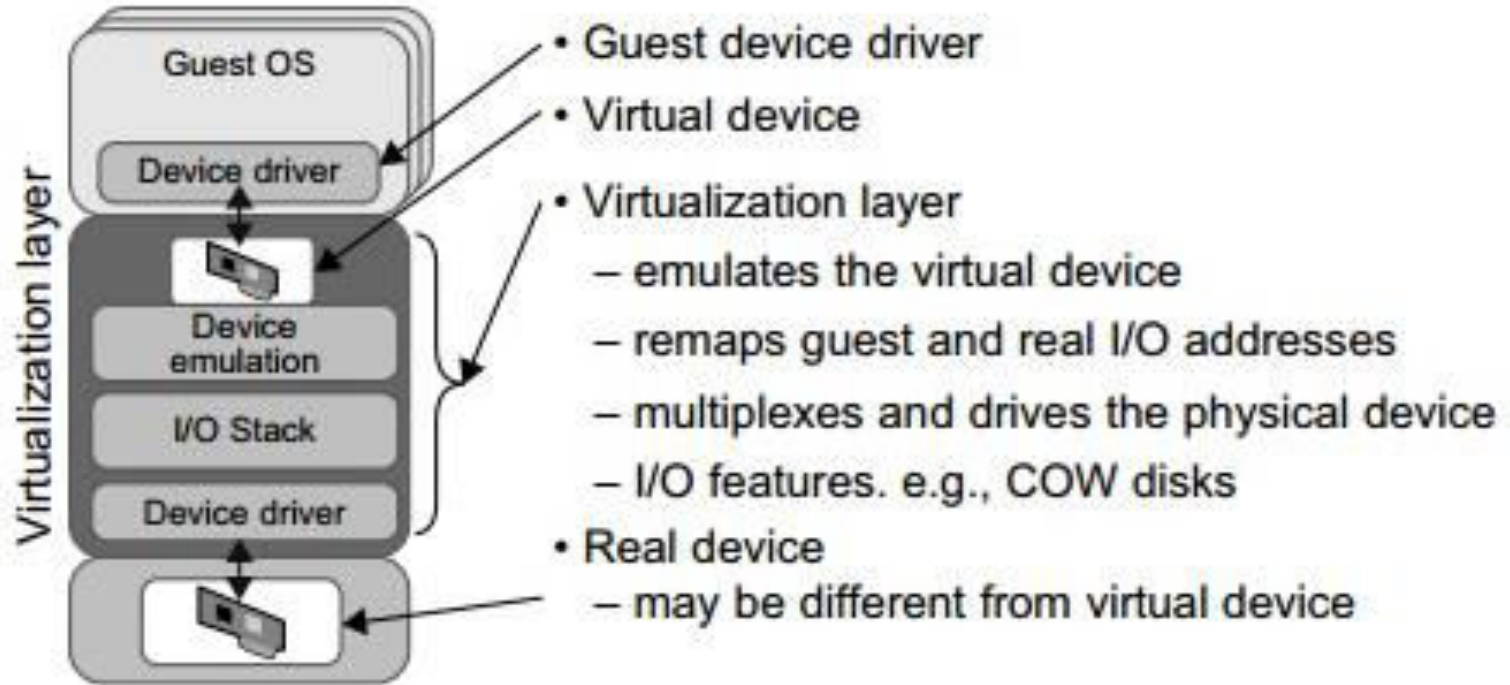
There are three ways to implement I/O virtualization: full device emulation, para-virtualization, and direct I/O.

Full device approach emulates well-known, real-world devices. All the functions of a device or bus infrastructure, such as device enumeration, identification, interrupts, and DMA, are replicated in software.

This software is located in the VMM and acts as a virtual device.

The I/O access requests of the guest OS are trapped in the VMM which interacts with the I/O devices.

I/O Virtualization



The full device emulation approach

I/O Virtualization

Direct I/O virtualization lets the VM access devices directly.

It can achieve close-to-native performance without high CPU costs. However, current direct I/O virtualization implementations focus on networking for mainframes.

There are a lot of challenges for commodity hardware devices.

For example, when a physical device is reclaimed (required by workload migration) for later reassign-ment, it may have been set to an arbitrary state (e.g., DMA to some arbitrary memory locations) that can function incorrectly or even crash the whole system.

Since software-based I/O virtualization requires a very high overhead of device emulation, hardware-assisted I/O virtualization is critical.

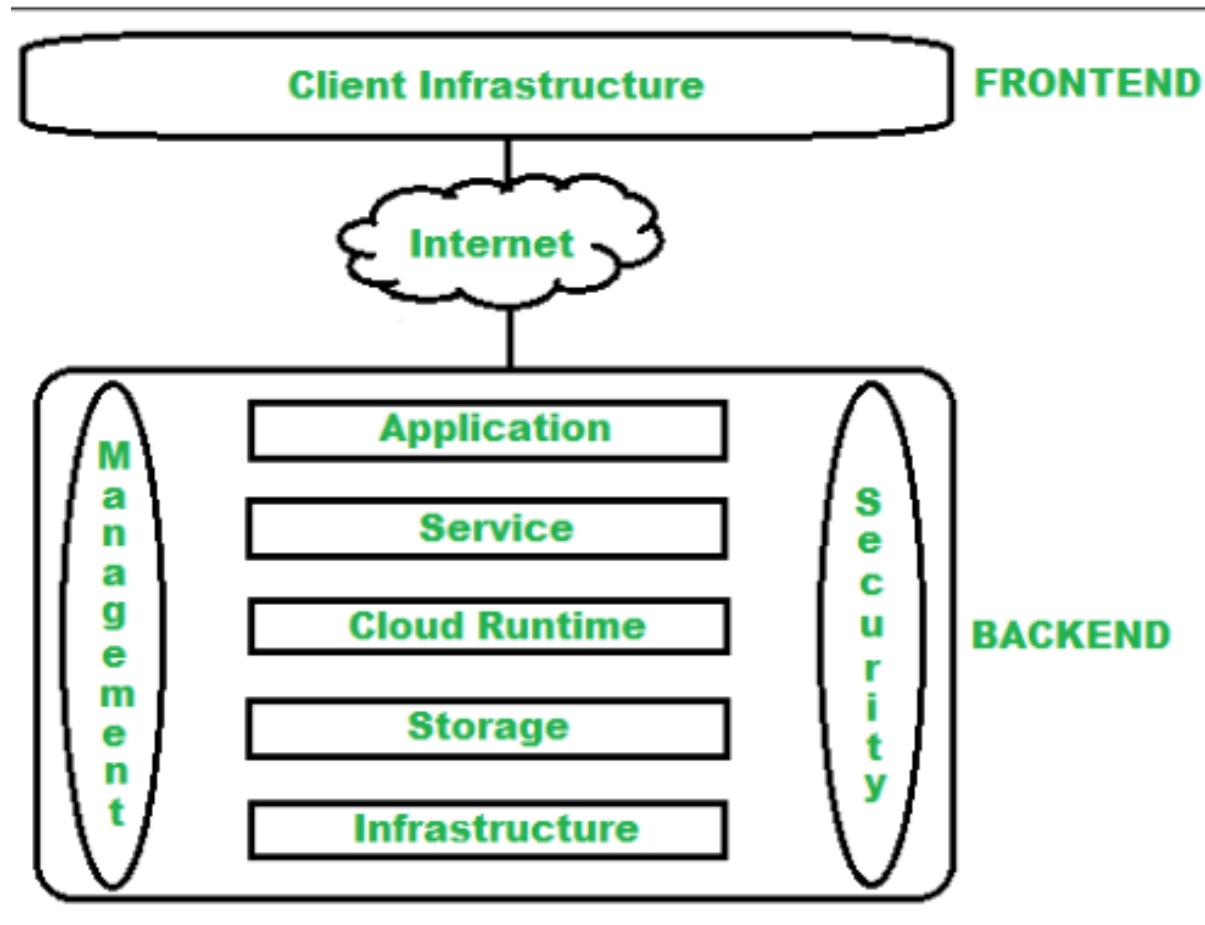
Cloud Infrastructure



Architecture of Cloud Computing

The cloud architecture is divided into 2 parts i.e.

- **Frontend**
- **Backend**



Architecture of Cloud Computing

1. Frontend :

Frontend of the cloud architecture refers to the client side of cloud computing system. Means it contains all the user interfaces and applications which are used by the client to access the cloud computing services/resources. For example, use of a web browser to access the cloud platform.

Client Infrastructure – Client Infrastructure is a part of the frontend component. It contains the applications and user interfaces which are required to access the cloud platform.

In other words, it provides a GUI(Graphical User Interface) to interact with the cloud.

Architecture of Cloud Computing

2. Backend :

Backend refers to the cloud itself which is used by the service provider. It contains the resources as well as manages the resources and provides security mechanisms. Along with this, it includes huge storage, virtual applications, virtual machines, traffic control mechanisms, deployment models, etc.

Application – Application in backend refers to a software or platform to which client accesses. Means it provides the service in backend as per the client requirement.

Service – Service in backend refers to the major three types of cloud based services like SaaS, PaaS and IaaS. Also manages which type of service the user accesses.

Runtime Cloud- Runtime cloud in backend provides the execution and Runtime platform/environment to the Virtual machine.

Storage – Storage in backend provides flexible and scalable storage service and management of stored data.

Infrastructure – Cloud Infrastructure in backend refers to the hardware and software components of cloud like it includes servers, storage, network devices, virtualization software etc.

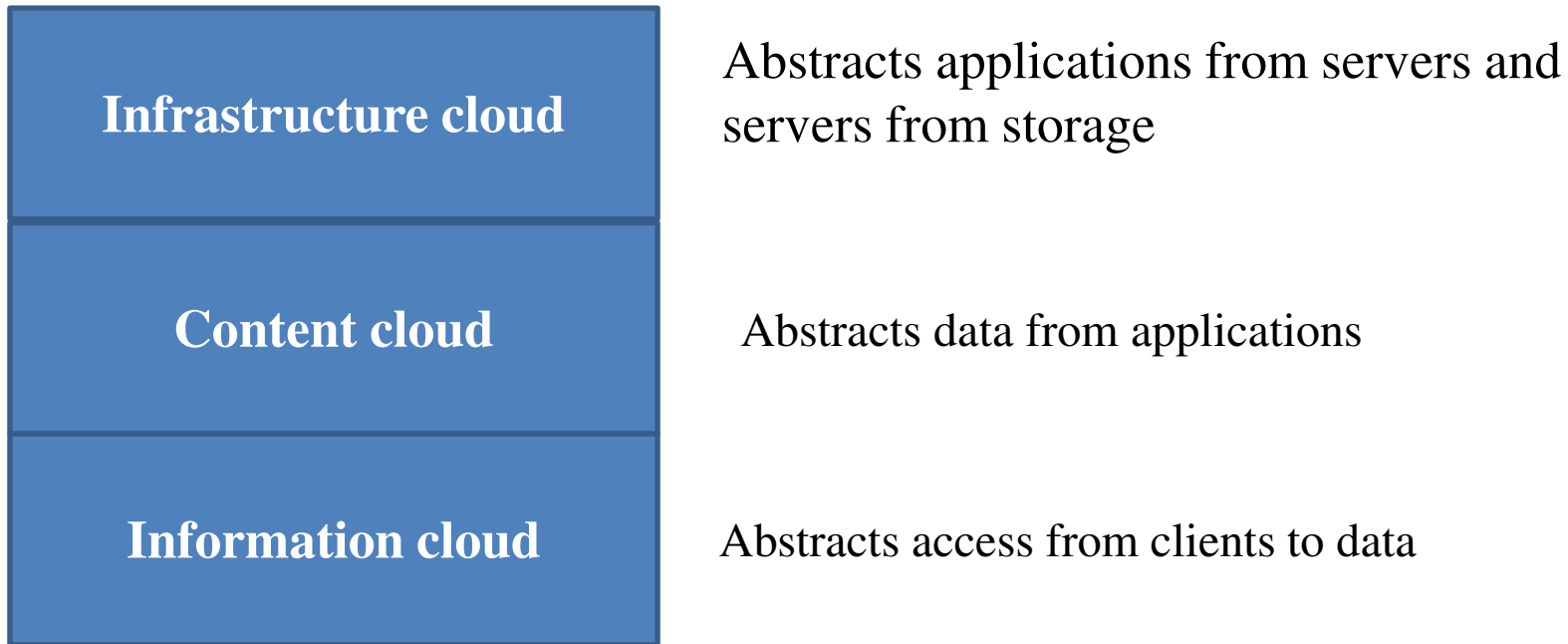
Management – Management in backend refers to management of backend components like application, service, runtime cloud, storage, infrastructure, and other security mechanisms etc.

Security – Security in backend refers to implementation of different security mechanisms in the backend for secure cloud resources, systems, files, and infrastructure to end-users.

Internet – Internet connection acts as the medium or a bridge between frontend and backend and establishes the interaction and communication between frontend and backend.

Layered Cloud Architecture Development

The three cloud layers are:



Infrastructure cloud

An **infrastructure cloud** includes the physical components that run applications and store data.

Virtual servers are created to run applications, and virtual storage pools are created to house new and existing data into dynamic tiers of storage based on performance and reliability requirements.

Virtual abstraction is employed so that servers and storage can be managed as logical rather than individual physical entities.

Content cloud

The **content cloud** implements metadata and indexing services over the infrastructure cloud to provide abstracted data management for all content.

The goal of a content cloud is to abstract the data from the applications so that different applications can be used to access the same data, and applications can be changed without worrying about data structure or type.

The content cloud transforms data into objects so that the interface to the data is no longer tied to the actual access to the data, and the application that created the content in the first place can be long gone while the data itself is still available and searchable.

Information cloud

The **information cloud** is the ultimate goal of cloud computing and the most common from a public perspective.

The information cloud abstracts the client from the data. For example, a user can access data stored in a database in Singapore via a mobile phone in Atlanta, or watch a video located on a server in Japan from his a laptop in the U.S.

The information cloud abstracts everything from everything.

The Internet is an information cloud.

Cloud Computing Design Challenges

- **Data Security and Privacy**
- **Cost Management**
- **Multi-Cloud Environments**
- **Performance Challenges**
- **Interoperability and Flexibility**
- **High Dependence on Network**
- **Lack of Knowledge and Expertise**

Data Security and Privacy

User or organizational data stored in the cloud is critical and private.

Even if the cloud service provider assures data integrity, it is your responsibility to carry out user authentication and authorization, identity management, data encryption, and access control.

Security issues on the cloud include identity theft, data breaches, malware infections, and a lot more which eventually decrease the trust amongst the users of your applications.

This can in turn lead to potential loss in revenue alongside reputation and stature.

Also, dealing with cloud computing requires sending and receiving huge amounts of data at high speed, and therefore is susceptible to data leaks.

Cost Management

Even as almost all cloud service providers have a “Pay As You Go” model, which reduces the overall cost of the resources being used, there are times when there are huge costs incurred to the enterprise using cloud computing.

When there is under optimization of the resources, let's say that the servers are not being used to their full potential, add up to the hidden costs.

If there is a degraded application performance or sudden spikes or overages in the usage, it adds up to the overall cost.

Unused resources are one of the other main reasons why the costs go up. If you turn on the services or an instance of cloud and forget to turn it off during the weekend or when there is no current use of it, it will increase the cost without even using the resources.

Multi-Cloud Environments

Due to an increase in the options available to the companies, enterprises not only use a single cloud but depend on multiple cloud service providers.

Most of these companies use hybrid cloud tactics and close to 84% are dependent on multiple clouds.

This often ends up being hindered and difficult to manage for the infrastructure team.

The process most of the time ends up being highly complex for the IT team due to the differences between multiple cloud providers.

Performance Challenges

If the performance of the cloud is not satisfactory, it can drive away users and decrease profits.

Even a little latency while loading an app or a web page can result in a huge drop in the percentage of users.

This latency can be a product of inefficient load balancing, which means that the server cannot efficiently split the incoming traffic so as to provide the best user experience.

Challenges also arise in the case of fault tolerance, which means the operations continue as required even when one or more of the components fail.

Interoperability and Flexibility

When an organization uses a specific cloud service provider and wants to switch to another cloud-based solution, it often turns up to be a tedious procedure since applications written for one cloud with the application stack are required to be re-written for the other cloud.

There is a lack of flexibility from switching from one cloud to another due to the complexities involved.

Handling data movement, setting up the security from scratch and network also add up to the issues encountered when changing cloud solutions, **thereby reducing flexibility.**

High Dependence on Network

Since cloud computing deals with provisioning resources in real-time, it deals with enormous amounts of data transfer to and from the servers. *This is only made possible due to the availability of the high-speed network.*

Although these data and resources are exchanged over the network, this can prove to be highly vulnerable in case of limited bandwidth or cases when there is a sudden outage.

Even when the enterprises can cut their hardware costs, they need to ensure that the internet bandwidth is high as well there are zero network outages, or else it can result in a potential business loss.

It is therefore a major challenge for smaller enterprises that have to maintain network bandwidth that comes with a high cost.

Lack of Knowledge and Expertise

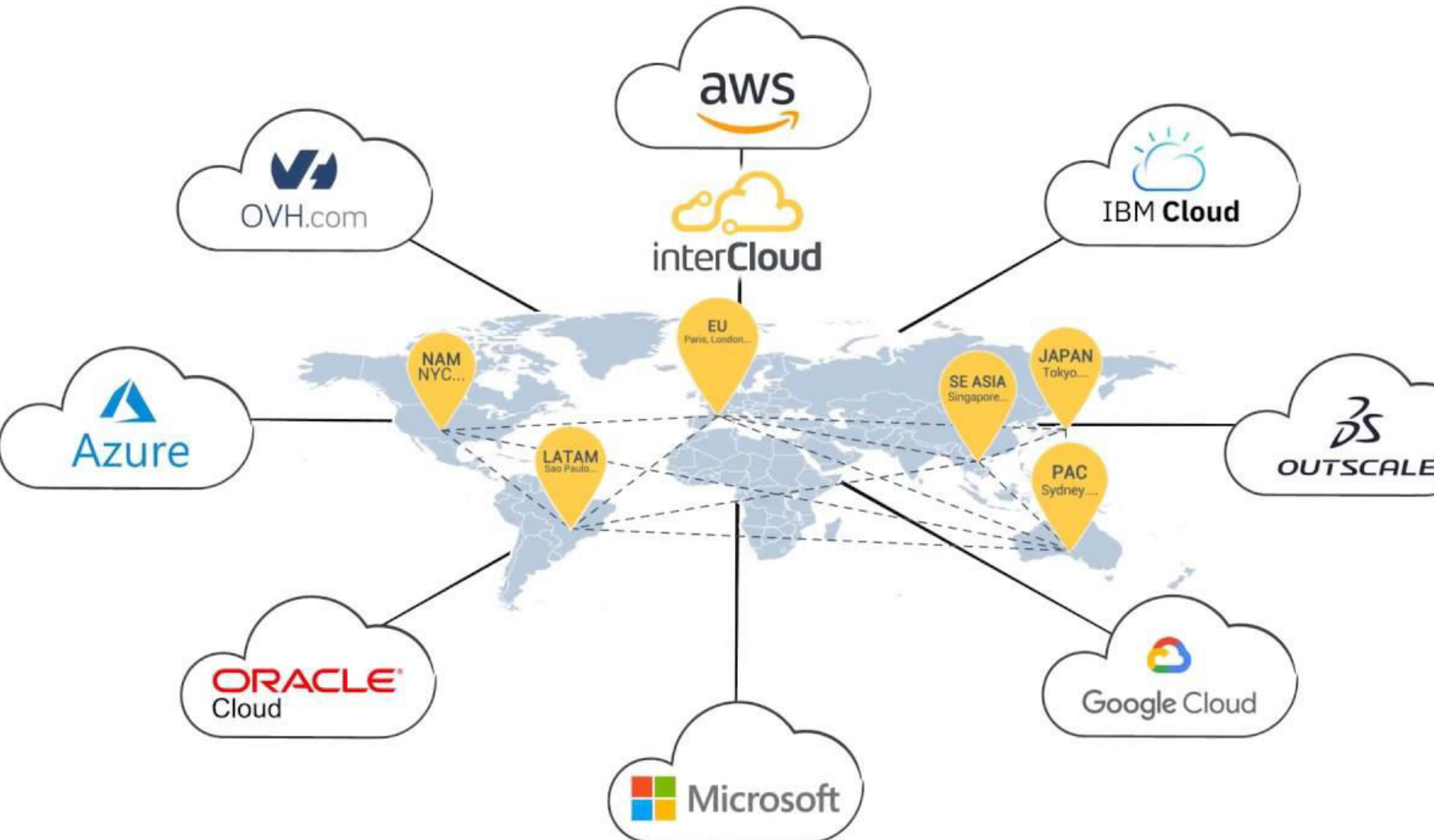
Due to the complex nature and the high demand for research working with the cloud often ends up being a highly tedious task.

It requires immense knowledge and wide expertise on the subject. Although there are a lot of professionals in the field they need to constantly update themselves.

Cloud computing is a highly paid job due to the extensive gap between demand and supply. There are a lot of vacancies but very few talented cloud engineers, developers, and professionals.

Therefore, there is a need for upskilling so these professionals can actively understand, manage and develop cloud-based applications with minimum issues and maximum reliability.

Cloud of Clouds (Intercloud)



Cloud of Clouds (Intercloud)

Intercloud or ‘cloud of clouds’ is a term refer to a theoretical model for **cloud computing services based on the idea of combining many different individual clouds into one seamless mass** in terms of on-demand operations.

The intercloud would simply make sure that **a cloud could use resources beyond its reach**, by taking advantage of pre-existing contracts with other cloud providers.

The Intercloud scenario is based on the key concept that **each single cloud does not have infinite physical resources or ubiquitous geographic footprint**.

If a cloud saturates the computational and storage resources of its infrastructure, or is requested to use resources in a geography where it has no footprint, it would still be able to satisfy such requests for service allocations sent from its clients.

Cloud of Clouds (Intercloud)

The Intercloud scenario would address such situations where each cloud would use the computational, storage, or any kind of resource (through semantic resource descriptions, and open federation) of the infrastructures of other clouds.

This is analogous to the way the Internet works, in that a service provider, to which an endpoint is attached, will access or deliver traffic from/to source/destination addresses outside of its service area by using Internet routing protocols with other service providers with whom it has a pre-arranged exchange or **peering relationship**.

It is also analogous to the way mobile operators implement roaming and inter-carrier interoperability. Such forms of cloud exchange, peering, or roaming may introduce **new business opportunities among cloud providers if they manage to go beyond the theoretical framework**.

Inter Cloud Resource Management

The cloud infrastructure layer can be divided as Data as a Service (DaaS) and Communication as a Service (CaaS) in addition to compute and storage in IaaS.

Cloud players are divided into three classes:

(1) cloud service providers and IT administrators,

(2) software developers or vendors, and

(3) end users or business users.

- These cloud players vary in their roles under the IaaS, PaaS, and SaaS models.

Inter Cloud Resource Management

From the software vendors' perspective, **application performance** on a given cloud platform is most important.

From the providers' perspective, **cloud infrastructure performance** is the primary concern.

From the end users' perspective, **the quality of services, including security**, is the most important

Developers have to consider how to design the system to meet critical requirements such as **high throughput, HA, and fault tolerance**.

Even the operating system might be modified to meet the special requirement of cloud data processing.

Inter Cloud Resource Management

Runtime Support Services

- As in a cluster environment, there are also some runtime supporting services in the cloud computing environment.
- **Cluster monitoring** is used to collect the runtime status of the entire cluster.
- The **scheduler** queues the tasks submitted to the whole cluster and assigns the tasks to the processing nodes according to node availability.
- Runtime support is software needed in **browser-initiated** applications applied by thousands of cloud customers.

Inter-cloud Challenges

Identification: A system should be created where each cloud can be identified and accessed by another cloud, similar to how devices connected to the internet are identified by IP addresses.

Communication: A universal language of the cloud should be created so that they are able to verify each other's available resources.

Payment: When one provider uses the assets of another provider, a question arises on how the second provider will be compensated, so a proper payment process should be developed.

.

Global Exchange of Cloud Resources

Enterprises currently employ Cloud services in order to improve the scalability of their services

□ Enterprise service consumers with global operations require faster response time. Save time by distributing workload requests to multiple Clouds in various locations at the same time

Cloud computing to mature, it is required that the services follow standard interfaces,

□ BECAUSE,

- No standard interfaces/ compatibility between vendor's systems
- No uniform rate

Global Exchange of Cloud Resources

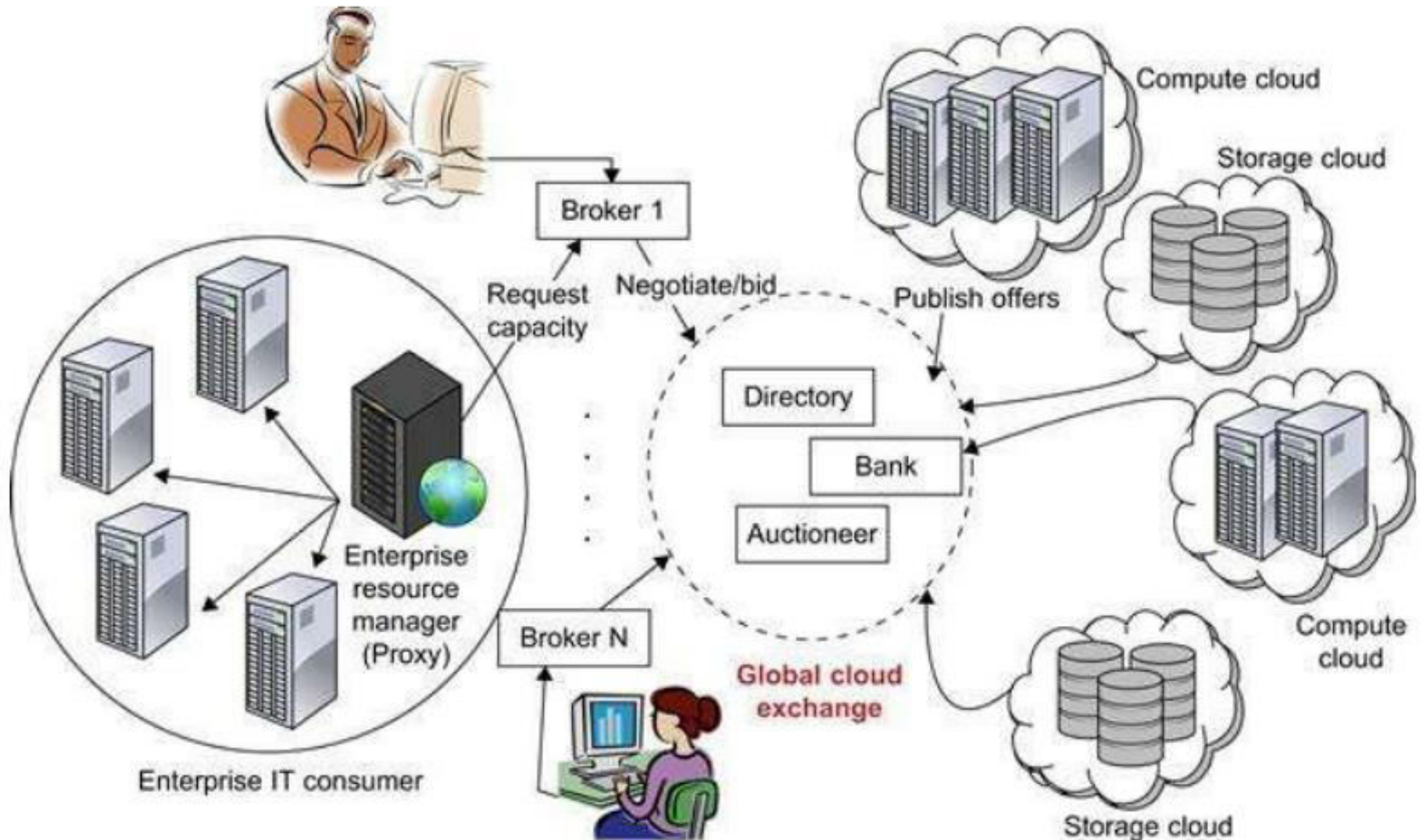
University of Melbourne proposed InterCloud architecture (Cloudbus)

- Supports brokering and exchange of cloud resources for scaling applications across multiple clouds
- Cloud Exchange (CEX) acts as a market maker for bringing together service producers and consumers

It aggregates the infrastructure demands from application brokers and evaluates them against the available supply.

- It supports trading of cloud services based on competitive economic models such as commodity markets and auctions

Market oriented Cloud



Cloud Exchange Components

Market directory allows participants to locate providers or consumers with the right offers

- Auctioneers periodically clear bids and asks received from market participants
- Banking system ensures that financial transactions pertaining to agreements between participants are carried out

Brokers mediate between consumers and providers by buying capacity from the provider and sub-leasing these to the consumers

- Has to choose those users whose applications can provide it maximum utility
- Interacts with resource providers and other brokers to gain or to trade resource shares. Equipped with a negotiation module that is informed by the current conditions of the resources and the current demand to make its decisions

Cloud Exchange Components

Consumers, brokers and providers are bound to their requirements and related compensations through **SLAs** ([Service Level Agreements](#))

Consumer participates in the utility market through a resource management proxy that selects a set of brokers based on their offering

- Then forms SLAs with the brokers that bind the latter to provide the guaranteed resources

- then deploys his own environment on the leased resources or uses the provider's interfaces in order to scale his applications

Cloud Exchange Components

Providers can use the markets in order to perform effective capacity planning

- A provider is equipped with a price-setting mechanism which sets the current price for there source based on market conditions, user demand, and current level of utilization of the resource
- Negotiation process proceeds until an SLA is formed or the participants decide to break off

An admission-control mechanism at a provider's end selects the auctions to participate in or the brokers to negotiate with, based on an initial estimate of the utility.

- Resource management system provides functionalities such as advance reservations that enable guaranteed provisioning of resource capacity



REFERENCES

<https://www.bmc.com/blogs/cloud-infrastructure/>

<https://www.youtube.com/watch?v=54hM9LB72fQ>

https://www.youtube.com/watch?v=QAhr0_SPMIY

<https://www.youtube.com/watch?v=m1WKONoS0XQ>

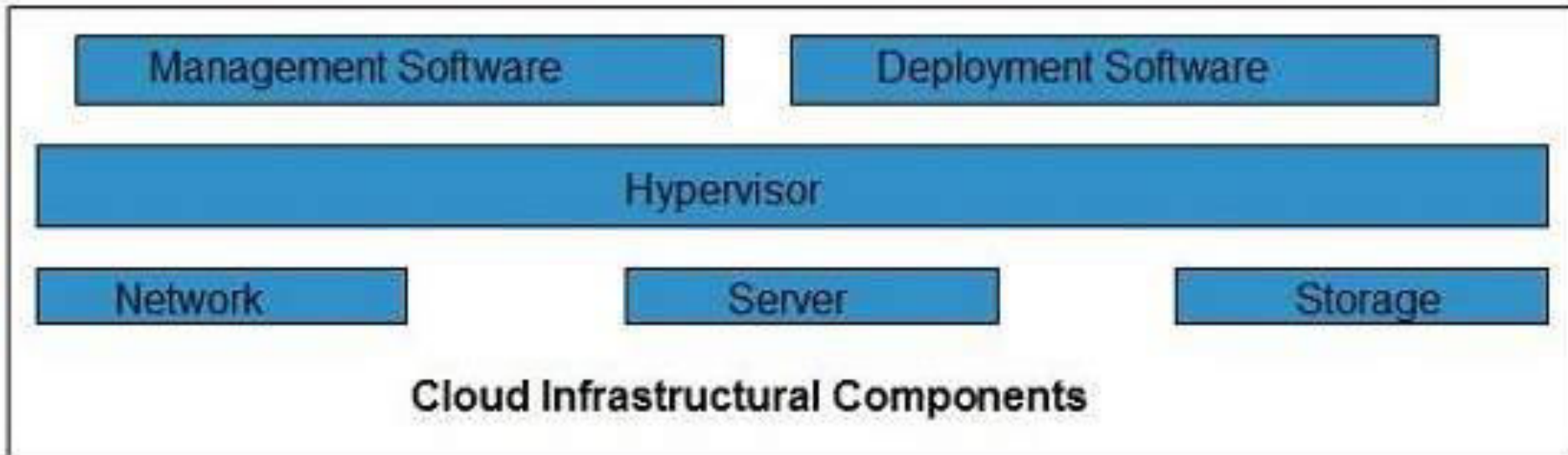
<https://www.youtube.com/watch?v=n-YHBJwA3yk>

Cloud Computing Technology

Module 4

Cloud Computing Infrastructure

Cloud infrastructure consists of servers, storage devices, network, cloud management software, deployment software, and platform virtualization.



Cloud infrastructure

Hypervisor is a **firmware** or **low-level program** that acts as a Virtual Machine Manager. It allows to share the single physical instance of cloud resources between several tenants.

Management Software helps to maintain and configure the infrastructure.

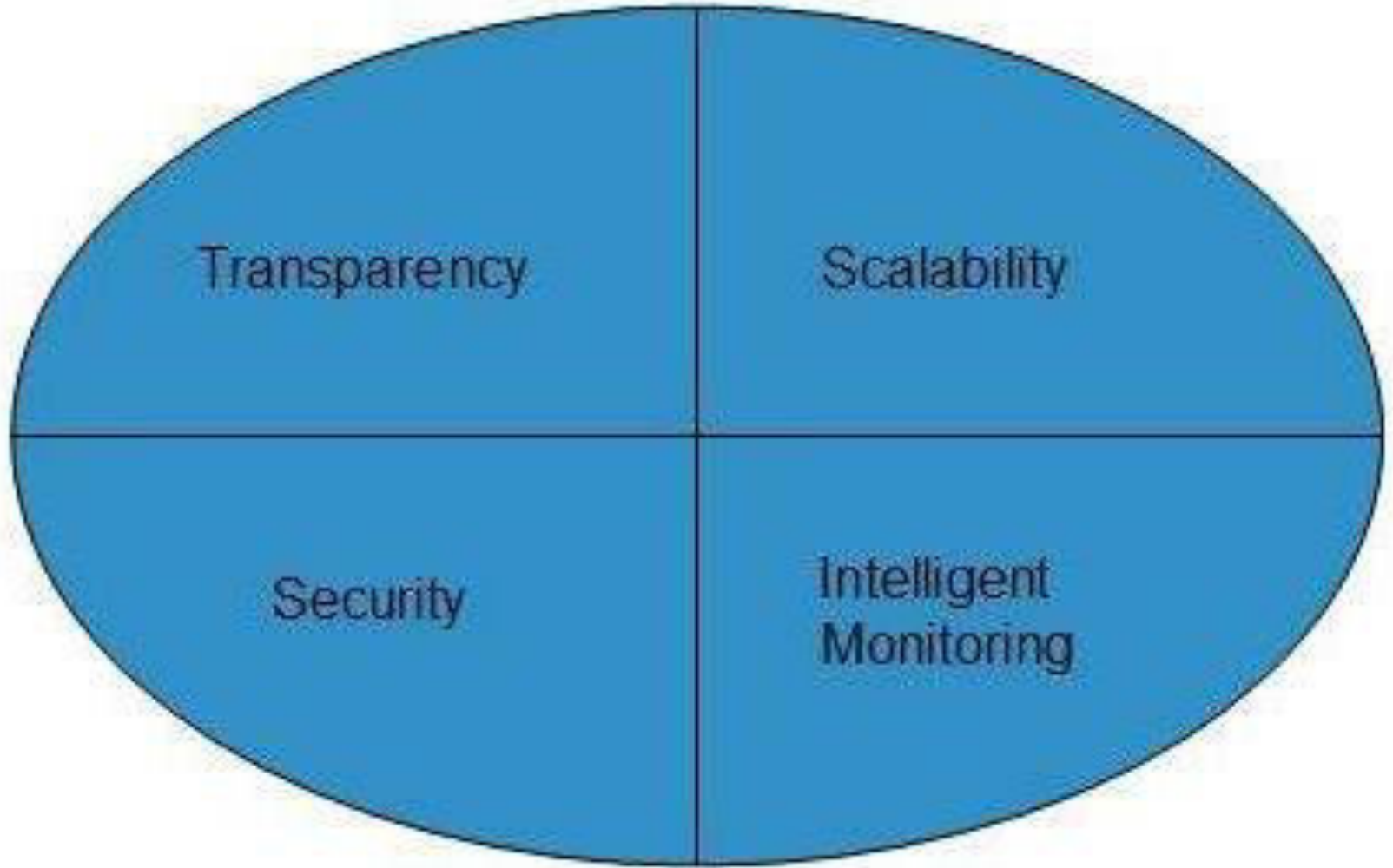
Deployment Software helps to deploy and integrate the application on the cloud.

Network allows to connect cloud services over the Internet. It is also possible to deliver network as a utility over the Internet, which means, the customer can customize the network route and protocol.

Server helps to compute the resource sharing and offers other services such as resource allocation and de-allocation, monitoring the resources, providing security etc.

Storage: Cloud keeps multiple replicas of storage. If one of the storage resources fails, then it can be extracted from another one, which makes cloud computing more reliable.

Infrastructural Constraints



Infrastructural Constraints

Transparency

Virtualization is the key to share resources in cloud environment. But it is not possible to satisfy the demand with single resource or server. Therefore, there must be transparency in resources, load balancing and application, so that we can scale them on demand.

Scalability

Scaling up an application delivery solution is not that easy as scaling up an application because it involves configuration overhead or even re-architecting the network. So, application delivery solution is need to be scalable which will require the virtual infrastructure such that resource can be provisioned and de-provisioned easily.

Intelligent Monitoring

To achieve transparency and scalability, application solution delivery will need to be capable of intelligent monitoring.

Security

The mega data center in the cloud should be securely architected. Also the control node, an entry point in mega data center, also needs to be secure.

Cloud Security Architecture

A cloud security architecture is defined by the security layers, design, and structure of the platform, tools, software, infrastructure, and best practices that exist within a cloud security solution.

A cloud security architecture provides the written and visual model to define how to configure and secure activities and operations within the cloud, including such things as

- identity and access management;
- methods and controls to protect applications and data;
- approaches to gain and maintain visibility into compliance, threat posture, and overall security;
- processes for instilling security principles into cloud services development and operations;
- policies and governance to meet compliance standards; and
- physical infrastructure security components.

The Importance of Cloud Security

With so many business activities now being conducted in the cloud and with cyberthreats on the rise, cloud security becomes critically important.

A cyber or cybersecurity threat is a malicious act that seeks to damage data, steal data, or disrupt digital life in general. Cyber-attacks include threats like computer viruses, data breaches, and Denial of Service (DoS) attacks.

Common types of cyber attacks

Malware. Malware is a term used to describe malicious software, including spyware, ransomware, viruses, and worms. ...

Phishing. ...

Man-in-the-middle attack. ...

Denial-of-service attack. ...

SQL injection. ...

Zero-day exploit. ...

DNS Tunneling.

The Importance of Cloud Security

Cloud security solutions offer protection from different types of threats, such as:

Malicious activity caused by cybercriminals and nation states, including ransomware, malware, phishing; denial of service (DoS) attacks; advanced persistent threats; cryptojacking; brute force hacking; man-in-the-middle attacks; zero-day attacks; and data breaches.

Insider threats caused by rogue or disgruntled staff members; human error; staff negligence; and infiltration by external threat actors who have obtained legitimate credentials without authorization.

Vulnerabilities such as flaws or weaknesses in systems, applications, procedures, and internal controls that can accidentally or intentionally trigger a security breach or violation.

Technologies for Cloud Security

Cloud Access Security Broker (CASB): A cloud access security broker sits between cloud service providers and cloud users to monitor activity and enforce security policies. *A CASB combines multiple types of security policy enforcement, such as authentication, authorization, single sign-on, credential mapping, device profiling, encryption tokenization, and logging.*

Cloud Encryption: Cloud encryption involves using algorithms to change data and information stored in the cloud to make it indecipherable to anyone not possessing access to the encryption keys. Encryption protects information from being accessed and read, even if it is intercepted during a transfer from a device to the cloud or between different cloud systems. *Cloud Service Providers and cloud vendors offer unique ways to provide encryption into an environment with Key Management tools.*

Technologies for Cloud Security

Cloud Firewalls: A firewall is a type of security solution that blocks or filters out malicious traffic. Cloud firewalls form virtual barriers between corporate assets that exist in the cloud and untrusted, external, internet-connected networks and traffic. Cloud firewalls are also sometimes called a firewall-as-a-service (FWaaS).

Cloud Security Posture Management: CSPM security solutions continuously monitor the dynamic and ever-changing cloud environments and identify gaps between security policy and security posture. CSPM solutions comply with specific frameworks (e.g. SOC2, CIS v1.1, HIPAA) and can be used to consolidate misconfigurations, safeguard the flow of data between internal IT architecture and the cloud, and broaden security policies to extend beyond in-house infrastructure.

Cloud Computing Platforms and Technologies

Cloud Computing Top Platforms



Google
App Engine



Microsoft Azure



Cloud Computing Platforms and Technologies

Amazon Web Services ([AWS](#)) –

AWS provides different wide-ranging clouds IaaS services, which ranges from virtual compute, storage, and networking to complete computing stacks.

AWS is well known for its storage and compute on demand services, named as *Elastic Compute Cloud (EC2) and Simple Storage Service (S3)*.

EC2 offers customizable virtual hardware to the end user which can be utilize as the base infrastructure for deploying computing systems on the cloud.

Either the AWS console, which is a wide-ranged Web portal for retrieving AWS services, or the web services API available for several programming language is used to deploy the EC2 instances

Amazon Web Services (AWS)

EC2 also offers the capability of saving an explicit running instance as image, thus allowing users to create their own templates for deploying system.

S3 stores these templates and delivers persistent storage on demand. S3 is well ordered into buckets which contains objects that are stored in binary form and can be grow with attributes.

End users can store objects of any size, from basic file to full disk images and have them retrieval from anywhere.

In addition, EC2 and S3, a wide range of services can be leveraged to build virtual computing system including: networking support, caching system, DNS, database support, and others.

Google AppEngine –

Google AppEngine is a scalable runtime environment frequently dedicated to executing web applications.

These utilize benefits of the large computing infrastructure of Google to dynamically scale as per the demand.

AppEngine offers both a secure execution environment and a collection of which simplifies the development of scalable and high-performance Web applications.

These services include: in-memory caching, scalable data store, job queues, messaging, and cron tasks.

Developers and Engineers can build and test applications on their own systems by using the AppEngine SDK, which replicates the production runtime environment, and helps test and profile applications

Microsoft Azure –

Microsoft Azure is a Cloud operating system and a platform in which user can develop the applications in the cloud. Generally, a scalable runtime environment for web applications and distributed applications is provided.

Application in Azure are organized around the fact of roles, which identify a distribution unit for applications and express the application's logic.

Azure provides a set of additional services that complement application execution such as support for storage, networking, caching, content delivery, and others.

Hadoop –

Apache Hadoop is an open source framework that is appropriate for processing large data sets on commodity hardware.

Hadoop is an implementation of MapReduce, an application programming model which is developed by Google. This model provides two fundamental operations for data processing: map and reduce.

Yahoo! Is the sponsor of the Apache Hadoop project, and has put considerable effort in transforming the project to an enterprise-ready cloud computing platform for data processing. Hadoop is an integral part of the Yahoo! Cloud infrastructure and it supports many business processes of the corporates.

Web API

API stands for Application Programming Interface. API is actually some kind of interface which is having a set of functions. These set of functions will allow programmers to acquire some specific features or the data of an application

Web API is an API as the name suggests, it can be accessed over the web using the HTTP protocol.

The web API can be developed by using different technologies such as java, ASP.NET, etc.

Web API is used in either a web server or a web browser. Basically Web API is a web development concept. It is limited to Web Application's client-side and also it does not include a web server or web browser details

Where to use Web API?

Web APIs are very useful in implementation of [RESTFUL web services](#) using .NET framework.

Web API helps in enabling the development of HTTP services to reach out to client entities like browser, devices or tablets.

[ASP.NET Web API](#) can be used with MVC for any type of application

.
A web API can help you develop ASP.NET application via [AJAX](#).

Hence, web API makes it easier for the developers to build an ASP.NET application that is compatible with any browser and almost any device.

Why to Choose Web API?

A Web API services are preferable over other services to use with a native application that does not support SOAP but require web services.

For creating resource-oriented services, the web API services are the best to choose. By using HTTP or restful service, these services are established.

If you want good performance and fast development of services, the web API services are very helpful.

For developing light weighted and maintainable web services, web API services are really helpful to develop that service. It supports any text pattern like [JSON](#), [XML](#) etc.

The devices that have tight bandwidth or having a limitation in bandwidth, then the Web API services are the best for those devices.

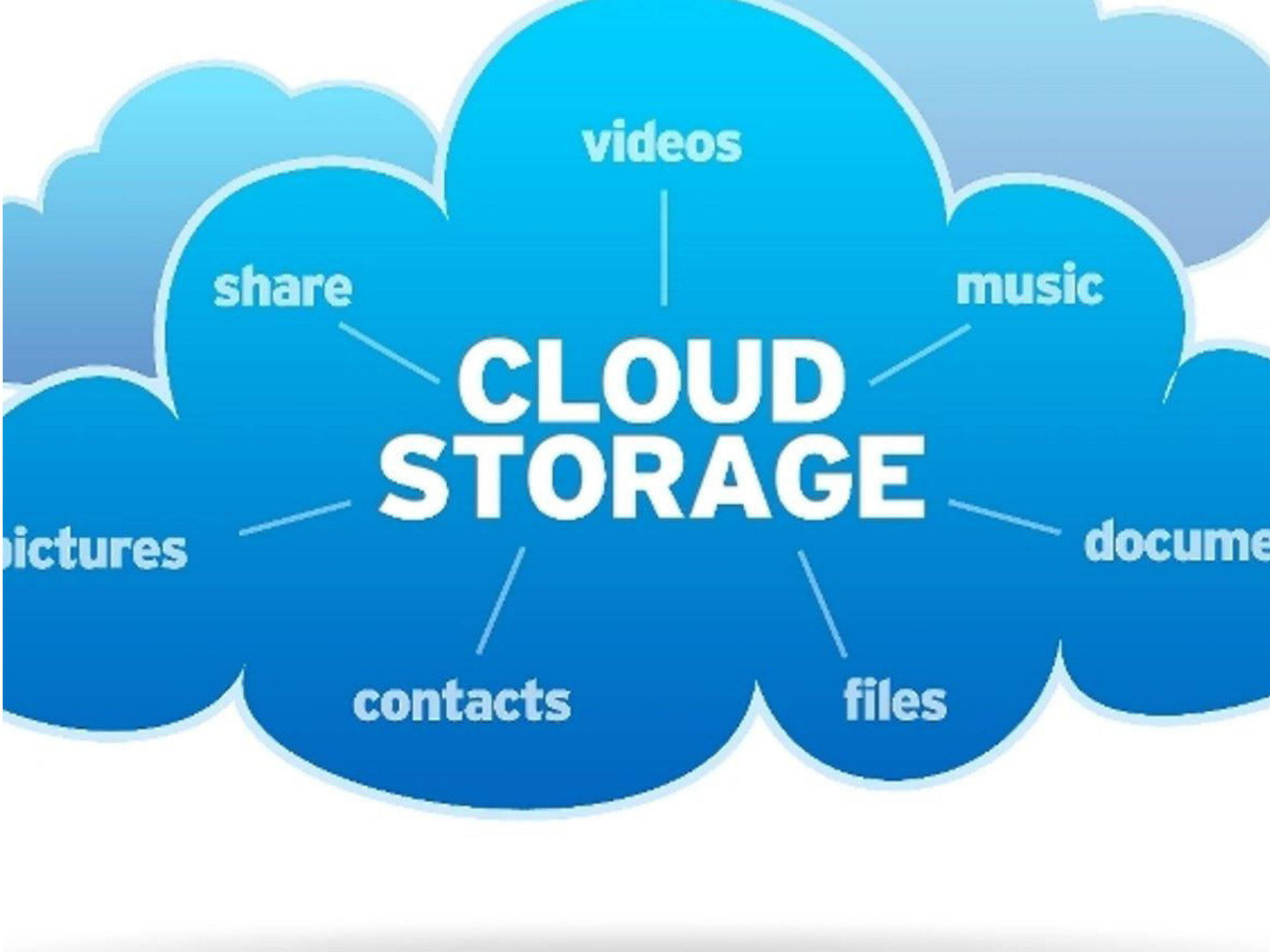
How to use Web API?

Web API receives requests from different types of client devices like mobile, laptop, etc, and then sends those requests to the webserver to process those requests and returns the desired output to the client.

Web API is a System to System interaction, in which the data or information from one system can be accessed by another system, after the completion of execution the resultant data or we can say as output is shown to the viewer.

API provides data to its programmers which is made available to outside users. When programmers decide to make some of their data available to the public, they “expose endpoints, ” meaning they publish a portion of the language they have used to build their program.

Other programmers can then extract the data from the application by building URLs or using HTTP clients to request data from those endpoints.



videos

share

music

**CLOUD
STORAGE**

ictures

docume

contacts

files

Cloud Storage

The cloud storage is a computer data storage model in which the data that is digital in format is stored, and hence it is said to be on the cloud, in logical pools.

This physical storage consists of multiple servers which can be located in the different parts of the country or may be in different parts of the world depends on many factors. The maintenance of these servers is owned by some private companies.

The cloud storage services are also responsible for keeping the data available and accessible 24x7, and it also safeguard the data and run the physical environment.

In order to store user, entity, or application information, individuals and organizations purchase or lease storage capacity from providers.

Cloud storage

It consists of several distributed resources, but still functions as one, either in a cloud architecture of federated or cooperative storage.

Highly fault-tolerant via redundancy and data distribution.

Extremely durable through the manufacture of copies of versions.

Companies just need to pay for the storage they actually use, normally an average of a month's consumption. This does not indicate that cloud storage is less costly, rather that operating costs are incurred rather than capital expenses.

Cloud storage companies can cut their energy usage by up to 70 percent, making them a greener company.

The provision of storage and data security is inherent in the architecture of object storage, so the additional infrastructure, effort and expense to incorporate accessibility and security can be removed depending on the application.

.

Cloud storage provider

A cloud storage provider, also known as a managed service provider ([MSP](#)), is a company that offers organizations and individuals the ability to place and retain data in an off-site storage system. Customers can lease cloud storage capacity per month or on demand.

A cloud storage provider hosts a customer's data in its own data center, providing fee-based computing, networking and storage infrastructure. Both individual and corporate customers can get unlimited storage capacity on a provider's servers at a low per-gigabyte price.

Rather than store data on local storage devices, such as a hard disk drive, flash storage or tape, customers choose a cloud storage provider to host data on a system in a remote data center. Users can then access those files using an internet connection.

Cloud storage provider



iCloud



OneDrive



sync.com

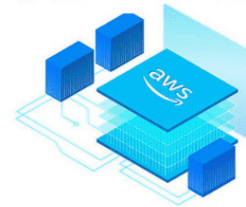


Dropbox



icedrive

POLARBACKUP



Cloud standards

Common goals for cloud standards include portability, migration, and security. They also enable interoperability across multiple platforms, balance workloads and underpin security and data protection.

Security and assurance standards are a significant force in cloud computing when it comes to these kinds of objectives, but products and vendors are still fiercely innovating and standards won't emerge until their technologies stabilize.

Open standards are formalised protocols, accepted by multiple parties as a basis for joint (usually commercial) action, usually under the jurisdiction of some governing authority. They may have multiple implementations or different ways of achieving the same outcome.

The **NIST standard** describes important aspects of cloud computing and serves as a benchmark for comparing cloud services and deployment strategies.

Cloud standards

Cloud Standards Customer Council (CSCC)

CSCC is an end-user support group focused on the adoption of cloud technology and examining cloud standards and security and interoperability issues. It has been superseded by the Cloud Working Group and addresses cloud standards issues via its Cloud Working Group.

DMTF

DMTF supports the management of existing and new technologies, such as cloud, by developing appropriate standards. Its working groups, such as Open Cloud Standards Incubator, Cloud Management Working Group and Cloud Auditing Data Federation Working Group, address cloud issues in greater detail.

European Telecommunications Standards Institute (ETSI)

ETSI primarily develops telecommunications standards. Among its cloud-focused activities are Technical Committee CLOUD, the Cloud Standards Coordination initiative and Global Inter-Cloud Technology Forum, each of which addresses cloud technology issues.

Cloud Computing Applications

**Cloud
Computing
Application**

Art

Business

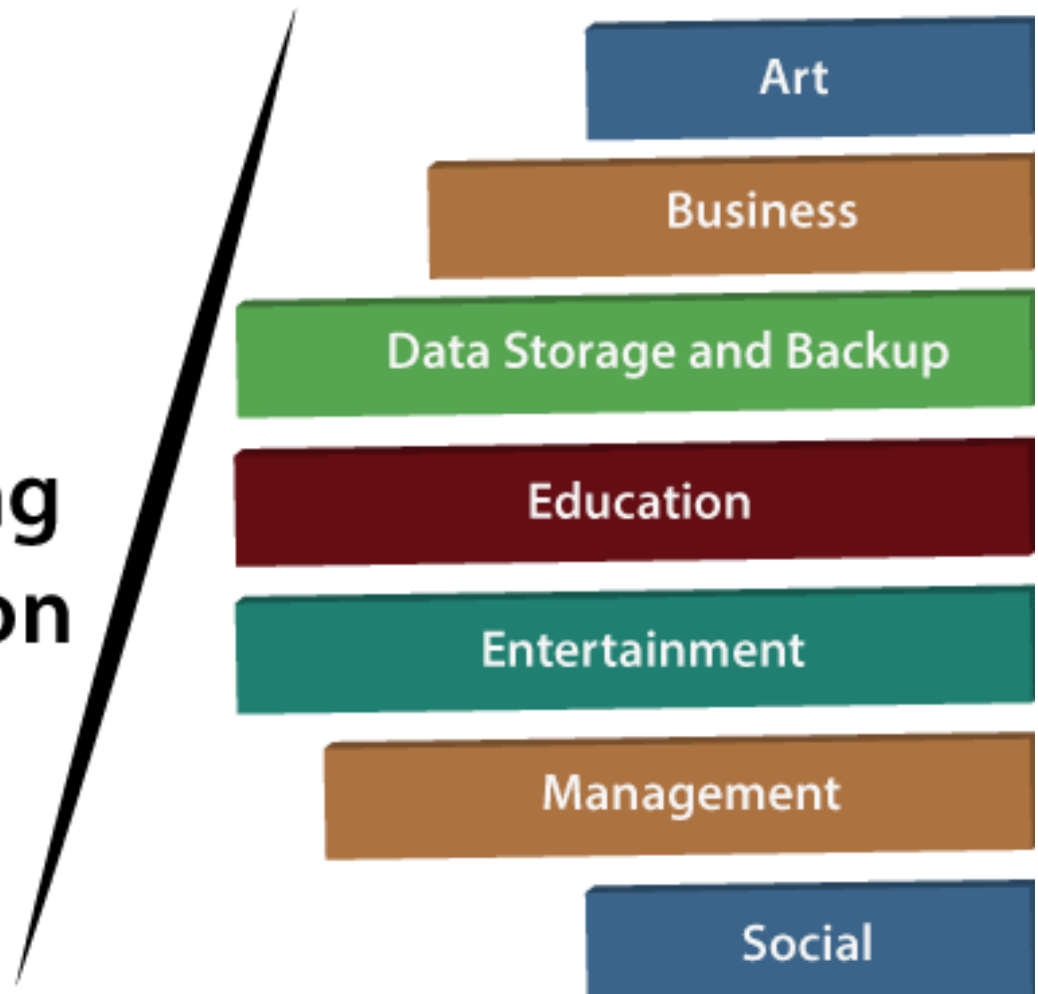
Data Storage and Backup

Education

Entertainment

Management

Social



Art Applications

Cloud computing offers various art applications for quickly and easily design **attractive cards, booklets, and images**. Some most commonly used cloud art applications are given below:

Vistaprint

Vistaprint allows us to easily design various printed marketing products such as business cards, Postcards, Booklets, and wedding invitations cards.

Adobe Creative Cloud

Adobe creative cloud is made for designers, artists, filmmakers, and other creative professionals. It is a suite of apps which includes PhotoShop image editing programming, Illustrator, InDesign, TypeKit, Dreamweaver, XD, and Audition.

Can you list other applications?

Business Applications

Business applications are based on cloud service providers. Today, every organization requires the cloud business application to grow their business. It also ensures that business applications are 24*7 available to users.

There are the following business applications of cloud computing

i. MailChimp is an email publishing platform which provides various options to design, send, and save templates for emails.

ii. Salesforce

Salesforce platform provides tools for sales, service, marketing, e-commerce, and more. It also provides a cloud development platform.

iv. Chatter

Chatter helps us to share important information about the organization in real time.

Data Storage and Backup Applications

Cloud computing allows us to store information (data, files, images, audios, and videos) on the cloud and access this information using an internet connection. As the cloud provider is responsible for providing security, so they offer various backup recovery application for retrieving the lost data.

A list of data storage and backup applications in the cloud are given below –

Box.com : Box provides an online environment for secure content management, workflow, and collaboration. It allows us to store different files such as Excel, Word, PDF, and images on the cloud.

The main advantage of using box is that it provides drag & drop service for files and easily integrates with Office 365, G Suite, Salesforce, and more than 1400 tools.

ii. Mozy

Mozy provides powerful **online backup solutions** for our personal and business data. It schedules automatically back up for each day at a specific time.

iii. Joukuu

Joukuu provides the simplest way to **share** and **track cloud-based backup files**. Many users use joukuu to search files, folders, and collaborate on documents.

iv. Google G Suite

Google G Suite is one of the best **cloud storage** and **backup** application. It includes Google Calendar, Docs, Forms, Google+, Hangouts, as well as cloud storage and tools for managing cloud apps.

The most popular app in the Google G Suite is Gmail. Gmail offers free email services to users.

Education Applications

Cloud computing in the education sector becomes very popular. It offers various **online distance learning platforms** and **student information portals** to the students.

i. Google Apps for Education

Google Apps for Education is the most widely used platform for free web-based email, calendar, documents, and collaborative study.

ii. Chromebooks for Education

Chromebook for Education is one of the most important Google's projects. It is designed for the purpose that it enhances education innovation.

iii. AWS in Education

AWS cloud provides an education-friendly environment to universities, community colleges, and schools.

What is a Cloud Client?

A cloud client is a hardware device or software used to access a cloud service. Computer systems, tablets, navigation devices, home automation devices, mobile phones and other smart devices, operating systems, and browsers can all be cloud clients.

Cloud clients have the basic processing and software capabilities needed to access specified cloud services.

A cloud client is essential to accessing cloud services, but may also be completely functional without those services. There are three main service models of cloud computing: SaaS, PaaS, and IaaS.

Users securely access cloud services by connecting with cloud client devices such as desktop computers, laptops, tablets and smartphones.

Some cloud clients rely on cloud computing for all or a majority of their applications and may be essentially useless without their cloud services