

Introduction to Cloud Computing

Trends in computing - Grid Computing, Cluster Computing, Distributed Computing, Utility Computing, Cloud Computing Defining a Cloud ,Vision of Cloud, Cloud Computing Reference Model, Characteristics and benefits ,Challenges of Cloud

22/8/2022

Computing

Distributed computing

- It is a composition of multiple independent systems but serve as a single entity to the users.
- Share resources and also use them effectively and efficiently.
- scalability, concurrency, continuous availability, heterogeneity, and independence in failures.
- main problem - used for same geographical location.
- **mobile systems, social media networks, weather monitoring systems, and e-commerce systems**

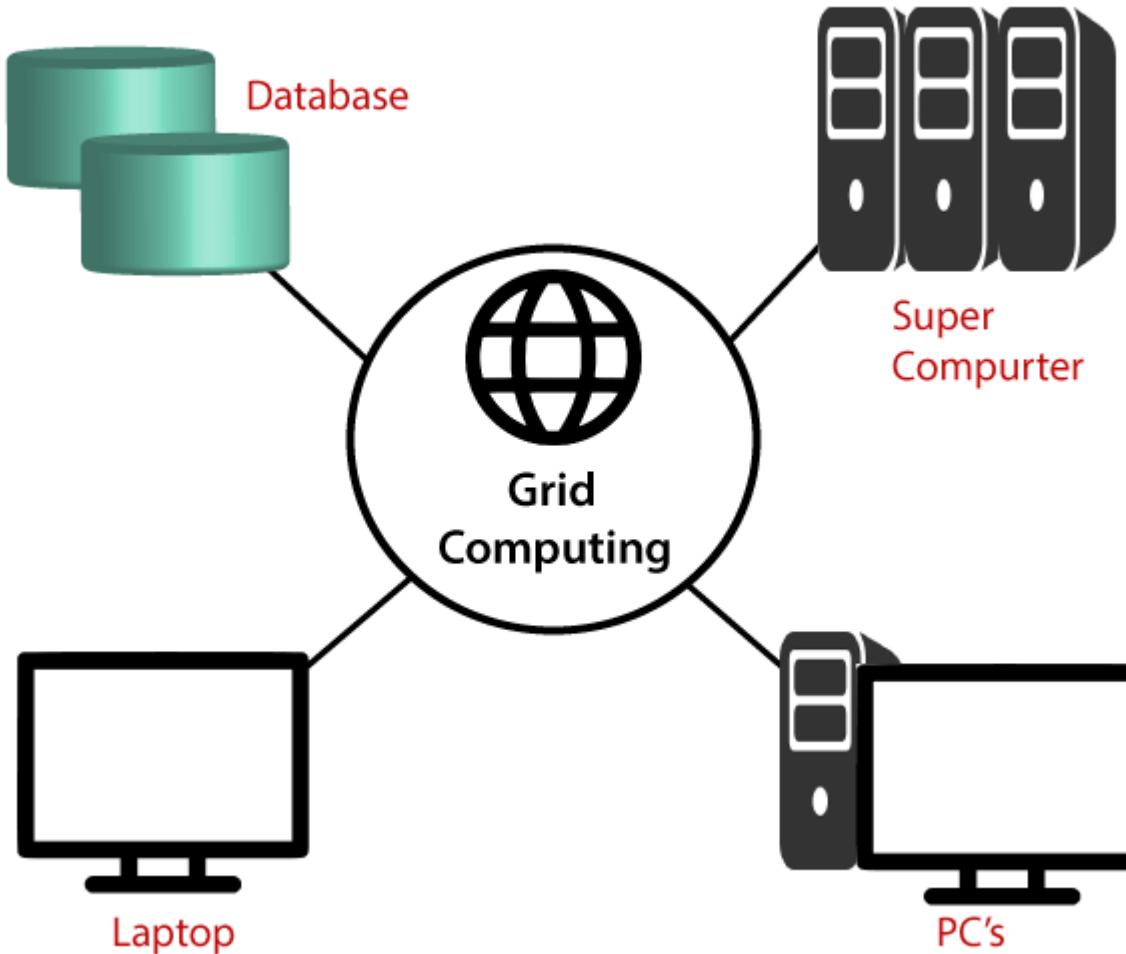
Mainframe computing

- Highly powerful and reliable computing machines.
- Responsible for handling large data such as massive input-output operations.
- These systems have almost no downtime with high fault tolerance., these increased the processing capabilities of the system.
- Very expensive.
- To reduce this cost, cluster computing came as an alternative to mainframe technology.

Cluster computing

- Each machine in the cluster was connected to each other by a network with high bandwidth.
- These were way cheaper than those mainframe systems.
- These were equally capable of high computations.
- Also, new nodes could easily be added to the cluster if it was required.

Grid Computing



Virtualization

- Virtualization is the process of creating a virtual environment to run multiple applications and operating systems on the same server.
- E.g. Amazon EC2, VMware vCloud, etc work on.
- Types of Virtualization
- Hardware virtualization
- Operating system virtualization
- Server virtualization
- Storage virtualization

Hardware Virtualization

- Hardware virtualization is **the method used to create virtual versions of physical desktops and operating systems.**
- It uses a virtual machine manager (VMM) called a hypervisor to provide abstracted hardware to multiple guest operating systems, which can then share the physical hardware resources more efficiently.
- **Types of Hardware Virtualization**
- Full virtualization
- Emulation virtualization
- Para-virtualization

Operating system virtualization

- the kernel enables the existence of various isolated user-space instances.
- It is installed over a pre-existing operating system and that operating system is called the host operating system.
- In this virtualization, a user installs the virtualization software in the operating system of his system like any other program and utilizes this application to operate and generate various virtual machines.
- Here, the virtualization software allows direct access to any of the created virtual machines to the user.
- As the host OS can provide hardware devices with the mandatory support, operating system virtualization may affect compatibility issues of hardware even when the hardware driver is not allocated to the virtualization software.

operating system-based services are :

- Backup and Recovery.
- Security Management.
- Integration to Directory Services.

Server Virtualization

- Server virtualization *means partitioning of a physical server into various virtual servers.*
- Each virtual private server can run independently.
- It is used to expand the server resources.
- Advantages of Server Virtualization
 - Independent Restart
 - Low Cost
 - Disaster Recovery<
 - Faster deployment of resources
 - Security
- Disadvantages of Server Virtualization
 - The biggest disadvantage of server virtualization is that when the server goes offline, all the websites that are hosted by the server will also go down.
 - There is no way to measure the performance of virtualized environments.
 - It requires a huge amount of RAM consumption.
 - It is difficult to set up and maintain.
 - Some core applications and databases are not supported virtualization.
 - It requires extra hardware resources.
 - Uses of Server Virtualization
- A list of uses of server virtualization is given below -
 - Server Virtualization is used in the testing and development environment.
 - It improves the availability of servers.
 - It allows organizations to make efficient use of resources.
 - It reduces redundancy without purchasing additional hardware components.

Software Virtualization

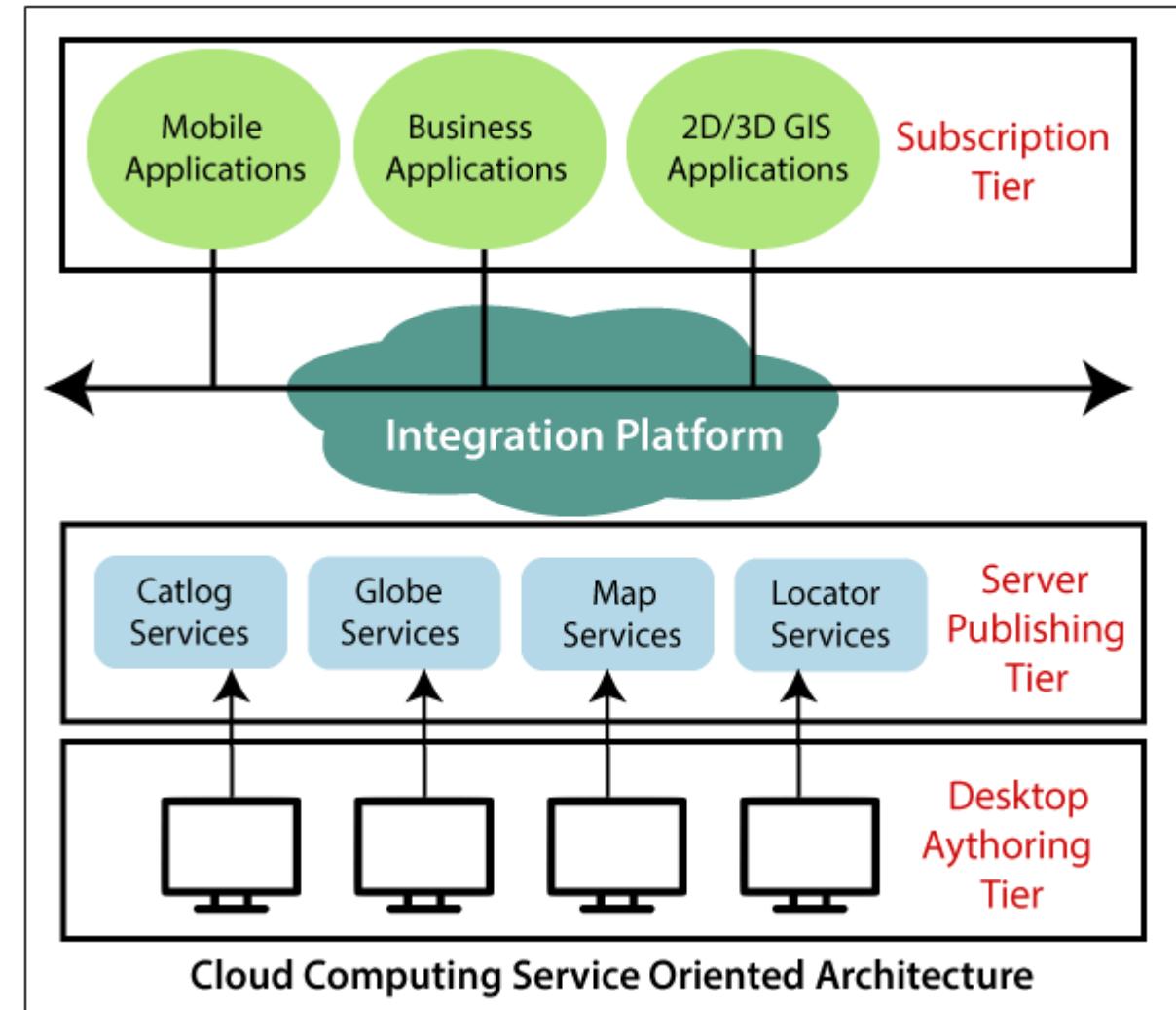
- **Software virtualization** is just like a virtualization but *able to abstract the software installation procedure and create virtual software installations.*
- **Virtualized software** is an application that will be "installed" into its own self-contained unit.
- Example of software virtualization is *VMware software, virtual box* etc.

Web 2.0

- Web 2.0 also **gave rise to web apps, self-publishing platforms like WordPress, as well as social media sites.**
- A Web 2.0 website **allows users to interact and collaborate with each other through social media dialogue as creators of user-generated content in a virtual community.**
- E.g. Wikipedia, Facebook, Twitter, and various blogs.

Service-Oriented Architecture (SOA)

- It supports low-cost, flexible, and evolvable applications.
- These were Quality of Service (QoS) which also includes the SLA (Service Level Agreement) and Software as a Service (SaaS).



Utility Computing

- It is a computing model that defines service provisioning techniques for services such as compute services along with other major services such as storage, infrastructure, etc which are provisioned on a pay-per-use basis.
- Utility computing is a **model in which computing resources are provided to the customer based on specific demand**. The service provider charges exactly for the services provided, instead of a flat rate.
- Examples of these IT services are **storage, computing power, and applications**.

Cloud Computing

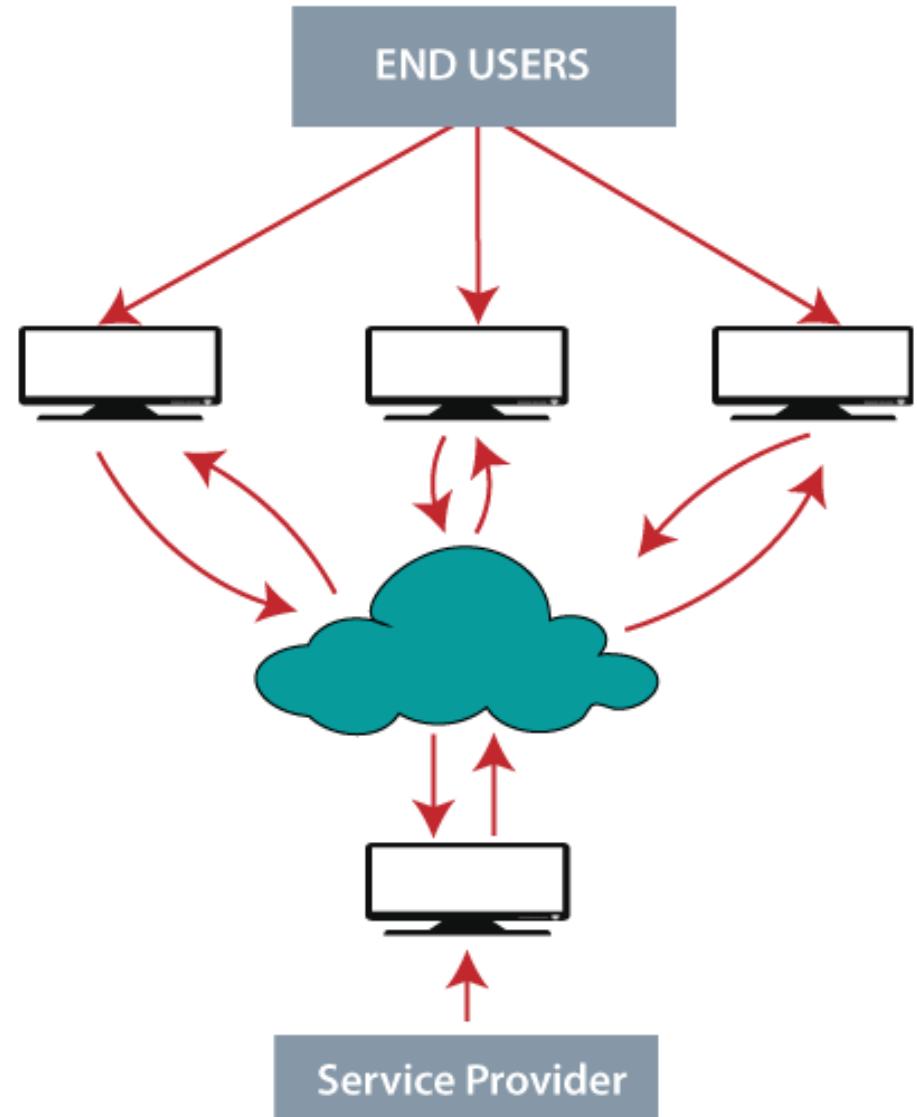


Why Cloud Computing?

- Cloud computing offers **platform independency**, as the software is not required to be installed locally on the PC.
- Hence, the Cloud Computing is making our business applications **mobile** and **collaborative**.

Why named Cloud?

- The term **Cloud** refers to a **Network or Internet**.
- In other words, we can say that Cloud is something, which is present at remote location.



Features of Cloud Computing

- The features of cloud computing are:

1.On-demand self-services:-

2.Broad network access:-

3.Resource pooling:-

4.Rapid Elasticity:-

5.Multi-sharing:-

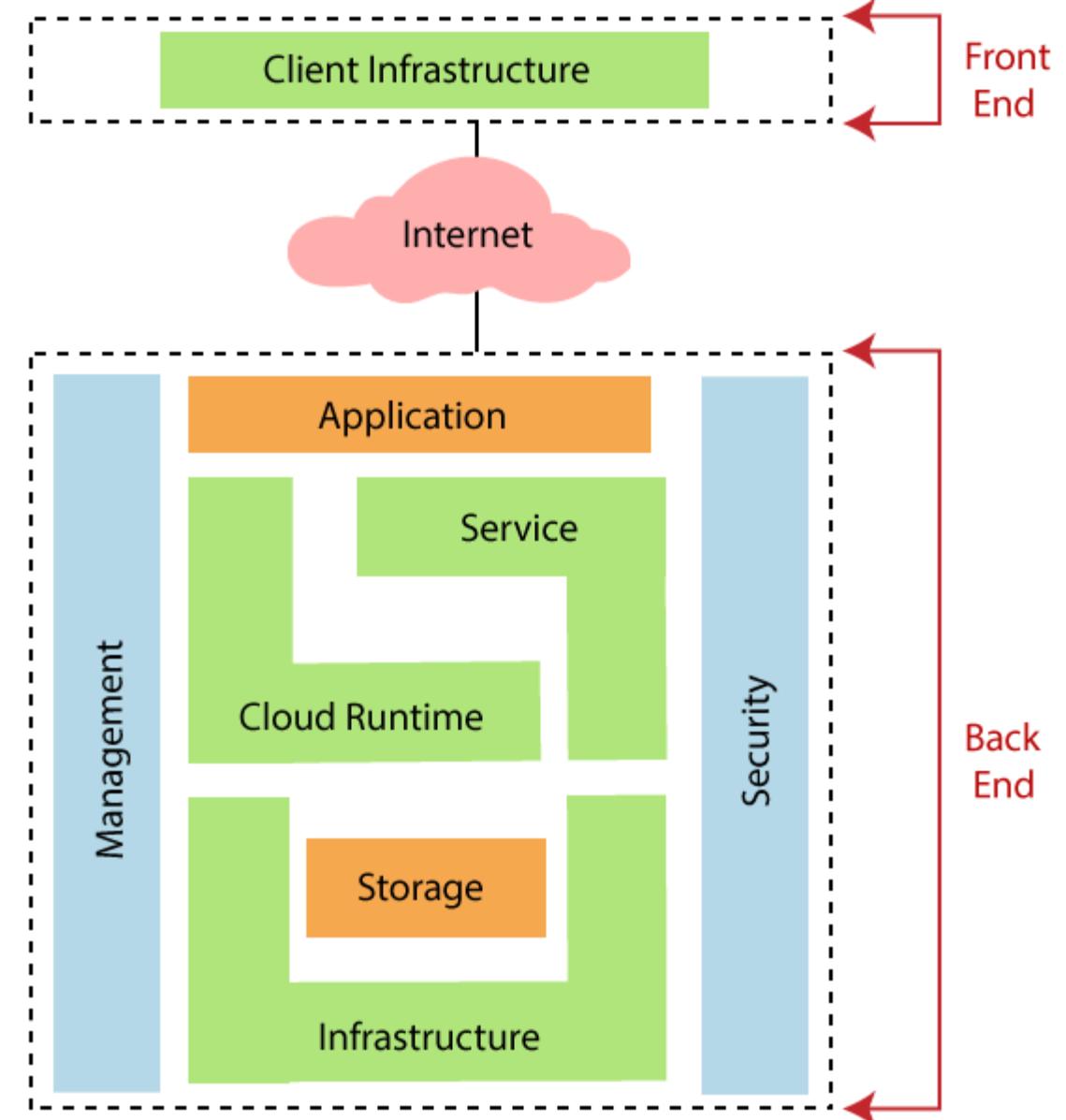
6.Maintenance:-

Benefits of Cloud Computing

- 1. Wide applications:**
- 2. Reduced upfront licensing cost:**
- 3. Lower cost of hardware:**
- 4. Allows working from any computer anywhere:**
- 5. Saves money:**
- 6. Drives sales:**

Architecture of Cloud Computing

- The architecture of cloud computing is broadly divided into two parts.
 1. Front end
 2. Back end



Components of Cloud Computing Architecture

1. Client infrastructure:

2. Application:

3. Service:

4. Runtime cloud:

5. Storage:

6. Infrastructure:

7. Management:

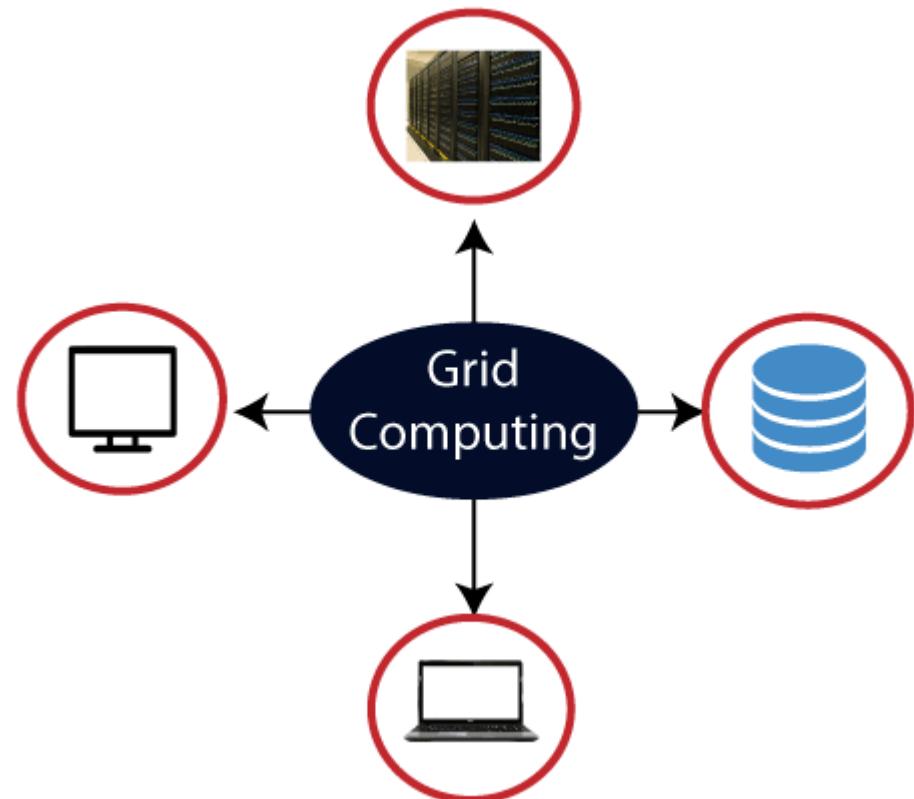
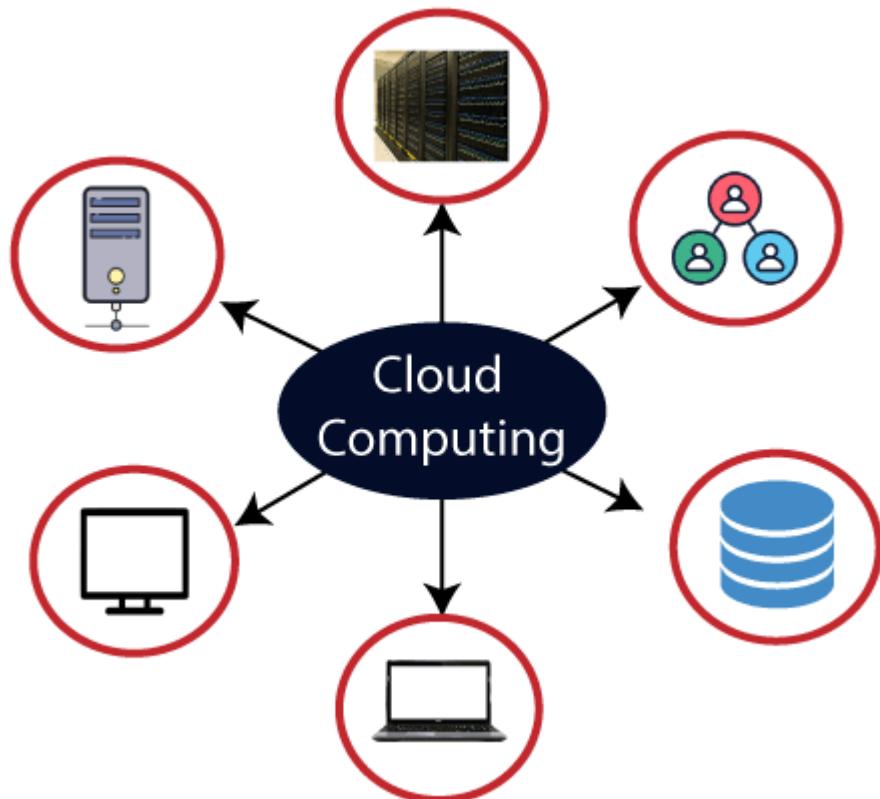
8. Security:

9. Internet:

Cloud Computing Technologies

- A list of cloud computing technologies are given below -
 - Virtualization
 - Service-Oriented Architecture (SOA)
 - Grid Computing
 - Utility Computing

Difference between Cloud Computing and Grid Computing

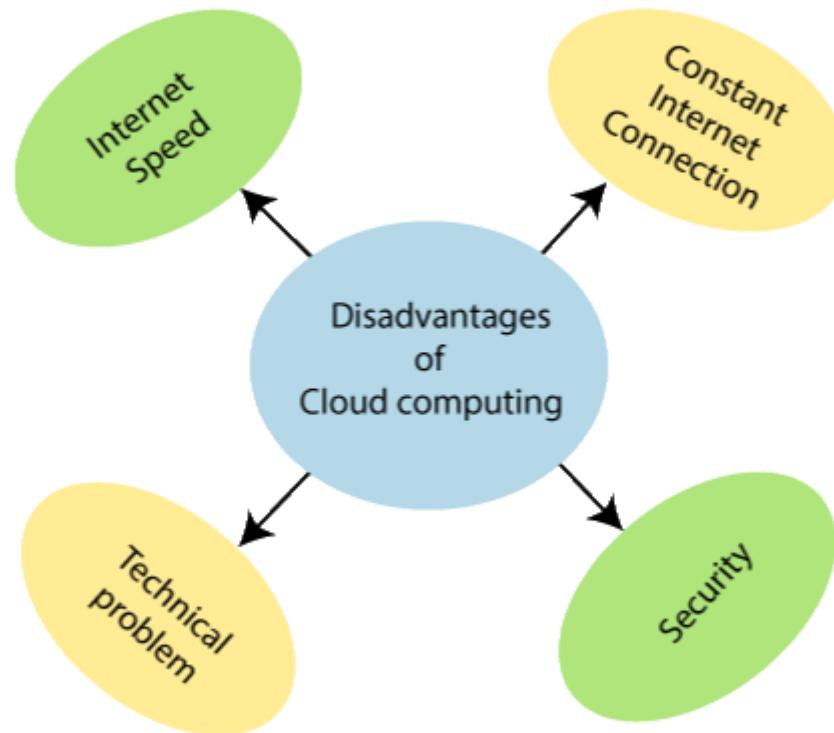


Cloud Computing	Grid Computing
Cloud Computing follows client-server computing architecture.	Grid computing follows a distributed computing architecture.
Scalability is high.	Scalability is normal.
Cloud Computing is more flexible than grid computing.	Grid Computing is less flexible than cloud computing.
Cloud operates as a centralized management system.	Grid operates as a decentralized management system.
In cloud computing, cloud servers are owned by infrastructure providers.	In Grid computing, grids are owned and managed by the organization.
Cloud computing uses services like IaaS, PaaS, and SaaS.	Grid computing uses systems like distributed computing, distributed information, and distributed pervasive.
Cloud Computing is Service-oriented.	Grid Computing is Application-oriented.
It is accessible through standard web protocols.	It is accessible through grid middleware.

How does Cloud Computing Works



Disadvantages of Cloud Computing



Cloud Deployment Models

1. public clouds
2. private/enterprise clouds
3. hybrid clouds

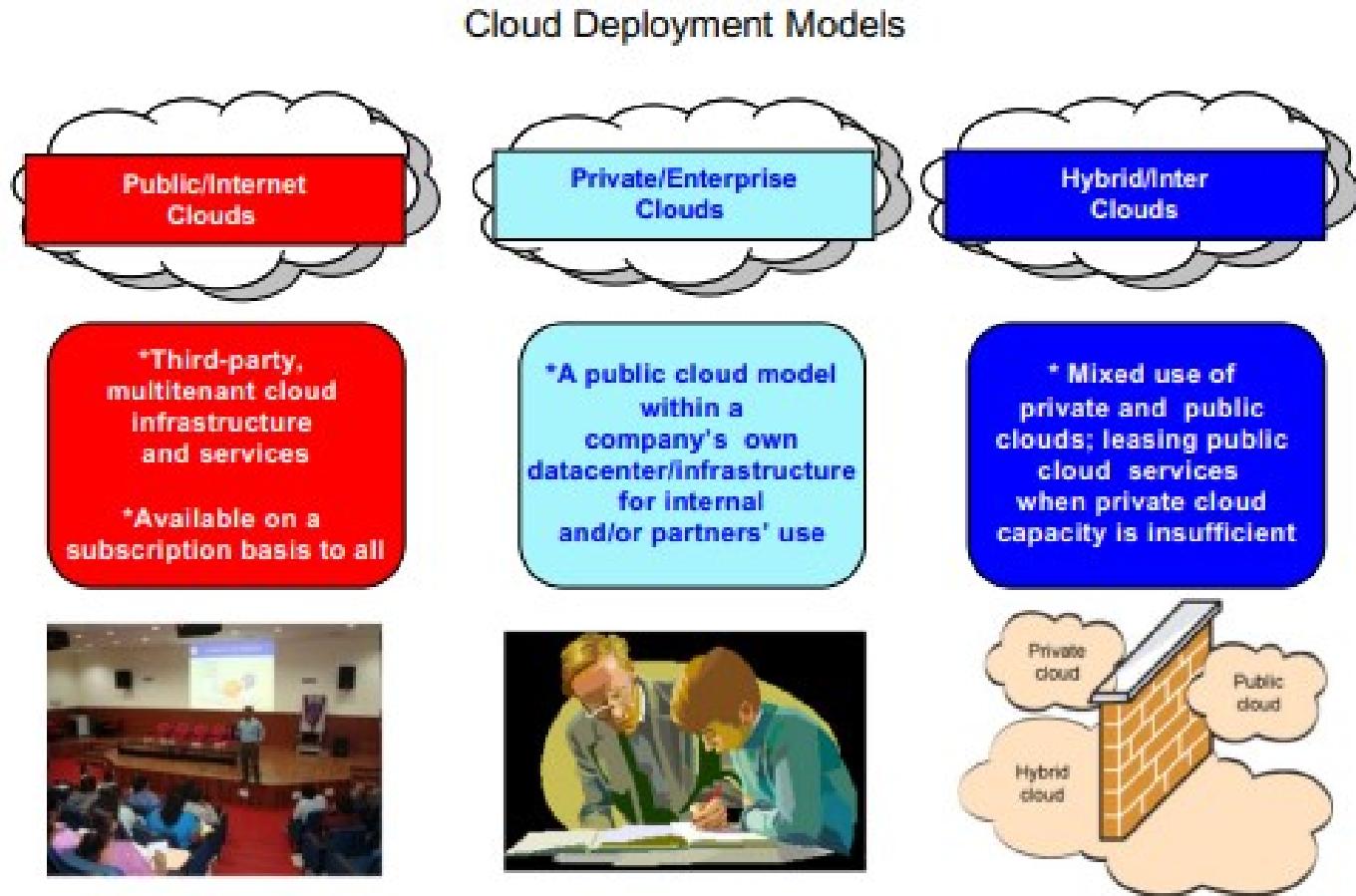


FIGURE 1.4

Major deployment models for cloud computing.

Public Cloud

- A public cloud is a [cloud service](#) offered to multiple customers by a cloud provider.
- The term "public cloud" is used to differentiate between the original cloud model of services accessed over the Internet and the private cloud model.
- Public clouds include [SaaS](#), [PaaS](#), and [IaaS](#) services.



When to use the public cloud

- The public cloud is most suitable for these types of environments:
- Predictable computing needs, such as communication services for a specific number of users
- Apps and services necessary to perform IT and business operations
- Additional resource requirements to address varying peak demands
- Software development and test environments

Advantages of public cloud

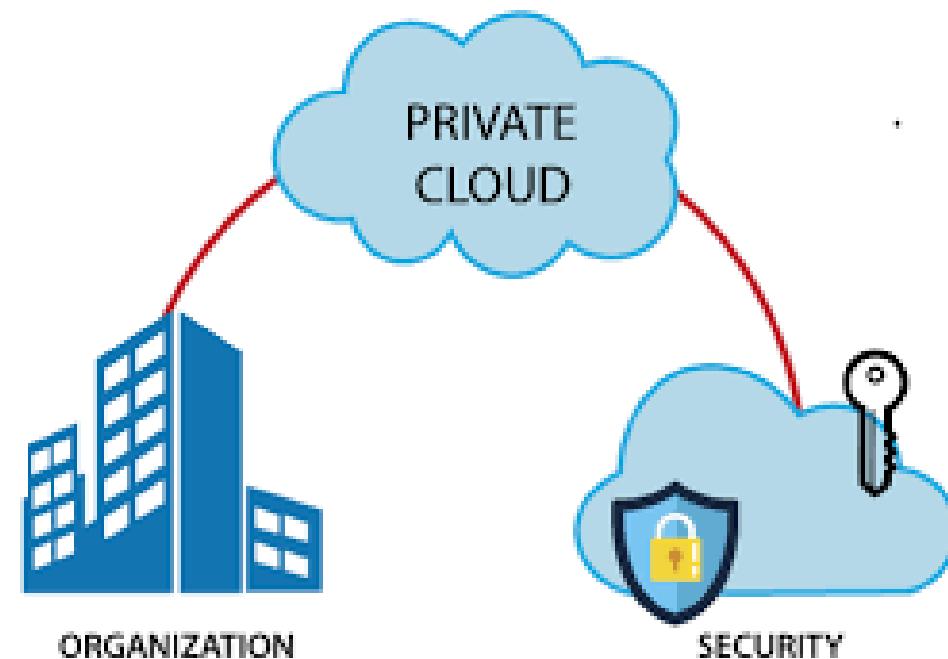
- People appreciate these public cloud benefits:
 1. **No CapEx.**
 2. **Technical agility.**
 3. **Business focus. .**
 4. **Affordability.**
 5. **Cost agility.**

Drawbacks of public cloud

- The public cloud does come with limitations:
 1. **Lack of cost control.**
 2. **Lack of security.**
 3. **Minimal technical control.**

What is the private cloud?

- Private cloud is a cloud computing environment dedicated to a single customer.
- It combines many of the benefits of cloud computing with the security and control of on-premises IT infrastructure.



When to use the private cloud

- The private cloud is best suited for:
 1. Highly regulated industries and government agencies
 2. Sensitive data
 3. Companies that require strong control and security over their IT workloads and the underlying infrastructure
 4. Large enterprises that require advanced data center technologies to operate efficiently and cost-effectively
 5. Organizations that can afford to invest in high performance and availability technologies

Advantages of private cloud

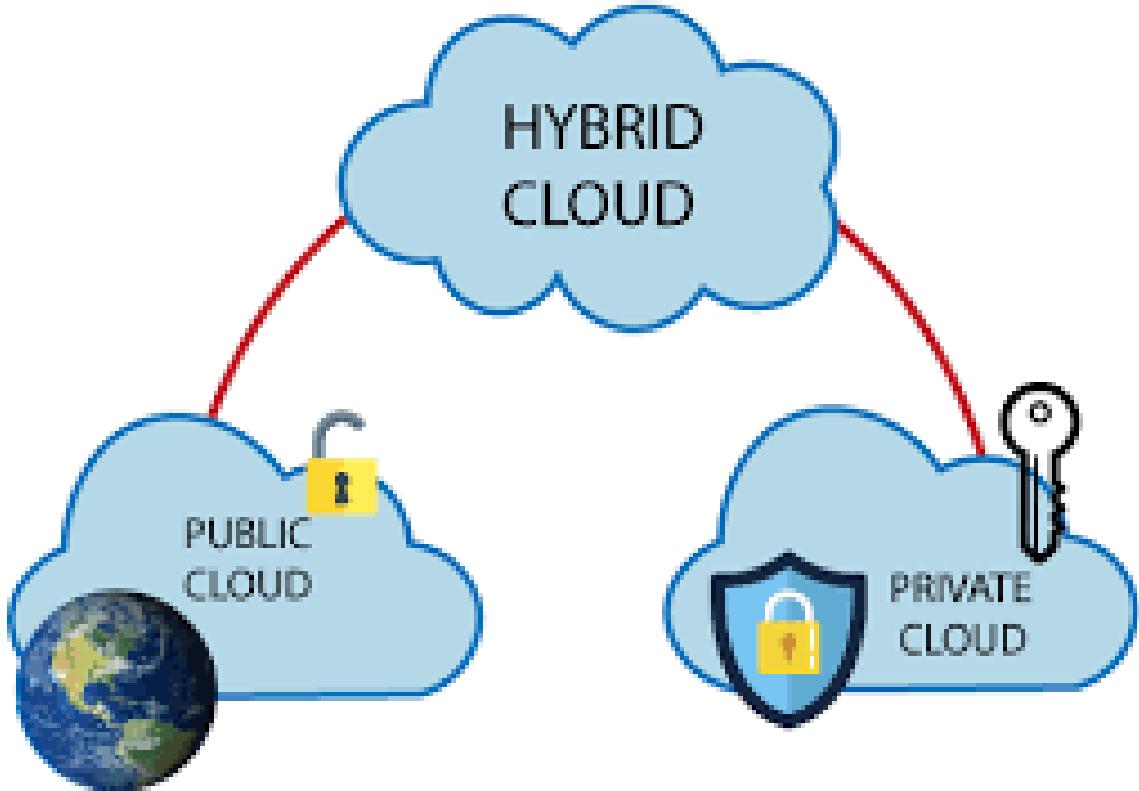
- The most popular benefits of private cloud include:
 1. Exclusive environments.
 2. Custom security.
 3. Scalability without tradeoffs.
 4. Efficient performance.
 5. Flexibility.

Drawbacks of private cloud

- The private cloud has drawbacks that might limit use cases:
 1. Price.
 2. Mobile difficulty.
 3. Scalability depends.

What is hybrid cloud?

- The hybrid cloud is any cloud infrastructure environment that combines both public and private cloud solutions.



When to use the hybrid cloud

- hybrid cloud might suit best:
 - Organizations serving multiple verticals facing different IT security, regulatory, and performance requirements
 - Optimizing cloud investments without compromising on the value that public or private cloud technologies can deliver
 - Improving security on existing cloud solutions such as SaaS offerings that must be delivered via secure private networks
 - Strategically approaching cloud investments to continuously switch and tradeoff between the best cloud service delivery model available in the market

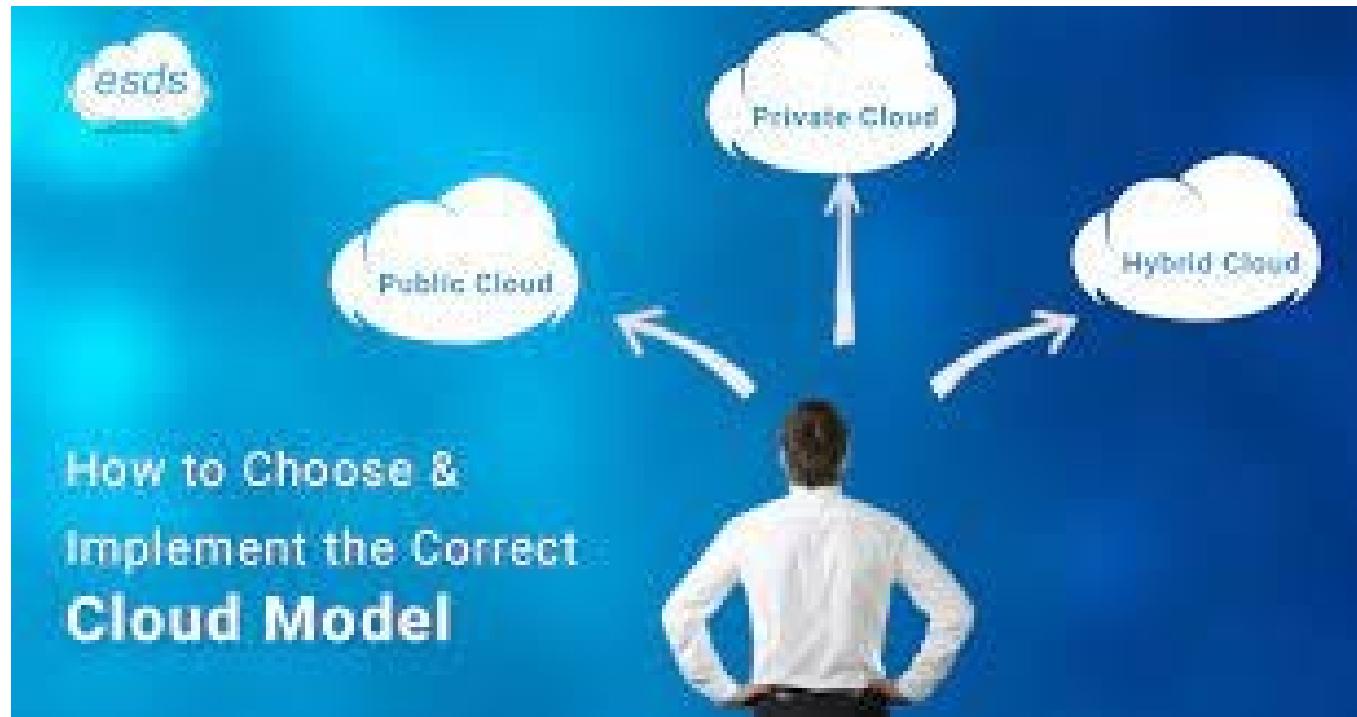
Advantages of hybrid cloud

- **Policy-driven option.**
- **Scale with security.**
- **Reliability.**
- **Cost control.**

Drawbacks of hybrid cloud

- Common drawbacks of the hybrid cloud include:
 - **Price.**
 - **Management.**
 - **Added complexity.**

Which cloud to choose?



The cloud computing reference model

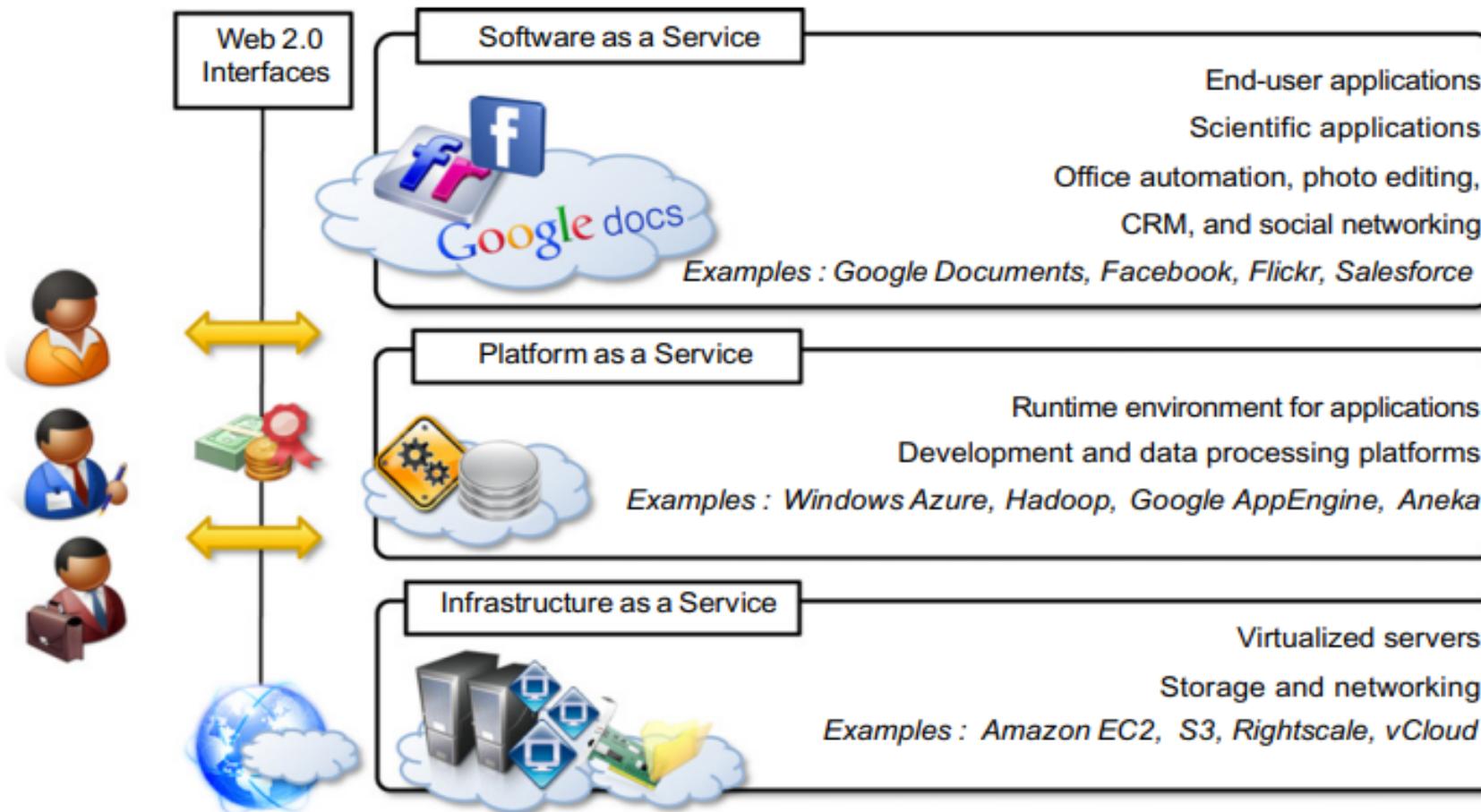


FIGURE 1.5

The Cloud Computing Reference Model.

Infrastructure-as-a-Service

Trends in computing

Virtualization in Cloud

Introduction & benefit of Virtualization, Implementation Levels of Virtualization, Types: Full and para virtualization Taxonomy of virtualization techniques - Execution Virtualization, Virtualization and cloud computing, Pros and cons of virtualization

Introduction & benefit of Virtualization

- Virtualization is a technique of how to separate a service from the underlying physical delivery of that service.
- it is a technique that allows multiple users or organizations to make use of a single resource thread or an application among themselves.
- It involves using specialized software to create a virtual or software-created version of a computing resource rather than the actual version of the same resource.
- 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 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.

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.

Virtualization gained interest

- Increased performance and computing capacity - PCs are having immense computing power.
- Underutilized hardware and software resources - Limited use of increased performance & computing capacity. –
- Greening initiatives
 - Reduce carbon footprints
 - Reducing the number of servers, reduce power consumption.
 - Lack of space - Continuous need for additional capacity.
- Rise of administrative costs - Power and cooling costs are higher than IT equipments.

Architecture of Virtual Machines

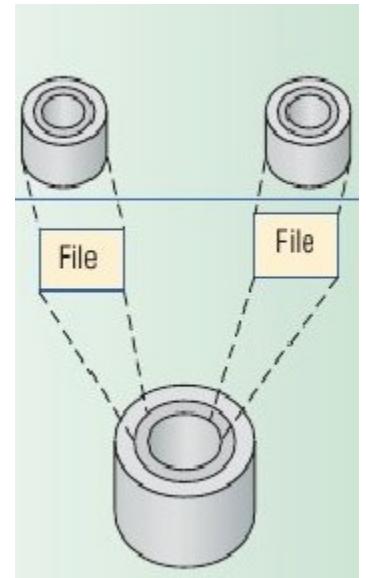
- VM can support individual processes or a complete system
- Virtualization can be from OS to programming languages to processor architecture.
- VMs enhance
 - Software interoperability (to work together)
 - System impregnability (having strength)
 - Platform versatility

Abstraction and Virtualization

- Computer system is complex, and yet it continue to evolve.
- Computer is designed as hierarchies of well-defined interfaces that separate level of abstraction
- Simplifying abstractions hide lower-level implementation details

Abstraction

- Ex. Disk storage
- Hides hard-disk addressing details (sectors and tracks)
- It appears to application software as a variable sized files.
- User can create, write and read files without knowing the underneath details.

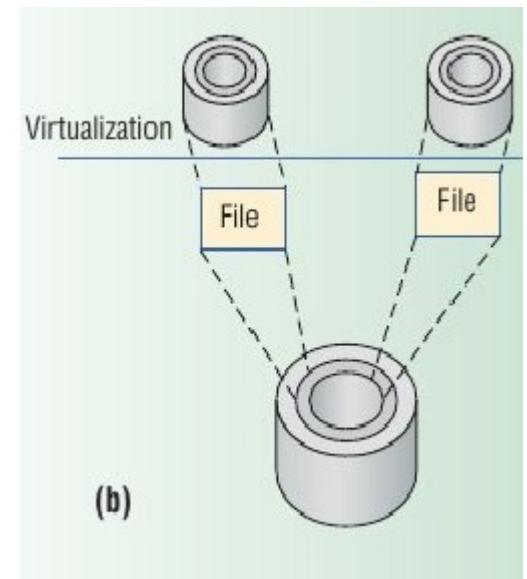


Pros and cons of Abstraction

- Well-defined interfaces permit development of interacting computer subsystems not only in different organization but also at different time.
- Limitation of well-defined interfaces , designed specification to one interface will not work for other.

Virtualization

- Virtualization of system or components like – processor, memory or an I/O device – at a given abstraction level.
- It transforms a entire system or components of the system
Ex. disk storage



Virtual Machine

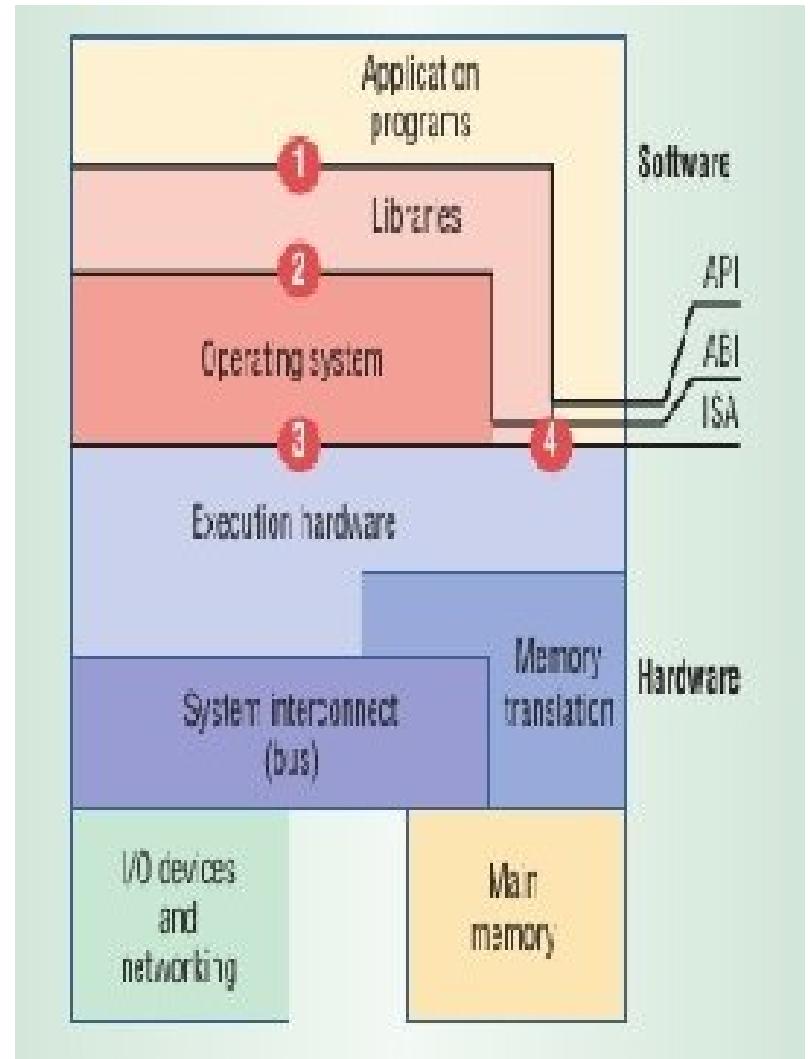
- Virtualization can be applied to entire machine.
- VM can be implemented by adding a software layer to a real machine to support desired architecture.
- VM implementation lie at architected interfaces

Architected Interfaces

- Architecture, as applied to computer systems, refer to a formal specification to an interface in the system, including the logical behavior of the resources managed via the interface.
- Implementation describes the actual embodiment of an architecture.
- Abstraction levels correspond to implementation layers, having its own interface or architecture.

Computer System Architecture

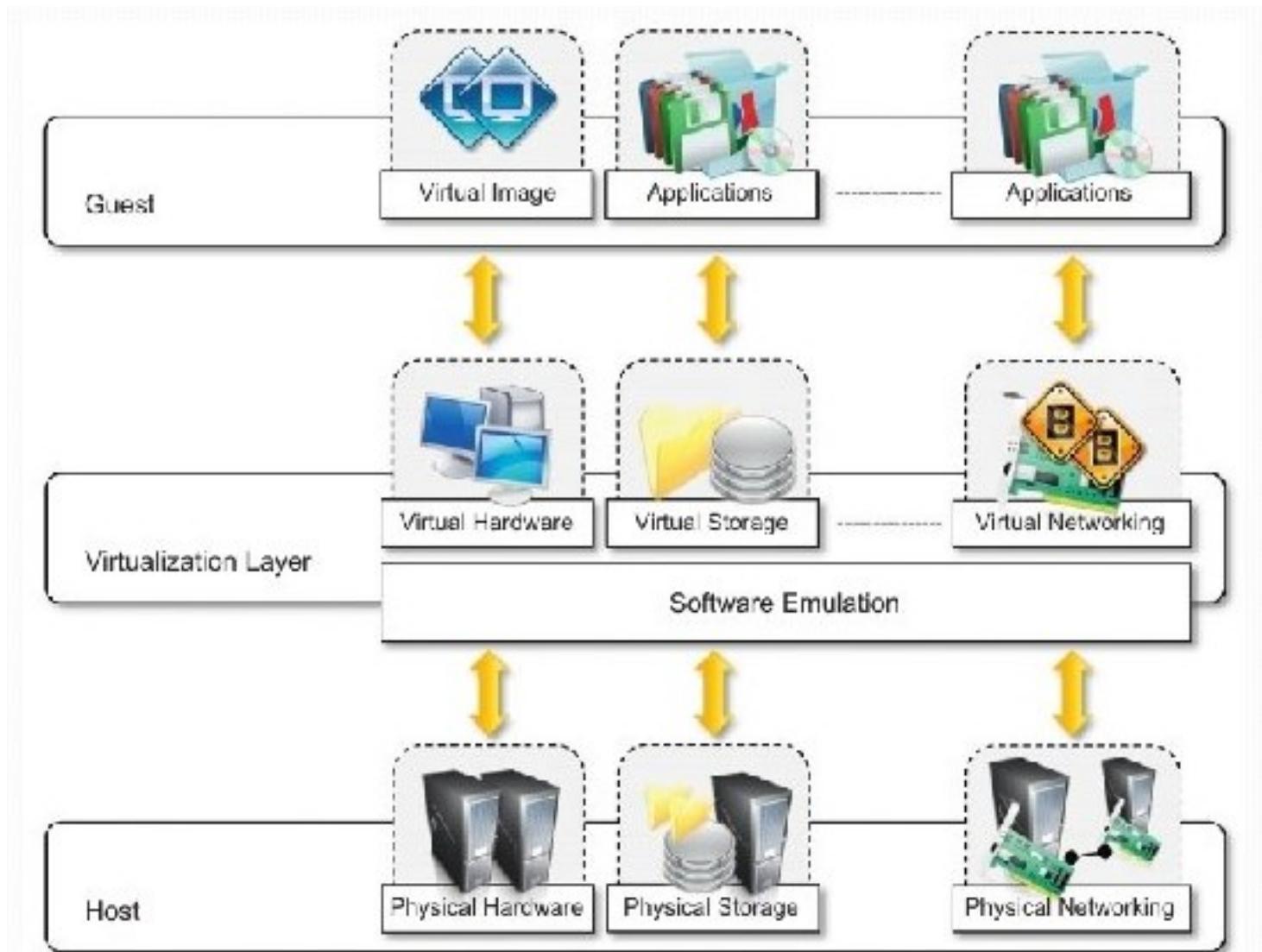
- Interfaces at or near the H/w S/w boundary :
- – ISA – Instruction Set Architecture.
- – API – Application Program Interface
- – ABI – Application Binary Interface



Virtualized Environments

- Three major components of Virtualized Environments
 - Guest – system component that interacts with Virtualization Layer.
 - Host – original environment where guest runs.
 - Virtualization Layer – recreate the same or different environment where guest will run.

Virtualization Reference Model



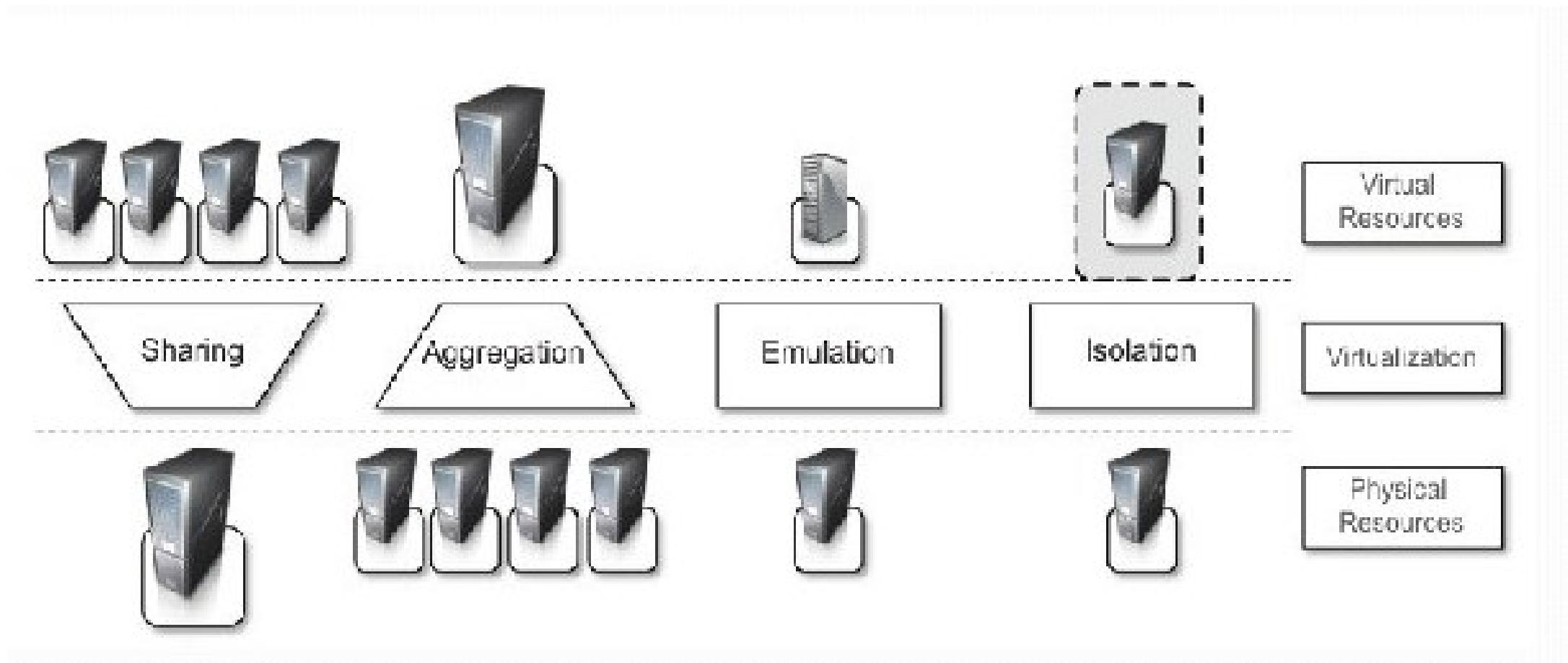
Advantages of Virtualization

- Increased Security
 - Ability to control the execution of a guest
 - Guest is executed in emulated environment.
 - Virtual Machine Manager control and filter the activity of the guest.
 - Hidding of resources.
 - Having no effect on other users/guest environment.

Advantages of Virtualization

- Managed Execution types :
 - Sharing - Creating separate computing environment within the same host.
Underline host is fully utilized.
 - Aggregation – A group of separate hosts can be tied together and represented as single virtual host.
 - Emulation –Controlling & Tuning the environment exposed to guest.
 - Isolation Complete separate environment for guests.

Managed Execution

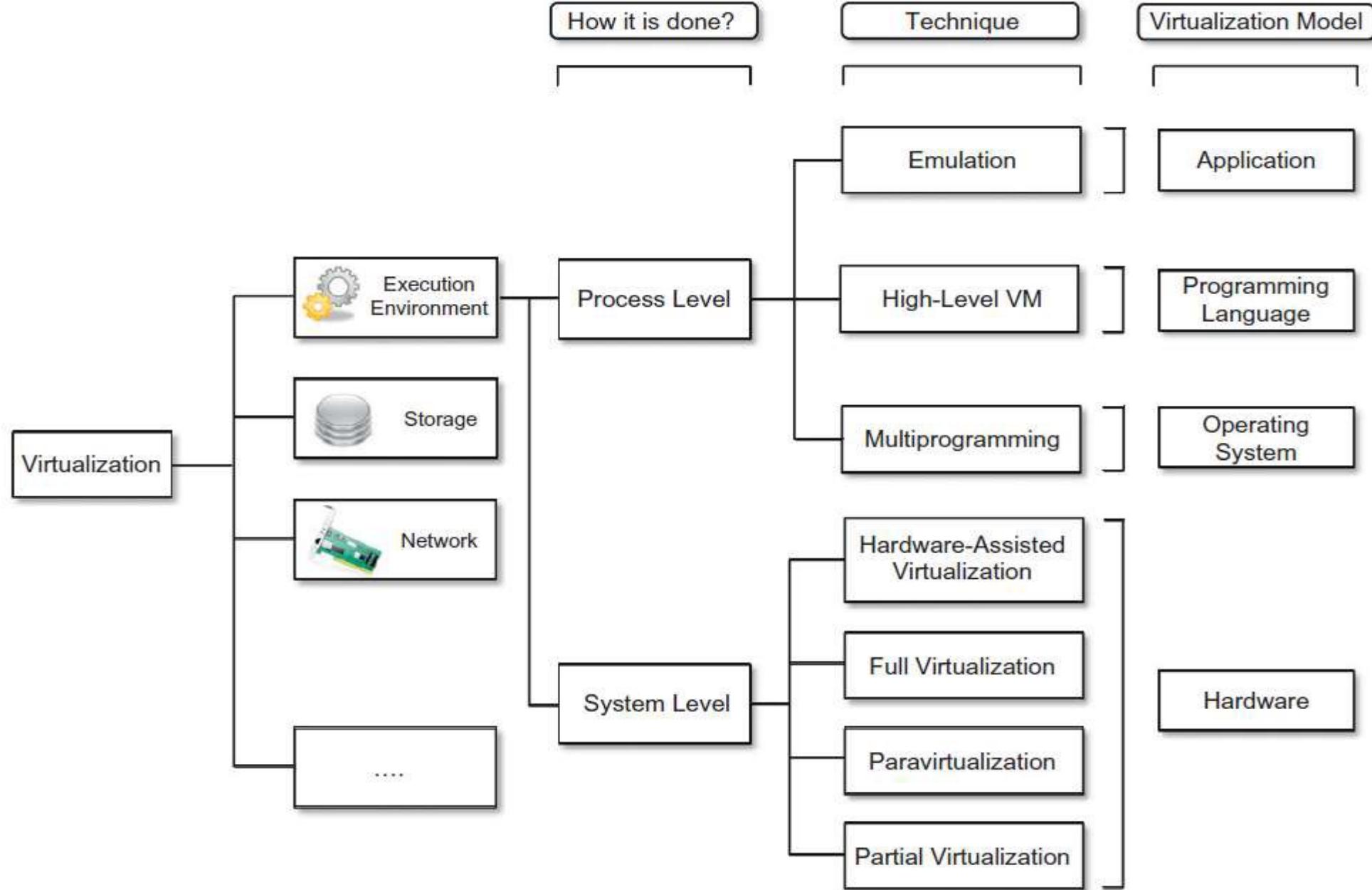


Advantages of Virtualization

- Performance Tuning – control the performance of guest.
- Virtual Machine Migration – move virtual image into another machine.
- Portability – safely moved and executed on top of different virtual machine. Availability of system is with you.

Taxonomy of Virtualization Techniques

- Virtualization is mainly used to emulate execution environment , storage and networks.
- Execution Environment classified into two :
 - Process-level – implemented on top of an existing operating system.
 - System-level – implemented directly on hardware and do not or minimum requirement of existing operating system

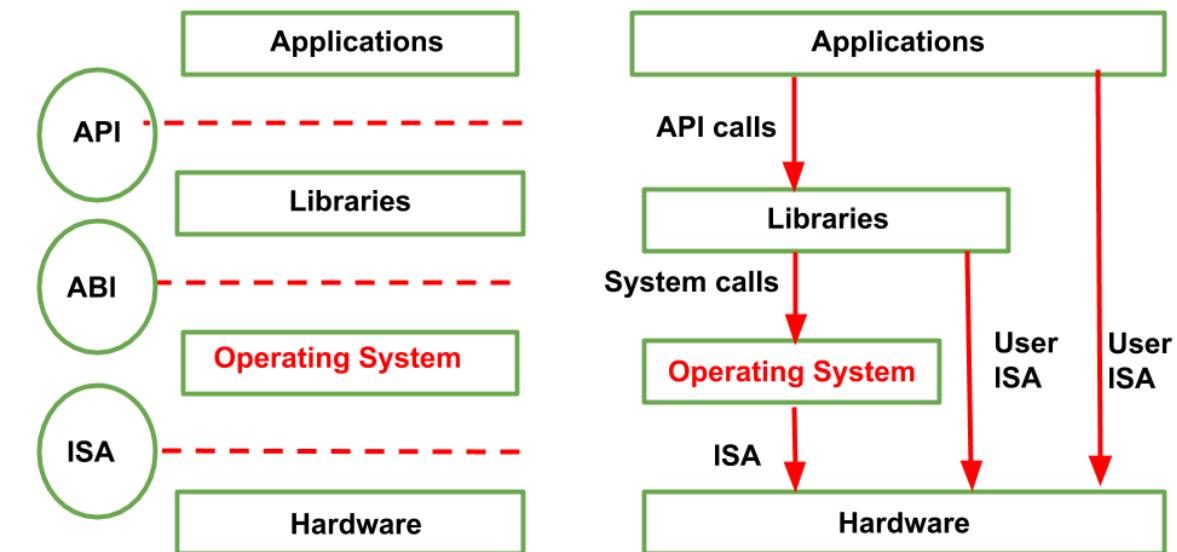


Machine Reference Model

- It defines the interfaces between the levels of abstractions, which hide implementation details.
- Virtualization techniques actually replace one of the layers and intercept the calls that are directed towards it.

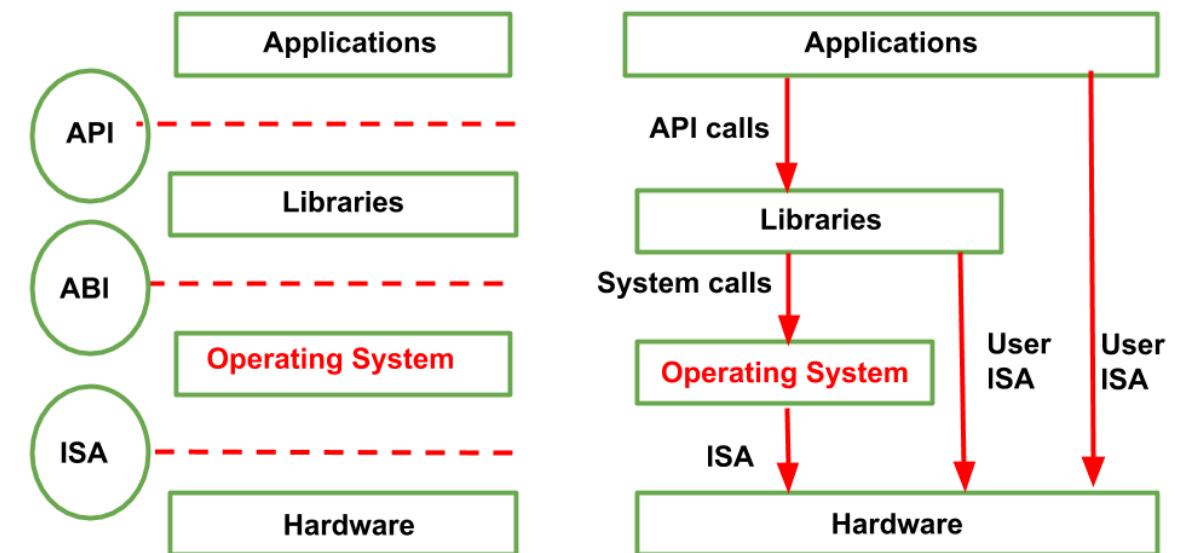
Machine Reference Model

- Hardware is expressed in terms of the Instruction Set Architecture (ISA).
- ISA for processor, registers, memory and the interrupt management.
- Application Binary Interface (ABI) separates the OS layer from the application and libraries which are managed by the OS.
 - System Calls defined
 - Allows portabilities of applications and libraries across OS.



Machine Reference Model

- API – it interfaces applications to libraries and/or the underlying OS.
- Layered approach simplifies the development and implementation of computing system.
- ISA has been divided into two security classes:
 - Privileged Instructions
 - Nonprivileged Instructions

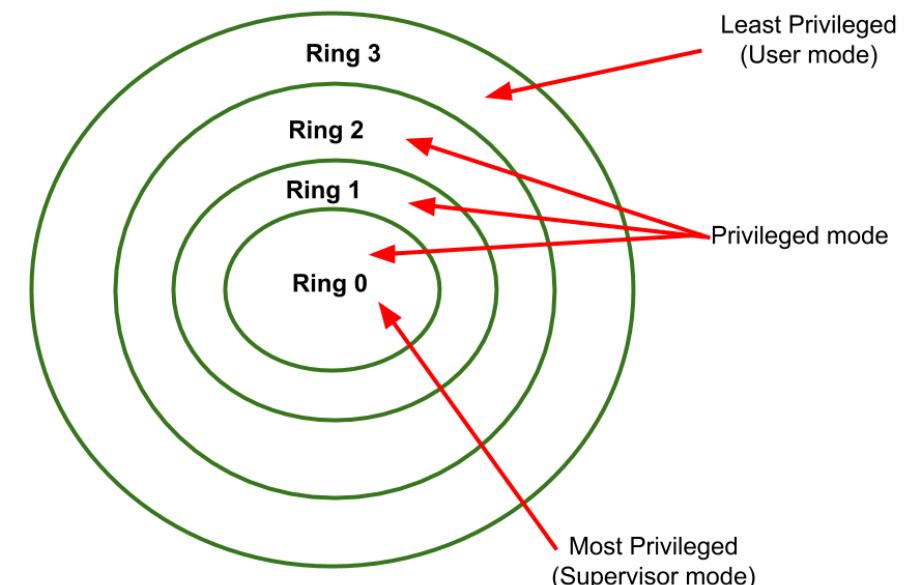


ISA: Security Classes

- Nonprivileged instructions –
 - That can be used without interfering with other tasks because they do not access shared resources. Ex. Arithmetic , floating & fixed point.
- Privileged instructions
 - That are executed under specific restrictions and are mostly used for sensitive operations, which expose (behavior-sensitive) or modify (control-sensitive) the privileged state.
 - Behavior-sensitive – operate on the I/O
 - Control-sensitive – alter the state of the CPU register.

Privileged Hierarchy: Security Ring

- Ring-0 is in most privileged level , used by the kernel.
- Ring-1 & 2 used by the OS-level services and ,
- R 3 in the least privileged level , used by the user.
- Recent system support two levels :
 - Ring 0 – supervisor mode
 - Ring 3 – user mode



Types of Virtualization

1. Application Virtualization.
2. Network Virtualization.
3. Desktop Virtualization.
4. Storage Virtualization.
5. Server Virtualization.
6. Data virtualization.
7. Hardware virtualization.

1. 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.
- Example of this would be a user who needs to run two different versions of the same software.
- Technologies that use application virtualization are hosted applications and packaged applications.

benefits of virtualized applications are:

- Simplified management
- Scalability
- Security
- Simple Installation
- Easy deployment
- Better Portability
- Simple to get rid of applications
- Reduced conflicts between applications
- Easier Rollback
- Improved Security
- Easier updates
- Simplified Support
- Independence from the Operating System

Drawbacks of Application Virtualization

- Graphics-intensive applications can slow down (lag) during the rendering process.
- A steady and reliable connection to the server is required to provide users with a solid UX with the applications.
- Some applications require device drivers and 16-bit applications running in the memory.
- Some applications, such as anti-virus programs, must integrate with the local OS, as they require continuous access to local data.
- The use of peripheral devices, like printers, can get more complicated with app virtualization.
- System monitoring tools can have trouble with virtualized applications, making it tricky to troubleshoot and isolate performance concerns.

Use Cases for Application Virtualization

- **Cost Control:** If you have a huge number of employees or end-users, then purchasing expensive PCs for everyone can turn out to be drastically expensive. Application virtualization comes to the rescue in such a situation as it allows you to deliver critical applications to any endpoint.
- **Application Mobility:** Enterprise applications should be accessible from any kind of mobile device for ease of use. Application virtualization offers application mobility by allowing applications to be delivered to any endpoint.
- **Secure Remote Access Capabilities:** Application virtualization allows employees to access critical applications from anywhere and that too in a secure manner. Application virtualization is useful for work-from-home scenarios that not only provide ease but also security.
- **Simplified Migrations:** Since application virtualization separates applications from the underlying operating system, there is no need to carry out extensive migrations from one kind of OS to the other.
- **Deploying In-House Applications :** Another important use case of application virtualization is the deployment of in-house applications which are updated frequently by developers. The updates, installation, and delivery of these applications are made remote and quick using application virtualization. Application virtualization is equally important for organizations that deploy in-house applications.

2. 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.
- **Tools for Network Virtualization :**
- **Physical switch OS** – It is where the OS must have the functionality of network virtualization.
- **Hypervisor** – It is which uses third-party software or built-in networking and the functionalities of network virtualization.
- The basic functionality of the OS is to give the application or the executing process with a simple set of instructions. System calls that are generated by the OS and executed through the libc library are comparable to the service primitives given at the interface between the application and the network through the SAP (Service Access Point).
- The hypervisor is used to create a virtual switch and configuring virtual networks on it. The third-party software is installed onto the hypervisor and it replaces the native networking functionality of the hypervisor. A hypervisor allows us to have various VMs all working optimally on a single piece of computer hardware.

Advantages of Network Virtualization

- Improves manageability
- Reduces CAPEX
- Improves utilization
- Enhances performance
- Enhances security

Disadvantages of Network Virtualization

- It needs to manage IT in the abstract.
- It needs to coexist with physical devices in a cloud-integrated hybrid environment.
- Increased complexity.
- Upfront cost.
- Possible learning curve.

3. 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.

Types of Desktop Virtualization

- Hosted Desktop Virtualization: In this desktop virtualization model, a computer server living in a data center basically hosts the virtual machines, enabling users to connect to the server via standard protocols like Remote Desktop Protocol (RDP) or connection brokers.
- Hosted desktop virtualization is executed in three main formats:

1. Virtual Desktop Infrastructure (VDI): Here, the operating system (OS) operates a virtual machine that contains desktop images on a server. Consequently, a hypervisor is employed to split the server into disparate desktop images that users can remotely access through their endpoint devices.

2. Remote Desktop Services (RDS)

This variation of hosted desktop virtualization provides users remote access via shared desktops and applications on Microsoft Windows Server OS.

3. Desktop-as-a-Service (DaaS)

This variation operates in a manner similar to that of VDI, as end-users can access their desktops and computer applications from any endpoint device or platform. However, the key difference with DaaS is that one has to purchase, deploy, and manage all the hardware components themselves.

- Client Virtualization
- This desktop virtualization model revolves around the installation of a hypervisor on a client device to run multiple operating systems, thus eliminating the need for users to maintain their own dedicated hardware and software.
- Principally, client virtualization deployment has two key variants:

1. Presentation virtualization

This avails a web-based portal that users can leverage to interact with desktops and apps.

2. Application virtualization

This client virtualization approach enables apps to run on other platforms. For instance, running Windows apps on the Linux OS.

Advantages of Desktop Virtualization

- 1. Flexibility**
- 2. Cost efficiency**
- 3. Enhanced security**
- 4. Environmentally Friendly**
- 5. Centralized management**
- 6. Disaster recovery**
- 7. Increase Employee Productivity and Onboarding**

Disadvantages of Desktop Virtualization

- Desktop Virtualization is **Cap-ex intensive**. One needs to buy the Desktop Virtualization Software/Licenses, Servers, Centralized Storage infrastructure, Upgrade Network infrastructure to support more bandwidth, etc in addition to buying computers/ thin-clients for each user.
- There is **no reduction** in the number of end-user client machines (computers) that are needed in the network.
- The **licenses** for Operating Systems, applications etc, still needs to be bought for each user (mostly) and there is no reduction of costs there.
- The **thin-clients** are sometimes as expensive/ more expensive than individual computers as, with huge volumes computer prices plummet drastically as they are manufactured and distributed in bulk quantities.
- The network infrastructure needs to handle all that **extra bandwidth** that Desktop Virtualization is going to introduce. Otherwise, it has to be upgraded. The WAN links need to have sufficient bandwidth to handle all those remote DV users, as well.
- If the bandwidth on the remote end is not sufficient/ if there is congestion in LAN, the **display quality** may not be as good (when images are streamed from server) as processing and viewing applications right from a desktop.
- Its difficult to handle **graphics/ high-definition video** with Desktop Virtualization. But there are some work-around methods that vendors follow to overcome this limitation (Including having local graphic acceleration cards, rendering graphical applications on the desktop, etc).
- Some vendor Desktop Virtualization solutions work only with their **Server Virtualization** counterparts, hence limiting the choices for the customers.
- There is a limit to the **number of Operating Systems** that can be supported by Desktop Virtualization products.

4. 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 software maintains smooth operations, consistent performance and a continuous suite of advanced functions despite changes, break down and differences in the underlying equipment.

Advantages of Storage Virtualization

- It allows for migrations to be performed quickly.
- It creates better workflows.
- It allows more than one type of storage array.
- It is a cheaper option for storage.
- It allows for the costs to become more predictable.
- It provides better access to your data.
- It allows anyone to create business opportunities for themselves.

Disadvantages of Storage Virtualization

- It requires you to deal with multiple vendors.
- It can make upgrades challenging to process.
- It does not always scale to some areas.
- It still has limitations which must be considered.
- It does not eliminate data security risks.
- It may create availability issues.
- It creates more links for data access instead of less

5. Server Virtualization

- 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.

Benefits of Server Virtualization

- Higher server ability
- Cheaper operating costs
- Eliminate server complexity
- Increased application performance
- Deploy workload quicker

Three Kinds of Server Virtualization

- **Full Virtualization:** Full virtualization uses a hypervisor, a type of software that directly communicates with a physical server's disk space and CPU. The hypervisor monitors the physical server's resources and keeps each virtual server independent and unaware of the other virtual servers. It also relays resources from the physical server to the correct virtual server as it runs applications. The biggest limitation of using full virtualization is that a hypervisor has its own processing needs. This can slow down applications and impact server performance.
- **Para-Virtualization:** Unlike full virtualization, para-virtualization involves the entire network working together as a cohesive unit. Since each operating system on the virtual servers is aware of one another in para-virtualization, the hypervisor does not need to use as much processing power to manage the operating systems.
- **OS-Level Virtualization:** Unlike full and para-virtualization, OS-level visualization does not use a hypervisor. Instead, the virtualization capability, which is part of the physical server operating system, performs all the tasks of a hypervisor. However, all the virtual servers must run that same operating system in this server virtualization method.

Advantage of Server Virtualization

- Cost reduction: One of the most significant benefits of server virtualization is cost reduction. This cost reduction can be realized in four distinct areas:
 - **1. Hardware savings.** Without the need to purchase or upgrade costly server hardware, companies can reallocate their funds back into growing their business.
 - **2. Operational savings.** With a large server array, it is necessary to employ dedicated technicians to maintain it all. Salaries for full-time tech staff are a considerable drain on spending for many companies, but with virtualization, these resources can be reduced to part-time or even channeled into less-costly managed services.
 - **3. Energy savings.** An extensive collection of physical servers is a massive drain on energy resources. Not only do they need to be kept running, they need to be kept cool—meaning there will be extra energy involved to maintain climate stability in the server room. Virtualization eliminates this need, potentially reducing your energy consumption by a significant degree.
 - **4. Real estate savings.**
 - Virtualization promotes automation
 - Virtualization makes backup and recovery more efficient and reliable
 - Cyberattacks, power surges, and terrestrial disasters
 - Server redundancies reduce downtime to a minimum and prevent outages
 - Virtualization supports scale

Disadvantages of server virtualization

- The cost of entry can be prohibitive
- Virtualization, like any other technological initiative, is **pay-to-play**. For instance, the physical servers that can be virtualized cost more than their traditional counterparts.
- You will also bear the cost of software licensing. Fortunately, the cost savings you will see as a result of virtualization should balance that cost out in the end.
- Not all applications can be virtualized
- You may still need to maintain a hybrid system to ensure all of your applications keep working as they should. Today, **most applications support virtualization**, but if you are running proprietary software, you may want to look at its capabilities before moving forward.
- Security risks
- Data security is one of the **most significant issues** we face today. Virtualization carries an added security risk, so additional spending will be required to ensure **data safety and integrity**. This can extend to insurance, security hardware, software, and a more robust monitoring policy.

6. 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, At scale, Cdata, etc.

Advantages of data virtualization -

- It allows users to access the data without worrying about where it resides on the memory.
- It offers better customer satisfaction, retention, and revenue growth.
- It provides various security mechanism that allows users to safely store their personal and professional information.
- It reduces costs by removing data replication.
- It provides a user-friendly interface to develop customized views.
- It provides various simple and fast deployment resources.
- It increases business user efficiency by providing data in real-time.
- It is used to perform tasks such as data integration, business integration, Service-Oriented Architecture (SOA) data services, and enterprise search

Disadvantages of Data Virtualization

- It creates availability issues, because availability is maintained by third-party providers.
- It required a high implementation cost.
- It creates the availability and scalability issues.
- Although it saves time during the implementation phase of virtualization but it consumes more time to generate the appropriate result.

Use of data virtualization

- Analyze performance: Data virtualization is used to analyze the performance of the organization compared to previous years.
- Search and discover interrelated data: Data Virtualization (DV) provides a mechanism to easily search the data which is similar and internally related to each other.
- 3. Agile Business Intelligence
- It is one of the most common uses of Data Virtualization. It is used in agile reporting, real-time dashboards that require timely aggregation, analyze and present the relevant data from multiple resources. Both individuals and managers use this to monitor performance, which helps to make daily operational decision processes such as sales, support, finance, logistics, legal, and compliance.
- 4. Data Management
- Data virtualization provides a secure centralized layer to search, discover, and govern the unified data and its relationships.

Industries that use Data Virtualization

- **Communication & Technology**

In Communication & Technology industry, data virtualization is used to increase revenue per customer, create a real-time ODS for marketing, manage customers, improve customer insights, and optimize customer care, etc.

- **Finance**

In the field of finance, DV is used to improve trade reconciliation, empowering data democracy, addressing data complexity, and managing fixed-risk income.

- **Government**

In the government sector, DV is used for protecting the environment.

- **Healthcare**

Data virtualization plays a very important role in the field of healthcare. In healthcare, DV helps to improve patient care, drive new product innovation, accelerating M&A synergies, and provide a more efficient claims analysis.

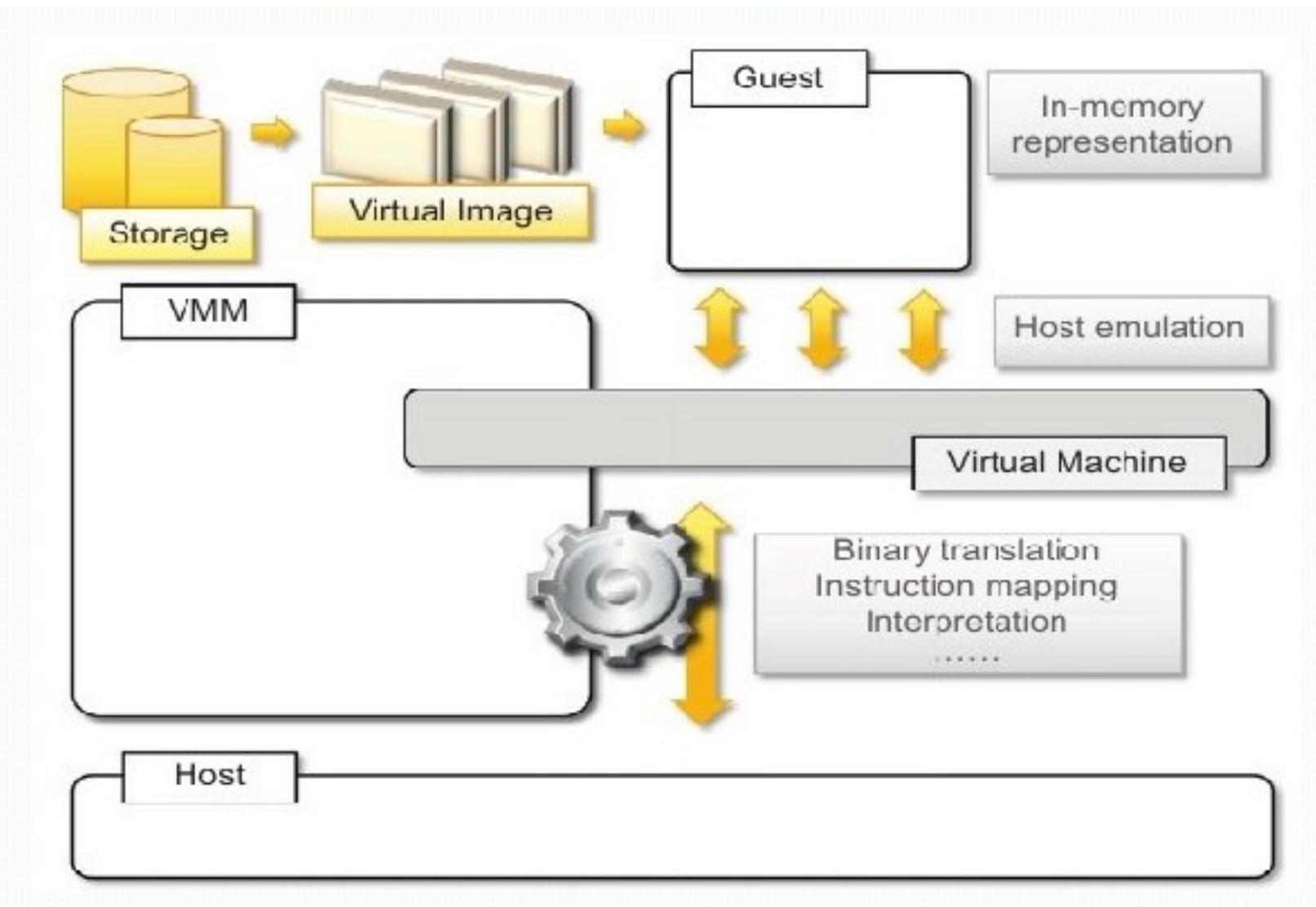
- **Manufacturing**

In manufacturing industry, data virtualization is used to optimize a global supply chain, optimize factories, and improve IT assets utilization.

7. Hardware Virtualization

- In Cloud Computing, hardware virtualization is used in server platforms since it offers more flexibility as opposed to physical machines.
- When it comes to hardware virtualization, VM software gets installed within the hardware system, known as hardware virtualization.
- It also comprises a hypervisor that controls and monitors the process, hardware resources and memory of the system.
- After the completion of the hardware virtualization process, the concerned user can install a different OS in it and different applications can be used simultaneously.

7. Hardware Virtualization



Advantages of Hardware Virtualization

- It reduces the maintenance overhead of para-virtualization as it reduces (ideally, eliminates) the modification in the guest operating system.
- It is also significantly convenient to attain enhanced performance.
- A practical benefit of hardware-based virtualization has been mentioned by VMware engineers and Virtual Iron.

Disadvantages of hardware-based virtualization

Hardware-based virtualization requires explicit support in the host CPU, which may not be available on all x86/x86_64 processors.

A “pure” hardware-based virtualization approach, including the entire unmodified guest operating system, involves many VM traps, and thus a rapid increase in CPU overhead occurs which limits the scalability and efficiency of server consolidation.

This performance hit can be mitigated by the use of para-virtualized drivers; the combination has been called “hybrid virtualization”.

Why should virtualization be even considered?

- Machines can run multiple instances of a single application simultaneously.
- Streamlining the IT cost and minimizing the IT administration structure.
- Effective management of large scale installation and server farms.
- Enhanced reliability, security, scalability, device dependence.
- Options to enable multiple OS use on the same hardware.

Terms that are associated with virtualization

- **Hypervisor** : It is the OS that runs on actual hardware and the Virtual counterpart is a part of this OS as a running process. Hypervisors are often seen as Domain 0 or Dom0.
- **Virtual Machine (VM)**: It is a virtual computer that runs under a hypervisor.
- **Container** : These are light-weight VMs that are part of the same OS instance as its hypervisor. So, containers are nothing but a group of processes that are running with their respective namespace for process identifiers.
- **Virtualization Software** : It is a software that aids in implementing virtualization on any computer. It can either be a part of a software application package or an OS or a special variant of that OS.
- **Virtual Network** : Virtual Network is a logically separate network within servers that can be extended to other servers or across multiple servers.

Implementation Levels of Virtualization

- **1.) 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. With ISA, it is possible to make the virtual machine hardware agnostic.
- 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.

Implementation Levels of Virtualization

- **2.) 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.

Implementation Levels of Virtualization

- **3.) 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.

Implementation Levels of Virtualization

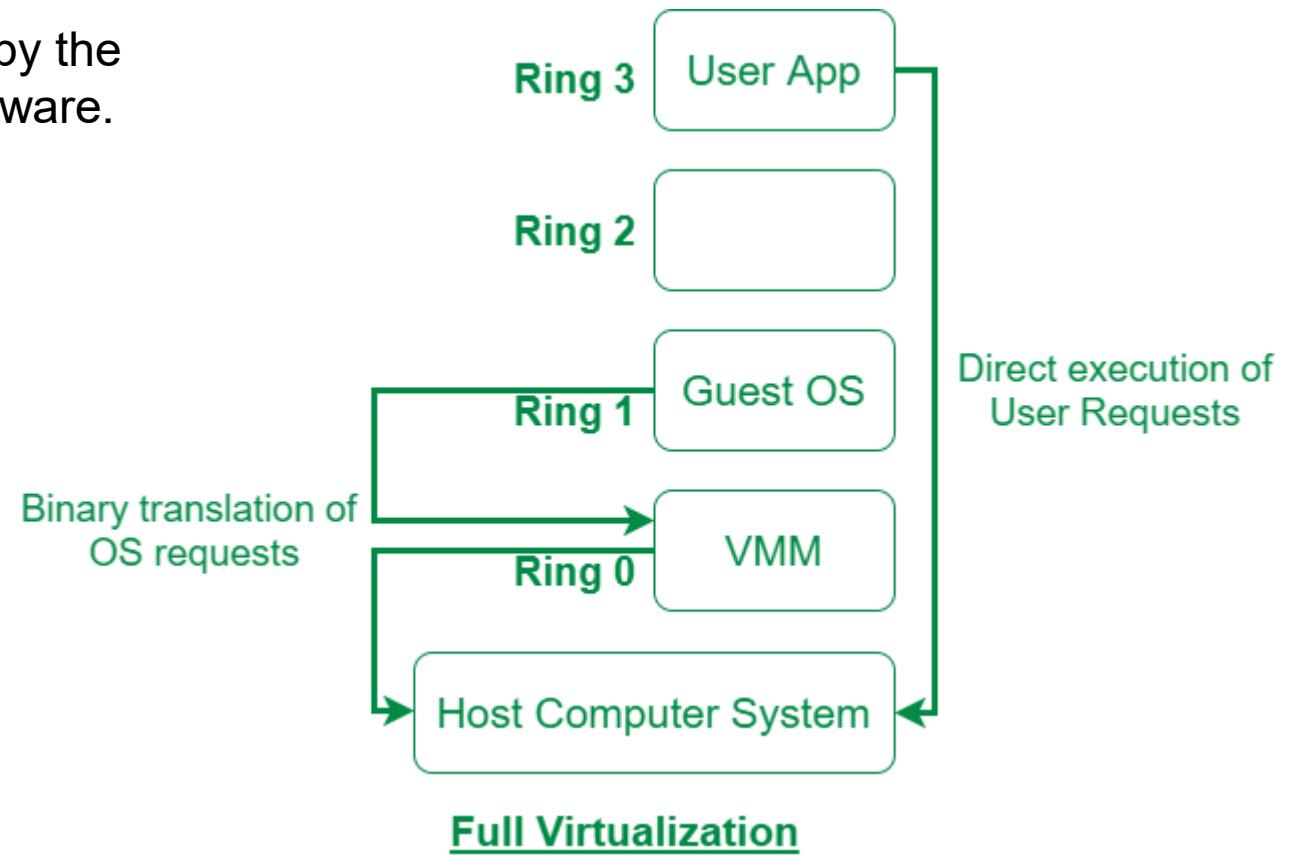
- **4.) Library Level**
- The operating system is cumbersome, and this is when the applications make use of the API that is from the libraries at a user level. These APIs are documented well, and this is why the library virtualization level is preferred in these scenarios. API hooks make it possible as it controls the link of communication from the application to the system.

Implementation Levels of Virtualization

- **5.) 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.

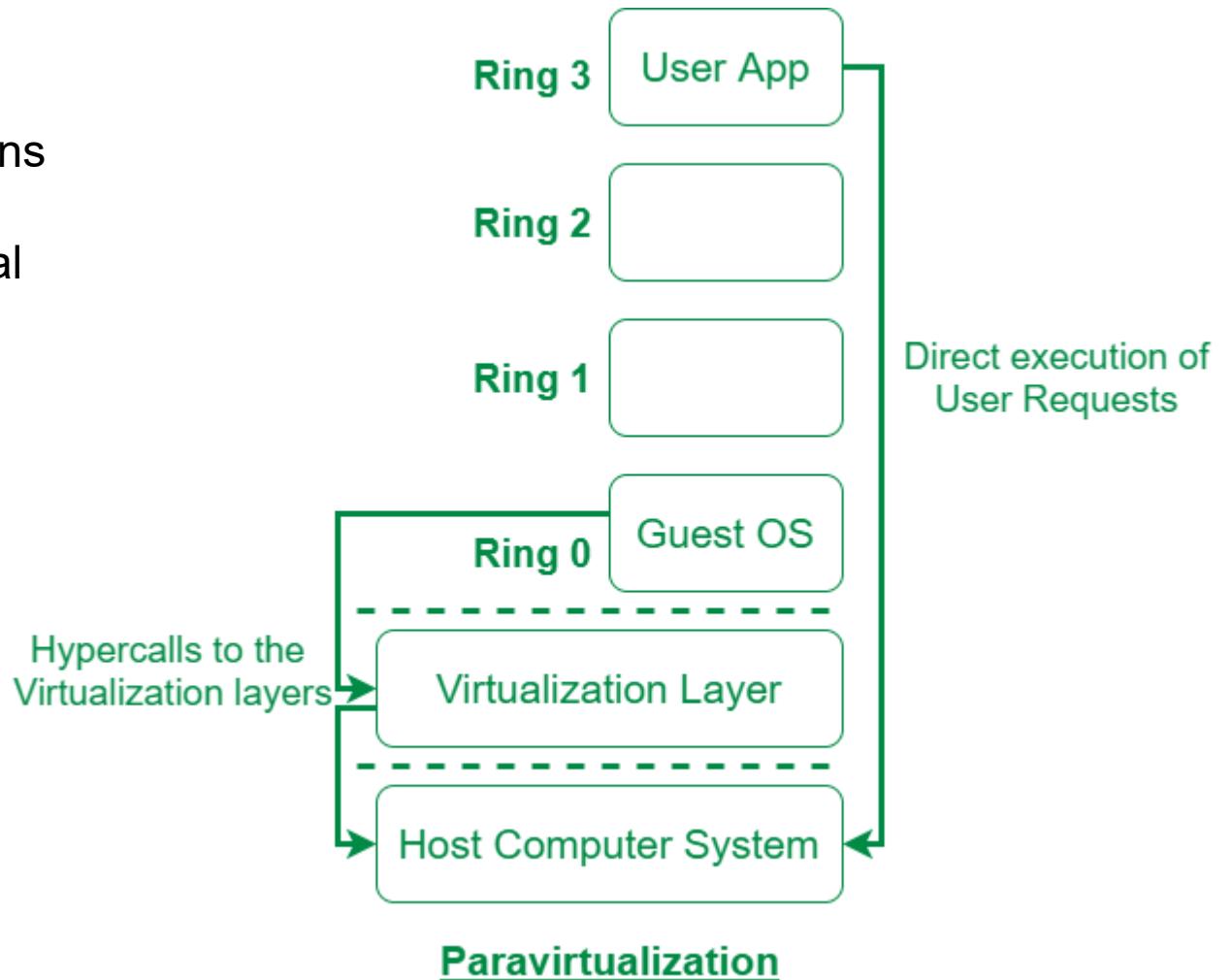
Type of Virtualization: Full Virtualization

In full virtualization, guest OS is completely isolated by the virtual machine from the virtualization layer and hardware. Microsoft and Parallels systems are examples of full virtualization.



Type of Virtualization: Para Virtualization

Paravirtualization is the category of CPU virtualization which uses hypercalls for operations to handle instructions at compile time. In paravirtualization, guest OS is not completely isolated but it is partially isolated by the virtual machine from the virtualization layer and hardware. VMware and Xen are some examples of paravirtualization.



Difference between Full and Para Virtualization

S.No.	Full Virtualization	Paravirtualization
1.	In Full virtualization, virtual machines permit the execution of the instructions with the running of unmodified OS in an entirely isolated way.	In paravirtualization, a virtual machine does not implement full isolation of OS but rather provides a different API which is utilized when OS is subjected to alteration.
2.	Full Virtualization is less secure.	While the Paravirtualization is more secure than the Full Virtualization.
3.	Full Virtualization uses binary translation and a direct approach as a technique for operations.	While Paravirtualization uses hypercalls at compile time for operations.
4.	Full Virtualization is slow than paravirtualization in operation.	Paravirtualization is faster in operation as compared to full virtualization.
5.	Full Virtualization is more portable and compatible.	Paravirtualization is less portable and compatible.
6.	Examples of full virtualization are Microsoft and Parallels systems.	Examples of paravirtualization are Microsoft Hyper-V, Citrix Xen, etc.
7.	It supports all guest operating systems without modification.	The guest operating system has to be modified and only a few operating systems support it.
8.	The guest operating system will issue hardware calls.	Using the drivers, the guest operating system will directly communicate with the hypervisor.
9.	It is less streamlined compared to para-virtualization.	It is more streamlined.
10.	It provides the best isolation.	It provides less isolation compared to full virtualization.

Pros of Virtualization in Cloud Computing

- **Utilization of Hardware Efficiently –**

With the help of Virtualization Hardware is Efficiently used by user as well as Cloud Service Provider. In this the need of Physical Hardware System for the User is decreases and this results in less costly. In Service Provider point of View, they will vitalize the Hardware using Hardware Virtualization which decrease the Hardware requirement from Vendor side which are provided to User is decreased. Before Virtualization, Companies and organizations have to set up their own Server which require extra space for placing them, engineer's to check its performance and require extra hardware cost but with the help of Virtualization the all these limitations are removed by Cloud vendor's who provide Physical Services without setting up any Physical Hardware system.

- **Availability increases with Virtualization –**

One of the main benefit of Virtualization is that it provides advance features which allow virtual instances to be available all the times. It also has capability to move virtual instance from one virtual Server another Server which is very tedious and risky task in Server Based System. During migration of Data from one server to another it ensures its safety. Also, we can access information from any location and any time from any device.

- **Disaster Recovery is efficient and easy –**

With the help of virtualization Data Recovery, Backup, Duplication becomes very easy. In traditional method , if somehow due to some disaster if Server system Damaged then the surety of Data Recovery is very less. But with the tools of Virtualization real time data backup recovery and mirroring become easy task and provide surety of zero percent data loss.

- **Virtualization saves Energy –**

Virtualization will help to save Energy because while moving from physical Servers to Virtual Server's, the number of Server's decreases due to this monthly power and cooling cost decreases which will Save Money as well. As cooling cost reduces it means carbon production by devices also decreases which results in Fresh and pollution free environment.

- **Quick and Easy Set up –**

In traditional methods Setting up physical system and servers are very time-consuming. Firstly Purchase them in bulk after that wait for shipment. When Shipment is done then wait for Setting up and after that again spend time in installing required software etc. Which will consume very time. But with the help of virtualization the entire process is done in very less time which results in productive setup.

- **Cloud Migration becomes easy –**

Most of the companies those who already have spent a lot in the server have a doubt of Shifting to Cloud. But it is more cost-effective to shift to cloud services because all the data that is present in their server's can be easily migrated into the cloud server and save something from maintenance charge, power consumption, cooling cost, cost to Server Maintenance Engineer etc.

Cons of Virtualization

- **Data can be at Risk –**
Working on virtual instances on shared resources means that our data is hosted on third party resource which put's our data in vulnerable condition. Any hacker can attack on our data or try to perform unauthorized access. Without Security solution our data is in threaten situation.
- **Learning New Infrastructure –**
As Organization shifted from Servers to Cloud. They required skilled staff who can work with cloud easily. Either they hire new IT staff with relevant skill or provide training on that skill which increase the cost of company.
- **High Initial Investment –**
It is true that Virtualization will reduce the cost of companies but also it is truth that Cloud have high initial investment. It provides numerous services which are not required and when unskilled organization will try to set up in cloud they purchase unnecessary services which are not even required to them.

Hypervisor

- Understanding the importance of Hypervisors, Type I & Type II Hypervisors.

Hypervisor

- A hypervisor is a form of virtualization software used in Cloud hosting to divide and allocate the resources on various pieces of hardware. The program which provides partitioning, isolation, or abstraction is called a virtualization hypervisor. The hypervisor is a hardware virtualization technique that allows multiple guest operating systems (OS) to run on a single host system at the same time. A hypervisor is sometimes also called a virtual machine manager(VMM).

Types of Hypervisor – 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. Examples of Type 1 hypervisors include VMware ESXi, Citrix XenServer, and Microsoft Hyper-V hypervisor.
- **Pros & Cons of Type-1 Hypervisor:**
- **Pros:** Such kinds of hypervisors are very efficient because they have direct access to the physical hardware resources(like Cpu, Memory, Network, and Physical storage). This causes the empowerment of the security because there is nothing any kind of the third party resource so that attacker couldn't compromise with anything.
- **Cons:** One problem with Type-1 hypervisors is that they usually need a dedicated separate machine to perform their operation and to instruct different VMs and control the host hardware resources.

Types of Hypervisor – 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). Basically, the software is installed on an operating system. Hypervisor asks the operating system to make hardware calls. An example of a Type 2 hypervisor includes VMware Player or Parallels Desktop. Hosted hypervisors are often found on endpoints like PCs. The type-2 hypervisor is very useful for engineers, and security analysts (for checking malware, or malicious source code and newly developed applications).
- **Pros & Cons of Type-2 Hypervisor:**
- **Pros:** Such kind of hypervisors allows quick and easy access to a guest Operating System alongside the host machine running. These hypervisors usually come with additional useful features for guest machines. Such tools enhance the coordination between the host machine and the guest machine.
- **Cons:** Here there is no direct access to the physical hardware resources so the efficiency of these hypervisors lags in performance as compared to the type-1 hypervisors, and potential security risks are also there an attacker can compromise the security weakness if there is access to the host operating system so he can also access the guest operating system.

Choosing the right hypervisor

- **Type 1 hypervisors offer much better performance than Type 2 ones** because there's no middle layer, making them the logical choice for mission-critical applications and workloads. But that's not to say that hosted hypervisors don't have their place – they're much simpler to set up, so they're a good bet if, say, you need to deploy a test environment quickly. One of the best ways to determine which hypervisor meets your needs is to compare their performance metrics. These include CPU overhead, the amount of maximum host and guest memory, and support for virtual processors. The following factors should be examined before choosing a suitable hypervisor:
 - **1. Understand your needs:** The company and its applications are the reason for the data center (and your job). Besides your company's needs, you (and your co-workers in IT) also have your own needs. Needs for a virtualization hypervisor are:
 - a. Flexibility
 - b. Scalability
 - c. Usability
 - d. Availability
 - e. Reliability
 - f. Efficiency
 - g. Reliable support
 - **2. The cost of a hypervisor:** For many buyers, the toughest part of choosing a hypervisor is striking the right balance between cost and functionality. While a number of entry-level solutions are free, or practically free, the prices at the opposite end of the market can be staggering. Licensing frameworks also vary, so it's important to be aware of exactly what you're getting for your money.
 - **3. Virtual machine performance:** Virtual systems should meet or exceed the performance of their physical counterparts, at least in relation to the applications within each server. Everything beyond meeting this benchmark is profit.

HYPERVERISOR REFERENCE MODEL

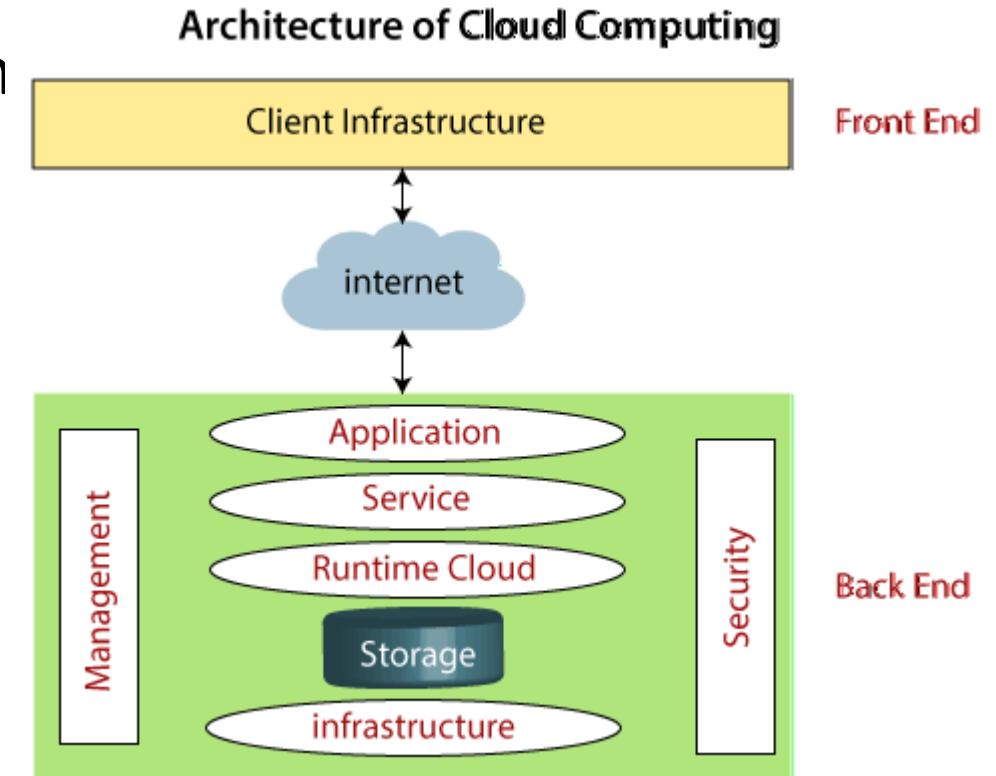
- There are 3 main modules coordinates in order to emulate the underlying hardware:
- **DISPATCHER:**
The dispatcher behaves like the entry point of the monitor and reroutes the instructions of the virtual machine instance to one of the other two modules.
- **ALLOCATOR:**
The allocator is responsible for deciding the system resources to be provided to the virtual machine instance. It means whenever a virtual machine tries to execute an instruction that results in changing the machine resources associated with the virtual machine, the allocator is invoked by the dispatcher.
- **INTERPRETER:**
The interpreter module consists of interpreter routines. These are executed, whenever a virtual machine executes a privileged instruction.

Cloud Architecture

Cloud Types: Private Cloud, Public cloud, Hybrid cloud, community cloud. Cloud as a service : Infrastructure as a service, Platform as a service, Software as a service, XaaS

Intro

- Cloud computing architecture is a combination of **service-oriented architecture** and **event-driven architecture**.
- Cloud computing architecture is divided into the following two parts -
 - Front End
 - Back End
- The below diagram shows the architecture of cloud computing -



Cloud Types

- Every cloud abstracts, pools, and shares scalable computing resources across a network. Every cloud type also enables cloud computing, which is the act of running workloads within that system.
- And every cloud is created using a unique mix of technologies, which almost always includes an operating system, some kind of management platform, and application programming interfaces (APIs).
- Virtualization and automation software can also be added to every kind of cloud for additional capabilities or increased efficiencies.

Cloud Types: Public Cloud

- Public clouds are a type of cloud computing run by a third-party cloud provider.
- These cloud providers deliver cloud services to their clients over the public internet.
- A cloud provider keeps ownership and control of the cloud storage, hardware, infrastructure and resources.
- This means that the cloud provider typically handles any updates or issues that require troubleshooting.
- Some public clouds give their clients free use of their cloud services. Other public clouds use a tiered subscription system where clients can choose how much storage and other cloud resources they need.
- Since public clouds offer shared resources to multiple clients, they're usually the most cost-efficient type of cloud deployment.

Cloud Types: Public Cloud

- Public Cloud provides a shared platform that is accessible to the general public through an Internet connection.
- Public cloud operated on the pay-as-per-use model and administrated by the third party, i.e., Cloud service provider.
- In the Public cloud, the same storage is being used by multiple users at the same time.
- Public cloud is owned, managed, and operated by businesses, universities, government organizations, or a combination of them.
- Amazon Elastic Compute Cloud (EC2), Microsoft Azure, IBM's Blue Cloud, Sun Cloud, and Google Cloud are examples of the public cloud.
- public cloud model

Cloud Types: Public Cloud

Advantages of Public Cloud

- Low Cost: Public cloud has a lower cost than private, or hybrid cloud, as it shares the same resources with a large number of consumers.
- Location Independent: Public cloud is location independent because its services are offered through the internet.
- Save Time: In Public cloud, the cloud service provider is responsible for the manage and maintain data centers in which data is stored, so the cloud user can save their time to establish connectivity, deploying new products, release product updates, configure, and assemble servers.
- Quickly and easily set up: Organizations can easily buy public cloud on the internet and deployed and configured it remotely through the cloud service provider within a few hours.
- Business Agility: Public cloud provides an ability to elastically re-size computer resources based on the organization's requirements.
- Scalability and reliability: Public cloud offers scalable (easy to add and remove) and reliable (24*7 available) services to the users at an affordable cost.
- Disadvantages of Public Cloud
- 1) Low Security: Public Cloud is less secure because resources are shared publicly.
- 2) Performance: In the public cloud, performance depends upon the speed of internet connectivity.
- 3) Less customizable: Public cloud is less customizable than the private cloud.

Cloud Types: Private Cloud

- Only one individual or business uses the resources and storage of a private cloud.
- Users access private cloud services over a private network that others can't access from the public internet.
- Private clouds can be physically located on a company's premises.
- Some third-party cloud providers may also offer clients a private cloud option for a higher price than a public cloud.
- Since private clouds don't share their resources with multiple clients over the internet, private clouds can offer organizations greater security than a public cloud.
- However, a private cloud typically costs more than a public one.

Cloud Types: Hybrid Cloud

- A hybrid cloud combines services of both public and private clouds. With a hybrid cloud, organizations can typically choose to combine various elements of both types of clouds.
- Since organizations can often customize hybrid clouds, this type of cloud deployment gives companies greater flexibility in their infrastructure and operations.
- For example, an organization may use aspects of a private cloud to keep their confidential business data secure, but use elements of public clouds for normal file-sharing or document collaboration.
- As another example, a company might create a hybrid cloud with primarily public cloud functions for their marketing and sales staff, but that mainly operates as a private cloud for their IT and accounting departments.

Cloud Types: Hybrid Cloud

- Hybrid cloud is a combination of **public and private** clouds.
Hybrid cloud = public cloud + private cloud
- The main aim to combine these cloud (Public and Private) is to create a unified, automated, and well-managed computing environment.
- In the Hybrid cloud, **non-critical activities** are performed by the **public cloud** and **critical activities** are performed by the **private cloud**.
- Mainly, a hybrid cloud is used in finance, healthcare, and Universities.
- The best hybrid cloud provider companies are **Amazon, Microsoft, Google, Cisco, and NetApp**.
-

Cloud Types: Hybrid Cloud

- Advantages of Hybrid Cloud
 - 1) Flexible and secure: It provides flexible resources because of the public cloud and secure resources because of the private cloud.
 - 2) Cost effective: Hybrid cloud costs less than the private cloud. It helps organizations to save costs for both infrastructure and application support.
 - 3) Cost effective: It offers the features of both the public as well as the private cloud. A hybrid cloud is capable of adapting to the demands that each company needs for space, memory, and system.
 - 4) Security: Hybrid cloud is secure because critical activities are performed by the private cloud.
 - 5) Risk Management: Hybrid cloud provides an excellent way for companies to manage the risk.

Disadvantages of Hybrid Cloud

- 1) Networking issues: In the Hybrid Cloud, networking becomes complex because of the private and the public cloud.
- 2) Infrastructure Compatibility: Infrastructure compatibility is the major issue in a hybrid cloud. With dual-levels of infrastructure, a private cloud controls the company, and a public cloud does not, so there is a possibility that they are running in separate stacks.
- 3) Reliability: The reliability of the services depends on cloud service providers.

Cloud Types: Community Cloud

- Community cloud is a cloud infrastructure that allows systems and services to be accessible by a group of several organizations to share the information. It is owned, managed, and operated by one or more organizations in the community, a third party, or a combination of them.
- **It is owned, managed, and operated by one or more organizations in the community, a third party, or a combination of them.** Example: Our government organization within India may share computing infrastructure in the cloud to manage data.
- **Example:** Our government organization within India may share computing infrastructure in the cloud to manage data.

Cloud Types: Community Cloud

Advantages of Community Cloud

- There are the following advantages of Community Cloud -
- Cost effective
- Community cloud is cost effective because the whole cloud is shared between several organizations or a community.
- Flexible and Scalable
- The community cloud is flexible and scalable because it is compatible with every user. It allows the users to modify the documents as per their needs and requirement.
- Security
- Community cloud is more secure than the public cloud but less secure than the private cloud.
- Sharing infrastructure
- Community cloud allows us to share cloud resources, infrastructure, and other capabilities among various organizations.

Disadvantages of Community Cloud

- There are the following disadvantages of Community Cloud -
- Community cloud is not a good choice for every organization.
- Slow adoption to data
- The fixed amount of data storage and bandwidth is shared among all community members.
- Community Cloud is costly than the public cloud.
- Sharing responsibilities among organizations is difficult.

Cloud Types: multi-Cloud

- A multi-cloud system refers to when a business uses multiple third-party cloud providers. Some organizations choose to use multiple cloud providers to improve their cybersecurity systems. Multi-cloud environments can also help maintain separate clouds for different workflows, departments or branches within their company. However, with a multi-cloud system, all of your cloud resources and data operate on separate infrastructures, which can make it more challenging to share resources between clouds.

Cloud Types: High-performance computing (HPC)

Cloud

- HPC clouds specifically provide cloud services for high-performing computer applications and devices, sometimes referred to as supercomputers. Some organizations use supercomputers to perform complex computational tasks, such as forecasting the weather or modeling chemical molecules. An HPC cloud offers enough data space and server power to help ensure that supercomputers continue to run efficiently while providing organizations with the services they need.

Cloud as a service

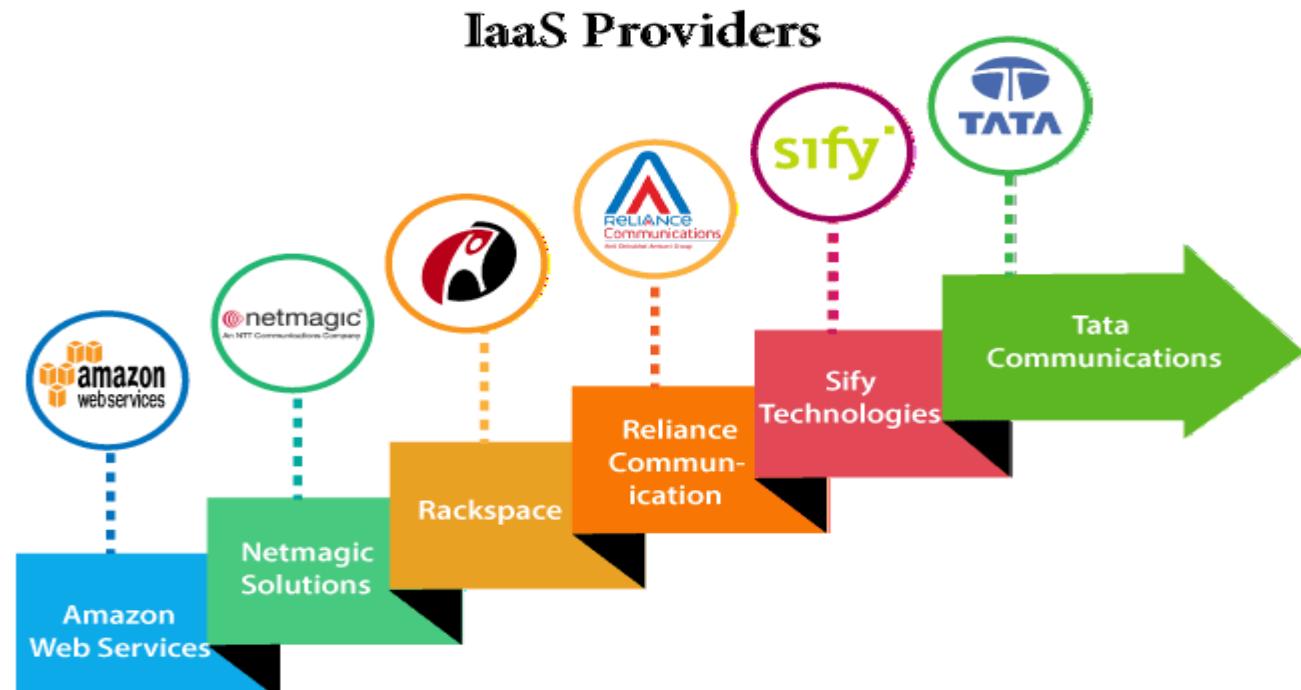
Infrastructure as a service

- IaaS is also known as **Hardware as a Service (HaaS)**
- It allows customers to outsource their IT infrastructures such as servers, networking, processing, storage, virtual machines, and other resources. Customers access these resources on the Internet using a pay-as-per use model.
- In traditional hosting services, IT infrastructure was rented out for a specific period of time, with pre-determined hardware configuration. The client paid for the configuration and time, regardless of the actual use. With the help of the IaaS cloud computing platform layer, clients can dynamically scale the configuration to meet changing requirements and are billed only for the services actually used.
- IaaS cloud computing platform layer eliminates the need for every organization to maintain the IT infrastructure.
- IaaS is offered in three models: public, private, and hybrid cloud.
 - The private cloud implies that the infrastructure resides at the customer-premise.
 - public cloud, it is located at the cloud computing platform vendor's data center
 - hybrid cloud is a combination of the two in which the customer selects the best of both public cloud or private cloud.

Infrastructure as a service

IaaS provider provides the following services -

1. **Compute:** Computing as a Service includes virtual central processing units and virtual main memory for the Vms that is provisioned to the end- users.
2. **Storage:** IaaS provider provides back-end storage for storing files.
3. **Network:** Network as a Service (NaaS) provides networking components such as routers, switches, and bridges for the Vms.
4. **Load balancers:** It provides load balancing capability at the infrastructure layer.



Infrastructure as a service

Advantages of IaaS cloud computing layer

- 1. Shared infrastructure:** IaaS allows multiple users to share the same physical infrastructure.
- 2. Web access to the resources:** IaaS allows IT users to access resources over the internet.
- 3. Pay-as-per-use model:** IaaS providers provide services based on the pay-as-per-use basis. The users are required to pay for what they have used.
- 4. Focus on the core business:** IaaS providers focus on the organization's core business rather than on IT infrastructure.
- 5. On-demand scalability:** On-demand scalability is one of the biggest advantages of IaaS. Using IaaS, users do not worry about to upgrade software and troubleshoot the issues related to hardware components.

Disadvantages of IaaS cloud computing layer

- 1. Security** Security is one of the biggest issues in IaaS. Most of the IaaS providers are not able to provide 100% security.
- 2. Maintenance & Upgrade:** Although IaaS service providers maintain the software, but they do not upgrade the software for some organizations.
- 3. Interoperability issues:** It is difficult to migrate VM from one IaaS provider to the other, so the customers might face problem related to vendor lock-in.

Infrastructure as a service

IaaS Vendor	Iaas Solution	Details
Amazon Web Services	Elastic, Elastic Compute Cloud (EC2) MapReduce, Route 53, Virtual Private Cloud, etc.	The cloud computing platform pioneer, Amazon offers auto scaling, cloud monitoring, and load balancing features as part of its portfolio.
Netmagic Solutions	Netmagic IaaS Cloud	Netmagic runs from data centers in Mumbai, Chennai, and Bangalore, and a virtual data center in the United States. Plans are underway to extend services to West Asia.
Rackspace	Cloud servers, cloud files, cloud sites, etc.	The cloud computing platform vendor focuses primarily on enterprise-level hosting services.
Reliance Communications	Reliance Internet Data Center	RIDC supports both traditional hosting and cloud services, with data centers in Mumbai, Bangalore, Hyderabad, and Chennai. The cloud services offered by RIDC include IaaS and SaaS.
Sify Technologies	Sify IaaS	Sify's cloud computing platform is powered by HP's converged infrastructure. The vendor offers all three types of cloud services: IaaS, PaaS, and SaaS.
Tata Communications	InstaCompute	InstaCompute is Tata Communications' IaaS offering. InstaCompute data centers are located in Hyderabad and Singapore, with operations in both countries.

Platform as a service

Platform as a Service (PaaS) provides a runtime environment. It allows programmers to easily create, test, run, and deploy web applications. You can purchase these applications from a cloud service provider on a pay-as-per use basis and access them using the Internet connection. In PaaS, back end scalability is managed by the cloud service provider, so end-users do not need to worry about managing the infrastructure. PaaS includes infrastructure (servers, storage, and networking) and platform (middleware, development tools, database management systems, business intelligence, and more) to support the web application life cycle.

Example: Google App Engine, Force.com, Joyent, Azure.

Platform as a service

PaaS providers provide the Programming languages, Application frameworks, Databases, and Other tools: Platform as a Service

1. Programming languages: PaaS providers provide various programming languages for the developers to develop the applications. Some popular programming languages provided by PaaS providers are Java, PHP, Ruby, Perl, and Go.

2. Application frameworks: PaaS providers provide application frameworks to easily understand the application development. Some popular application frameworks provided by PaaS providers are Node.js, Drupal, Joomla, WordPress, Spring, Play, Rack, and Zend.

3. Databases: PaaS providers provide various databases such as ClearDB, PostgreSQL, MongoDB, and Redis to communicate with the applications.

4. Other tools: PaaS providers provide various other tools that are required to develop, test, and deploy the applications.

Platform as a service

Advantages of PaaS

1) Simplified Development

PaaS allows developers to focus on development and innovation without worrying about infrastructure management.

2) Lower risk

No need for up-front investment in hardware and software. Developers only need a PC and an internet connection to start building applications.

3) Prebuilt business functionality

Some PaaS vendors also provide already defined business functionality so that users can avoid building everything from very scratch and hence can directly start the projects only.

4) Instant community

PaaS vendors frequently provide online communities where the developer can get the ideas to share experiences and seek advice from others.

5) Scalability

Applications deployed can scale from one to thousands of users without any changes to the applications.

Platform as a service

Disadvantages of PaaS cloud computing layer

1) Vendor lock-in

One has to write the applications according to the platform provided by the PaaS vendor, so the migration of an application to another PaaS vendor would be a problem.

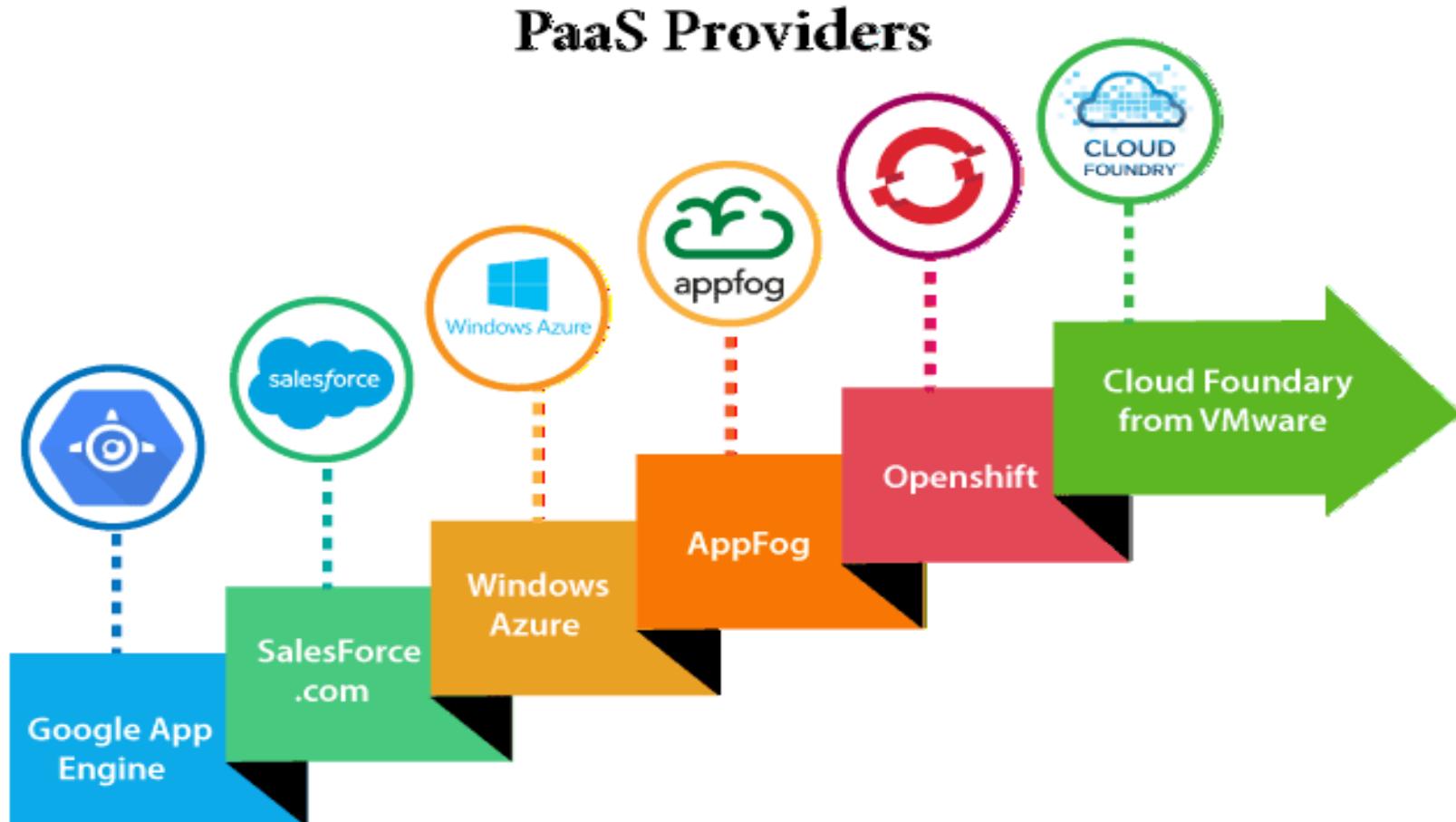
2) Data Privacy

Corporate data, whether it can be critical or not, will be private, so if it is not located within the walls of the company, there can be a risk in terms of privacy of data.

3) Integration with the rest of the systems applications

It may happen that some applications are local, and some are in the cloud. So there will be chances of increased complexity when we want to use data which is in the cloud with the local data.

Platform as a service



Platform as a service

Providers	Services
Google Engine (GAE)	App App Identity, URL Fetch, Cloud storage client library, Logservice
Salesforce.com	Faster implementation, Rapid scalability, CRM Services, Sales cloud, Mobile connectivity, Chatter.
Windows Azure	Compute, security, IoT, Data Storage.
AppFog	Justcloud.com, SkyDrive, GoogleDocs
Openshift	RedHat, Microsoft Azure.
Cloud Foundry from VMware	Data, Messaging, and other services.

Software as a service

SaaS is also known as "**On-Demand Software**". It is a software distribution model in which services are hosted by a cloud service provider. These services are available to end-users over the internet so, the end-users do not need to install any software on their devices to access these services.

There are the following services provided by SaaS providers -

Business Services - SaaS Provider provides various business services to start-up the business. The SaaS business services include **ERP** (Enterprise Resource Planning), **CRM** (Customer Relationship Management), **billing**, and **sales**.

Document Management - SaaS document management is a software application offered by a third party (SaaS providers) to create, manage, and track electronic documents.

Example: Slack, Samepage, Box, and Zoho Forms.

Social Networks - As we all know, social networking sites are used by the general public, so social networking service providers use SaaS for their convenience and handle the general public's information.

Mail Services - To handle the unpredictable number of users and load on e-mail services, many e-mail providers offering their services using SaaS.

Software as a service



Software as a service

- SaaS is also known as "**On-Demand Software**". It is a software distribution model in which services are hosted by a cloud service provider. These services are available to end-users over the internet so, the end-users do not need to install any software on their devices to access these services.
- There are the following services provided by SaaS providers -
- **Business Services** - SaaS Provider provides various business services to start-up the business. The SaaS business services include **ERP** (Enterprise Resource Planning), **CRM** (Customer Relationship Management), **billing**, and **sales**.
- **Document Management** - SaaS document management is a software application offered by a third party (SaaS providers) to create, manage, and track electronic documents.
Example: Slack, Samepage, Box, and Zoho Forms.
- **Social Networks** - As we all know, social networking sites are used by the general public, so social networking service providers use SaaS for their convenience and handle the general public's information.
- **Mail Services** - To handle the unpredictable number of users and load on e-mail services, many e-mail providers offering their services using SaaS.

Software as a service- Advantages

- SaaS is easy to buy: SaaS pricing is based on a monthly fee or annual fee subscription, so it allows organizations to access business functionality at a low cost, which is less than licensed applications. SaaS providers are generally pricing the applications using a subscription fee
- One to Many: SaaS services are offered as a one-to-many model means a single instance of the application is shared by multiple users.
- Less hardware required for SaaS: The software is hosted remotely, so organizations do not need to invest in additional hardware.
- Low maintenance required for SaaS: Software as a service removes the need for installation, set-up, and daily maintenance for the organizations. The initial set-up cost for SaaS is typically less than the enterprise software. SaaS vendors are pricing their applications based on some usage parameters, such as a number of users using the application. So SaaS does easy to monitor and automatic updates.
- No special software or hardware versions required: All users will have the same version of the software and typically access it through the web browser. SaaS reduces IT support costs by outsourcing hardware and software maintenance and support to the IaaS provider.
- Multidevice support: SaaS services can be accessed from any device such as desktops, laptops, tablets, phones, and thin clients.
- API Integration: SaaS services easily integrate with other software or services through standard APIs.
- No client-side installation: SaaS services are accessed directly from the service provider using the internet connection, so do not need to require any software installation.

Software as a service

Disadvantages of SaaS cloud computing layer

- 1) Security: Actually, data is stored in the cloud, so security may be an issue for some users. However, cloud computing is not more secure than in-house deployment.
- 2) Latency issue: Since data and applications are stored in the cloud at a variable distance from the end-user, there is a possibility that there may be greater latency when interacting with the application compared to local deployment. Therefore, the SaaS model is not suitable for applications whose demand response time is in milliseconds.
- 3) Total Dependency on Internet: Without an internet connection, most SaaS applications are not usable.
- 4) Switching between SaaS vendors is difficult: Switching SaaS vendors involves the difficult and slow task of transferring the very large data files over the internet and then converting and importing them into another SaaS also.

Software as a service

Provider	Services
Salseforce.com	On-demand CRM solutions
Microsoft Office 365	Online office suite
Google Apps	Gmail, Google Calendar, Docs, and sites
NetSuite	ERP, accounting, order management, CRM, Professionals Services Automation (PSA), and e-commerce applications.
GoToMeeting	Online meeting and video-conferencing software
Constant Contact	E-mail marketing, online survey, and event marketing
Oracle CRM	CRM applications
Workday, Inc	Human capital management, payroll, and financial management.

XaaS

"Anything as a service" (XaaS) describes a general category of cloud computing and remote access services. It recognizes the vast number of products, tools, and technologies now delivered to users as a service over the Internet.

Essentially, any IT function can be a service for enterprise consumption. The service is paid for in a flexible consumption model rather than an advance purchase or license.

XaaS

benefits of XaaS?

- **Expenditure model improvements.** With XaaS, businesses can cut costs by purchasing services from providers on a subscription basis. Before XaaS and cloud services, businesses had to buy separate products-software, hardware, servers, security, infrastructure-install them on-site, and then link everything together to form a network. With XaaS, businesses buy what they need and pay on the go. The previous capital expenditure now becomes an operating expense.
- **Speed up new apps and business processes.** This model allows businesses to adopt new apps or solutions to changing market conditions. Using multi-tenant approaches, cloud services can provide much-needed flexibility. Resource pooling and rapid elasticity support mean that business leaders can add or subtract services. When users need innovative resources, a company can use new technologies, automatically scaling up the infrastructure.
- **Transferring IT resources to high-value projects.** Increasingly, IT organizations are turning to a XaaS delivery model to streamline operations and free up resources for innovation. They are also harnessing the benefits of XaaS to transform digitally and become more agile. XaaS gives more users access to cutting-edge technology, democratizing innovation. In a recent survey by Deloitte, 71% of companies report that XaaS now constitutes more than half of their company's enterprise IT.

XaaS

What are the disadvantages of XaaS?

- **Possible downtime.** The Internet sometimes breaks down, and when this happens, your XaaS provider can be a problem too. With XaaS, there can be issues of Internet reliability, flexibility, provisioning, and management of infrastructure resources. If XaaS servers go down, users will not be able to use them. XaaS providers can guarantee services through SLAs.
- **Performance issues.** As XaaS becomes more popular, bandwidth, latency, data storage, and recovery times can be affected. If too many clients use the same resources, the system may slow down. Apps running in virtualized environments can also be affected. Integration issues can occur in these complex environments, including the ongoing management and security of multiple cloud services.
- **Complexity effect.** Advancing technology for XaaS can relieve IT workers from day-to-day operational headaches; however, it can be difficult to troubleshoot if something goes wrong.
- Internal IT staff still needs to stay updated on new technology. The cost of maintaining a high-performance, a robust network can add up - although the overall cost savings of the XaaS model are usually enormous. Nonetheless, some companies want to maintain visibility into their XaaS service provider's environment and infrastructure. Furthermore, a XaaS provider that gets acquired shuts down a service or changes its roadmap can profoundly impact XaaS users.

XaaS

Hardware as a Service (HaaS) –

Managed Service Providers (MSP) provide and install some hardware on the customer's site on demand. The customer uses the hardware according to service level agreements. This model is very similar to IaaS as computing resources present at MSP's site are provided to users substituted for physical hardware.

XaaS

Communication as a Service (CaaS) –

This model comprises solutions for different communication like IM, VoIP, and video conferencing applications which are hosted in the provider's cloud. Such a method is cost-effective and reduces time expenses.

XaaS

Desktop as a Service (DaaS) –

DaaS provider mainly manages storing, security, and backing up user data for desktop apps. And a client can also work on PCs using third-party servers.

Security as a Service (SECaaS) –

In this method, the provider integrates security services with the company's infrastructure through the internet which includes anti-virus software, authentication, encryption, etc.

XaaS

Healthcare as a Service (HaaS) –

The healthcare industry has opted for the model HaaS service through electronic medical records (EMR). IoT and other technologies have enhanced medical services like online consultations, health monitoring 24/7, medical service at the doorstep e.g. lab sample collection from home, etc.

Transport as a Service (TaaS) –

Nowadays, there are numerous apps that help in mobility and transport in modern society. The model is both convenient and ecological friendly e.g. Uber taxi services is planning to test flying taxis and self-driving planes in the future

Benefits in XaaS :

Cost Saving – When an organization uses XaaS then it helps in cost-cutting and simplifies IT deployments.

Scalability –XaaS can easily handle the growing amount of work by providing the required resources/service.

Accessibility – It helps in easy accessing and improving accessibility as long as the internet connection is there.

Faster Implementation –It provides faster implementation time to various activities of the organization.

Quick Modification – It provides updates for modification as well as undergoes quick updating by providing quality services.

Better Security –It contains improved security controls and is configured to the exact requirements of the business.

Boost innovation – While XaaS is used it Streamlines the operations and frees up resources for innovation.

Flexibility – XaaS provides flexibility by using cloud services and multiple advanced approaches.

Benefits in XaaS :

Disadvantages in XaaS :

Internet Breakage – Internet breaks sometimes for XaaS service providers where there can also be issues in internet reliability, provisioning, and managing the infrastructure resources.

Slowdown – When too many clients are using the same resources at the same time, the system can slow down.

Difficult in Troubleshoot – XaaS can be a solution for IT staff in day-to-day operational headaches, but if anywhere problem occurs it is harder to troubleshoot it as in XaaS multiple services are included with various technologies and tools.

Change brings problems – If a XaaS provider discontinues a service or alters it gives an impact on XaaS users.

XaaS is on the rise now :

Public cloud services are growing at an exponential rate. Researchers assumed that global cloud computing revenue is going to reach \$342 billion dollars by 2025. Through the XaaS model by servitization, products and services are combined through which businesses innovate faster and enhance their relationship with customers which further increases their revenue.

Future of XaaS :

A combination of cloud computing and good internet access allows accessing good quality XaaS services and better improvement of XaaS. Some companies are not confident to take XaaS because of security and business governance concerns. But service providers increasingly reveal these concerns which allow organizations which put additional workloads into the cloud.

Function as a Service

- FaaS is a type of cloud computing service. It provides a platform for its users or customers to develop, compute, run and deploy the code or entire application as functions. It allows the user to entirely develop the code and update it at any time without worrying about the maintenance of the underlying infrastructure. The developed code can be executed with response to the specific event. It is also **as same as PaaS**.
- FaaS is an event-driven execution model. It is implemented in the serverless container. When the application is developed completely, the user will now trigger the event to execute the code. Now, the triggered event makes response and activates the servers to execute it. The servers are nothing but the Linux servers or any other servers which is managed by the vendor completely. Customer does not have clue about any servers which is why they do not need to maintain the server hence it is **serverless architecture**.
- Both PaaS and FaaS are providing the same functionality but there is still some differentiation in terms of Scalability and Cost.
- FaaS, provides auto-scaling up and scaling down depending upon the demand. PaaS also provides scalability but here users have to configure the scaling parameter depending upon the demand.
- In FaaS, users only have to pay for the number of execution time happened. In PaaS, users have to pay for the amount based on pay-as-you-go price regardless of how much or less they use

Function as a Service

Advantages of FaaS :

- **Highly Scalable:** Auto scaling is done by the provider depending upon the demand.
- **Cost-Effective:** Pay only for the number of events executed.
- **Code Simplification:** FaaS allows the users to upload the entire application all at once. It allows you to write code for independent functions or similar to those functions.
- Maintenance of code is enough and no need to worry about the servers.
- Functions can be written in any programming language.
- Less control over the system.
- The various companies providing Function as a Service are Amazon Web Services – Firecracker, Google – Kubernetes, Oracle – Fn, Apache OpenWhisk – IBM, OpenFaaS,

Cloud Security

Identity and access management, security challenges, Storage basics, Storage as a service providers, aspects of data security AAA model, SSO model, Threat Agents

- Anonymous Attacker, Malicious Service Agent, Trusted Attacker, Malicious Insider Cloud Security

Threats - Traffic Eavesdropping, Malicious Intermediary, Denial of Service, Insufficient Authorization, Virtualization Attack, Overlapping Trust Boundaries, Common Attacks, Cloud-Specific Attacks, Flawed Implementations, Risk Management

Best practices for security in Cloud

- Encrypt data at rest and in motion
- Use multifactor authentication
- Adopt firewalls, IPSes and antimalware
- Isolate cloud data backups
- Ensure data location visibility and control
- Log and monitor all aspects of data access

Identity and access management

- Historically, defense-in-depth was mostly performed through network-layer controls. Advanced threat prevention tools are able to recognize the applications that traverse the network and determine whether or not they should be allowed. This type of security is still very much required in cloud native environments, but it's no longer sufficient on its own.
- Public cloud providers offer a rich portfolio of services, and the only way to govern and secure many of them is through identity and access management (IAM).
- Identity and access management (IAM) **ensures that the right people and job roles in your organization (identities) can access the tools they need to do their jobs.** Identity management and access systems enable your organization to manage employee apps without logging into each app as an administrator.
- If the IAM profile follows the “least-privilege” principle and only allows the function to put items in the table, the blast radius will be greatly reduced in the case of an incident.
- Managing a large number of privileged users with access to an ever-expanding set of services is challenging. Managing separate IAM roles and groups for these users and resources adds yet another layer of complexity.

Cloud IAM includes the following features

- Single Access Control Interface. Cloud IAM solutions provide a clean and consistent access control interface for all cloud platform services. The same interface can be used for all cloud services.
- Enhanced Security. You can define increased security for critical applications.
- Resource-level Access Control. You can define roles and grant permissions to users to access resources at different granularity levels.

NEED IDENTITY AND ACCESS MANAGEMENT

- Identity and Access Management technology can be used to initiate, capture, record, and manage user identities and their access permissions. All users are authenticated, authorized, and evaluated according to policies and roles.
- Poorly controlled IAM processes may lead to regulatory non-compliance; if the organization is audited, management may not be able to prove that company data is not at risk of being misused.

IAM solution very valuable

- The ability to spend less on enterprise security by relying on the centralized trust model to deal with Identity Management across third-party and own applications.
- It enables your users to work from any location and any device.
- You can give them access to all your applications using just one set of credentials through Single Sign-On.
- You can protect your sensitive data and apps: Add extra layers of security to your mission-critical apps using Multifactor Authentication.
- It helps maintain compliance of processes and procedures. A typical problem is that permissions are granted based on employees' needs and tasks, and not revoked when they are no longer necessary, thus creating users with lots of unnecessary privileges.

Getting IAM Security Right

- **Don't use root accounts** - Always create individual IAM users with relevant permissions, and don't give your root credentials to anyone.
- **Adopt a role-per-group model** - Assign policies to groups of users based on the specific things those users need to do. Don't "stack" IAM roles by assigning roles to individual users and then adding them to groups. This will make it hard for you to understand their effective permissions.
- **Grant least-privilege** - Only grant the least amount of permissions needed for a job, This will ensure that if a user or resource is compromised, the blast radius is reduced to the one or few things that entity was permitted to do. This is an ongoing task. As your application is constantly changing, you need to make sure that your permissions adapt accordingly.
- **Leverage cloud provider tools** - Managing many permission profiles at scale is challenging. Leverage the platforms you are already using to generate least-privilege permission sets and analyze your existing services. Remember that the cloud provider recommendation is to always manually review the generated profiles before implementing them.

Identity and Access Challenges

- **IAM and SSO:** Most businesses today use some form of single sign-on (SSO), such as Okta, to manage the way users interact with cloud services. This is an effective way of centralizing access across a large number of users and services. While using SSO to log into public cloud accounts is definitely the best practice, the mapping between SSO users and IAM roles can become challenging, as users can have multiple roles that span several cloud accounts.
- **Effective permissions:** Considering that users and services have more than one permission set attached to them, understanding the effective permissions of an entity becomes difficult.
- **Multi-cloud:** According to RightScale, more than 84% of organizations use a multi-cloud strategy. Each provider has its own policies, tools and terminology. There is no common language that helps you understand relationships and permissions across cloud providers.

security challenges

- **1. Data Breaches :** There is no concern more palpable than a data breach. It's something every organization is focused on. However, few have the resources and strategies in place to truly tackle it in a worthy manner. This makes it a critical concern (and something that has to be dealt with in a proactive and preventative way). Failure to deal with data properly (through deliberate encryption) opens your business up to huge compliance risks – not to mention data breach penalties, fines, and serious violations of customer trust. The onus is on you to protect your customer and employee data, regardless of what any Service-Level Agreement (SLA) says.
- **2. Compliance With Regulatory Mandates :** It's commonplace for organizations – particularly smaller and mid-size companies – to assume that they're getting maximum protection simply by working with a cloud solutions provider. But there's more to it than meets the eye. Compliance goes beyond international and federal regulations. There are also additional industry mandates that must be addressed. Examples include EU data protection, PCI DSS, FISMA, GLBA, HIPAA, and FERPA – to name a few. The right cloud security solutions provide the technical capacity to abide by regulatory mandates, but there has to be regular oversight and granular attention to detail. Under the responsibility model, the cloud provider offers security *of* the cloud, while the end user provides security *in* the cloud.
- **3. Lack of IT Expertise :** According to the Cloud Security Alliance "Cloud Adoption Practices & Priorities Survey Report," [34 percent of companies](#) are currently avoiding the cloud because they don't believe their IT and business managers have the knowledge and experience to handle the demands of cloud computing. This makes it one of the top-four concerns businesses have in regards to cloud security. The average enterprise now has between three and four clouds. This creates added layers of complexity that require technical competence and relevant experience. This speaks to a larger trend that we'll expect to see emerge in the coming months and years. Rather than just having managerial experience and financial literacy, IT and business managers will be required to bring technical cloud competency to the table. This doesn't mean they'll have to be cloud experts, but basic understanding and the ability to lead targeted initiatives becomes integral.
- **4. Cloud Migration Issues :** Cloud migration is happening in droves, but it has to be handled properly (otherwise, it exposes the business to unnecessary risk). [According to one report](#), the four biggest challenges facing businesses are visibility into infrastructure security (43 percent), compliance (38 percent), setting security policies (35 percent), and security failing to keep up with the pace of change in applications (35 percent). As a result, security professionals and IT pros are feeling overwhelmed by everything that's asked of them. Simpler and more straight forward migration strategies will help businesses manage this transition flawlessly. Trying to accomplish everything at once is a major mistake. The migration process should be broken down into stages to reduce the risk of critical errors that could corrupt data and/or lead to vulnerabilities.

security challenges

- **4. Cloud Migration Issues:** Cloud migration is happening in droves, but it has to be handled properly (otherwise, it exposes the business to unnecessary risk). According to one report, the four biggest challenges facing businesses are visibility into infrastructure security (43 percent), compliance (38 percent), setting security policies (35 percent), and security failing to keep up with the pace of change in applications (35 percent). As a result, security professionals and IT pros are feeling overwhelmed by everything that's asked of them.
- Simpler and more straight forward migration strategies will help businesses manage this transition flawlessly. Trying to accomplish everything at once is a major mistake. The migration process should be broken down into stages to reduce the risk of critical errors that could corrupt data and/or lead to vulnerabilities.
- **5. Unsecured APIs:** The difficult thing about the cloud is that there are so many different possible entry points for attacks. So while the surface attack area may be smaller in totality, it's much more fragmented. Perhaps this can be seen most clearly when it comes to micro-service architecture and the increasing trend around serverless functions. APIs are great, but you have to consider how they impact the larger system. Even if the cloud is technically safe and sound, intruders can hijack data by hacking into less-secure APIs. This is problematic! The proper cloud security solutions can help you carefully vet each application to protect against weak points like these.
- **6. Insider Threats:** It's a good business practice to trust your employees. Unfortunately, many businesses take this trust too far – or fail to vet the driving factors behind their trust on the front end. According to research from Intel, insider threats are responsible for an incredible 43 percent of all breaches. Half are intentional and half are accidental. More specifically, businesses need to think about access management and limiting who can access what and when. Access to cloud applications and data sources should be given on an as-required basis. Nobody should have more access than is needed to complete their job-related responsibilities.
- **7. Open Source:** Use of open source to develop applications. Open source packages are vulnerable. Most often hackers poison the well in the Git repo, waiting for developers to use the packages and later compromise the application through a well prepared attack vector.

Storage Basic

- **Cloud storage is a cloud computing model that stores data on the Internet through a cloud computing provider who manages and operates data storage as a service.** It's delivered on demand with just-in-time capacity and costs, and eliminates buying and managing your own data storage infrastructure.
- Cloud storage is purchased from a third party cloud vendor who owns and operates data storage capacity and delivers it over the Internet in a pay-as-you-go model. These cloud storage vendors manage capacity, security and durability to make data accessible to your applications all around the world.
- Applications access cloud storage through traditional storage protocols or directly via an API. Many vendors offer complementary services designed to help collect, manage, secure and analyze data at massive scale.

Storage Basic

- Storing data in the cloud lets IT departments transform three areas:
- Total Cost of Ownership. With cloud storage, there is no hardware to purchase, storage to provision, or capital being used for "someday" scenarios. You can add or remove capacity on demand, quickly change performance and retention characteristics, and only pay for storage that you actually use. Less frequently accessed data can even be automatically moved to lower cost tiers in accordance with auditable rules, driving economies of scale.
- Time to Deployment. When development teams are ready to execute, infrastructure should never slow them down. Cloud storage allows IT to quickly deliver the exact amount of storage needed, right when it's needed. This allows IT to focus on solving complex application problems instead of having to manage storage systems.
- Information Management. Centralizing storage in the cloud creates a tremendous leverage point for new use cases. By using cloud storage lifecycle management policies, you can perform powerful information management tasks including automated tiering or locking down data in support of compliance requirements.

Cloud Storage Requirements

- Durability. Data should be redundantly stored, ideally across multiple facilities and multiple devices in each facility. Natural disasters, human error, or mechanical faults should not result in data loss.
- Availability. All data should be available when needed, but there is a difference between production data and archives. The ideal cloud storage will deliver the right balance of retrieval times and cost.
- Security. All data is ideally encrypted, both at rest and in transit. Permissions and access controls should work just as well in the cloud as they do for on premises storage.

Types of Cloud Storage

- Object Storage - Applications developed in the cloud often take advantage of object storage's vast scalability and metadata characteristics. Object storage solutions like Amazon Simple Storage Service (S3) are ideal for building modern applications from scratch that require scale and flexibility, and can also be used to import existing data stores for analytics, backup, or archive.
- File Storage - Some applications need to access shared files and require a file system. This type of storage is often supported with a Network Attached Storage (NAS) server. File storage solutions like Amazon Elastic File System (EFS) are ideal for use cases like large content repositories, development environments, media stores, or user home directories.
- Block Storage - Other enterprise applications like databases or ERP systems often require dedicated, low latency storage for each host. This is analogous to direct-attached storage (DAS) or a Storage Area Network (SAN). Block-based cloud storage solutions like Amazon Elastic Block Store (EBS) are provisioned with each virtual server and offer the ultra low latency required for high performance workloads.

Ways to Use Cloud Storage

- **Backup and Recovery:** Backup and recovery is a critical part of ensuring data is protected and accessible, but keeping up with increasing capacity requirements can be a constant challenge. Cloud storage brings low cost, high durability, and extreme scale to backup and recovery solutions.
- **Software Test and Development:** Software test and development environments often require separate, independent, and duplicate storage environments to be built out, managed, and decommissioned. In addition to the time required, the up-front capital costs required can be extensive.
- **Cloud Data Migration:** The availability, durability, and cost benefits of cloud storage can be very compelling to business owners, but traditional IT functional owners like storage, backup, networking, security, and compliance administrators may have concerns around the realities of transferring large amounts of data to the cloud.
- **Compliance:** Storing data in the cloud can raise concerns about regulation and compliance, especially if this data is already stored in compliant storage systems.
- **Big Data and Data Lakes:** Traditional on-premises storage solutions can be inconsistent in their cost, performance, and scalability — especially over time. Big data projects demand large-scale, affordable, highly available, and secure storage pools that are commonly referred to as data lakes.

Storage as a service providers

- Storage as a Service or STaaS is cloud storage that you rent from a Cloud Service Provider (CSP) and that provides basic ways to access that storage.
- Enterprises, small and medium businesses, home offices, and individuals can use the cloud for multimedia storage, data repositories, data backup and recovery, and disaster recovery.
- There are also higher-tier managed services that build on top of STaaS, such as Database as a Service, in which you can write data into tables that are hosted through CSP resources.
- The key benefit to STaaS
 - offloading the cost and effort to manage data storage infrastructure and technology to a third-party CSP.
 - more effective to scale up storage resources without investing in new hardware or taking on configuration costs.
 - respond to changing market conditions faster.

Example

- Acronis Cyber Infrastructure supplements the company's on-premises software with a cloud storage, backup, DR, file sync-and-share suite that's typically delivered by an independent service provider.
- Box has evolved into an enterprise-focused service offering a variety of storage-centric content management, collaboration, compliance and development services designed for business process transformation.
- ClearSky Data has on-demand primary storage, offsite backup and DR as a service using metro-area POPs and caching at the customer edge.
- Dropbox is still best known for pioneering consumer sync-and-share services, but it has developed enterprise products for collaboration, file backup and recovery and data management.
- StorOne service delivers on-premises hardware with a cloud subscription model (note that many storage vendors are emulating this model -- see below).
- Zadara Cloud Platform is a hybrid, distributed software-defined storage that works on local hardware and public cloud infrastructure (AWS, Google Cloud Platform, Microsoft Azure, Oracle and VMware on AWS) and offers block, file and object services.

Aspects of data security

- no insight into how cloud providers are storing and securing their data
- Integrity Protection :
- Enabling public audit for cloud data storage security
- Security issues with virtualization
- Data Segregation
- applications and data housed on third-party infrastructure
- Different cloud providers have varying capabilities, which can result in inconsistent cloud data protection and security
- Application vulnerabilities and malware propagation

What Are the Major Security Risks in the Cloud?

- A Broader Threat Landscape
- Lack of Control Over Cloud Host's Security Services
- Automation in DevOps
- Weak Access Management
- Inconsistent Security in Complex Environments
- Compliance Requirements

Benefits of cloud data protection

- Secure applications and data across multiple environments while maintaining complete visibility into all user, folder and file activity.
- Proactively identify and mitigate risks, such as security threats, suspicious user behavior, malware and others.
- Better govern access.
- Define policies.
- Prevent and detect data loss and disruption.

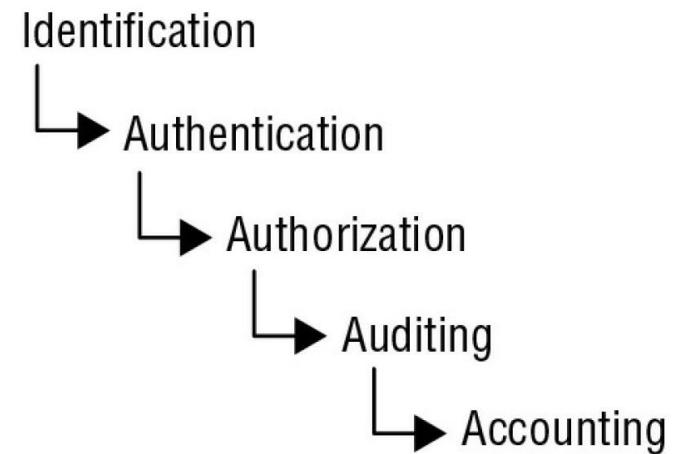
six pillars of cloud security

1. Secure Access Controls: A good security framework starts by implementing secure Identity Access Management (IAM) protocols. Assure that team members have the minimal amount of access necessary to systems, assets, and APIs that they need to do their job. As privileges increase, so should the level of authentication required to gain access. Employees should take ownership as well through enforced password policies.
2. Zero-Trust Network Security Controls: Keep your mission-critical assets and applications in strategically isolated portions of your cloud network. For example, in a virtual private cloud through AWS or vNET through Microsoft Azure. Segregate secure workloads from those that don't require data security protocols and enforce these micro-segments with strict security policies.
3. Change Management: Use change management protocols offered by your cloud security provider to govern change and enforce compliance controls any time a change is requested, a new server is provisioned, or sensitive assets are moved or changed. Change management applications will provide auditing functionality that can monitor for unusual behavior and deviation from protocol so that you can investigate, or can trigger automatic mitigation to correct the issue.

six pillars of cloud security

4. Web Application Firewall: A web application firewall (WAF) will scrutinize traffic into and out of your web application and servers to monitor and alert the administrator of any unusual behavior to prevent breaches and strengthen endpoint security.
5. Data Protection: To provide enhanced data security, your organization should encrypt data at every transport layer. Additionally, there should be security protocols applied to any file sharing, communication applications, and any other area within your environment where data might be held, used, or transmitted.
6. Continuous Monitoring: Many cloud security providers can offer insight into your cloud-native logs by comparing them against internal logs from your other security tools such as asset management, change management, vulnerability scanners as well as external threat intelligence insight.

AAA model

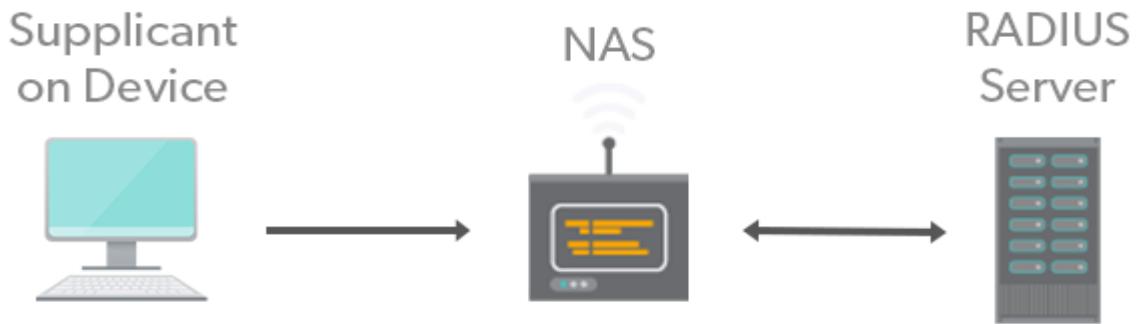


- 1: Identification : Claiming to be an identity when attempting to access a secured area or system
- 2: Authentication: Proving that you are that identity
- 3: Authorization: Defining the permissions (i.e., allow/grant and/or deny) of a resource and object access for a specific identity
- 4: Auditing: Recording a log of the events and activities related to the system and subjects
- 5: Accounting (aka accountability): Reviewing log files to check for compliance and violations in order to hold subjects accountable for their actions

AAA protocols include

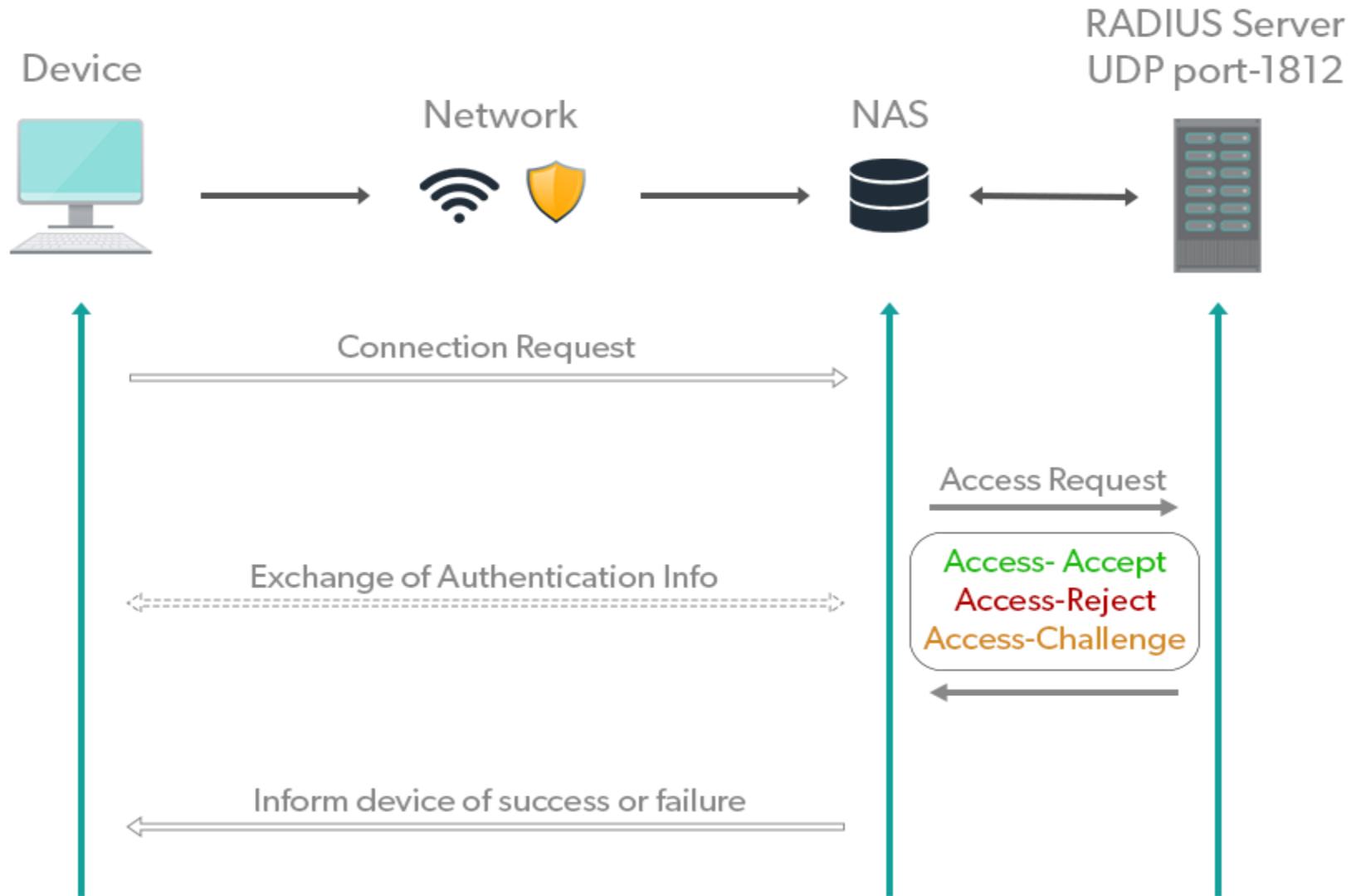
- Diameter, a successor to Remote Authentication Dial-In User Service (RADIUS)
- Terminal Access Controller Access-Control System (TACACS)
- Terminal Access Controller Access-Control System Plus (TACACS+) a proprietary Cisco Systems protocol that provides access for network servers, routers and other network computing devices.

RADIUS Component



- RADIUS uses a client-server model, and its three primary components include the:
- Supplicant: The supplicant is generally software built-in or installed ad hoc on a user's operating system that passes information about a user (username, password, etc.) to a second component, the network access server (NAS), along with an Access-Request query. An Access-Request query is just that, a request for access from a client to a server to utilize a resource like a network.
- Network Access Server: In the client/server architecture, the NAS acts as the client. NAS devices can be switches, routers, VPNs, or wireless access points (WAPs), among other things. The client/supplicant asks the server to determine if a user is allowed access to a particular resource — also called authentication.
- RADIUS server: The RADIUS server waits for requests from NAS devices. The benefit of RADIUS is that no matter what type of NAS you're trying to connect to, the RADIUS server centralizes authentication and simplifies the process. Authentication server that ensures the user is allowed to access the network with the proper permission levels. This server can also provide accounting functions for the purposes of billing, time tracking, and device/connection details.

RADIUS Authentication Process



- Advantages of the RADIUS Protocol
- Increased **network security** and control.
- Simplified password management.
- Centralized point for user and device authentication.
- Ideal tool for large networks that are managed by multiple IT personnel.
- Reduction in manual IT labor.
- Modern cloud and hosted RADIUS options exist for cloud-forward organizations.

Disadvantages of the RADIUS Protocol

- Traditionally implemented on-prem, but many modern IT environments don't fit this model.
- Setting up a RADIUS server can be difficult and time consuming.
- Configuration options are widespread, making setup complicated.
- The spread of options for implementing RADIUS can feel overwhelming and confusing.

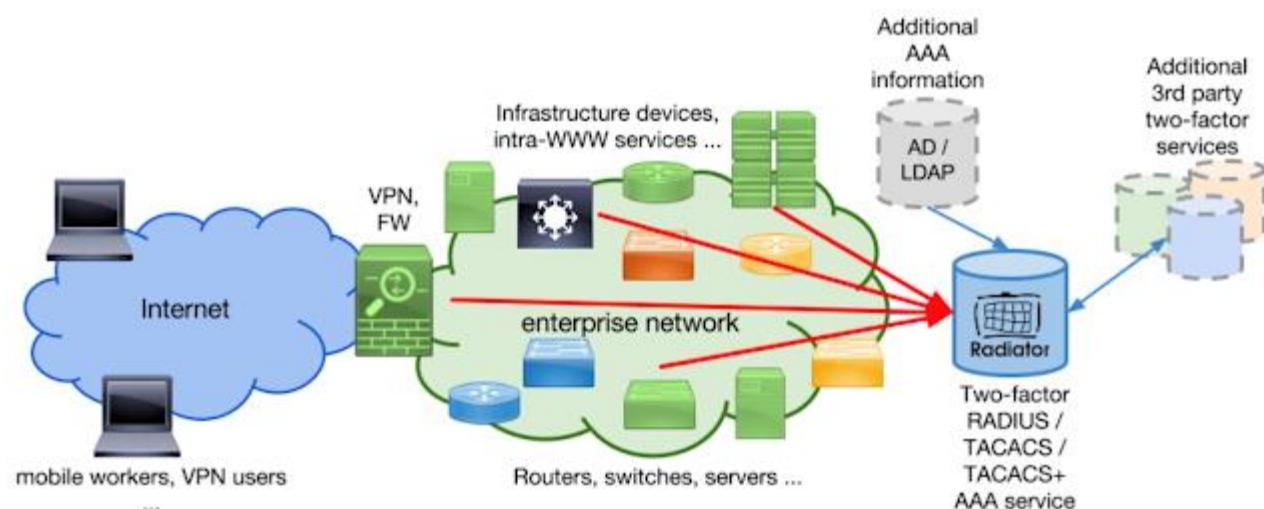
Tool: Cloud RADIUS

AAA protocols - RADIUS

- Remote Authentication Dial-In User Service (RADIUS) is a client/server protocol developed by the IETF. The RADIUS client is typically a NAS, and the RADIUS server is usually a daemon process running on a UNIX or Windows server.
- The RADIUS client (that is, the NAS) passes user information to designated RADIUS servers and acts on the returned response. RADIUS servers receive user connection requests via the NAS, authenticate the user, and then provide the NAS with configuration information necessary for it to deliver a specific service to the user.
- Transactions between the RADIUS client and RADIUS server are authenticated with a shared secret key, which is never sent over the network. In addition, user passwords are sent encrypted between the RADIUS client and RADIUS server to eliminate the possibility that someone snooping on an insecure network could determine a user's password.
- The accounting features of the RADIUS protocol can be used independently of RADIUS authentication or authorization. The NAS, which provides a service to the dial-in user (such as PPP or Telnet) is responsible for passing user accounting information to a designated RADIUS accounting server.

AAA protocol- TACACS

- The TACACS protocol, which has a very simple working mechanism, accepts a user query from a remote server and forwards this query for necessary action to the authentication server. The authentication server may allow or deny a user query on the host's behalf. The host is a system or platform that runs on the server. The query result is sent to the query initiator as a feedback response. The routing node used in dialup connections during the user login process allows or denies user access based on the query's response.



- TACACS+ which stands for Terminal Access Controller Access Control Server is a security protocol used in the AAA framework to provide centralized authentication for users who want to gain access to the network.
- **Features** – Some of the features of TACACS+ are:
 - Cisco developed protocol for AAA framework i.e it can be used between the Cisco device and Cisco ACS server.
 - It uses TCP as a transmission protocol.
 - It uses TCP port number 49.
 - If the device and ACS server are using TACACS+ then all the AAA packets exchanged between them are encrypted.
 - It separates AAA into distinct elements i.e authentication, authorization, and accounting are separated.
 - It provides greater granular control (than RADIUS) as the commands that are authorized to be used by the user can be specified.
 - It provides accounting support but is less extensive than RADIUS.

AAA protocol- TACACS+

- TACACS+ is an enhancement to the TACACS security protocol. TACACS+ improves on TACACS by separating the functions of authentication, authorization, and accounting (AAA) and by encrypting all traffic between the Vendor device and the TACACS+ server. TACACS+ allows for arbitrary length and content authentication exchanges, which allow any authentication mechanism to be utilized with your device.
- TACACS+ is extensible to provide for site customization and future development features. The protocol allows the device to request very precise access control and allows the TACACS+ server to respond to each component of that request.
-

Types of AAA servers

- **Access Network AAA (AN-AAA):** Communicates with the RNC in the Access Network (AN) to enable authentication and authorization functions to be performed at the AN. The interface between AN and AN-AAA is known as the A12 interface.
- **Broker AAA (B-AAA):** Acts as an intermediary to proxy AAA traffic between roaming partner networks (i.e., between the H-AAA server in the home network and V-AAA server in the serving network). B-AAA servers are used in CRX networks to enable CRX providers to offer billing settlement functions.
- **Home AAA (H-AAA):** The AAA server in the roamer's home network. The H-AAA is similar to the HLR in voice. The H-AAA stores user profile information, responds to authentication requests, and collects accounting information.
- **Visited AAA (V-AAA):** The AAA server in the visited network from which a roamer is receiving service. The V-AAA in the serving network communicates with the H-AAA in a roamer's home network. Authentication requests and accounting information are forwarded by the V-AAA to the H-AAA, either directly or through a B-AAA.

- **Working –**
The client of the TACACS+ is called Network Access Device (Nad) or Network Access Server (NAS). Network Access Device will contact the TACACS+ server to obtain a username prompt through **CONTINUE** message. The user then enters a username and the Network Access Device again contacts the TACACS+ server to obtain a password prompt (Continue message) displaying the password prompt to the user, the user enters a password, and the password is then sent to the TACACS+ server.
- The server can respond with one of the following reply messages:
- If the credentials entered are valid then the TACACS+ server will respond with an **ACCEPT** message.
- If the credentials entered are not valid then the TACACS+ server will respond with a **REJECT** message.
- If the link between the TACACS+ server and NAS or TACACS+ server is not working properly then it will respond with an **ERROR** message.
- If TACACS+ authorization is required, the TACACS+ server is again contacted and it returns an **ACCEPT** or **REJECT** authorization response. If the **ACCEPT** message is returned, it contains attributes that are used to determine services that a user is allowed to do.

- **Advantage –**
- Provides greater granular control than RADIUS.TACACS+ allows a network administrator to define what commands a user may run.
- All the AAA packets are encrypted rather than just passwords (in the case of Radius).
- TACACS+ uses TCP instead of UDP. TCP guarantees communication between the client and server.
- **Disadvantage –**
- As it is Cisco proprietary, therefore it can be used between the Cisco devices only. TACAS+ is an open standard RFC8907
- Less extensive support for accounting than RADIUS.

Tool: JunosE Software

SSO model

- Single sign-on (SSO) is an authentication method that enables users to securely authenticate with multiple applications and websites by using just one set of credentials.
- A user browses to the application or website they want access to, aka, the Service Provider.
- The Service Provider sends a token that contains some information about the user, like their email address, to the SSO system, aka, the Identity Provider, as part of a request to authenticate the user.
- The Identity Provider first checks to see whether the user has already been authenticated, in which case it will grant the user access to the Service Provider application and skip to step 5.
- If the user hasn't logged in, they will be prompted to do so by providing the credentials required by the Identity Provider. This could simply be a username and password or it might include some other form of authentication like a One-Time Password (OTP).
- Once the Identity Provider validates the credentials provided, it will send a token back to the Service Provider confirming a successful authentication.
- This token is passed through the user's browser to the Service Provider.
- The token that is received by the Service Provider is validated according to the trust relationship that was set up between the Service Provider and the Identity Provider during the initial configuration.
- The user is granted access to the Service Provider.

Is SSO Secure?

- “It depends.”
- A single sign-on solution can simplify username and password management for both users and administrators. Users no longer have to keep track of different sets of credentials and can simply remember a single more complex password. SSO often enables users to just get access to their applications much faster.
- SSO can also cut down on the amount of time the help desk has to spend on assisting users with lost passwords. Administrators can centrally control requirements like password complexity and multi-factor authentication (MFA). Administrators can also more quickly relinquish login privileges across the board when a user leaves the organization.
- Single Sign-On does have some drawbacks. For example, you might have applications that you want to have locked down a bit more. For this reason, it would be important to choose an SSO solution that gives you the ability to, say, require an additional authentication factor before a user logs into a particular application or that prevents users from accessing certain applications unless they are connected to a secure network.

Type of SSO

- Central Authentication Service (CAS): Developed by Shawn Bayern at Yale University, CAS differs from typical SAML SSO by enacting Server-to-Server communication. The Client Machine is used to initiate the token request, but the final verification is handled by a back-end communication between the CAS server and the Service Provider. CAS is a typical SSO protocol used in education organizations because of reliance on that extra, more direct verification. Like SAML, no passwords are exchanged through the SSO token. CAS is a common SSO protocol for higher education. Check out the SSO for Education page for more details.
- Shibboleth SSO: Shibboleth is another SSO protocol typically seen in educational organizations – specifically where a high number of institutions are federated to share applications and/or services. Shibboleth is built with SAML as a foundation but uses Discovery Service to improve upon SAML's organization of data from a large number of sources. Additionally, Shibboleth helps to automate the parsing of metadata to handle security certificate updates and other configurations that may be set by individual institutions within a federation
- Cookie-Based SSO: Works by using Web based HTTP Cookies to transport user credentials from browser to server without input from the user. Existing credentials on the client machine are gathered and encrypted before being stored in the cookie and sent to the destination server. The server receives the cookie, extracts and decrypts the credentials, and validates them against the internal server directory of users.

Type of SSO

- Claims-Based SSO: Claims (aka “assertions”) are created by a claims issuer that is trusted by multiple parties. Claims are typically packaged into a digitally signed token that can be sent over the network using Security Assertion Markup Language (SAML).
- NTLM-Based SSO: It is possible for a user to prove they know their password without actually providing the password itself. NTLM achieves this using a challenge and response protocol that first determines what type of NTLM and encryption mechanisms the client and server mutually support, then cryptographically hashes the user’s password and sends it to the server requiring authentication.
- Kerberos-based SSO: Kerberos enables users to log into their Windows domain accounts and then receive SSO to internal applications. Kerberos requires the user to have connectivity to a central Key Distribution Center (KDC). In Windows, each Active Directory domain controller acts as a KDC. Users authenticate themselves to services (e.g. web servers) by first authenticating to the KDC, then requesting encrypted service tickets from the KDC for the specific service they wish to use. This happens automatically in all major browsers using SPNEGO (see below).
- SPNEGO-based SSO: There are instances when the client application and remote server do not know what types of authentication the other one supports. This is when SPNEGO (Simple and Protected GSSAPI Negotiation Mechanism) can be used to find out what authentication mechanisms are mutually available. Some of these mechanisms can include Kerberos and NTLM authentication.

Type of SSO

- Reduced SSO: Reduced Single Sign-On is widely used for limiting the number of times a user will be required to enter in their credentials to access different applications. With critical applications, reduced SSO also offers a technique to make sure that a user is not signed on without a second factor of authentication, having been provided by the user.
- Enrollment-Based SSO: A user logging into a website may choose to have their credentials permanently remembered for that site. This is accomplished by creating an encrypted cookie on the user's machine for that web browser that contains the user's credentials. This cookie persists across different browser sessions and restarts of the machine but will be set to expire after a set period. The next time the user accesses the website, the server recognizes the cookie, decrypts it to obtain the user's credentials, and completely bypasses the login screen after validating them successfully.
- Form-Filling SSO: Form-filling allows for the secure storage of information that is normally filled into a form. For users that repeatedly fill out forms (especially for security access), this technology will remember/store all relevant information and secure it with a single password. To access the information, the user only has to remember one password and the Form-filling technology can take care of filling in the forms.
- Banner XE/Banner 9: Banner XE/Banner 9 supports CAS SSO. While not the newest SSO protocol, CAS SSO improves Banner's usability. Additionally, CAS SSO simultaneously increases the integration points for Banner in various institutions. Higher education institutions that are looking for additional, more feasible options, may land on using Banner XE/Banner 9 to fill the gap. CAS SSO opens Banner XE/Banner 9 up to more unique configurations and deployments.

Threat Agents - Anonymous Attacker

The intent and method targeted at the intentional exploitation of a vulnerability or a situation and method that may accidentally trigger a vulnerability. Synonymous with Threat Agent.

Confidentiality

Integrity

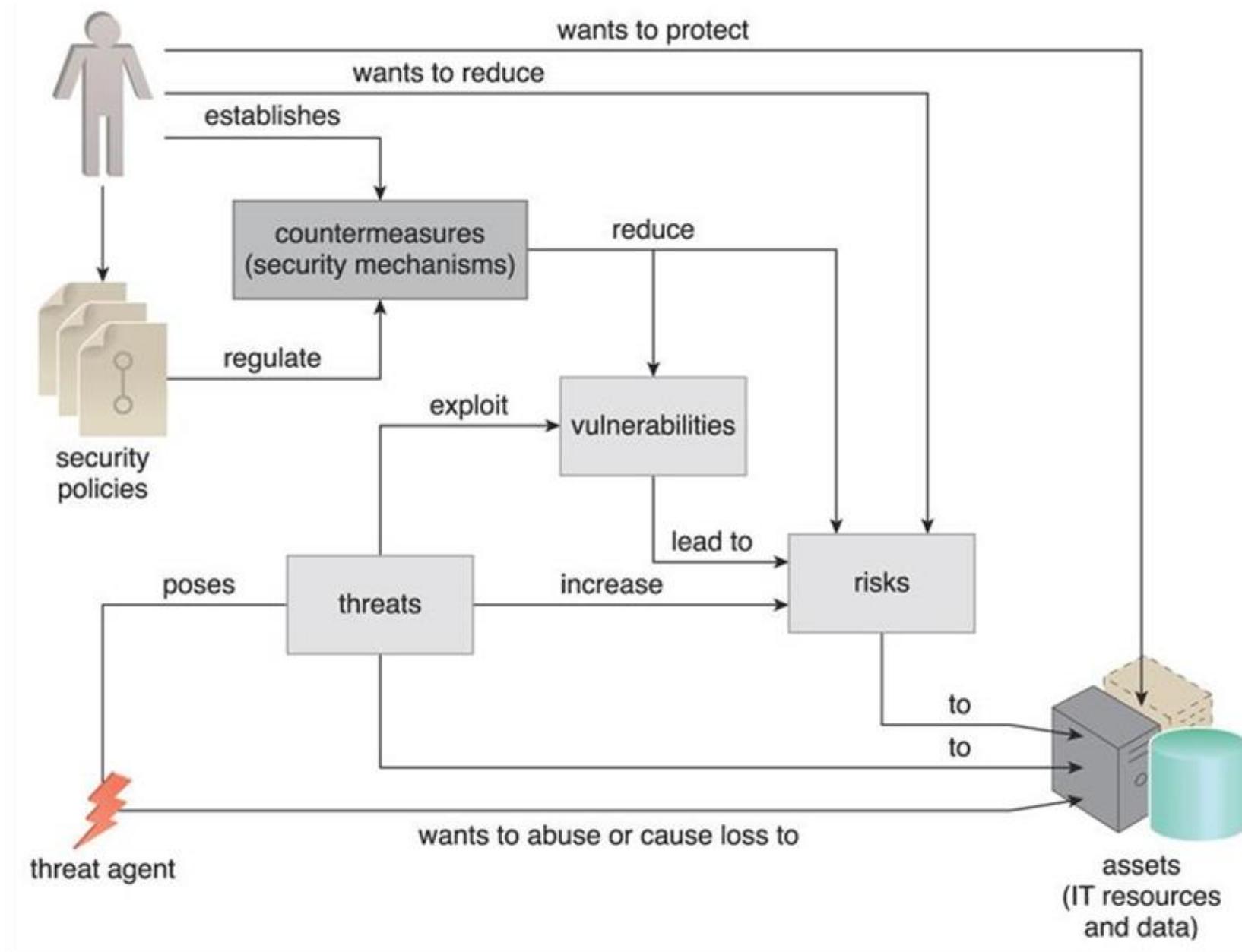
Authenticity

Availability

Threat

Vulnerability

Risk



Anonymous Attacker

An anonymous attacker is a non-trusted cloud service consumer without permissions in the cloud. It typically exists as an external software program that launches network-level attacks through public networks. When anonymous attackers have limited information on security policies and defenses, it can inhibit their ability to formulate effective attacks. Therefore, anonymous attackers often resort to committing acts like bypassing user accounts or stealing user credentials, while using methods that either ensure anonymity or require substantial resources for prosecution.

Malicious Service Agent

A malicious service agent is able to intercept and forward the network traffic that flows within a cloud (Figure 6.5). It typically exists as a service agent (or a program pretending to be a service agent) with compromised or malicious logic. It may also exist as an external program able to remotely intercept and potentially corrupt message contents.

Trusted Attacker

A trusted attacker shares IT resources in the same cloud environment as the cloud consumer and attempts to exploit legitimate credentials to target cloud providers and the cloud tenants with whom they share IT resources (Figure 6.6). Unlike anonymous attackers (which are nontrusted), trusted attackers usually launch their attacks from within a cloud's trust boundaries by abusing legitimate credentials or via the appropriation of sensitive and confidential information. Trusted attackers (also known as malicious tenants) can use cloud-based IT resources for a wide range of exploitations, including the hacking of weak authentication processes, the breaking of encryption, the spamming of e-mail accounts, or to launch common attacks, such as denial of service campaigns.

Malicious Insider

Malicious insiders are human threat agents acting on behalf of or in relation to the cloud provider. They are typically current or former employees or third parties with access to the cloud provider's premises. This type of threat agent carries tremendous damage potential, as the malicious insider may have administrative privileges for accessing cloud consumer IT resources.

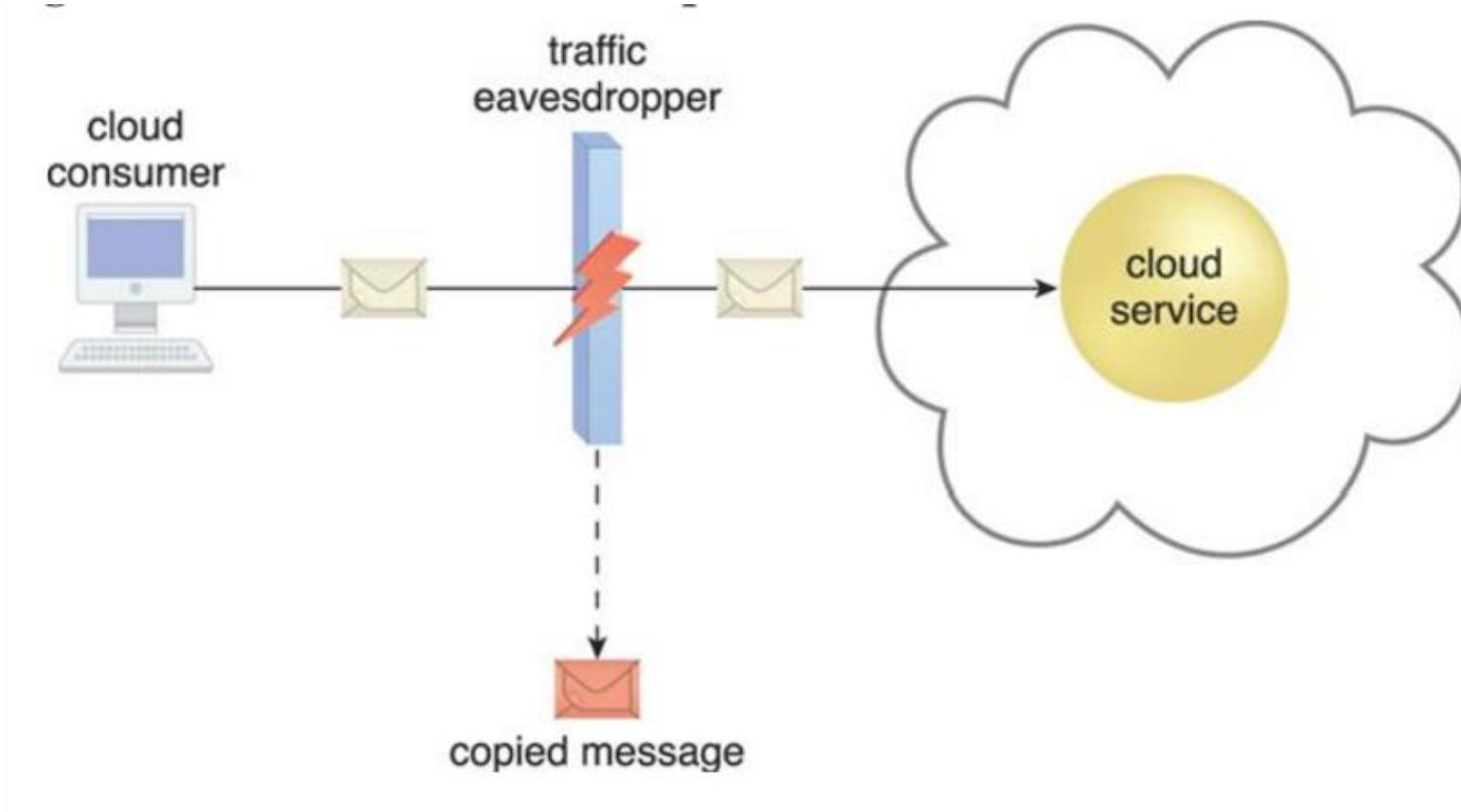
summary

- An anonymous attacker is a non-trusted threat agent that usually attempts attacks from outside of a cloud's boundary.
- A malicious service agent intercepts network communication in an attempt to maliciously use or augment the data.
- A trusted attacker exists as an authorized cloud service consumer with legitimate credentials that it uses to exploit access to cloud-based IT resources.
- A malicious insider is a human that attempts to abuse access privileges to cloud premises.

Cloud Security Threats

Traffic Eavesdropping Traffic eavesdropping occurs when data being transferred to or within a cloud (usually from the cloud consumer to the cloud provider) is passively intercepted by a malicious service agent for illegitimate information gathering purposes (Figure). The aim of this attack is to directly compromise the confidentiality of the data and, possibly, the confidentiality of the relationship between the cloud consumer and cloud provider. Because of the passive nature of the attack, it can more easily go undetected for extended periods of time.

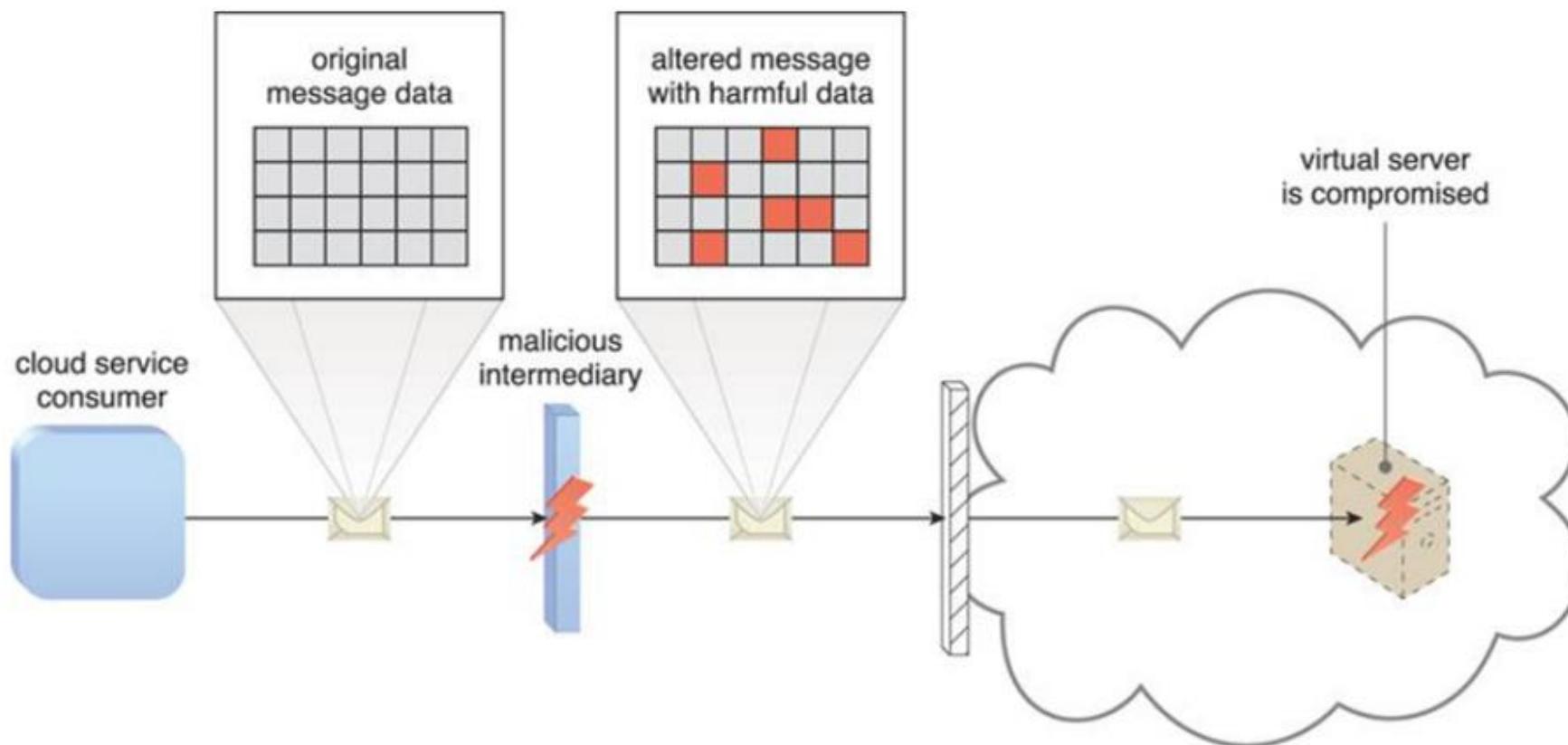
Traffic Eavesdropping



Malicious Intermediary

The malicious intermediary threat arises when messages are intercepted and altered by a malicious service agent, thereby potentially compromising the message's confidentiality and/or integrity. It may also insert harmful data into the message before forwarding it to its destination. Figure 6.9 illustrates a common example of the malicious intermediary attack.

Malicious Intermediary



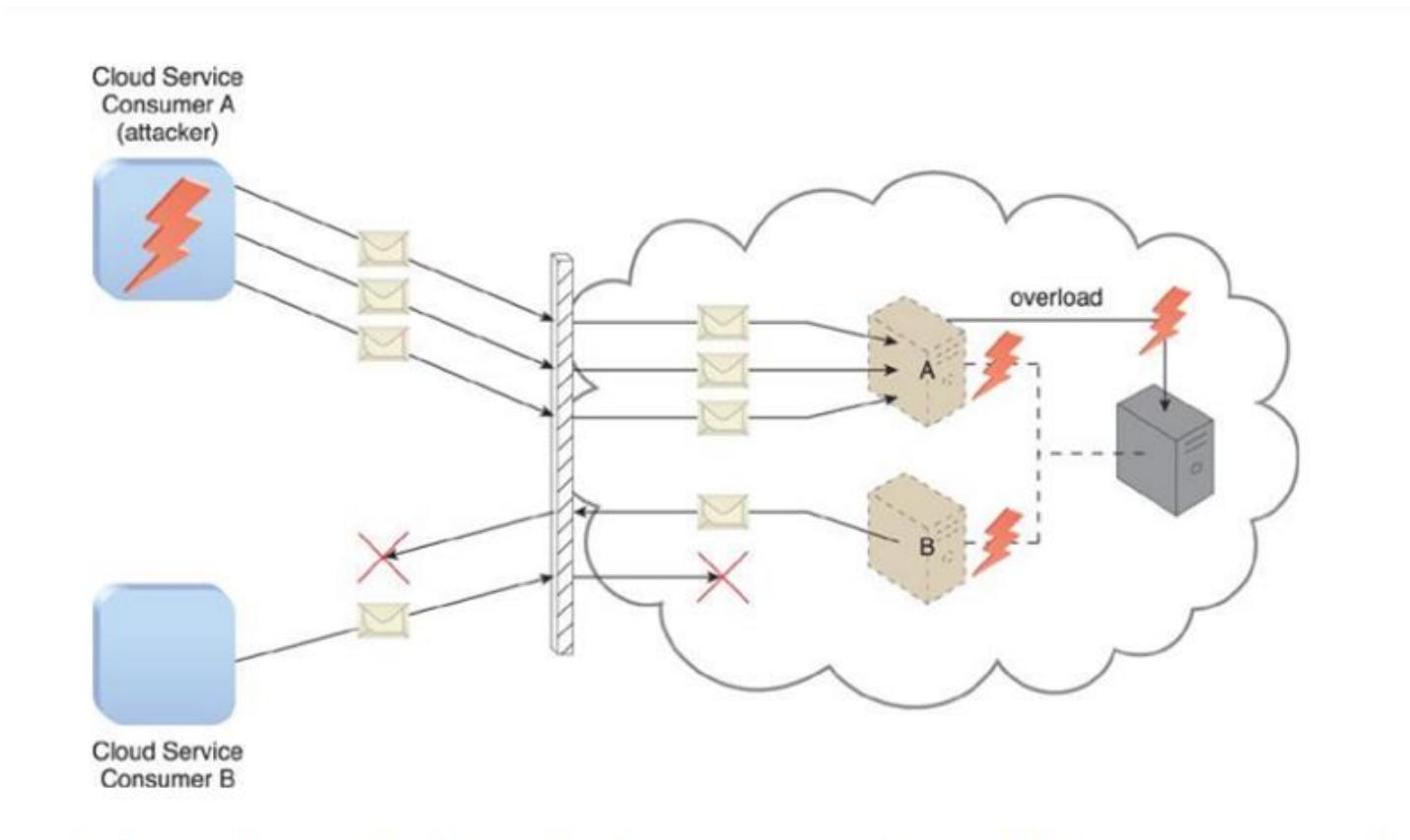
Denial of Service

The objective of the denial of service (DoS) attack is to overload IT resources to the point where they cannot function properly. This form of attack is commonly launched in one of the following ways:

- The workload on cloud services is artificially increased with imitation messages or repeated communication requests.
- The network is overloaded with traffic to reduce its responsiveness and cripple its performance.
- Multiple cloud service requests are sent, each of which is designed to consume excessive memory and processing resources.

Successful DoS attacks produce server degradation and/or failure, as illustrated in Figure 6.10.

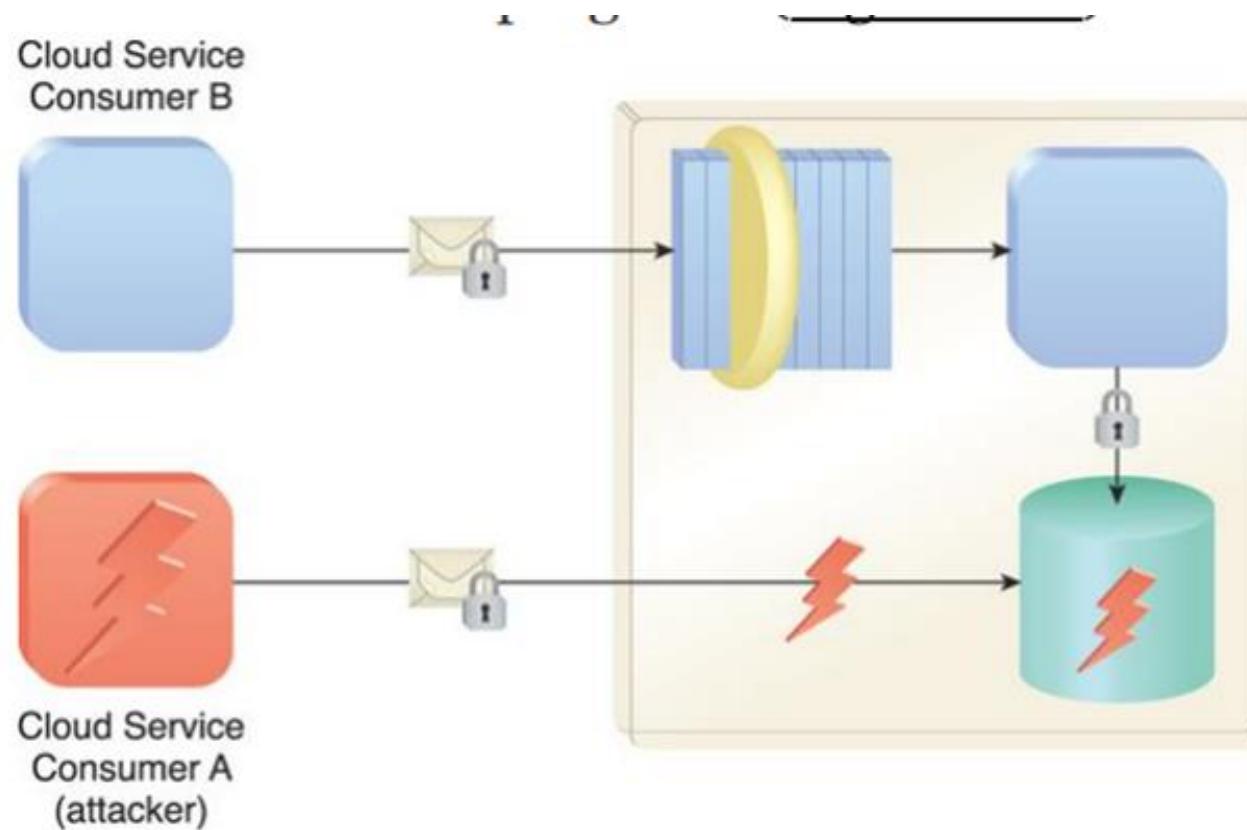
Denial of Service



Insufficient Authorization

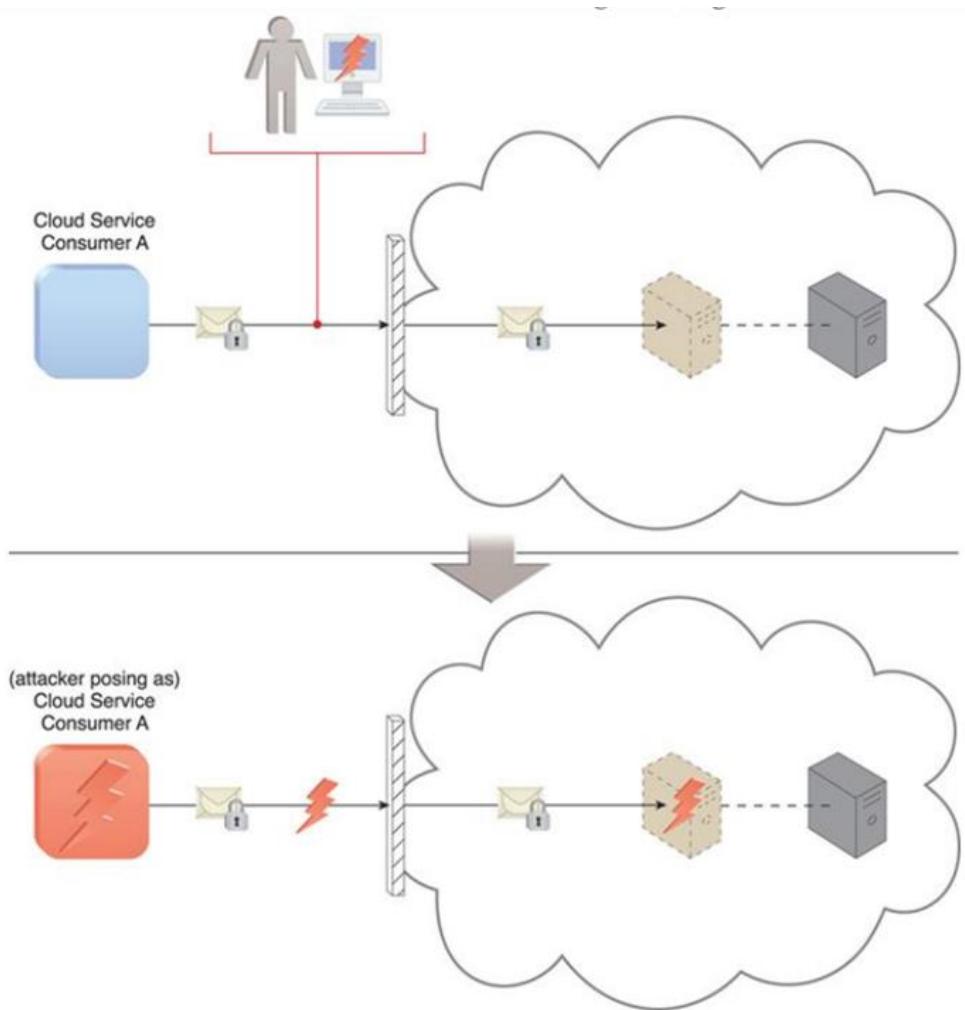
The insufficient authorization attack occurs when access is granted to an attacker erroneously or too broadly, resulting in the attacker getting access to IT resources that are normally protected. This is often a result of the attacker gaining direct access to IT resources that were implemented under the assumption that they would only be accessed by trusted consumer programs

Insufficient Authorization



A variation of this attack, known as weak authentication, can result when weak passwords or shared accounts are used to protect IT resources. Within cloud environments, these types of attacks can lead to significant impacts depending on the range of IT resources and the range of access to those IT resources the attacker gains

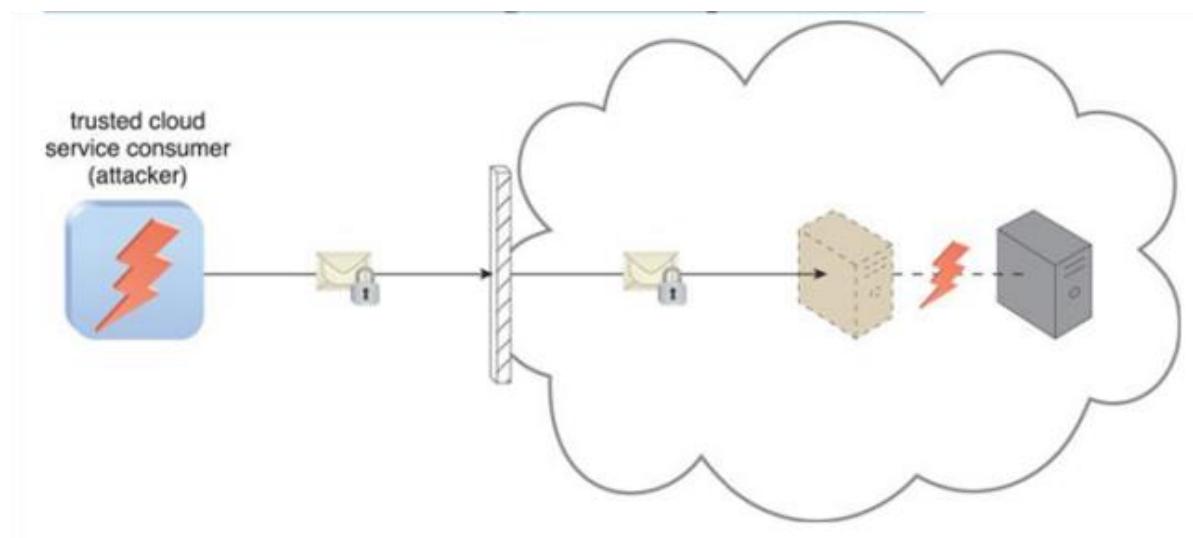
weak authentication



Virtualization Attack

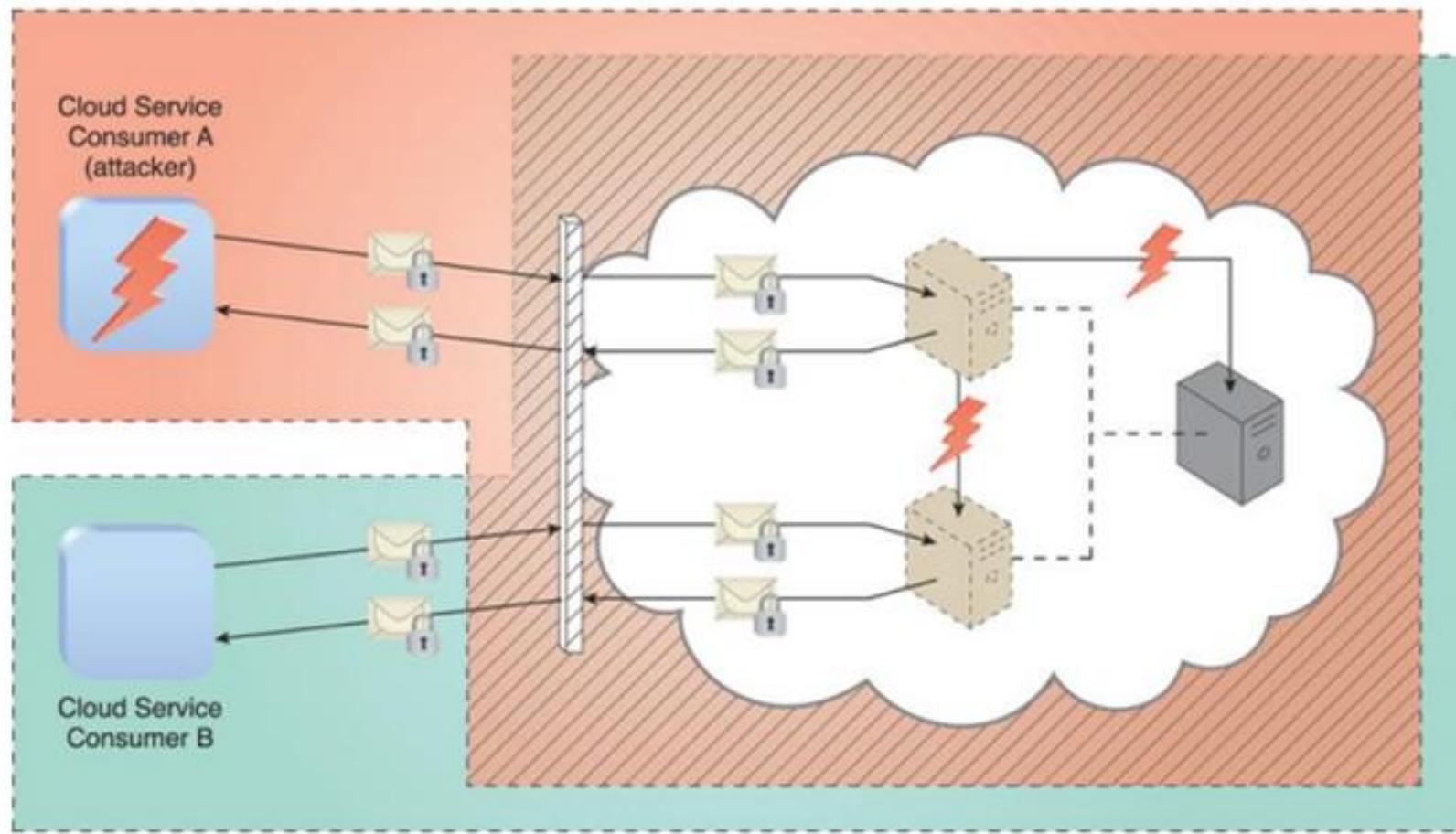
Virtualization provides multiple cloud consumers with access to IT resources that share underlying hardware but are logically isolated from each other. Because cloud providers grant cloud consumers administrative access to virtualized IT resources (such as virtual servers), there is an inherent risk that cloud consumers could abuse this access to attack the underlying physical IT resources. A virtualization attack exploits vulnerabilities in the virtualization platform to jeopardize its confidentiality, integrity, and/or availability. This threat is illustrated in Figure 6.13, where a trusted attacker successfully accesses a virtual server to compromise its underlying physical server. With public clouds, where a single physical IT resource may be providing virtualized IT resources to multiple cloud consumers, such an attack can have significant repercussions.

Virtualization Attack



Overlapping Trust Boundaries

If physical IT resources within a cloud are shared by different cloud service consumers, these cloud service consumers have overlapping trust boundaries. Malicious cloud service consumers can target shared IT resources with the intention of compromising cloud consumers or other IT resources that share the same trust boundary. The consequence is that some or all of the other cloud service consumers could be impacted by the attack and/or the attacker could use virtual IT resources against others that happen to also share the same trust boundary. Figure 6.14 illustrates an example in which two cloud service consumers share virtual servers hosted by the same physical server and, resultantly, their respective trust boundaries overlap.



Summary of Key Points

- Traffic eavesdropping and malicious intermediary attacks are usually carried out by malicious service agents that intercept network traffic.
- A denial of service attack occurs when a targeted IT resource is overloaded with requests in an attempt to cripple or render it unavailable. The insufficient authorization attack occurs when access is granted to an attacker erroneously or too broadly, or when weak passwords are used.
- A virtualization attack exploits vulnerabilities within virtualized environments to gain unauthorized access to underlying physical hardware. Overlapping trust boundaries represent a threat whereby attackers can exploit cloud-based IT resources shared by multiple cloud consumers.

Cloud Malware

Cloud security is complex. While cloud providers take responsibility for security of the infrastructure they manage, cloud users are responsible for configuring cloud security correctly, and securing their applications and workloads.

Misconfiguration and lack of security at the application level can lead to many security issues, and one of the most severe is malware infection in your cloud computing environment.

Malware in the cloud is a relatively new phenomenon, but cybercriminals quickly realized that cloud systems are an ideal media for spreading malware. Cloud-based systems are:

- Typically open to the Internet.

- Standardized and easy to learn for an attacker.

- Composed of a large number of entities, like virtual machines (VMs), containers or storage buckets, each of which can be a weak link for attackers to exploit.

In this article, you will learn:

- [The Rise of Cloud Malware](#)

- [5 Types of Cloud Malware Attacks](#)

 - [DDoS Attacks](#)

 - [Hypercall Attacks](#)

 - [Hypervisor DoS](#)

 - [Hyperjacking](#)

 - [Exploiting Live Migration](#)

- [3 Ways to Keeps your Cloud Malware-Free](#)

 - [Employee Education](#)

 - [Strengthen Access Control](#)

 - [Contain the Spread of Viruses with User Segmentation](#)

- [Cloud Security with NetApp Cloud Insights](#)

5 Types of Cloud Malware Attacks

DDoS Attacks

Large-scale botnets, composed of millions of compromised devices, are becoming widely available to attackers. Threat actors are offering botnets as a service for low prices, lowering the barrier of entry to anyone who wants to wage a DDoS attack.

In the cloud, a DDoS attack against your organization or any of your “neighbours” in the public cloud can affect the entire “neighborhood”, and the underlying cloud infrastructure. In addition, there is a constant risk that unattended VMs or containers will be compromised by attackers, and your cloud computing resources will be used for criminal activity.

Hypercall Attacks

In a hypercall attack, an attacker compromises an organization’s VMs using the hypercall handler. This is part of the virtual machine manager (VMM), deployed on every cloud machine in services like Amazon EC2. The attack grants attackers access to VMM permissions, and in some cases lets them execute malicious code on the VM.

5 Types of Cloud Malware Attacks

Hypervisor DoS

A hypervisor attack is an attack in which an attacker exploits the hypervisor, which controls multiple VMs on a virtual host. When the hypervisor is infected, malware can affect any of the VMs running on the host.

One possible consequence of an infected hypervisor is that virtual machine resource usage increases, resulting in denial of service to the entire host or even multiple hosts. Because hosts are typically interconnected, and do not always require authentication to connections from another host, they can easily infect other hosts, making the problem much more serious.

Hyperjacking

A hyperjacking attack is an attempt by an attacker to take control of the hypervisor, using a rootkit installed on a virtual machine. If the attacker is successful, they gain access to the entire host, and are able to modify the behavior of virtual machines, cause damage to running VMs, and even run new VMs for malicious activity.

Exploiting Live Migration

Attackers have learned that migration to the cloud or between clouds represents a major opportunity. When the organization performs an automated live migration, attackers can compromise the cloud management system, and manipulate it in several ways:

- Create multiple fake migrations, which becomes a DoS attack

- Migrate resources to a virtual network or cloud subscription under the attacker's control

- Make changes to migrated systems to make them vulnerable to future attacks

3 Ways to Keeps your Cloud Malware-Free

Employee Education

Strengthen Access Control

Multi-factor authentication—helps prevent account takeover, by requiring at least two authentication methods, one of which must be physically possessed by the user.

Least privilege—both users and integrated systems should only have access to resources they really need, and should have the exact level of permission they require for their role.

Contain the Spread of Viruses with User Segmentation

A cloud ecosystem has the following characteristics:

- Broad network connectivity
- Cloud consumers have less visibility and control.
- Changing system boundaries and overlapping roles/responsibilities between cloud Consumers and cloud providers.
- Multiple tenancies
- Data retention
- Measurable service
- Significant expansion in size (on demand), dynamics (elasticity, cost optimization), and complexity (automation, virtualization).

Cloud Infrastructure Mechanisms

Logical Network Perimeter, Virtual Server, Cloud Storage Device, Cloud Usage Monitor, Resource Replication Ready-Made Environment. Specialized Cloud Mechanisms - Automated Scaling Listener, Load Balancer, SLA Monitor, Pay-Per-Use Monitor Monitor, Pay-Per-Use Monitor, Audit Monitor, Failover System, Hypervisor, Resource Cluster, MultiDevice Broker, State Management Database. Types of Data Center – Enterprise Data Centers; managed Services Data Centers; Colocation; Cloud Data CentersDesign consideration for Private Cloud (Enterprise Data Centers), On Premise vs. Cloud propositions

CLOUD COMPUTING MECHANISMS

CLOUD COMPUTING MECHANISMS CATEGORIES

CLOUD INFRASTRUCTURE MECHANISMS

implement lower level virtual infrastructure components like Virtual Machines, Networks, Storage etc.

SPECIALIZED CLOUD MECHANISMS

ensure cloud is always running as per the expected and desired cloud characteristics
eg. load balancing, SLA monitoring, hypervisor etc.

CLOUD MANAGEMENT MECHANISMS

software to help coordinate the resources as per the management actions from users such as cloud consumers
eg VIM, Billing software etc.

CLOUD SECURITY MECHANISMS

Security techniques to keep the cloud secure

CLOUD COMPUTING MECHANISMS

CLOUD INFRASTRUCTURE MECHANISMS

- 1. Logical Network Perimeter
- 2. Virtual Server
- 3. Cloud Storage Device
- 4. Cloud Usage Monitor
- 5. Resource Replication
- 6. Ready-Made Environment

CLOUD MANAGEMENT MECHANISMS

- 17. Remote Administration System
- 18. Resource Management System
- 19. SLA Management System
- 20. Billing Management System

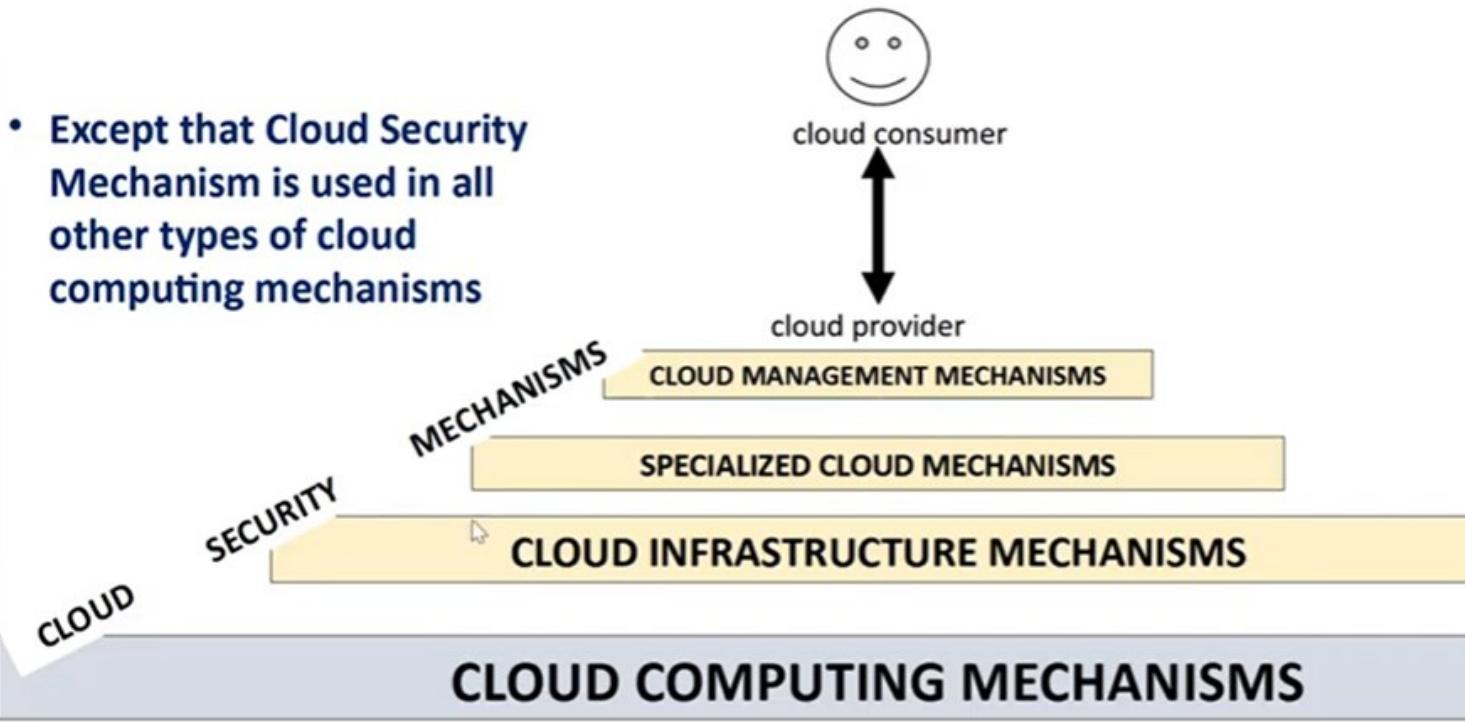
SPECIALIZED CLOUD MECHANISMS

- 7. Automated Scaling Listener
- 8. Load Balancer
- 9. SLA Monitor
- 10. Pay-Per-Use Monitor
- 11. Audit Monitor
- 12. Failover System
- 13. Hypervisor
- 14. Resource Cluster
- 15. Multi-Device Broker
- 16. State Management Database

CLOUD SECURITY MECHANISMS

- 21. Encryption
- 22. Hashing
- 23. Digital Signature
- 24. Digital Certificate
- 25. Public Key Infrastructure
- 26. Identity and Access Management
- 27. Single Sign-on
- 28. Cloud Security Groups
- 29. Hardened Virtual Server Images

- Note that each mechanism is supporting the underlying mechanism



Foundational building blocks of cloud environment

Infrastructure core components

6 pillars for infrastructure mechanism

1. Logical Network Perimeter - Techniques to implement networks in cloud
2. Virtual Server - techniques to implement machines in cloud
3. Cloud storage device - techniques to implement storage in cloud
4. Cloud usage monitor - techniques to monitor usage data of cloud resource
5. Resource replication - techniques to replicate/ duplicate resources such as networks, machines, softwares etc
6. Ready made environment- techniques to provide a ready made platform solution to do something

Logical Network Perimeter

The logical network perimeter establishes a virtual network boundary that can encompass and isolate a group of related cloud-based IT resources that may be physically distributed. It is defined as the isolation of a network environment from the rest of a communications network.

The logical network perimeter can be implemented to:

- isolate IT resources in a cloud from non-authorized users
- isolate IT resources in a cloud from non-users
- isolate IT resources in a cloud from cloud consumers
- control the bandwidth that is available to isolated IT resources

Logical Network Perimeter

Logical network perimeters are typically established via network devices that supply and control the connectivity of a data center and are commonly deployed as virtualized IT environments that include:

- *Virtual Firewall* – An IT resource that actively filters network traffic to and from the isolated network while controlling its interactions with the Internet.
- *Virtual Network* – Usually acquired through VLANs, this IT resource isolates the network environment within the data center infrastructure.

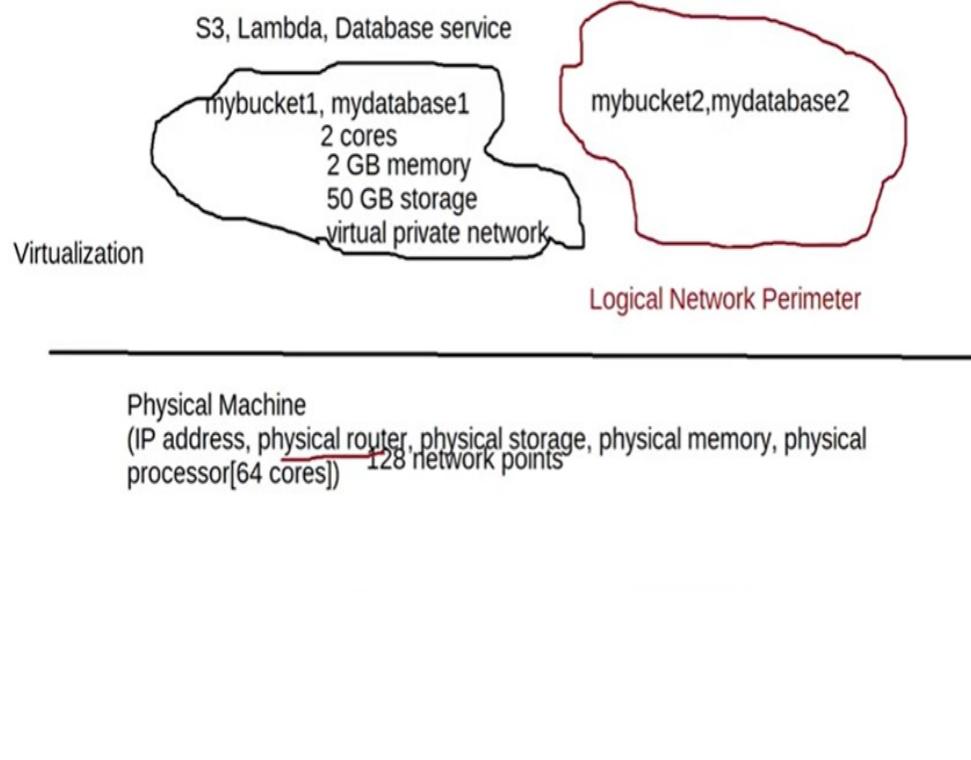
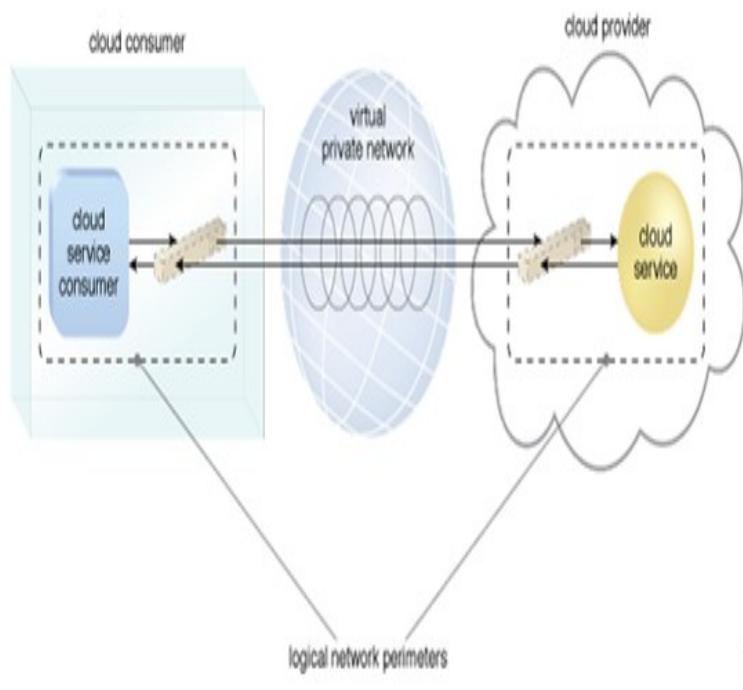
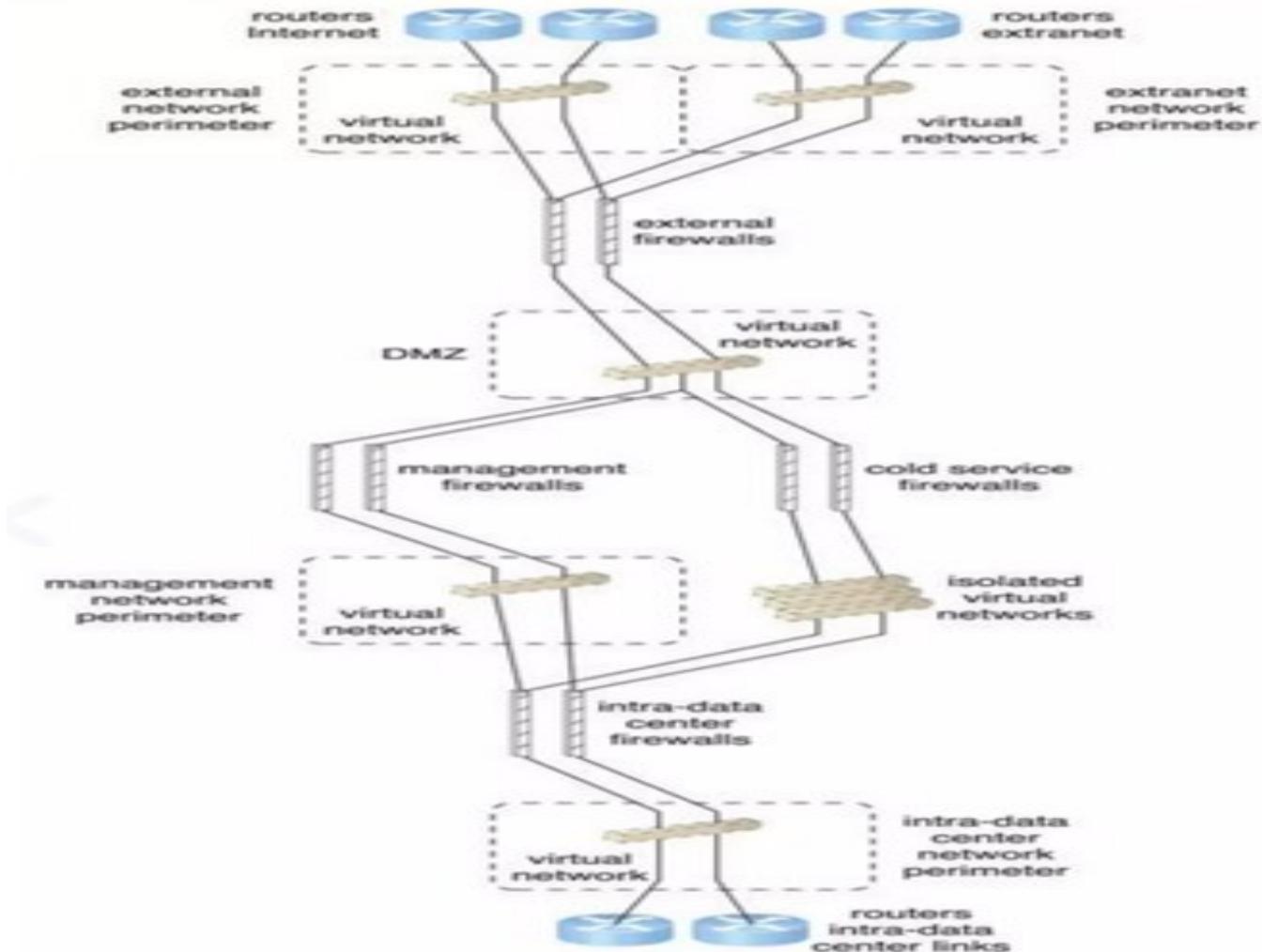


Figure 2 - Two logical network perimeters surround the cloud consumer and cloud provider environments.



Logical Network Perimeter

□ Definition

- The isolation of a network environment from the rest of a general communication network
- A virtual network boundary that can encompass and isolate a group of related cloud-based IT resources that may be physically distributed

□ Implementation mechanisms

- Isolate IT resources in a cloud from non-authorized users / non-users / other cloud consumers
- Control the bandwidth that is available to isolated IT resources

□ Isolation mechanisms

- Isolated via network devices that supply and control the connectivity of a data center and are commonly deployed as virtualized IT environments that include:
 - **Virtual firewall:** an IT resource that actively filters network traffic to and from the isolated network while controlling its interactions with the Internet
 - **Virtual network:** usually acquired through **VLANs** – this IT resource isolates the network environment within the data center infrastructure
- The cloud consumer's IT environment and the cloud IT resources are connected via so called **VPN** implemented by point-to-point encryption of the data packets between two endpoints



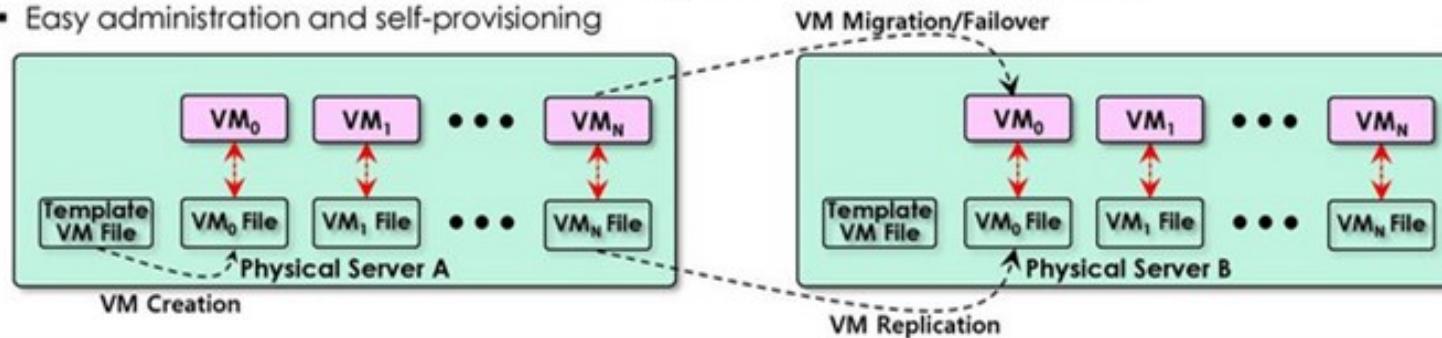
Virtual Server

□ Definition

- A logical server resource created by virtualization software to emulate a physical server resource
- A virtual resource that can be instantly provided regardless of the availability of physical resource

□ Virtualization mechanisms

- A core technology enabling cloud computing service platform & the most fundamental building block of cloud environment
- Allow multiple cloud consumers to share the same physical server with the illusion that each cloud consumer solely own the server (multitenancy)
- Instant VM creation by copying template VM image file (on-demand resource provisioning)
- On-line scaling up/down (by allocating more/less cores) or out/in (by adding/removing VM instances)
- On-line server migration by replicating VM image file to the other physical host and switching over
- Seamless service failover by reinstating the same VM image file within the same physical host or by replicating the VM mage file and reinstating it between different physical hosts
- Effective load balancing by even provisioning and real-time on-line migration
- Easy administration and self-provisioning



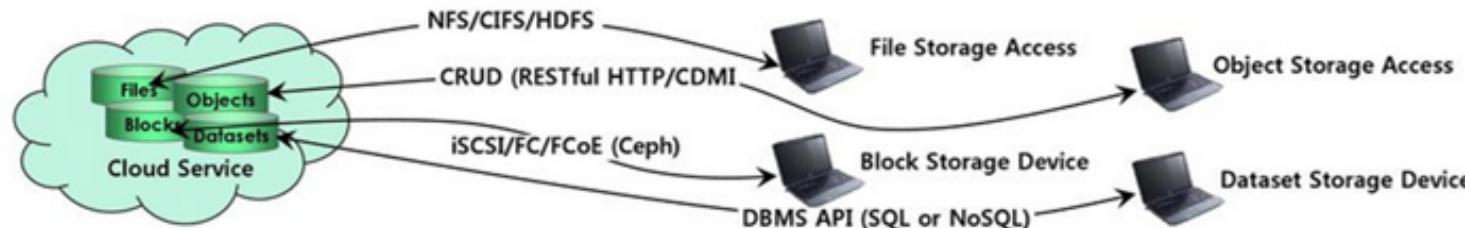
Cloud Storage Device – 1/3

□ Definition & concerns

- Storage devices that are designed specially for cloud-based provisioning
- Possibly virtualized and/or distributed in general
- Usually upper-bounded due to capacity allocation in support of the **pay-per-use** mechanism
- Open to remote access via cloud storage services (via **Representational State Transfer** or **RESTful APIs**)
- Main concern: the security, integrity, and confidentiality of data
- Legal and regulatory issues for relocating data across geographical or national boundaries
- Performance issues as well due to remote and/or large data access

□ Cloud storage levels

- Cloud storage device mechanisms provide common logical units of data storages:
 - **Files:** collections of data are grouped into files that are located in folders.
 - **Blocks:** the lowest level of storage and the closest to the hardware – a block is the smallest unit of data that is still individually accessible.
 - **Datasets:** sets of data are organized into a table-based, delimited or record format.
 - **Objects:** data and its associated metadata are organized as Web-based resources.
- Each data storage levels associated with a certain type of technical interface or APIs



Cloud Storage Device – 2/3

❑ Network Storage Interfaces

- Storage devices in compliance with industry standard protocols such as **SCSI** for storage blocks, the server message block (**SMB**), common Internet file system (**CIFS**) and network file system (**NFS**) for file and network storage
- Destitute storage devices for large data sets such as **HDFS** or **Ceph**
- File interfaces: data file format with different size and complex store/retrieve mechanism – less optimal in terms of performance
- Block interface: data block format with fixed size and simple (block number/LUN) store/retrieve mechanism – optimal in terms of performance, but need for filesystem mechanism on top of it anyway (except for raw device usage)

❑ Object Storage Interfaces

- Data referenced and stored as Web resources – object (all or nothing access – **put/get/post/delete**)
- Accessed via **REST** or Web service-based cloud services using **HTTP** as the prime protocol
- Defined by Storage Networking Industry Association's Cloud Data Management Interface (**SNIA's CDMI**)

❑ Database Storage Interfaces

- Cloud storage device mechanism typically supporting a query language in addition to basic storage operations
- Storage management carried out using a standard API or an administrative user interface divided into two main categories: **relational data storage & non-relational data storage**

➤ Relational data storage

- Traditional structured data storage stored in **tables** composed of **rows** and **columns** – **RDBMS**
- Related data connected to each other via **index** through **normalization** process in order to protect data **integrity** and to eliminate data **redundancy**

Cloud Storage Device – 3/3

- Accessed via industry standard **Structured Query Language (SQL)**
 - Accessed via a number of different commercial or free database products such as **Oracle, DB2, MSSQL, MySQL, Tibero**, etc. with their own access protocols
 - Challenges with scaling and performance due to the nature of virtualization, remote access and distributed aspect of the underlying storage device or mechanism
 - Higher processing overhead and latency for database with complex relational or large data set, especially when remotely accessed via cloud interface
 - Required to satisfy **ACID** (Atomicity, Consistency, Isolation, Durability) properties
- Non-relational data storage
- Commonly known as **NoSQL** storage with non-traditional & unstructured data
 - Weakening consistency model with **BASE** (Basically Available, Soft states, Eventual consistency) properties
 - Purpose of avoiding the potential complexity and processing overhead that can be imposed by relational database
 - More horizontally scalable than relational storages – scale-out rather than scale-up (vertical)
 - Trade-off: possible consistency violation with no traditional relational database functions such as **transactions or joins**
 - Data denormalized while being exported from relational storage to non-relational storage resulting in size increase due mainly to increased data redundancy

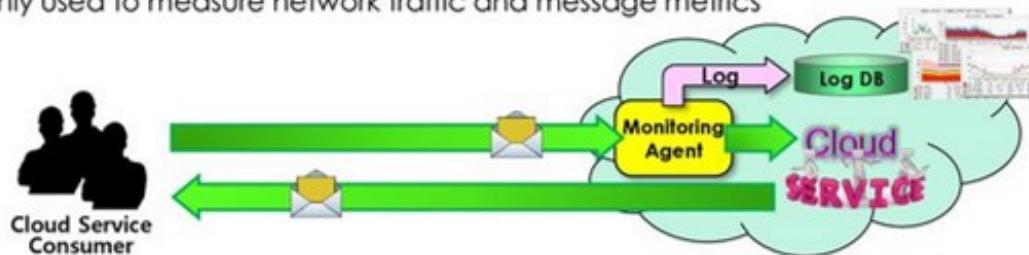
Cloud Usage Monitor – 1/2

Definition

- A lightweight and autonomous software program responsible for collecting and processing IT resource usage data
- Different formats depending on the type of usage metrics and the way usage data needs to be collected
- Agent-based implementation in which usage data are collected and forwarded to a log database for post processing and reporting purpose

Monitoring agent

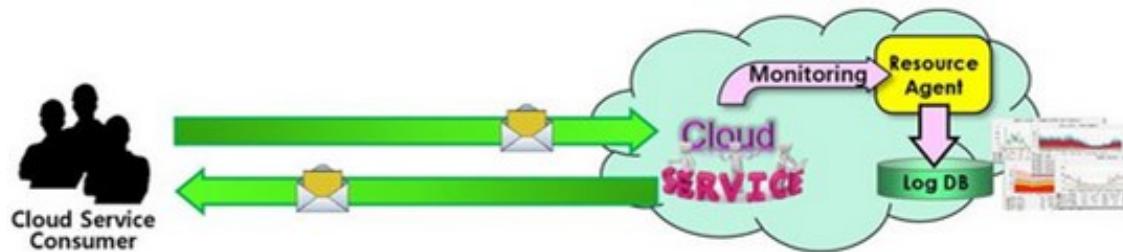
- An intermediary, event-driven program that exists as a service agent and resides along existing communication paths to transparently monitor and analyze dataflow
- Commonly used to measure network traffic and message metrics



Resource agent

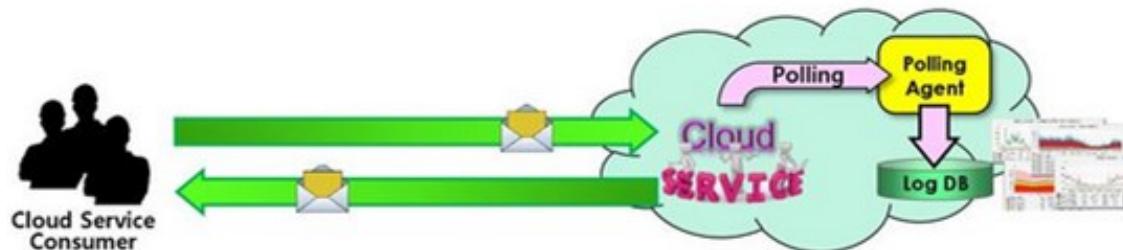
- A processing module that collects usage data by having **event-driven** interactions with specialized resource software
- Commonly used to monitor usage metrics based on pre-defined, observable events at the resource software level such as **initiating, suspending, resuming** and **vertical scaling**

Cloud Usage Monitor – 2/2



Polling agent

- A processing module that collects cloud service usage data by polling IT resource
- Commonly used to **periodically** monitor IT resource status such as **uptime** and **downtime**



Consumer consideration

- How do consumers trust the cloud provider's usage statistics?
- What if those agents work too hard?

Resource Replication

□ Definition

- Periodic replication of cloud resource including data
- Primarily to enhance the **availability** and the **performance** of IT resource

□ Nature of virtualization technology

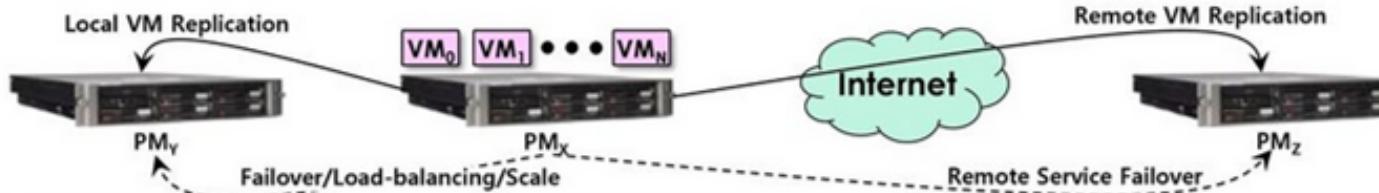
- Everything – VM, configuration, memory status and data – stored in **image files** in virtualized environment
- Resource replication easily done via replicating image files

□ Availability enhancement

- Each VM image files **periodically replicated** locally (within the same data center) and/or remotely (to another data center)
- A new VM instance activated with the same service access point (IP) from the replica locally or remotely in case of a VM/PM failure minimizing VM downtime

□ Performance enhance

- A VM image files replicated onto another physical machine with enough processing power when the performance of a VM is degraded
- VM service switched over by deactivating the current VM instance and activating a new VM instance from the replica on another physical machine
- VM migration mainly for the purpose of load balancing and on-line VM scale-up/down



Ready-made Environment

Definition

- A PaaS cloud delivery model that represents a pre-defined, cloud-based platform comprised of a set of already installed IT resources
- Ready to be used and customized by a cloud consumer

User-defined environment

- Utilized by cloud consumers to remotely develop and deploy their own services and applications with a cloud
- Many IT resources pre-installed such as **database**, **middleware**, **development tools** and **governance tools** as specified by consumers in general (master or template VM image files)
- Typically equipped with a complete software development kit (**SDK**) that provides cloud consumers with programmatic access to the development technologies that comprise their preferred programming stacks
- Middleware for multitenant platforms to support the development and deployment of Web application

Specialized Cloud Mechanisms

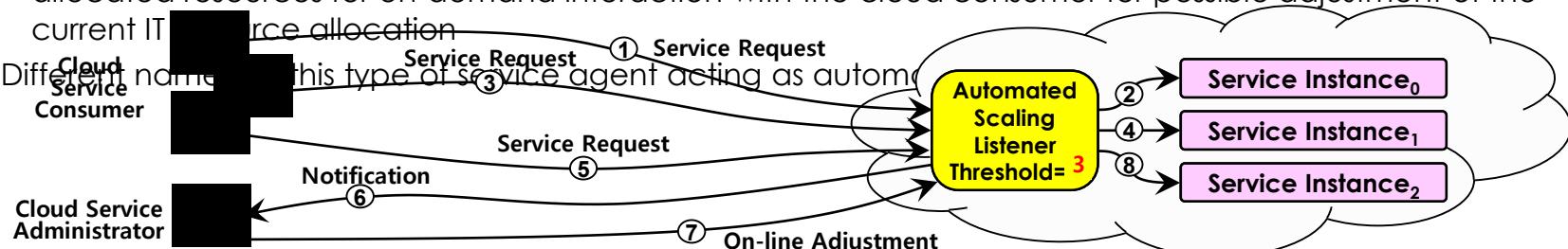
Automated Scaling Listener – 1/2

□ Definition

- A service agent that monitors and tracks communications between cloud service consumers and cloud services for dynamic scaling purpose
- Deployed within the cloud, typically near the firewall from where these agents automatically track workload status information

□ Implementation mechanisms

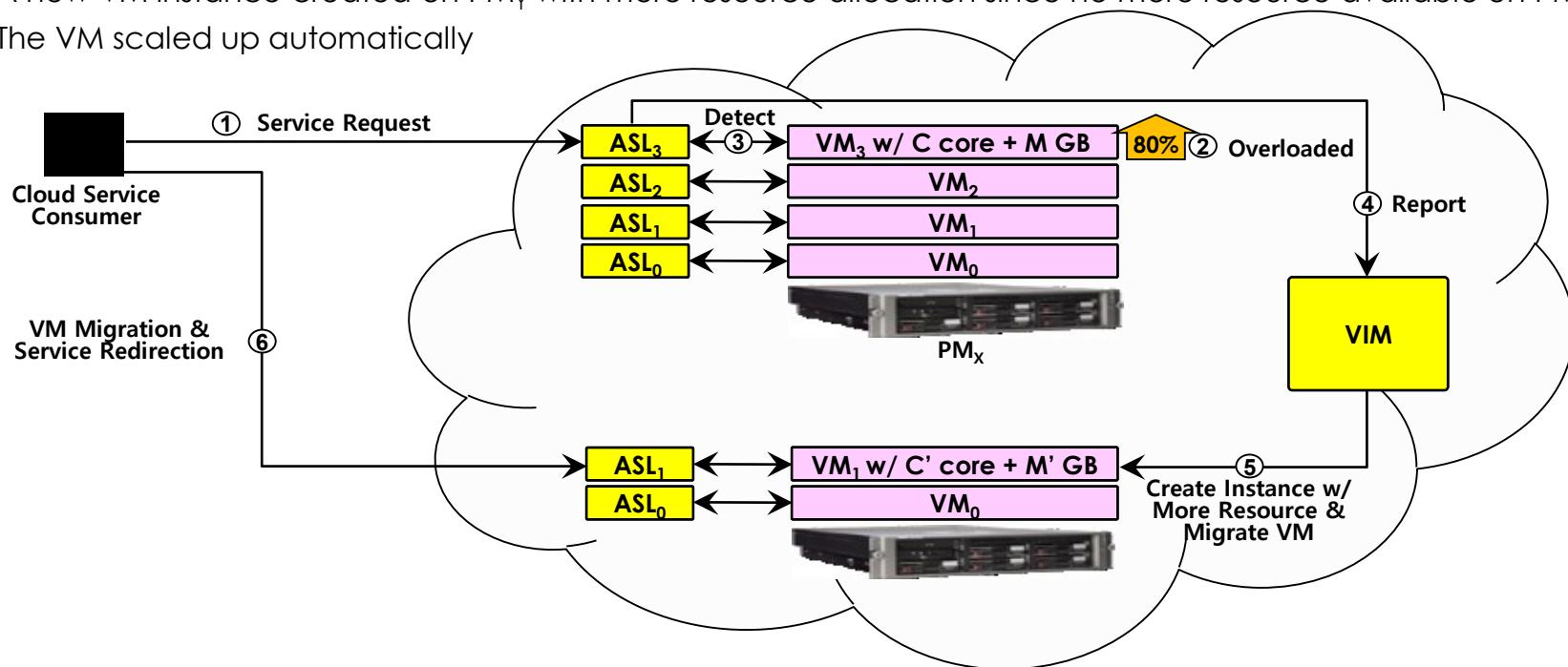
- Typical cloud workload measure = requests generated by cloud consumers + backend processing demands triggered by certain types of request – e.g., long and heavy processing demand triggered by a single short request
- Different types of responses to workload fluctuation conditions such as:
 - Automatically scaling IT resources out or in based on parameters previously defined by the cloud consumer (commonly referred to as **auto-scaling**)
 - Automatic notification of the cloud consumer when workloads exceed current threshold or fall below allocated resources for on-demand interaction with the cloud consumer for possible adjustment of the current IT resource allocation



Automated Scaling Listener – 2/2

□ Automated scaling example

- A VM instance created on PM_x by request and overloaded (over X CPU utilization for Y seconds)
- Detected by Automated Scaling Listener and reported to VIM
- A new VM instance created on PM_y with more resource allocation since no more resource available on PM_x
- The VM scaled up automatically



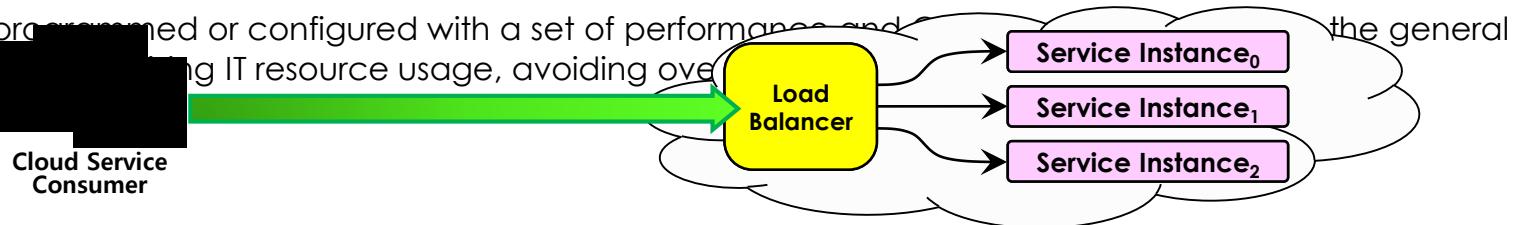
Load Balancer – 1/2

□ Definition

- A runtime agent to balance a workload across two or more IT resources to increase performance and capacity beyond what a single IT resource can provide
- An attempt to distribute overall workload as evenly as possible across all available IT resources

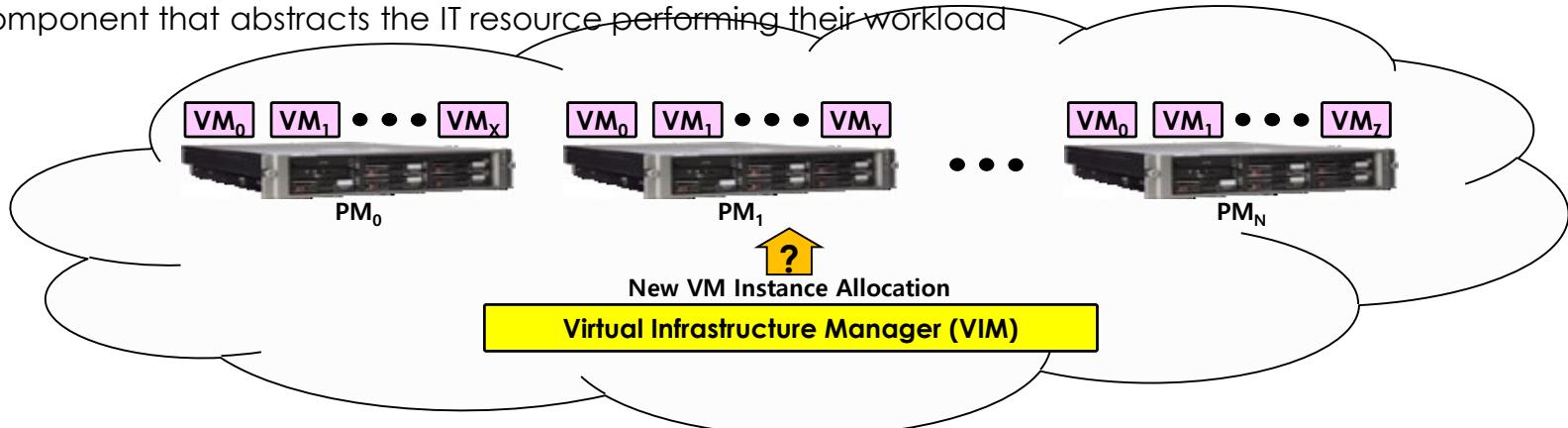
□ Implementation mechanisms

- Two different levels of load balancing: for each consumer service or for overall cloud service
- A range of specialized runtime workload distribution functions include:
 - **Round-robin distribution:** a simple division of labor distribution – one after another
 - **Less load first distribution:** assign a new request to one with the smallest current load
 - **Asymmetric distribution:** larger workloads are issued to IT resources with higher processing capacities
 - **Workload prioritization:** workloads are scheduled, queued, discarded and distributed according to their priority levels
 - **Content-aware distribution:** requests are distributed to different IT resources as dictated by the request content
- Typically programmed or configured with a set of performance objectives regarding IT resource usage, avoiding overutilization



Load Balancer – 2/2

- The load balancer mechanism typically resides:
 - Multi-layer network switch (layer 4 or higher)
 - Dedicated hardware appliance
 - Dedicated software-based system (common in server operating systems – Linux Virtual Server)
 - Service agent (usually controlled by cloud management software)
- Traditional load balancing logics are typically located on the communication path between the IT resources generating the workload and the IT resources performing the workload processing.
- Also embedded in all type of resource allocator module like VIM
- Implemented as a transparent agent that remains hidden from the cloud service consumers, or as a proxy component that abstracts the IT resource performing their workload



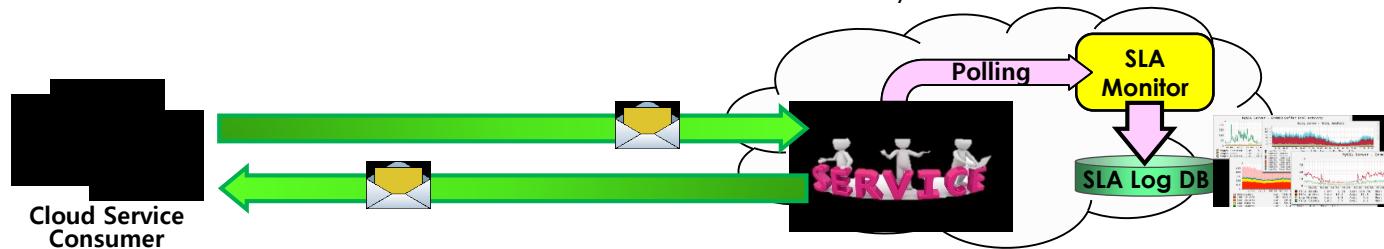
SLA Monitor

□ Definition

- A resource agent that monitors and keeps track of the runtime performance of cloud services to ensure that they are fulfilling the contractual QoS requirements that are published in SLAs

□ Implementation mechanisms

- The SLA monitor collects the runtime log data of the given cloud service by periodic **polling**.
- SLA management system processes the runtime log data collected by the SLA monitor and aggregates them into SLA reporting metrics as **uptime** and **downtime**.
- The system proactively repair or failover cloud services when exception condition occurs, such as when the SLA monitor reports a cloud service as "**down**."
- The SLA monitor plays a role of **health checker** or **heartbeat checker** in traditional HA systems for the give cloud services.
- The SLA monitor also collects any other runtime status statistics that are specified in SLA such as network service bandwidth or core allocation etc. in addition to service availability.



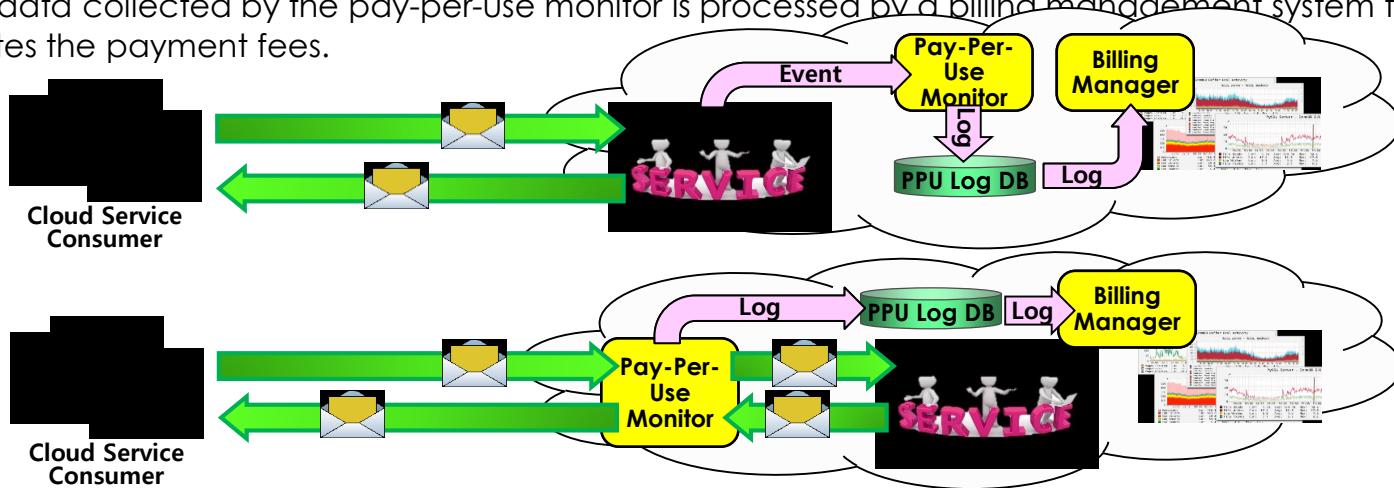
Pay-Per-Use Monitor

□ Definition

- An **event-driven** or **monitoring** resource agent that measures cloud-based IT resource usage in accordance with predefined pricing parameters and generates usage logs for fee calculations and billing purpose

□ Implementation mechanisms

- Typical monitoring variables include:
 - Request/response message quantity
 - Transmitted data volume
 - Bandwidth consumption
- The log data collected by the pay-per-use monitor is processed by a billing management system that calculates the payment fees.



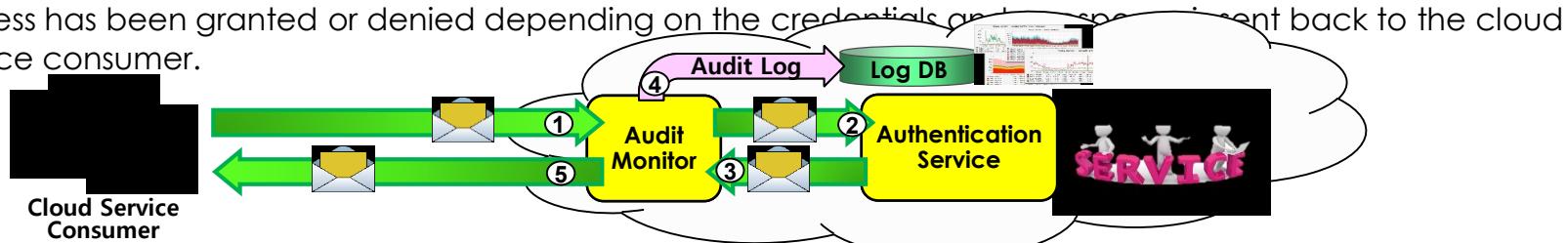
Audit Monitor

□ Definition

- A monitoring agent that collects audit tracking data for networks and IT resources in support of (or dictated by) regulatory and contractual obligations

□ Implementation mechanisms

- Intercepts “login” requests and stores the requestor’s credentials, as well as both failed or successful login attempt in a log database for future audit reporting purpose:
 - ① A cloud consumer requests access to a cloud service by sending a login request message with security credentials.
 - ② The audit monitor intercepts the message and forwards it to the authentication service.
 - ③ The authentication service processes the security credentials and generates a response message with the results from the login attempt – success or fail.
 - ④ The audit monitor intercepts the response message and stores the entire collected login event details in the log database in accordance with the organization’s audit policy requirements.
 - ⑤ Access has been granted or denied depending on the credentials and response sent back to the cloud service consumer.



Failover System – 1/3

❑ Definition

- A system to increase the reliability and availability of IT resources by using established clustering technology to provide redundant implementations

❑ Implementation mechanisms

- Basically, automatic switching over to a redundant or standby IT resource instance whenever the currently active IT resource becomes unavailable
- Commonly used for mission-critical systems and reusable services that can introduce a single point of failure for multiple applications
- Possibly system or service failover between more than one geographical region so that each location hosts one or more redundant implementations of the same IT resources
- Based on two different mechanisms – shared storage or resource replication:
 - Shared storage
 - Physically shared storage via SAN – configurable physical path between storage and multiple hosts
 - No need to replicate VM image files and data – no performance degradation & data/state loss
 - Expensive to implement – SAN storage, SAN switch, 2 HBA per each host
 - Resource replication
 - DAS-based & no physically shared storage – need to replicate all virtual IT resources and data
 - Periodic VM image files and data replication – performance degradation & data/state loss
 - Cost-effective, but less reliable implementation

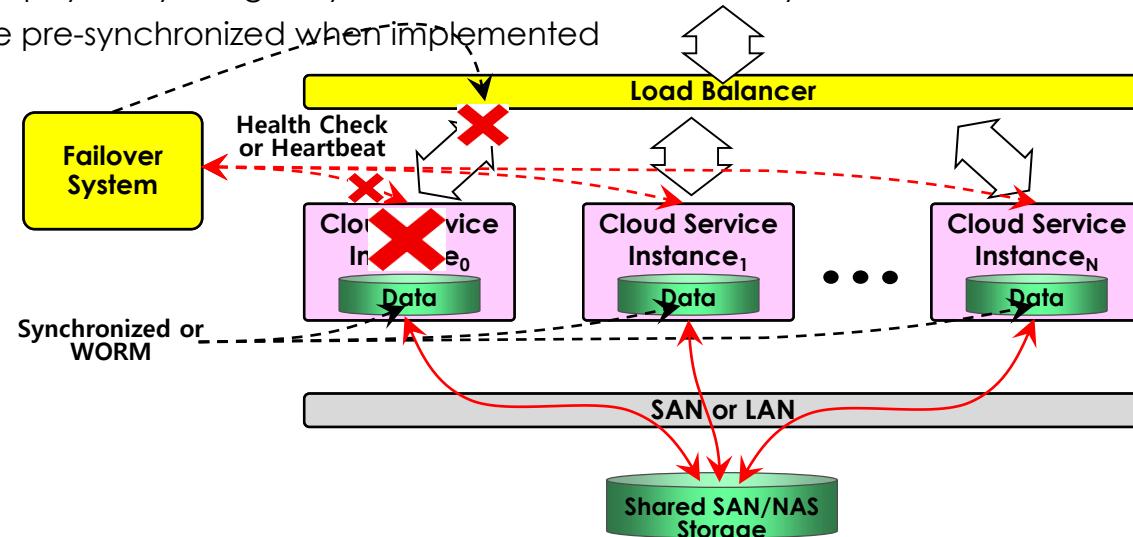
오상규

Basically redundant IT resource instances + close monitoring of active IT resource instance for the detection of failures or unavailability conditions + automatic service switchover mechanism

Failover System – 2/3

❑ Active-active / active-all / N-way active

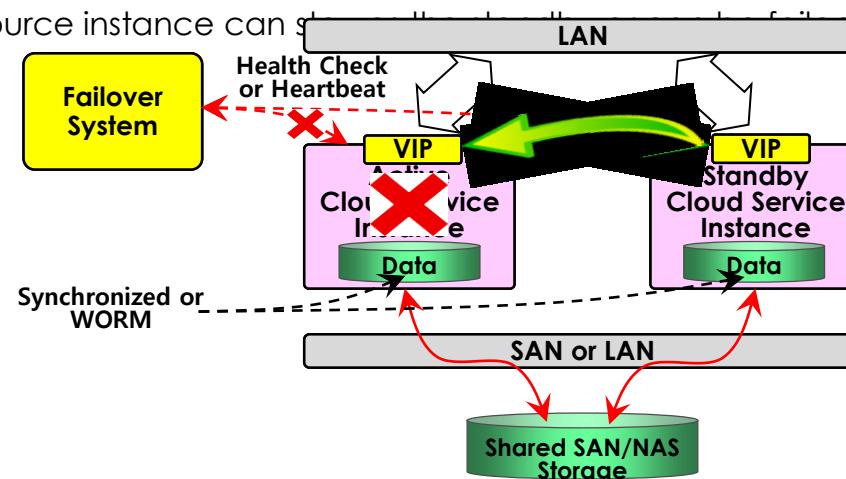
- 2 or more (N-way active-all) IT resource instances serving the workload synchronously
- Load balancing by distributing workload evenly across all available IT resources
- IT resource instance removed from load balancing logic when failed and reinstated when recovered
- DAS-based: periodic synchronization needed or **WORM** (Write Once Read Many) workload – i.e., Web
- LAN-based: network shared filesystem or **NAS** (Network Attached Storage) / distributed filesystem
- SAN-based: both physically & logically shared with SAN/cluster filesystem
- Other IT resource pre-synchronized when implemented



Failover System – 3/3

□ Active-passive / active-standby

- 1 active IT resource instance + 1 passive/standby redundant IT resource instance configuration
 - Active instance = 1 real IP for management + 1 virtual service IP
 - Passive instance = 1 real IP for management & taking over virtual service IP when failed over
- The virtual service IP dynamically reassigned to the standby IT resource instance during service failover
- DAS-based: periodic synchronization needed or **WORM** (Write Once Read Many) workload – i.e., Web
- LAN-based: network shared filesystem or **NAS** (Network Attached Storage) / distributed filesystem
- SAN-based: physically shared storage or both physically & logically shared with SAN/cluster filesystem
- The failed IT resource instance can ~~be recovered~~ back upon recovery.



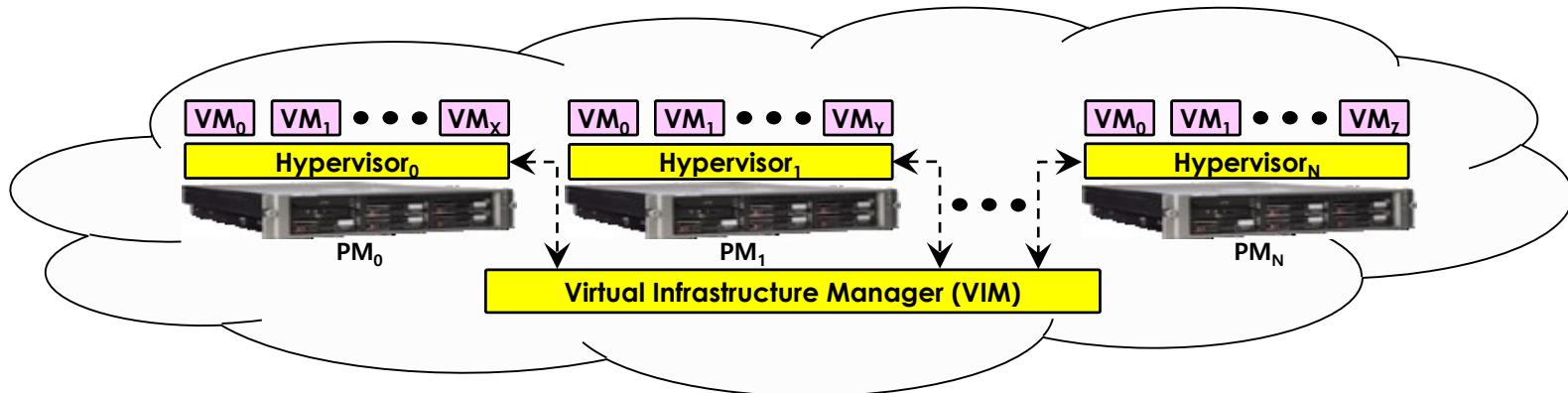
Hypervisor – 1/2

□ Definition

- A fundamental mechanism of virtualization infrastructure that is primarily used to generate virtual server instances of a physical server

□ Implementation mechanisms

- One hypervisor per a physical server in general creating multiple virtual images for the given physical server
- Responsible for creating VMs, increasing VM capacity, decreasing VM capacity or shutting VM down
- Installed in bare-metal (meaning no OS installed) server for controlling, sharing and scheduling the usage of hardware resources (processing power, memory, I/O, etc.) which appear as dedicated resources to each virtual server
- A single **VIM** (on any physical server) administrates multiple hypervisors across a physical server group.



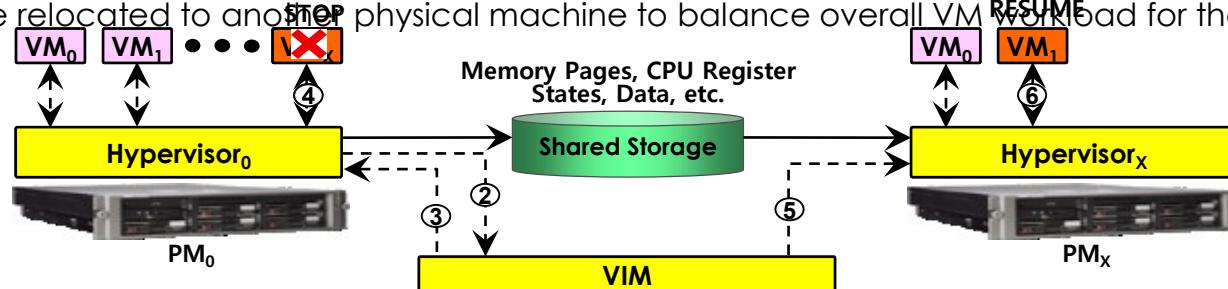
Hypervisor – 2/2

□ On-line VM scale-up or migration scenario

- ① VM_X becomes **overloaded** or needs to be **scaled-up**, but no more processing power available on PM_0 .
- ② Hypervisor₀ detects the situation and reports it to VIM.
- ③ VIM finds out that VM_X on PM_0 can not be scaled-up and sends a **STOP** signal for VM_X to Hypervisor₀ after deciding to migrate VM_X to PM_X with enough processing power available.
- ④ Hypervisor₀ suspends VM_X . \Rightarrow VM_1 image replication should be done first with no shared storage.
- ⑤ VIM sends a **RESUME** signal for VM_X to Hypervisor_X with more resources (processing power & memory etc.).
- ⑥ Hypervisor_X activates a new VM_1 on PM_X replacing VM_X of PM_0 .

□ VM failover and workload balancing

- The same mechanism applies to VM failover and workload balancing.
- Hypervisor₀ detects VM failure instead of VM becoming overloaded for VM failover process.
- For VM failover, VIM may restart failed VM on the same physical machine.
- A VM may be relocated to another **STOP** physical machine to balance overall VM **workload** for the given cloud.



Resource Cluster – 1/2

□ Definition

- A mechanism to group multiple IT resource instances together for them to act as **a single IT resource**
- A technology to logically combine multiple physical IT resources to improve the availability and to increase computing capacity

□ Implementation mechanisms

- Clustering target IT resources together via cluster-wide **distributed middleware** implementation relying on high-speed dedicated network connections and desirably physically shared storage devices
- Basic role of cluster middleware: workload distribution, task scheduling, data sharing, system synchronization, **SSI** (**S**ingle **S**ystem **I**mage) or **SAP** (**S**ingle **A**ccess **P**oint), etc.

□ Server cluster

- Clustering physical or virtual servers for increasing performance and availability
- Hypervisors (middleware role) on different physical servers configured to share virtual server execution state (memory pages, processor register states, etc.) in order to establish clustered virtual servers
- Possible live VM migration support between different physical servers with physically shared storage
- Front-end load balancer for linear scale-up for a single workload – WORM (Web) workload & scalable parallel application, etc.

□ Database cluster – RDBMS or SQL type

- Designed to improve data availability – not performance
- Data synchronization between different storage devices to maintain the consistency & availability

오상규

□ Large dataset cluster – NoSQL or bigdata processing

Resource Cluster – 2/2

❑ Load balanced resource cluster

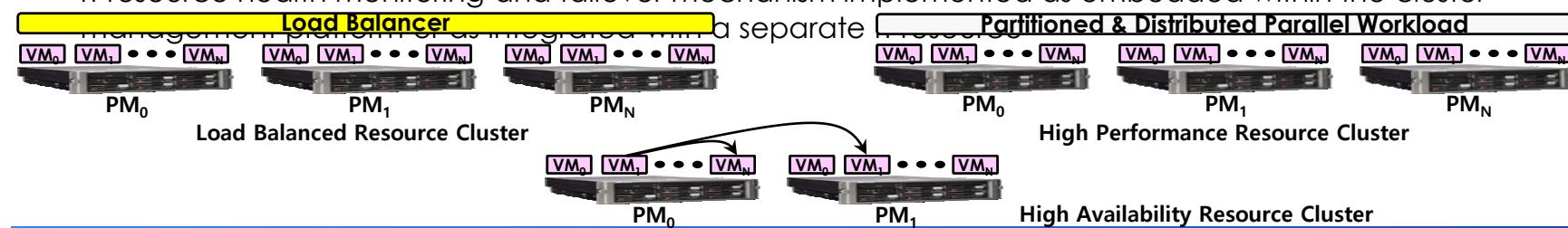
- Specialized for distributing workload among cluster nodes to increase IT resource capacity while preserving the centralization of IT resource management
- Load balancer mechanism implemented in general
- either embedded within the cluster management platform or set up as a separate IT resource

❑ High performance (HP) resource cluster

- Specialized for distributing workload among cluster nodes to increase IT resource capacity
- No load balancer mechanism implemented – pre-distributed or partitioned computing workload as a part of parallel processing platform – e.g., bigdata processing on Hadoop & Map reduce platform
- Theoretically possible implementation, but no particular reason to run HP workloads on virtualized environment in general – why bother loosing some processing power for virtualization when more performance is needed?

❑ High availability (HA) resource cluster

- Specialized for maintaining higher degree of IT resource availability
- IT resource health monitoring and failover mechanism implemented as embedded within the cluster



Multi-Device Broker

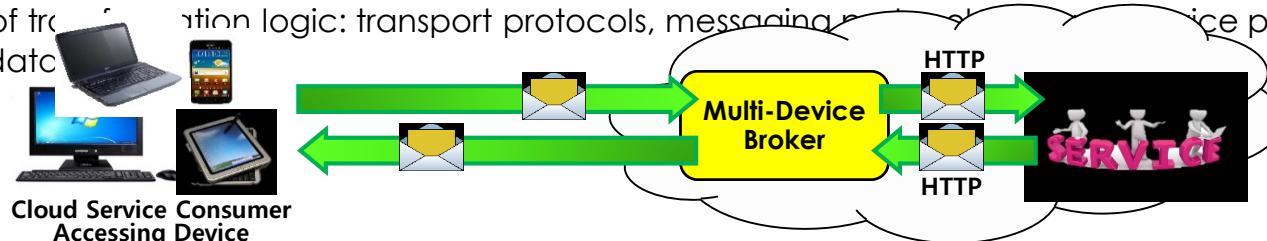
□ Definition

- A mechanism to facilitate runtime data transformation so as to make a cloud service accessible to a wider range of cloud server consumer programs and service

□ Implementation mechanisms

- Cloud service accessed by a range of cloud service consumers with a wide range of different hosting hardware devices or software programs using various communication protocols
- A mapping logic required to transform (or convert) information that is exchanged at runtime for overcoming incompatibility between a cloud service and a wide-range of different cloud service consumers
- Commonly implemented as a gateway or incorporate gateway components such as:
 - XML gateway – transmits and validates XML data
 - Cloud storage gateway – transforms cloud storage protocols and encodes storage devices to facilitate data transfer and storage
 - Mobile device gateway – transforms the communication protocols used by mobile devices into protocols that are compatible with a cloud service

- The levels of transformation logic: transport protocols, messaging protocols, service protocols, data schemas/data



What is a data center?

- A data center -- also known as a *datacenter* or *data centre* -- is a facility composed of networked computers, storage systems and computing infrastructure that organizations use to assemble, process, store and disseminate large amounts of data. A business typically relies heavily on the applications, services and data contained within a data center, making it a critical asset for everyday operations.
- Enterprise data centers increasingly incorporate facilities for securing and protecting cloud computing resources and in-house, on-site resources. As enterprises turn to cloud computing, the boundaries between cloud providers' data centers and enterprise data centers become less clear-cut.

How do data centers work?

- A data center facility, which enables an organization to collect its resources and infrastructure for data processing, storage and communications, includes the following:
- systems for storing, sharing, accessing and processing data across the organization;
- physical infrastructure for supporting data processing and data communications; and
- utilities such as cooling, electricity, network security access and uninterruptible power supplies ([UPSes](#)).
- Gathering all these resources in a data center enables the organization to do the following:
- protect proprietary systems and data;
- centralize IT and data processing employees, contractors and vendors;
- apply information security controls to proprietary systems and data; and
- realize economies of scale by consolidating sensitive systems in one place.

Data centers enable organizations to concentrate on the following:

- IT and data processing personnel;
- computing and network connectivity infrastructure; and
- computing facility security.

core components of data centers

- **Facility.** This includes the physical location with security access controls and sufficient square footage to house the data center's infrastructure and equipment. **Enterprise data storage.** A modern data center houses an organization's data systems in a well-protected physical and storage infrastructure along with servers, storage subsystems, networking switches, routers, firewalls, cabling and physical racks. **Support infrastructure.** This equipment provides the highest available sustainability related to uptime. Components of the support infrastructure include the following: power distribution and supplemental power subsystems;
- electrical switching;
- UPSes;
- backup generators;
- ventilation and data center cooling systems, such as in-row cooling configurations and computer room air conditioners; and
- adequate provisioning for network carrier, or telecom, connectivity.
- **Operational staff.** These employees are required to maintain and monitor IT and infrastructure equipment around the clock.

types of data centers

- **Enterprise data centers.** These proprietary data centers are built and owned by organizations for their internal end users. They support the IT operations and critical applications of a single organization and can be located both on-site and off-site.
- **Managed services data centers.** Managed by third parties, these data centers provide all aspects of data storage and computing services. Companies lease, instead of buy, the infrastructure and services.
- **Cloud-based data centers.** These off-site distributed data centers are managed by third-party or public cloud providers, such as Amazon Web Services, Microsoft Azure or Google Cloud. Based on an infrastructure-as-a-service model, the leased infrastructure enables customers to provision a virtual data center within minutes.

types of data centers

- **Colocation data centers.** These rental spaces inside colocation facilities are owned by third parties. The renting organization provides the hardware, and the data center provides and manages the infrastructure, including physical space, bandwidth, cooling and security systems. Colocation is appealing to organizations that want to avoid the large capital expenditures associated with building and maintaining their own data centers.
- **Edge data centers.** These are smaller facilities that solve the latency problem by being geographically closer to the edge of the network and data sources.
- **Hyperscale data centers.** Synonymous with large-scale providers, such as Amazon, Meta and Google, these hyperscale computing infrastructures maximize hardware density, while minimizing the cost of cooling and administrative overhead.

S.N	Cloud	Data Center
1.	Cloud is a virtual resource that helps businesses store, organize, and operate data efficiently.	Data Center is a physical resource that helps businesses store, organize, and operate data efficiently.
2.	The scalability of the cloud required less amount of investment.	The scalability of the Data Center is huge in investment compared to the cloud.
3.	Maintenance cost is less as compared to service providers.	Maintenance cost is high because the developers of the organization do the maintenance.
4.	The organization needs to rely on third parties to store its data.	The organization's developers are trusted for the data stored in the data centers.
5.	The performance is huge compared to the investment.	The performance is less than the investment.
6.	This requires a plan for optimizing the cloud.	It is easily customizable without any hard planning.
7.	It requires a stable internet connection to provide the function.	This may or may not require an internet connection.
8.	The cloud is easy to operate and is considered a viable option.	Data centers require experienced developers to operate and are not considered a viable option.

Cloud Data Centers

A cloud data center moves a traditional on-prem data center off-site. Instead of personally managing their own infrastructure, an organization leases infrastructure managed by a third-party partner and accesses data center resources over the Internet. Under this model, the cloud service provider is responsible for maintenance, updates, and meeting service level agreements (SLAs) for the parts of the infrastructure stack under their direct control.

Design consideration for Private Cloud

- Involve the stakeholders.
- Consider the use cases
- Metrics are key
- Avoid cloning a public cloud
- Focus on agility
- Think applications and not workloads
- Avoid round holes and square pegs
- Cloud portability is essential
- Use application best practices
- Plan for onboarding and migration

On Premise vs. Cloud propositions

- **Deployment**

On Premises: In an on-premises environment, resources are deployed in-house and within an enterprise's IT infrastructure. An enterprise is responsible for maintaining the solution and all its related processes.

Cloud: While there are different forms of cloud computing (such as public cloud, private cloud, and a hybrid cloud), in a public cloud computing environment, resources are hosted on the premises of the service provider but enterprises are able to access those resources and use as much as they want at any given time.

- **Cost**

On Premises: For enterprises that deploy software on premise, they are responsible for the ongoing costs of the server hardware, power consumption, and space.

Cloud: Enterprises that elect to use a cloud computing model only need to pay for the resources that they use, with none of the maintenance and upkeep costs, and the price adjusts up or down depending on how much is consumed.

- **Control**

On Premises: In an on-premises environment, enterprises retain all their data and are fully in control of what happens to it, for better or worse. Companies in highly regulated industries with extra privacy concerns are more likely to hesitate to leap into the cloud before others because of this reason.

Cloud: In a cloud computing environment, the question of ownership of data is one that many companies – and vendors for that matter, have struggled with. Data and encryption keys reside within your third-party provider, so if the unexpected happens and there is downtime, you maybe be unable to access that data.

On Premise vs. Cloud propositions

- **Security**
- **On Premises:** Companies that have extra sensitive information, such as government and banking industries must have a certain level of security and privacy that an on-premises environment provides. Despite the promise of the cloud, security is the primary concern for many industries, so an on-premises environment, despite some of its drawbacks and price tag, make more sense.
- **Cloud:** Security concerns remain the number one barrier to cloud computing deployment. There have been many publicized cloud breaches, and IT departments around the world are concerned. From personal information of employees such as login credentials to a loss of intellectual property, the security threats are real.
- **Compliance**
- **On Premises:** Many companies these days operate under some form of regulatory control, regardless of the industry. Perhaps the most common one is the Health Insurance Portability and Accountability Act (HIPAA) for private health information, but there are many others, including the Family Educational Rights and Privacy Act (FERPA), which contains detailed student records, and other government and industry regulations. For companies that are subject to such regulations, it is imperative that they remain compliant and know where their data is at all times.
- **Cloud:** Enterprises that do choose a cloud computing model must do their due diligence and ensure that their third-party provider is up to code and in fact compliant with all of the different regulatory mandates within their industry. Sensitive data must be secured, and customers, partners, and employees must have their privacy ensured

Synchronization in cloud environment

Clock synchronization protocols in cloud data centers, Leader Election protocols in cloud ,Gossip Protocols and its types

<https://www.enjoyalgorithms.com/blog/leader-election-system-design>

Problems

As a result of the difficulties managing time at smaller scales, there are problems associated with clock skew that take on more complexity in distributed computing in which several computers will need to realize the same global time.

Synchronization is required for accurate reproduction of streaming media. Clock synchronization is a significant component of audio over Ethernet systems.

Solutions

In a system with a central server, the synchronization solution is trivial; the server will dictate the system time.

In distributed computing, the problem takes on more complexity because a global time is not easily known.

The most used clock synchronization solution on the Internet is the Network Time Protocol (NTP) which is a layered client-server architecture based on User Datagram Protocol (UDP) message passing.

Lamport timestamps and vector clocks are concepts of the logical clock in distributed computing. In a wireless network, the problem becomes even more challenging due to the possibility of collision of the synchronization packets on the wireless medium and the higher drift rate of clocks on low-cost wireless devices

Berkeley algorithm

The Berkeley algorithm is suitable for systems where a radio clock is not present, this **system has no way of making sure of the actual time other than by maintaining a global average time as the global time**. A time server will **periodically fetch** the time from all the time clients, average the results, and then report back to the clients the adjustment that needs be made to their local clocks to achieve the average. This algorithm highlights the fact that **internal clocks may vary not only in the time they contain but also in the clock rate**.

Clock-sampling mutual network synchronization

Clock-sampling mutual network synchronization (CS-MNS) is suitable for **distributed and mobile applications**. It has been shown to be scalable over **mesh networks** that include **indirectly-linked non-adjacent nodes**, and is compatible with IEEE 802.11 and similar standards. It can be accurate to the order of few microseconds, but requires direct physical wireless connectivity with **negligible link delay** (less than 1 microsecond) on links between adjacent nodes, limiting the distance between neighboring nodes to a few hundred meters.

Cristian's algorithm

Cristian's algorithm relies on the **existence of a time server**. The **time server** maintains its clock by using a radio clock or other accurate time source, then all other computers in the system stay synchronized with it. A **time client** will maintain its clock by making a procedure call to the time server.

Satellite navigation systems

In addition to its use in navigation, the **Global Positioning System (GPS)** can also be used for clock synchronization. The **accuracy of GPS time signals is ± 10 nanoseconds**. Using GPS (or other satellite navigation systems) for synchronization requires a receiver connected to an antenna with unobstructed view of the sky.

Network Time Protocol

Network Time Protocol (NTP) is a **highly robust protocol**, widely deployed throughout the Internet in **distributed time synchronization protocols for unreliable networks**. It can reduce synchronization offsets to times of the order of a few milliseconds over the public Internet, and to sub-millisecond levels over local area networks.

A simplified version of the NTP protocol, Simple Network Time Protocol (SNTP), can also be used as a pure single-shot stateless primary/secondary synchronization protocol, but lacks the sophisticated features of NTP, and thus has much lower performance and reliability levels.

Precision Time Protocol

Precision Time Protocol (PTP) is a **master/slave protocol for delivery of highly accurate time over local area networks**.

Reference broadcast synchronization

The Reference Broadcast Time Synchronization (RBS) algorithm is often **used in wireless networks and sensor networks**. In this scheme, **an initiator broadcasts a reference message to urge the receivers to adjust their clocks**.

Reference Broadcast Infrastructure Synchronization

The Reference Broadcast Infrastructure Synchronization (RBIS) protocol is a **master/slave synchronization protocol**, like RBS, based on a **receiver/receiver synchronization paradigm**. It is specifically tailored to be used in IEEE 802.11 wireless networks configured in **infrastructure mode** (i.e., coordinated by an access point). **The protocol does not require any modification to the access point**.

Synchronous Ethernet

Synchronous Ethernet uses **Ethernet in a synchronous manner** such that when combined with synchronization protocols such as PTP in the case of the White Rabbit Project, sub-nanosecond synchronization accuracy is achieved.

Wireless ad hoc networks]

Synchronization is achieved in **wireless ad hoc networks through sending synchronization messages in a multi-hop manner and each node progressively synchronizing with the node that is the immediate sender of a synchronization message.**

Huygens

Huygens is implemented in software and thus can be deployed in **data centers or in public cloud environments**. By leveraging some key aspects of modern data centers, and applying **novel estimation algorithms and signal processing techniques**, the Huygens algorithm achieved an **accuracy of 10s of nanoseconds even at high network load**.

Data Synchronization

- **Data synchronization is the ongoing process of synchronizing data between two or more devices and updating changes automatically between them to maintain consistency within systems.**
- a business will see performance improvement in many areas, including:
 - Logistics and transportation
 - Sales team productivity
 - Order management
 - Invoice accuracy
 - Business systems
 - Cost efficiency
 - Reputation management

data synchronization methods

- **File Synchronization:** Faster and more error-proof than a manual copy technique, this method is most used for home backups, external hard drives, or updating portable data via flash drive. File synchronization ensures that two or more locations share the same data, occurs automatically, and prevents duplication of identical files.
- **Version Control:** This technique aims to provide synchronizing solutions for files that can be altered by more than one user at the same time.
- **Distributed File Systems:** When multiple file versions must be synced at the same time on different devices, those devices must always be connected for the distributed file system to work. A few of these systems allow devices to disconnect for short periods of time, as long as data reconciliation is implemented before synchronization.
- **Mirror Computing:** Mirror computing is used to provide different sources with an exact copy of a data set. Especially useful for backup, mirror computing provides an exact copy to just one other location — source to target.

data synchronization challenges

- Security
- Data quality
- Management
- Performance
- Data complexity

Network Time Protocol (NTP)

- Network Time Protocol (NTP) is an internet protocol used to synchronize with computer clock time sources in a network. It belongs to and is one of the oldest parts of the TCP/IP suite. The term *NTP* applies to both the protocol and the client-server programs that run on computers.

NTP time synchronization process

- The NTP client initiates a time-request exchange with the NTP server.
- The client is then able to calculate the link delay and its local offset and adjust its local clock to match the clock at the server's computer.
- As a rule, six exchanges over a period of about five to 10 minutes are required to initially set the clock.
- Once synchronized, the client updates the clock about once every 10 minutes, usually requiring only a single message exchange, in addition to client-server synchronization. This transaction occurs via User Datagram Protocol (UDP) on port 123. NTP also supports broadcast synchronization of peer computer clocks.
- Network acceleration and network management systems rely on the accuracy of timestamps to measure performance and troubleshoot problems.

NTP features

- They have **access to highly precise atomic clocks and Global Positioning System clocks.**
- **Specialized receivers are required to directly communicate with the NTP servers** for these time services.
- It is **not practical or cost-effective to equip every computer** with one of these receivers. Instead, **computers designated as primary time servers are outfitted with the receivers.**
- They use protocols such as **NTP to synchronize the clock times of networked computers.**
- **NTP uses Coordinated Universal Time (UTC) to synchronize computer clock times** with extreme precision. It offers greater accuracy on smaller networks -- down to 1 millisecond in a local area network (LAN) and within tens of milliseconds over the internet.
- **NTP does not account for time zones.** Instead, it relies on the host to perform such computations.

stratum levels

Degrees of separation from the UTC source are defined as *strata*. The various strata include the following:

- **Stratum 0.** A reference clock receives true time from a dedicated transmitter or satellite navigation system. It is categorized as stratum 0.
- **Stratum 1.** A device is directly linked to the reference clock.
- **Stratum 2.** A device receives its time from a stratum 1 computer.
- **Stratum 3.** A device receives its time from a stratum 2 computer.

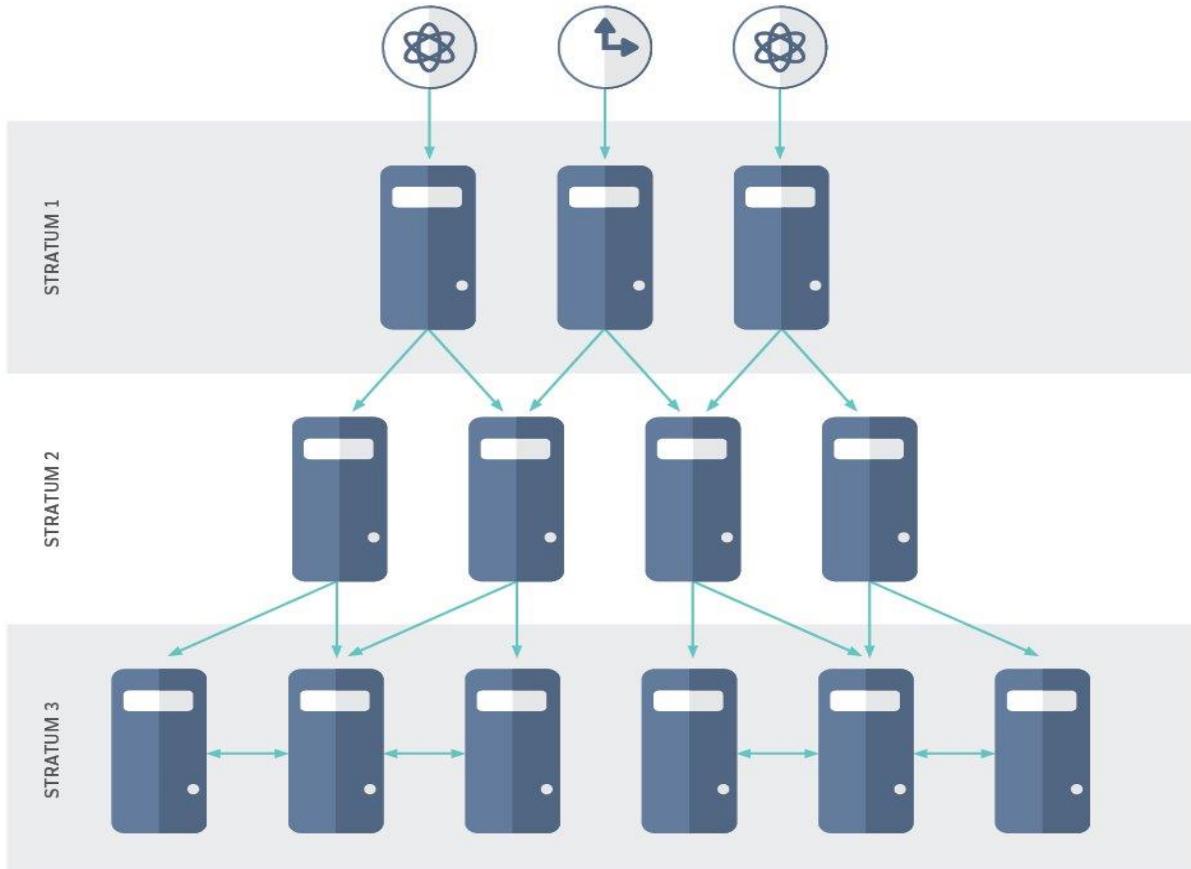
The stratum ranking continues from there. Accuracy is reduced with each additional degree of separation.

Security-wise, NTP has known vulnerabilities.

The protocol can be exploited and used in **denial-of-service attacks** for two reasons: First, **it replies to a packet with a spoofed source IP address**; second, **at least one of its built-in commands sends a long reply to a short request**.

Hierarchy of time servers

Synchronization of atomic or radio clocks in local and global networks



How is time information obtained?

Networking devices can poll host servers and listen for NTP broadcasts to get information on time.

Poll-based NTP associations

poll-based association modes are the client mode and the symmetric active mode.

They provide a high degree of accuracy and reliability for timing.

With the client mode, network devices are assigned time-serving hosts that they poll for the correct time.

It then chooses one host to synchronize with and doesn't provide any information back to the host.

This approach is best for clients, such as file servers and workstations, that aren't synchronizing with other clients.

With the symmetric active mode, a device polls its host for the correct time.

It also responds to polls from its hosts, which gather time-related information from networking devices.

This mode works best when several servers are interconnected using various network paths.

How is time information obtained?

Networking devices can poll host servers and listen for NTP broadcasts to get information on time.

Broadcast-based NTP associations

Broadcast-based NTP associations are somewhat less accurate and reliable than poll-based ones.

They are good for localized networks with limited bandwidth, memory or central processing unit (CPU) resources.

In the broadcast-based mode, a network device listens for NTP broadcast packets that broadcast time servers transmit. The time information flows only one direction.

Leader Election protocols in cloud

Leader election is the simple idea of **giving one thing (a process, host, thread, object, or human) in a distributed system some special powers.**

Those special powers could include **the ability to assign work, the ability to modify a piece of data, or even the responsibility of handling all requests in the system.**

Leader election is a powerful tool for **improving efficiency, reducing coordination, simplifying architectures, and reducing operations.**

On the other hand, leader election can **introduce new failure modes and scaling bottlenecks.**

In addition, leader election may make it more difficult for you **to evaluate the correctness of a system.**

For **data processing and workflows, workflow services like AWS Step Functions can achieve many of the same benefits as leader election and avoid many of its risks.**

For other systems, we often **implement idempotent APIs, optimistic locking, and other patterns** that make a single leader unnecessary.

leader election Algorithm

The problem of leader election is for each processor eventually to decide whether it is a leader or not, subject to the constraint that exactly one processor decides that it is the leader.

An algorithm solves the leader election problem if:

1. States of processors are divided into elected and not-elected states. Once elected, it remains as elected (similarly if not elected).
2. In every execution, exactly one processor becomes elected and the rest determine that they are not elected.

A valid leader election algorithm must meet the following conditions:^[4]

1. **Termination:** the algorithm should finish within a finite time once the leader is selected. In randomized approaches this condition is sometimes weakened (for example, requiring termination with probability 1).
2. **Uniqueness:** there is exactly one processor that considers itself as leader.
3. **Agreement:** all other processors know who the leader is.

leader election Algorithm

An algorithm for leader election may vary in the following aspects:

- Communication mechanism: the processors are either synchronous in which processes are synchronized by a clock signal or asynchronous where processes run at arbitrary speeds.
- Process names: whether processes have a unique identity or are indistinguishable (anonymous).
- Network topology: for instance, ring, acyclic graph or complete graph.
- Size of the network: the algorithm may or may not use knowledge of the number of processes in the system.

Leader election in rings

A ring network is a connected-graph topology in which each node is exactly connected to two other nodes, i.e., for a graph with n nodes, there are exactly n edges connecting the nodes.

A ring can be unidirectional, which means processors only communicate in one direction (a node could only send messages to the left or only send messages to the right), or bidirectional, meaning processors may transmit and receive messages in both directions (a node could send messages to the left and right).

Anonymous rings

A ring is said to be anonymous if every processor is identical.

More formally, the system has the same state machine for every processor.

There is no deterministic algorithm to elect a leader in anonymous rings, even when the size of the network is known to the processes.

This is due to the fact that there is no possibility of breaking symmetry in an anonymous ring if all processes run at the same speed.

The state of processors after some steps only depends on the initial state of neighbouring nodes.

So, because their states are identical and execute the same procedures, in every round the same messages are sent by each processor.

Therefore, each processor state also changes identically and as a result if one processor is elected as a leader, so are all the others.

Advantages Leader Election protocols in cloud

- A single leader makes systems easier for humans to think about. It puts all the concurrency in the system into a single place, reduces partial failure modes, and adds a single place to look for logs and metrics.
- A single leader can work more efficiently. It can often simply inform other systems about changes, rather than building consensus about the changes to be made.
- Single leaders can easily offer clients consistency because they can see and control all the changes made to the state of the system.
- A single leader can improve performance or reduce cost by providing a single consistent cache of data which can be used every time.
- Writing software for a single leader may be easier than other approaches like quorum. The single leader doesn't need to consider that other systems may be working on the same state at the same time.

Advantages Leader Election protocols in cloud

- A single leader is a single point of failure. If the system fails to detect or fix a bad leader, the whole system can be unavailable.
- A single leader means a single point of scaling, both in data size and request rate. When a leader-elected system needs to grow beyond a single leader, it requires a complete re-architecture.
- A leader is a single point of trust. If a leader is doing the wrong work with nobody checking it, it can quickly cause problems across the entire system. A bad leader has a high blast radius.
- Partial deployments may be hard to apply in leader-elected systems. Many software safety practices at Amazon depend on partial deployments, such as one-box, A-B testing, blue/green deployment, and incremental deployment with automatic rollback.

Gossip Protocol

A **gossip protocol** or **epidemic protocol** is a procedure or process of computer peer-to-peer communication that is based on the way epidemics spread.

Some distributed systems use peer-to-peer gossip to ensure that data is disseminated to all members of a group.

Some ad-hoc networks have no central registry and the only way to spread common data is to rely on each member to pass it along to their neighbors.

Gossip Protocol is a communication protocol, it is process computer to computer communication that works on the same principle as how information is shared on social networks.

There are 3 main types of gossip protocol:

- 1. Dissemination Protocols :** These protocols are also referred to as rumor-mongering protocols because they use gossip to spread information throughout the network, they flood the members of the network with gossips in a way that produces the worst-case load.
- 2. Anti-Entropy Protocols :**
These are used to repair the replicated data by comparing them and modifying the comparisons.
- 3. Protocols that compute aggregates :**
These protocols work by or compute an aggregate of the network by sampling information at the nodes and they combine the values to acquire a system-wide value – the largest value for some measurement nodes are making, smallest, etc.

Implementation of gossip protocol in Cloud Computing :

The Gossip protocol is used to repair the problems caused by multicasting; it is a type of communication where a piece of information or gossip in this scenario, is sent from one or more nodes to a set of other nodes in a network.

This is useful when a group of clients in the network require the same data at the same time. But there are many problems that occur during multicasting, if there are many nodes present at the recipient end, latency increases; the average time for a receiver to receive a multicast. To get this multicast message or gossip across the desired targets in the group, the gossip protocol sends out the gossip periodically to random nodes in the network, once a random node receives the gossip, it is said to be infected due to the gossip.

Now the random node that receives the gossip does the same thing as the sender, it sends multiple copies of the gossip to random targets.

This process continues until the target nodes get the multicast. This process turns the infected nodes to uninfected nodes after sending the gossip out to random nodes.

Chapter 6. Fundamental Cloud Security



6.1 Basic Terms and Concepts

6.2 Threat Agents

6.3 Cloud Security Threats

6.4 Additional Considerations

6.5 Case Study Example

This chapter introduces terms and concepts that address basic information security within clouds, and then concludes by defining a set of threats and attacks common to public cloud environments. The cloud security mechanisms covered in [Chapter 10](#) establish the security controls used to counter these threats.

6.1. Basic Terms and Concepts

Information security is a complex ensemble of techniques, technologies, regulations, and behaviors that collaboratively protect the integrity of and access to computer systems and data. IT security measures aim to

defend against threats and interference that arise from both malicious intent and unintentional user error.

The upcoming sections define fundamental security terms relevant to cloud computing and describe associated concepts.

Confidentiality

Confidentiality is the characteristic of something being made accessible only to authorized parties (Figure 6.1). Within cloud environments, confidentiality primarily pertains to restricting access to data in transit and storage.

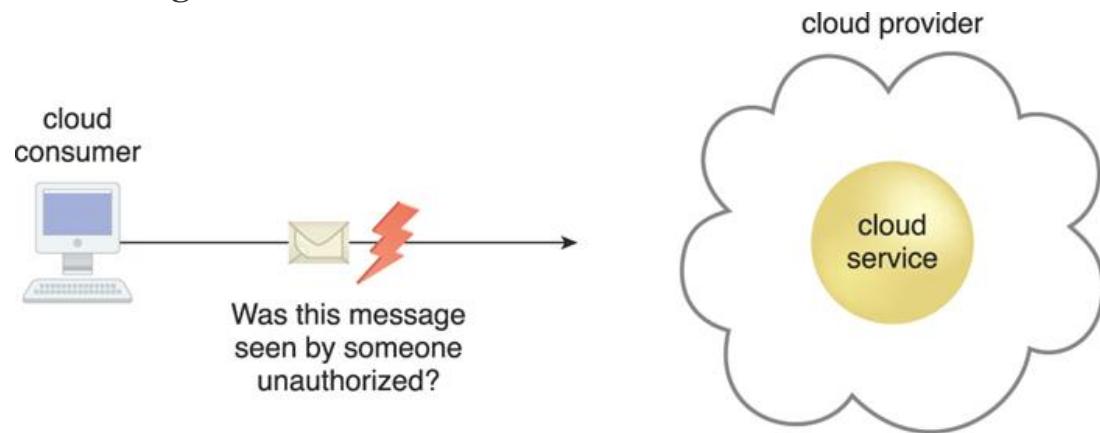


Figure 6.1. The message issued by the cloud consumer to the cloud service is considered confidential only if it is not accessed or read by an unauthorized party.

Integrity

Integrity is the characteristic of not having been altered by an unauthorized party (Figure 6.2). An important issue that concerns data integrity in the cloud is whether a cloud consumer can be guaranteed that the data it transmits to a cloud service matches the data received by that cloud service. Integrity can extend to how data is stored, processed, and retrieved by cloud services and cloud-based IT resources.

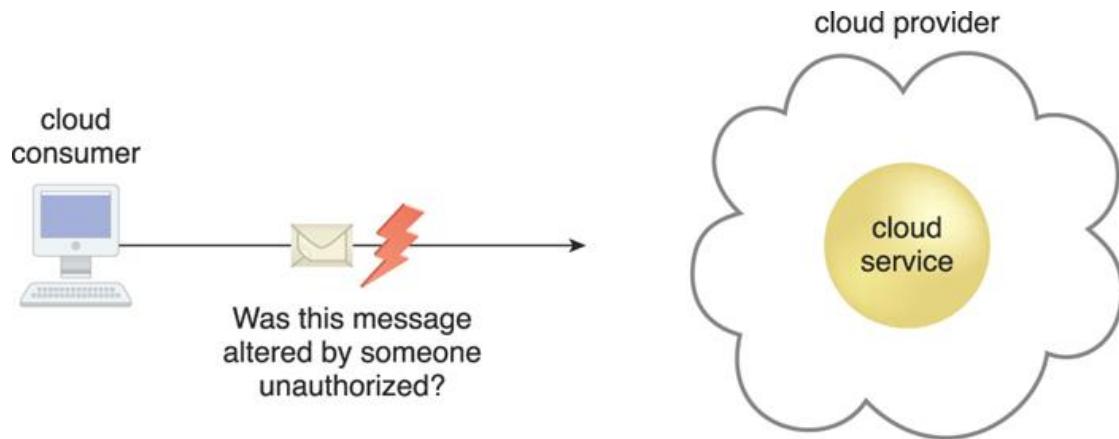


Figure 6.2. The message issued by the cloud consumer to the cloud service is considered to have integrity if it has not been altered.

Authenticity

Authenticity is the characteristic of something having been provided by an authorized source. This concept encompasses non-repudiation, which is the inability of a party to deny or challenge the authentication of an interaction. Authentication in non-repudiable interactions provides proof that these interactions are uniquely linked to an authorized source. For example, a user may not be able to access a non-repudiable file after its receipt without also generating a record of this access.

Availability

Availability is the characteristic of being accessible and usable during a specified time period. In typical cloud environments, the availability of cloud services can be a responsibility that is shared by the cloud provider and the cloud carrier. The availability of a cloud-based solution that extends to cloud service consumers is further shared by the cloud consumer.

Threat

A *threat* is a potential security violation that can challenge defenses in an attempt to breach privacy and/or cause harm. Both manually and automatically instigated threats are designed to exploit known weaknesses, also referred to as vulnerabilities. A threat that is carried out results in an *attack*.

Vulnerability

A *vulnerability* is a weakness that can be exploited either because it is protected by insufficient security controls, or because existing security controls are overcome by an attack. IT resource vulnerabilities can have a range of causes, including configuration deficiencies, security policy weaknesses, user errors, hardware or firmware flaws, software bugs, and poor security architecture.

Risk

Risk is the possibility of loss or harm arising from performing an activity. Risk is typically measured according to its threat level and the number of possible or known vulnerabilities. Two metrics that can be used to determine risk for an IT resource are:

- the probability of a threat occurring to exploit vulnerabilities in the IT resource
 - the expectation of loss upon the IT resource being compromised
- Details regarding risk management are covered later in this chapter.

Security Controls

Security controls are countermeasures used to prevent or respond to security threats and to reduce or avoid risk. Details on how to use security countermeasures are typically outlined in the security policy, which contains a set of rules and practices specifying how to implement a system, service, or security plan for maximum protection of sensitive and critical IT resources.

Security Mechanisms

Countermeasures are typically described in terms of security mechanisms, which are components comprising a defensive framework that protects IT resources, information, and services.

Security Policies

A security policy establishes a set of security rules and regulations. Often, security policies will further define how these rules and regulations are implemented and enforced. For example, the positioning and usage of security controls and mechanisms can be determined by security policies.

Summary of Key Points

- Confidentiality, integrity, authenticity, and availability are characteristics that can be associated with measuring security.
- Threats, vulnerabilities, and risks are associated with measuring and assessing insecurity, or the lack of security.
- Security controls, mechanisms, and policies are associated with establishing countermeasures and safeguards in support of improving security.

6.2. Threat Agents

A *threat agent* is an entity that poses a threat because it is capable of carrying out an attack. Cloud security threats can originate either internally or externally, from humans or software programs. Corresponding threat agents are described in the upcoming sections. [Figure 6.3](#) illustrates the role a threat agent assumes in relation to vulnerabilities, threats, and risks, and the safeguards established by security policies and security mechanisms.

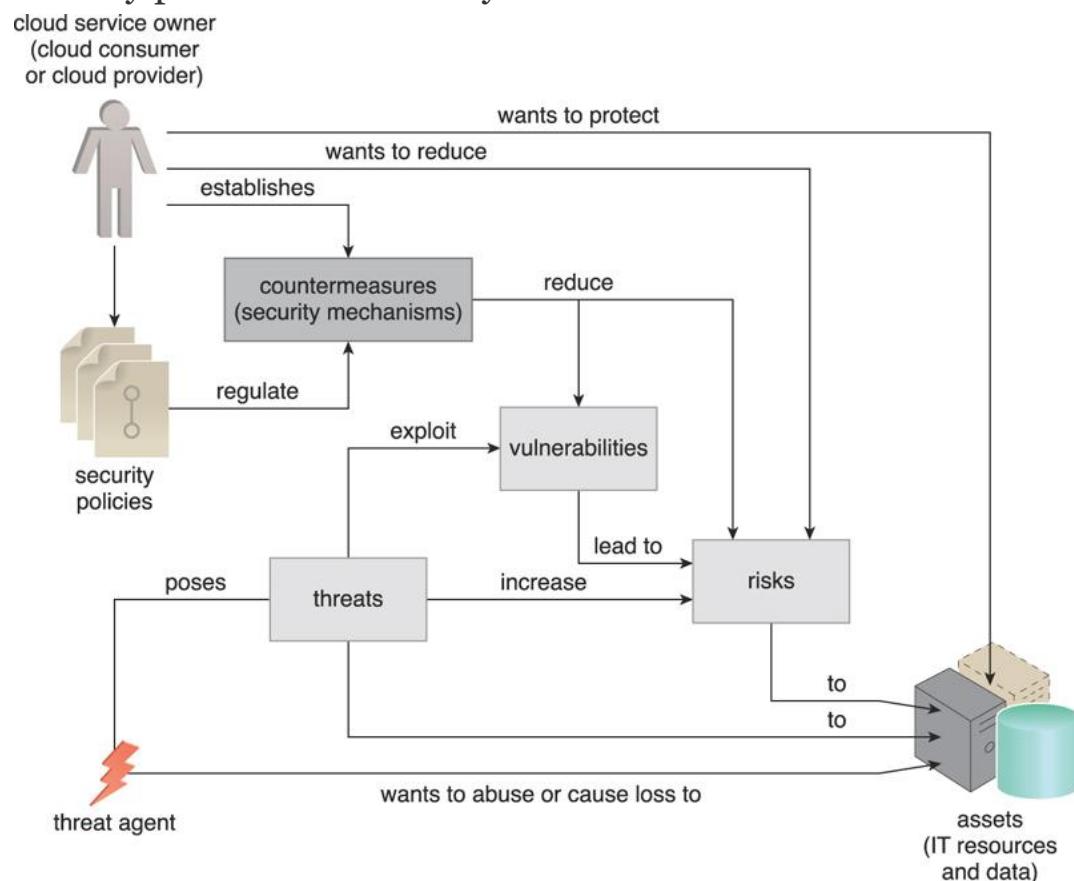


Figure 6.3. How security policies and security mechanisms are used to counter threats, vulnerabilities, and risks caused by threat agents.

Anonymous Attacker

An *anonymous attacker* is a non-trusted cloud service consumer without permissions in the cloud ([Figure 6.4](#)). It typically exists as an external software program that launches network-level attacks through public networks. When anonymous attackers have limited information on security policies and defenses, it can inhibit their ability to formulate effective attacks. Therefore, anonymous attackers often resort to committing acts like bypassing user accounts or stealing user credentials, while using methods that either ensure anonymity or require substantial resources for prosecution.



Figure 6.4. The notation used for an anonymous attacker.

Malicious Service Agent

A *malicious service agent* is able to intercept and forward the network traffic that flows within a cloud ([Figure 6.5](#)). It typically exists as a service agent (or a program pretending to be a service agent) with compromised or malicious logic. It may also exist as an external program able to remotely intercept and potentially corrupt message contents.



Figure 6.5. The notation used for a malicious service agent.

Trusted Attacker

A *trusted attacker* shares IT resources in the same cloud environment as the cloud consumer and attempts to exploit legitimate credentials to target cloud providers and the cloud tenants with whom they share IT resources ([Figure 6.6](#)). Unlike anonymous attackers (which are non-trusted), trusted attackers usually launch their attacks from within a cloud's trust boundaries by abusing legitimate credentials or via the appropriation of sensitive and confidential information.



Figure 6.6. The notation that is used for a trusted attacker.

Trusted attackers (also known as *malicious tenants*) can use cloud-based IT resources for a wide range of exploitations, including the hacking of weak authentication processes, the breaking of encryption, the spamming of e-mail accounts, or to launch common attacks, such as denial of service campaigns.

Malicious Insider

Malicious insiders are human threat agents acting on behalf of or in relation to the cloud provider. They are typically current or former employees or third parties with access to the cloud provider's premises. This type of threat agent carries tremendous damage potential, as the malicious insider may have administrative privileges for accessing cloud consumer IT resources.

Note

A notation used to represent a general form of human-driven attack is the workstation combined with a lightning bolt (Figure 6.7). This generic symbol does not imply a specific threat agent, only that an attack was initiated via a workstation.



Figure 6.7. The notation used for an attack originating from a workstation. The human symbol is optional.

Summary of Key Points

- An anonymous attacker is a non-trusted threat agent that usually attempts attacks from outside of a cloud's boundary.
- A malicious service agent intercepts network communication in an attempt to maliciously use or augment the data.
- A trusted attacker exists as an authorized cloud service consumer with legitimate credentials that it uses to exploit access to cloud-based IT resources.
- A malicious insider is a human that attempts to abuse access privileges to cloud premises.

6.3. Cloud Security Threats

This section introduces several common threats and vulnerabilities in cloud-based environments and describes the roles of the aforementioned

threat agents. Security mechanisms that are used to counter these threats are covered in [Chapter 10](#).

Traffic Eavesdropping

Traffic eavesdropping occurs when data being transferred to or within a cloud (usually from the cloud consumer to the cloud provider) is passively intercepted by a malicious service agent for illegitimate information gathering purposes ([Figure 6.8](#)). The aim of this attack is to directly compromise the confidentiality of the data and, possibly, the confidentiality of the relationship between the cloud consumer and cloud provider. Because of the passive nature of the attack, it can more easily go undetected for extended periods of time.

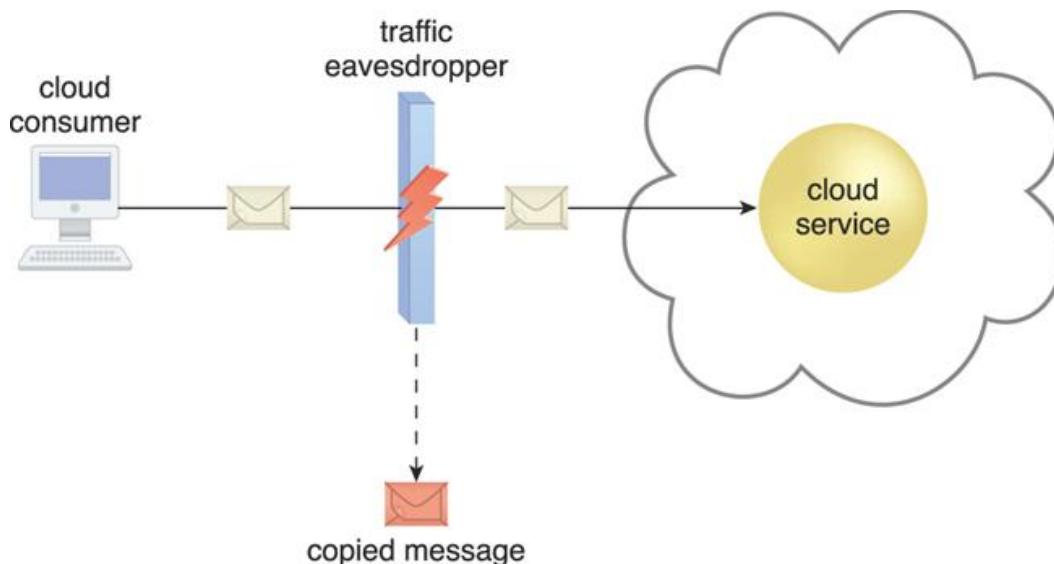


Figure 6.8. An externally positioned malicious service agent carries out a traffic eavesdropping attack by intercepting a message sent by the cloud service consumer to the cloud service. The service agent makes an unauthorized copy of the message before it is sent along its original path to the cloud service.

Malicious Intermediary

The *malicious intermediary* threat arises when messages are intercepted and altered by a malicious service agent, thereby potentially compromising the message's confidentiality and/or integrity. It may also insert harmful data into the message before forwarding it to its destination. [Figure 6.9](#) illustrates a common example of the malicious intermediary attack.

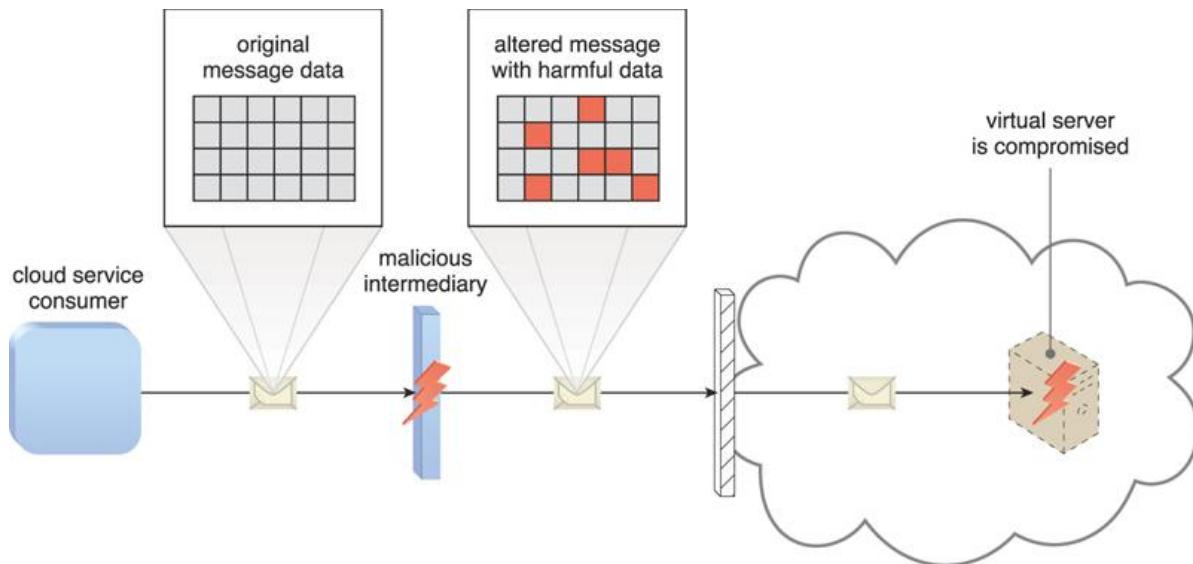


Figure 6.9. The malicious service agent intercepts and modifies a message sent by a cloud service consumer to a cloud service (not shown) being hosted on a virtual server. Because harmful data is packaged into the message, the virtual server is compromised.

Note

While not as common, the malicious intermediary attack can also be carried out by a malicious cloud service consumer program.

Denial of Service

The objective of the denial of service (DoS) attack is to overload IT resources to the point where they cannot function properly. This form of attack is commonly launched in one of the following ways:

- The workload on cloud services is artificially increased with imitation messages or repeated communication requests.
- The network is overloaded with traffic to reduce its responsiveness and cripple its performance.
- Multiple cloud service requests are sent, each of which is designed to consume excessive memory and processing resources.

Successful DoS attacks produce server degradation and/or failure, as illustrated in [Figure 6.10](#).

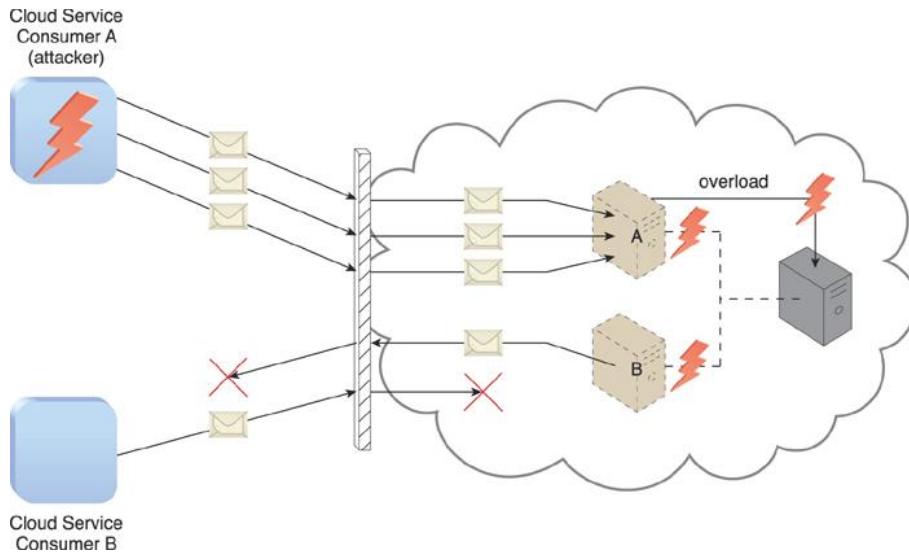


Figure 6.10. Cloud Service Consumer A sends multiple messages to a cloud service (not shown) hosted on Virtual Server A. This overloads the capacity of the underlying physical server, which causes outages with Virtual Servers A and B. As a result, legitimate cloud service consumers, such as Cloud Service Consumer B, become unable to communicate with any cloud services hosted on Virtual Servers A and B.

Insufficient Authorization

The insufficient authorization attack occurs when access is granted to an attacker erroneously or too broadly, resulting in the attacker getting access to IT resources that are normally protected. This is often a result of the attacker gaining direct access to IT resources that were implemented under the assumption that they would only be accessed by trusted consumer programs ([Figure 6.11](#)).

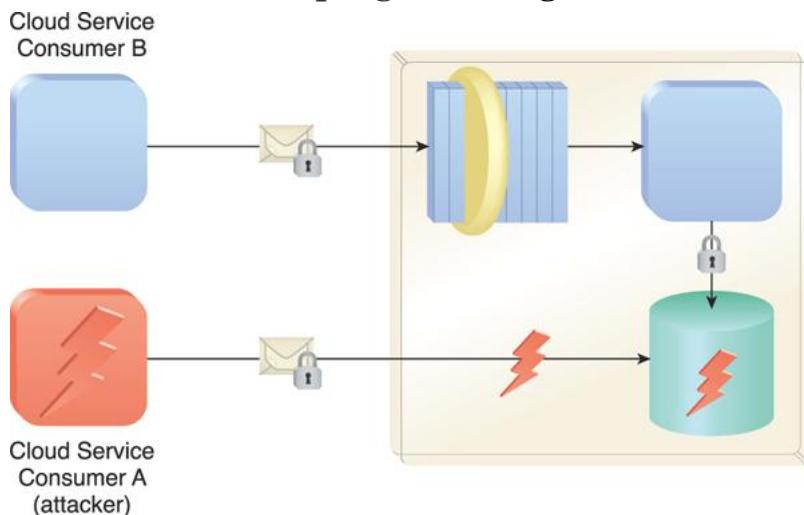


Figure 6.11. Cloud Service Consumer A gains access to a database that was implemented under the assumption that it would only be accessed through a Web service with a published service contract (as per Cloud Service Consumer B).

A variation of this attack, known as *weak authentication*, can result when weak passwords or shared accounts are used to protect IT resources. Within cloud environments, these types of attacks can lead to significant impacts depending on the range of IT resources and the range of access to those IT resources the attacker gains ([Figure 6.12](#)).

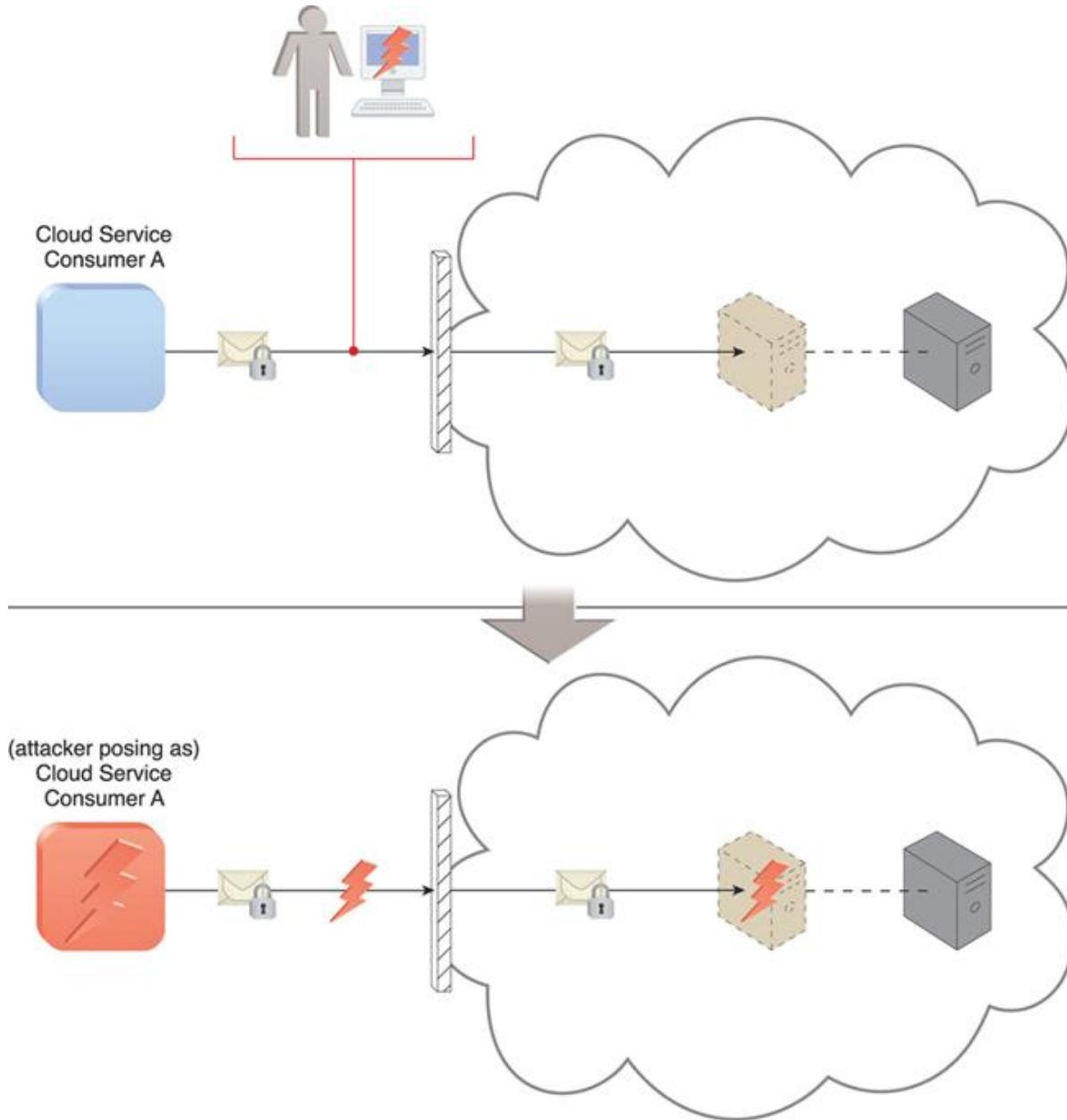


Figure 6.12. An attacker has cracked a weak password used by Cloud Service Consumer A. As a result, a malicious cloud service consumer (owned by the attacker) is designed to pose as Cloud Service Consumer A in order to gain access to the cloud-based virtual server.

Virtualization Attack

Virtualization provides multiple cloud consumers with access to IT resources that share underlying hardware but are logically isolated from each other. Because cloud providers grant cloud consumers administrative access to virtualized IT resources (such as virtual servers), there is an inherent risk that cloud consumers could abuse this access to attack the underlying physical IT resources.

A *virtualization attack* exploits vulnerabilities in the virtualization platform to jeopardize its confidentiality, integrity, and/or availability. This threat is illustrated in [Figure 6.13](#), where a trusted attacker successfully accesses a virtual server to compromise its underlying physical server. With public clouds, where a single physical IT resource may be providing virtualized IT resources to multiple cloud consumers, such an attack can have significant repercussions.

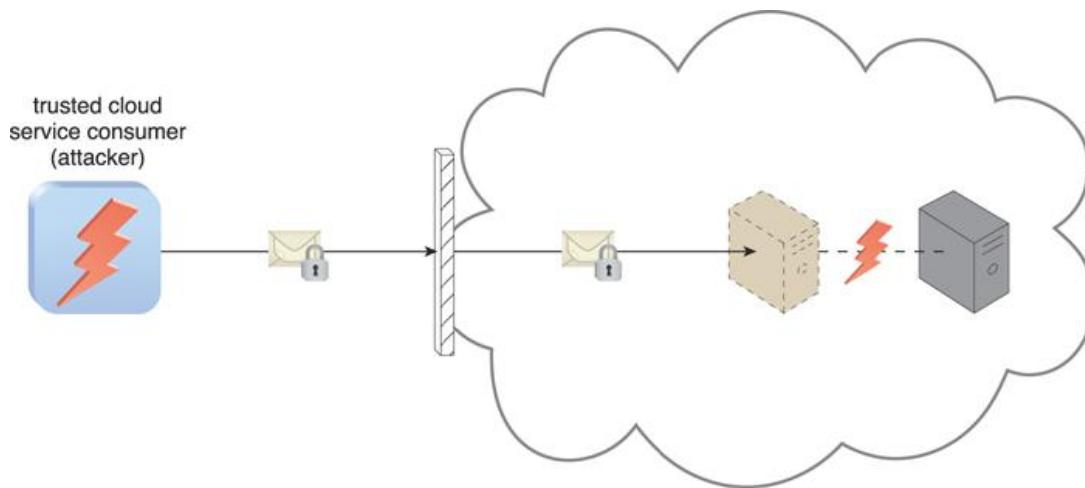


Figure 6.13. An authorized cloud service consumer carries out a virtualization attack by abusing its administrative access to a virtual server to exploit the underlying hardware.

Overlapping Trust Boundaries

If physical IT resources within a cloud are shared by different cloud service consumers, these cloud service consumers have overlapping trust boundaries. Malicious cloud service consumers can target shared IT resources with the intention of compromising cloud consumers or other IT resources that share the same trust boundary. The consequence is that some or all of the other cloud service consumers could be impacted by the attack and/or the attacker could use virtual IT resources against others that happen to also share the same trust boundary.

Figure 6.14 illustrates an example in which two cloud service consumers share virtual servers hosted by the same physical server and, resultantly, their respective trust boundaries overlap.

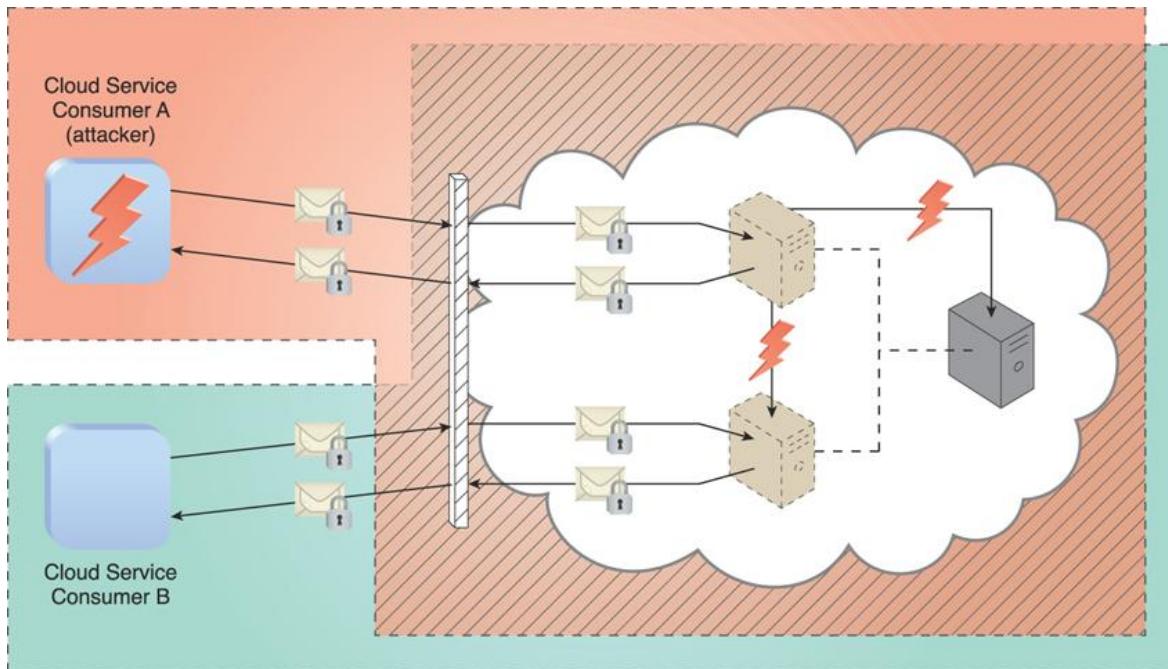


Figure 6.14. Cloud Service Consumer A is trusted by the cloud and therefore gains access to a virtual server, which it then attacks with the intention of attacking the underlying physical server and the virtual server used by Cloud Service Consumer B.

Summary of Key Points

- Traffic eavesdropping and malicious intermediary attacks are usually carried out by malicious service agents that intercept network traffic.
- A denial of service attack occurs when a targeted IT resource is overloaded with requests in an attempt to cripple or render it unavailable. The insufficient authorization attack occurs when access is granted to an attacker erroneously or too broadly, or when weak passwords are used.
- A virtualization attack exploits vulnerabilities within virtualized environments to gain unauthorized access to underlying physical hardware. Overlapping trust boundaries represent a threat whereby attackers can exploit cloud-based IT resources shared by multiple cloud consumers.

6.4. Additional Considerations

This section provides a diverse checklist of issues and guidelines that relate to cloud security. The listed considerations are in no particular order.

Flawed Implementations

The substandard design, implementation, or configuration of cloud service deployments can have undesirable consequences, beyond runtime exceptions and failures. If the cloud provider's software and/or hardware have inherent security flaws or operational weaknesses, attackers can exploit these vulnerabilities to impair the integrity, confidentiality, and/or availability of cloud provider IT resources and cloud consumer IT resources hosted by the cloud provider.

Figure 6.15 depicts a poorly implemented cloud service that results in a server shutdown. Although in this scenario the flaw is exposed accidentally by a legitimate cloud service consumer, it could have easily been discovered and exploited by an attacker.

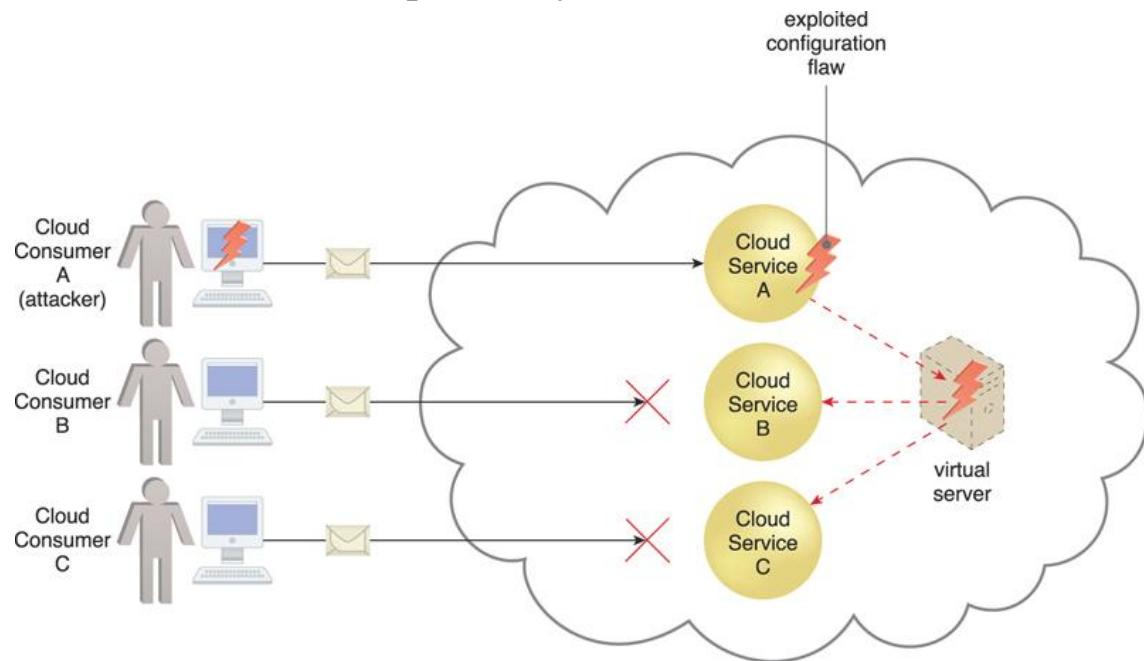


Figure 6.15. Cloud Service Consumer A's message triggers a configuration flaw in Cloud Service A, which in turn causes the virtual server that is also hosting Cloud Services B and C to crash.

Security Policy Disparity

When a cloud consumer places IT resources with a public cloud provider, it may need to accept that its traditional information security approach may not be identical or even similar to that of the cloud provider. This incompatibility needs to be assessed to ensure that any data or other IT assets being relocated to a public cloud are adequately protected. Even when leasing raw infrastructure-based IT resources, the cloud consumer may not be granted sufficient administrative control or influence over

security policies that apply to the IT resources leased from the cloud provider. This is primarily because those IT resources are still legally owned by the cloud provider and continue to fall under its responsibility. Furthermore, with some public clouds, additional third parties, such as security brokers and certificate authorities, may introduce their own distinct set of security policies and practices, further complicating any attempts to standardize the protection of cloud consumer assets.

Contracts

Cloud consumers need to carefully examine contracts and SLAs put forth by cloud providers to ensure that security policies, and other relevant guarantees, are satisfactory when it comes to asset security. There needs to be clear language that indicates the amount of liability assumed by the cloud provider and/or the level of indemnity the cloud provider may ask for. The greater the assumed liability by the cloud provider, the lower the risk to the cloud consumer.

Another aspect to contractual obligations is where the lines are drawn between cloud consumer and cloud provider assets. A cloud consumer that deploys its own solution upon infrastructure supplied by the cloud provider will produce a technology architecture comprised of artifacts owned by both the cloud consumer and cloud provider. If a security breach (or other type of runtime failure) occurs, how is blame determined? Furthermore, if the cloud consumer can apply its own security policies to its solution, but the cloud provider insists that its supporting infrastructure be governed by different (and perhaps incompatible) security policies, how can the resulting disparity be overcome?

Sometimes the best solution is to look for a different cloud provider with more compatible contractual terms.

Risk Management

When assessing the potential impacts and challenges pertaining to cloud adoption, cloud consumers are encouraged to perform a formal risk assessment as part of a risk management strategy. A cyclically executed process used to enhance strategic and tactical security, risk management is comprised of a set of coordinated activities for overseeing and controlling risks. The main activities are generally defined as risk assessment, risk treatment, and risk control ([Figure 6.16](#)).

- *Risk Assessment* – In the risk assessment stage, the cloud environment is analyzed to identify potential vulnerabilities and shortcomings that threats can exploit. The cloud provider can be asked to produce statistics and other information about past attacks (successful and unsuccessful) carried out in its cloud. The identified risks are quantified and qualified according to the probability of occurrence and the degree of impact in relation to how the cloud consumer plans to utilize cloud-based IT resources.

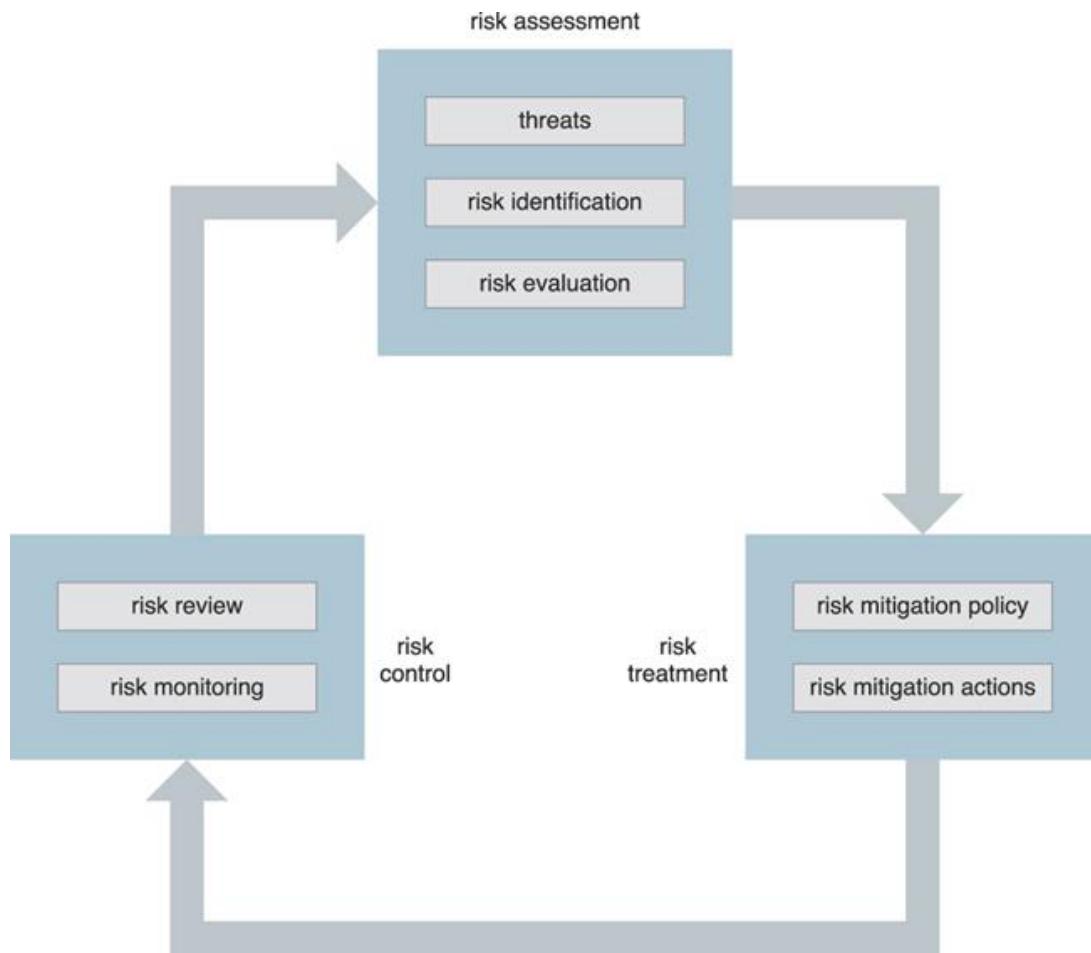


Figure 6.16. The on-going risk management process, which can be initiated from any of the three stages.

- *Risk Treatment* – Mitigation policies and plans are designed during the risk treatment stage with the intent of successfully treating the risks that were discovered during risk assessment. Some risks can be eliminated, others can be mitigated, while others can be dealt with via outsourcing or even incorporated into the insurance and/or operating loss budgets. The cloud provider itself may agree to assume responsibility as part of its contractual obligations.

- *Risk Control* – The risk control stage is related to risk monitoring, a three-step process that is comprised of surveying related events, reviewing these events to determine the effectiveness of previous assessments and treatments, and identifying any policy adjustment needs. Depending on the nature of the monitoring required, this stage may be carried out or shared by the cloud provider.

The threat agents and cloud security threats covered in this chapter (as well as others that may surface) can be identified and documented as part of the risk assessment stage. The cloud security mechanisms covered in [Chapter 10](#) can be documented and referenced as part of the corresponding risk treatment.

Summary of Key Points

- Cloud consumers need to be aware that they may be introducing security risks by deploying flawed cloud-based solutions.
- An understanding of how a cloud provider defines and imposes proprietary, and possibly incompatible, cloud security policies is a critical part of forming assessment criteria when choosing a cloud provider vendor.
- Liability, indemnity, and blame for potential security breaches need to be clearly defined and mutually understood in the legal agreements signed by cloud consumers and cloud providers.
- It is important for cloud consumers, subsequent to gaining an understanding of the potential security-related issues specific to a given cloud environment, to perform a corresponding assessment of the identified risks.

6.5. Case Study Example

Based on an assessment of its internal applications, ATN analysts identify a set of risks. One such risk is associated with the myTrendek application that was adopted from OTC, a company ATN recently acquired. This application includes a feature that analyzes telephone and Internet usage, and enables a multi-user mode that grants varying access rights. Administrators, supervisors, auditors, and regular users can therefore be assigned different privileges. The application's user-base encompasses internal users and external users, such as business partners and contractors.

The myTrendek application poses a number of security challenges pertaining to usage by internal staff:

- authentication does not require or enforce complex passwords
- communication with the application is not encrypted

- European regulations (ETelReg) require that certain types of data collected by the application be deleted after six months

ATN is planning to migrate this application to a cloud via a PaaS environment, but the weak authentication threat and the lack of confidentiality supported by the application make them reconsider. A subsequent risk assessment further reveals that if the application is migrated to a PaaS environment hosted by a cloud that resides outside of Europe, local regulations may be in conflict with ETelReg. Given that the cloud provider is not concerned with ETelReg compliance, this could easily result in monetary penalties being assessed to ATN. Based on the results of the risk assessment, ATN decides not to proceed with its cloud migration plan.

Unit 7

Economics of Cloud ,Challenges in Cloud, Fog Computing, Edge Computing,
Mobile Cloud Computing ,Business Transformation with Google Cloud
Superpowers of Cloud

Economics of Cloud

Economical background of the cloud is more useful for developers in the following ways:

- Pay as you go model offered by cloud providers.
- Scalable and Simple.

Cloud Computing Allows:

- Reduces the capital costs of infrastructure.
- Removes the maintenance cost.
- Removes the administrative cost.

What is Capital Cost?

It is cost occurred in the purchasing infrastructure or the assets that is important in the production of goods. It takes a long time to generate profit.

In the case of start-ups, there is no extra budget for the infrastructure and its maintenance. So cloud can minimizes expenses of any small organization in terms of economy. It leads to the developers can only focus on the development logic and not on the maintenance of the infrastructure.

There are three different **Pricing Strategies** that are introduced by Cloud Computing: Tiered Pricing, Per-unit Pricing, and Subscription-based Pricing. These are explained as following below.

1. **Tiered Pricing:** Cloud Services are offered in the various tiers. Each tier offers to fix service agreements at a specific cost. Amazon EC2 uses this kind of pricing.
2. **Per-unit Pricing:** The model is based upon the unit-specific service concept. Data transfer and memory allocation include in this model for specific units. GoGrid uses this kind of pricing in terms of RAM/hour.
3. **Subscription-based Pricing:** In this model, users are paying periodic subscription fees for the usage of the software.

So these models give more flexible solutions to the cloud economy.

- **Benchmarking:** Calculate the cost of operating your current data center, including capital costs over the equipment lifespan, labor costs and any other maintenance and operational costs, from licenses and software to spare parts.
- **Cloud costs:** Estimate the costs of the cloud infrastructure you're considering ([public cloud](#), [private cloud](#), hybrid cloud, etc.). You'll need a quote from your vendor, but look beyond this basic pricing structure to consider ongoing fees, labor and training costs, ongoing integration and testing of apps, as well as security and compliance.
- **Migration costs:** Determine the cost to migrate IT operations to the cloud or to switch cloud providers. These costs should include labor and expenses to integrate and test apps.

Challenges in Cloud

- 1. Data Security and Privacy**
- 2. Cost Management**
- 3. Multi-Cloud Environments**
- 4. Performance Challenges**
- 5. Interoperability and Flexibility**
- 6. High Dependence on Network**
- 7. Lack of Knowledge and Expertise**

Fog Computing

Fog computing is **a decentralized computing infrastructure in which data, compute, storage and applications are located somewhere between the data source and the cloud**. Like edge computing, fog computing brings the advantages and power of the cloud closer to where data is created and acted upon.

1. The devices comprising the fog infrastructure are known as fog nodes.
2. In fog computing, all the storage capabilities, computation capabilities, data along with the applications are placed between the cloud and the physical host.
3. All these functionalities are placed more towards the host. This makes processing faster as it is done almost at the place where data is created.
4. It improves the efficiency of the system and is also used to ensure increased security.

Fog Computing can be used in the following scenarios:

1. It is used when only selected data is required to send to the cloud. This selected data is chosen for long-term storage and is less frequently accessed by the host.
2. It is used when the data should be analyzed within a fraction of seconds i.e Latency should be low.
3. It is used whenever a large number of services need to be provided over a large area at different geographical locations.
4. Devices that are subjected to rigorous computations and processings must use fog computing.
5. Real-world examples where fog computing is used are in IoT devices (eg. Car-to-Car Consortium, Europe), Devices with Sensors, Cameras (IIoT-Industrial Internet of Things), etc.

Advantages of fog computing

- This approach reduces the amount of data that needs to be sent to the cloud.
- Since the distance to be traveled by the data is reduced, it results in saving network bandwidth.
- Reduces the response time of the system.
- It improves the overall security of the system as the data resides close to the host.
- It provides better privacy as industries can perform analysis on their data locally.

Disadvantages of fog computing

- Congestion may occur between the host and the fog node due to increased traffic (heavy data flow).
- Power consumption increases when another layer is placed between the host and the cloud.
- Scheduling tasks between host and fog nodes along with fog nodes and the cloud is difficult.
- Data management becomes tedious as along with the data stored and computed, the transmission of data involves encryption-decryption too which in turn release data.

Applications of fog computing

- It can be used to monitor and analyze the patients' condition. In case of emergency, doctors can be alerted.
- It can be used for real-time rail monitoring as for high-speed trains we want as little latency as possible.
- It can be used for gas and oils pipeline optimization. It generates a huge amount of data and it is inefficient to store all data into the cloud for analysis.

Edge computing

Edge computing optimizes Internet devices and web applications by bringing computing closer to the source of the data. This minimizes the need for long distance communications between client and server, which reduces latency and bandwidth usage.

Edge computing is a networking philosophy focused on bringing computing as close to the source of data as possible in order to reduce latency and bandwidth use. In simpler terms, edge computing means running fewer processes in the cloud and moving those processes to local places, such as on a user's computer, an IoT device, or an edge server. Bringing computation to the network's edge minimizes the amount of long-distance communication that has to happen between a client and server.

benefits of edge computing

Cost savings

Performance

New functionality

- Decreased latency
- Decrease in bandwidth use and associated cost
- Decrease in server resources and associated cost
- Added functionality

Mobile Cloud Computing

It is defined as a combination of mobile computing, cloud computing, and wireless network that come up together purpose such as rich computational resources to mobile users, network operators, as well as to cloud computing providers.

Mobile Cloud Computing is meant to make it possible for rich mobile applications to be executed on a different number of mobile devices.

In this technology, data processing, and data storage happen outside of mobile devices.

Advantages Mobile Cloud Computing

Extended battery life.

Improvement in data storage capacity and processing power.

Improved synchronization of data due to “store in one place, accessible from anywhere ” platform theme.

Improved reliability and scalability.

Ease of integration.

Factors Fostering Adoption Of Mobile Cloud Computing

Trends and demands: Customers expect convenience in using companies' websites or applications from anywhere and at any time. Mobile Cloud computing is meant for this purpose. Users always want to access business applications from anywhere, so that they can increase their productivity, even when they are on the commute.

Improved and increased broadband coverage: 3G and 4G along with WiFi, femtocells, are providing better connectivity for mobile devices cloud computing.

Enabling technologies: HTML5, CSS3, a hypervisor for mobile devices, cloudlets and Web 4.0 are enabling technologies that will drive adoption of mobile cloud computing.

Characteristics Of Mobile Cloud Computing Application

Cloud infrastructure: Cloud infrastructure is a specific form of information architecture that is used to store data.

Data cache: In this, the data can be locally cached.

User Accommodation: Scope of accommodating different user requirements in cloud app development is available in mobile Cloud Computing.

Easy Access: It is easily accessed from desktop or mobile devices alike.

Cloud Apps facilitate to provide access to a whole new range of services.

Mobile Cloud Computing Working

On a remote data center, Mobile Cloud Applications are operated generally by a third-party, data is stored, and compute cycles are carried out.

The uptime, integration, and security aspects are taken care of, by a backend, which also enables support to a multitude of access methods.

These apps can function online quite well, however, they need timely updating. These need not be permanently stored on the device but they do not always occupy any storage space on a computer or communications device.

Moreover, it offers the same experience as that of a desktop application, while offering the portability of a web application.

Benefits of Mobile Cloud Computing

Mobile Cloud Computing saves Business money.

Because of the portability which makes their work easy and efficient.

Cloud consumers explore more features on their mobile phones.

Developers reach greater markets through mobile cloud web services.

More network providers can join up in this field.

Challenges of Mobile Cloud Computing

Low bandwidth: This is one of the big issues in mobile cloud computing. Mobile cloud use radio waves which are limited as compared to wired networks. Available wavelength is distributed in different mobile devices. Therefore, it has been three times slower in accessing speed as compared to a wired network.

Security and Privacy: It is difficult to identify and manage threats on mobile devices as compared to desktop devices because in a wireless network there are more chances of the absence of the information from the network.

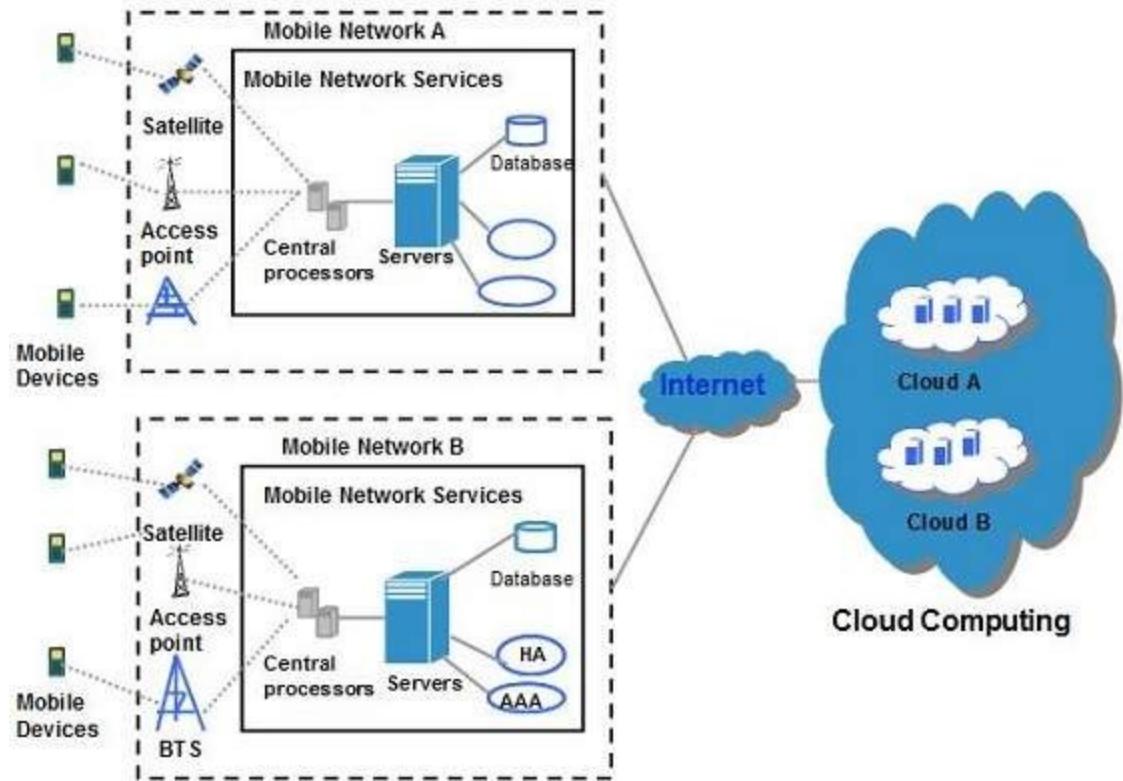
Service Availability: Users often find complaints like a breakdown of the network, transportation crowding, out of coverage, etc. Sometimes customers get a low-frequency signal, which affects the access speed and storage facility.

Alteration of Networks: Mobile cloud computing is used in a different operating system driven platforms like Apple iOS, Android, and Windows Phone. So it has to be compatible with different platforms. The performance of different mobile platform network is managed by the IRNA (Intelligent Radio Network Access) technique.

Limited Energy source: Mobile devices consume more energy and are less powerful. Mobile cloud computing increases battery usage of mobile devices which becomes an important issue. Devices should have a long-life battery to access applications and other operations. When the size of the altered code is small, the offloading consumes more energy than local processing.

MCC includes four types of cloud resources:

- Distant mobile cloud
- Distant immobile cloud
- Proximate mobile computing entities
- Proximate immobile computing entities
- Hybrid



Business Transformation with Google Cloud Superpowers of Cloud