MODULE - 1

INTRODUCTION

Computing is being transformed into a model consisting of services that are commoditized and delivered in a manner similar to utilities such as water, gas, electricity and telephony

Cloud computing is a technological advancement that focuses on the way we design computing systems, develop applications and leverage existing services for building software. It is based on the concept of dinamic provisioning, which is applied not only to services but also to compute capability, storage, networking and IT infrastructure

1.1 Cloud computing at a glance

Cloud computing allows renting infrastructure, runtime environments, and services on a pay-per-use basis. The principle finds several practical applications and then gives different images of cloud computing to different people.

Chief information and technology officers of large enterprises see opportunities for scaling their infrastructure on demand and sizing it according to their business needs. Endusers leveraging cloud computing services can access their documents & data anytime, anywhere and from any devices connected to Internet

The Vision of cloud computing

1. Cloud computing provides the facility to provision virtual hardware, runtime environment and services to a person having money

2. These all things can be used as long as they are needed by the user
3. The whole collection of computing system is transformed into collection of utilities, which can be provisioned and composed together to deploy systems in hours rather than days, with no maintanance cost
4. The long term vision ~~of~~ is that IT services are traded as utilities in an open market without technological and legal barriers
5. Due to the existence of global platform for trading cloud services will also help services providers to potentially increase their revenue.
6. A cloud provider can also become a consumer or buyer of a competitive service to fulfil its promises to customers

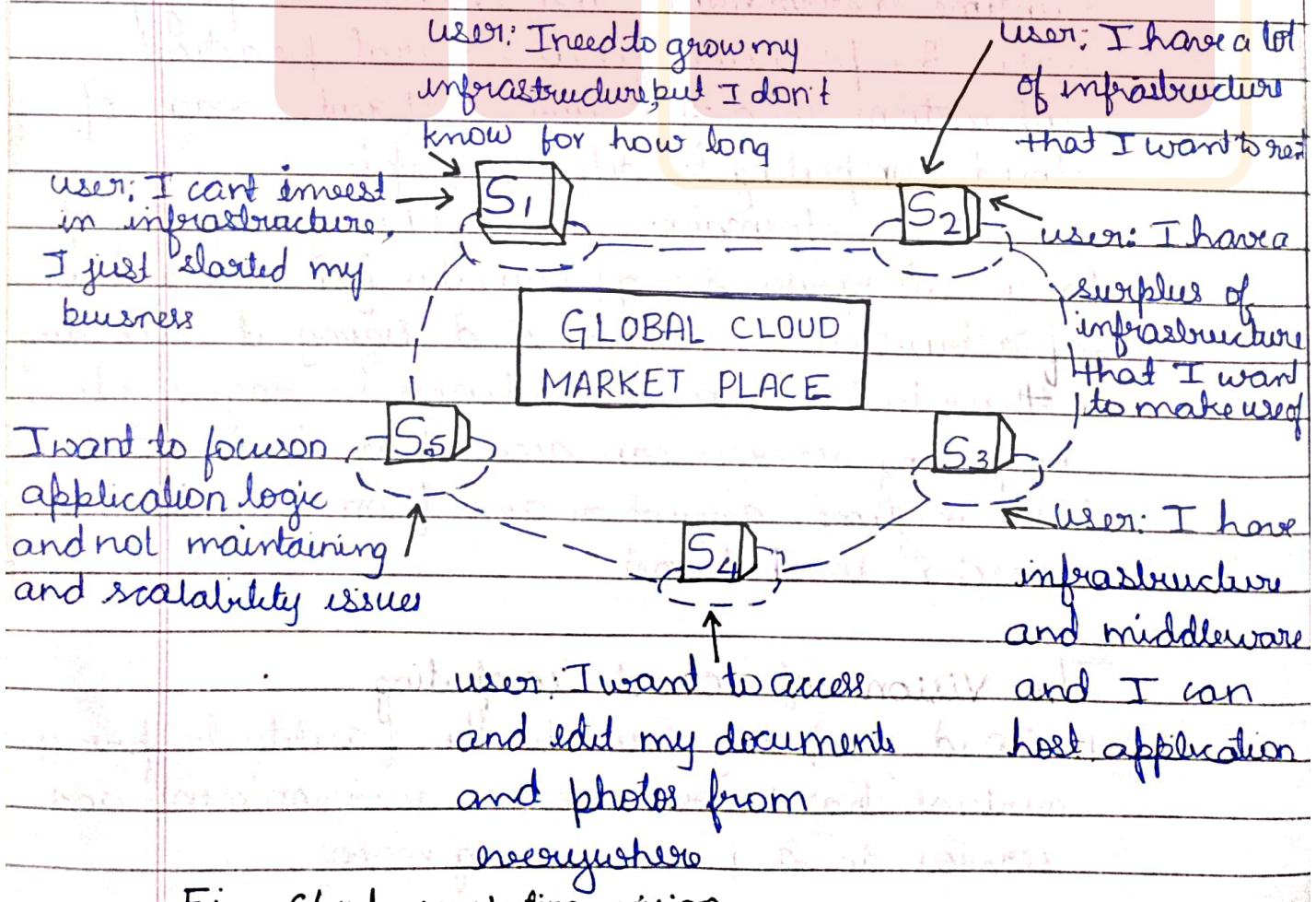


Fig: Cloud computing vision

Defining a cloud.

According to Armbrust et al:, "Cloud computing refers to both the applications delivered as services over the Internet and the hardware and system software in the data centers that provide those services"

U.S. National Institute of Standards and Technology defines "Cloud Computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly ~~shared~~ provisioned and released with minimal management effort or service provider interaction

A closure look:

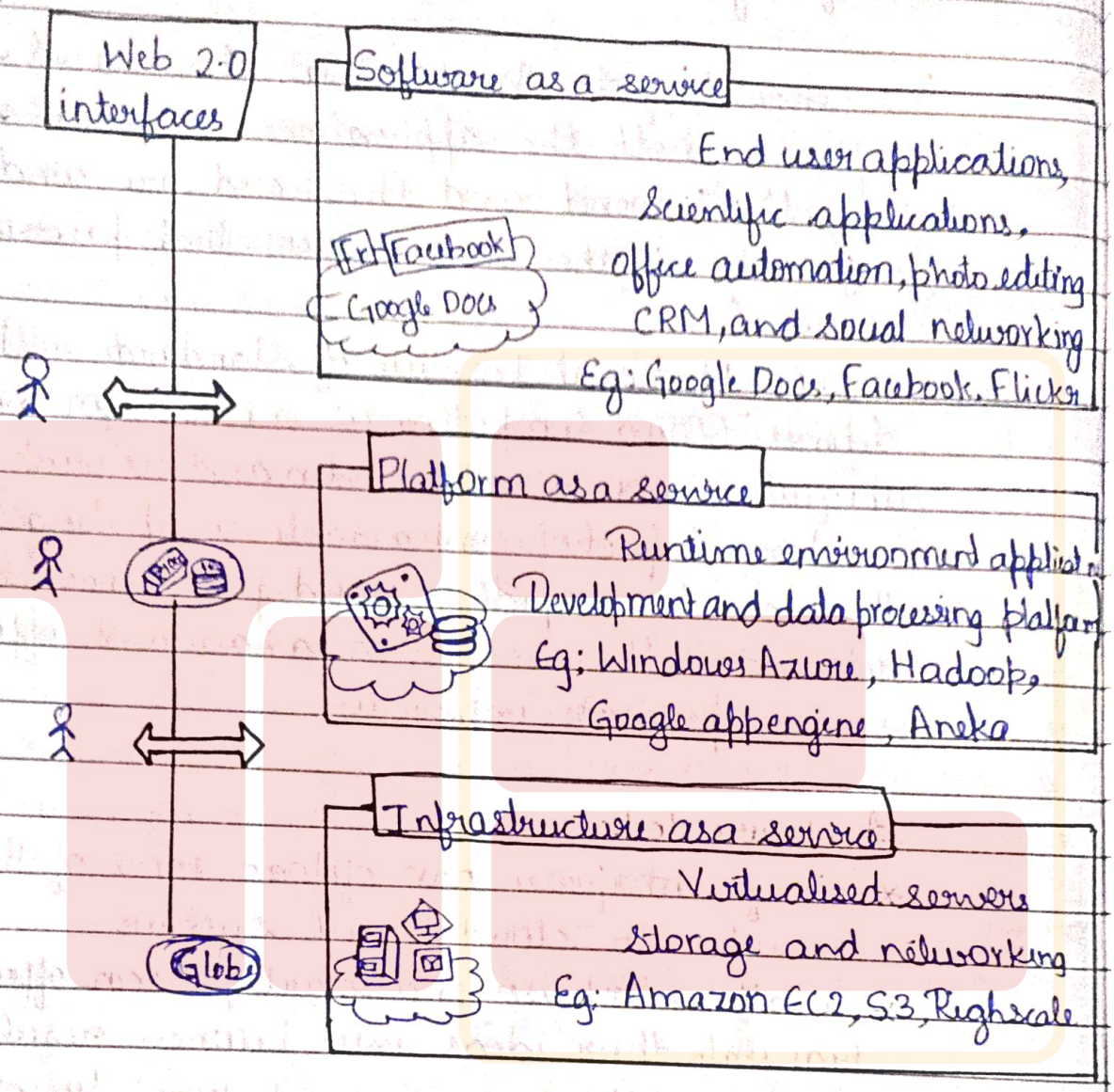
* Large enterprises can offload some of their activities to cloud-based systems
* Small enterprises and startups can afford to translate their ideas into business results more quickly, without excessive up-front ~~fo~~ costs.
* System developers can concentrate on the business logic rather than dealing with the complexity of infrastructure management and scalability
* End users can have their documents accessible from anywhere and any device

The three major models for deploying and accessing cloud computing environments are

- public clouds
- private / enterprise clouds
- hybrid clouds

VV. Imp

Cloud Computing Reference Model



The fundamental characteristics of cloud computing is the capability to deliver, on-demand, a variety of IT services that are quite diverse from each other. This variety creates different perceptions of what cloud computing is among users. Despite this lack of uniformity, it is possible to classify cloud computing services offering into 3 major categories

- Infrastructure as a service (IAAS)
- platform as a service (PAAS)
- Software as a service (SAAS)

IAAS : It often provides the infrastructure such as servers, virtual-machine networks, OS, storage and much more on a pay-as-you-use basis

* IAAS providers offer VM from small to extra large machines
* IAAS gives you complete freedom while choosing the instance type as your requirement
* ex: Amazon EC2, VCloud, Rightscale

PAAS : It is similar to IAAS, but it also provides additional tool such as DBMS, Business Intelligence services and so on.

* PAAS delivers a framework for developers that they can build upon and use to create customized applications
* ex: Windows Azure, Hadoop, Aneka

SAAS : It makes the users connect to the products through internet on a subscription basis model

* SAAS utilise the Internet to deliver applications, which are managed by a third party vendor, to its users.
* A majority of SAAS applications runs directly through your web browser, which means they do not require any downloads or installation on the client side
* eg: Google Document, Facebook, Flicker

Imp ** Characteristics and benifits

Cloud computing has some interesting characteristics that bring benifits to both cloud ~~computing~~ service consumers (CSC's) and providers (CSP's)

The characteristics are,

- No up-front commitments
- On-demand access
- Nice pricing
- Simplified application acceleration and scalability
- Efficient resource allocation
- Energy efficiency
- Seamless creation and use of third party service

Benifits of cloud computing are,

* Reduced IT costs: It may reduce the cost of managing and maintaining IT systems. Rather than purchasing expensive systems, you can use resources of cloud computing service provider
* Scalability: Business can scale up or scale down your operation and storage needs quickly to suit your situation, allowing flexibility as your needs change
* Business Continuity: Protecting your data and systems is an important part of business continuity planning
* Collaboration of efficiency: Collaboration in cloud environment gives your business the ability to communicate and share more easily outside of the traditional methods
* Flexibility of work practices: Cloud computing allows employees to be more flexible in their work practices
* Access to automatic updates for your IT requirement may be included in your service fee. This could include upto-date-versions of software, upgrades to servers and computer processing power.

Challenges Ahead

Challenges are regularly being posed to the cloud community, including IT practitioners, managers governments and regulators.

* Technical challenges arise for cloud service providers for the management of large computing infrastructures and the use of virtualisation technologies on top of them.
* Security in terms of confidentiality, secrecy and protection of data in a cloud environment is another important challenge. Organisations do not own the infrastructure they use to process data and store information. This condition poses challenges for confidential data.
* Legal issues may arise. Different legislations about privacy in different countries may potentially create disputes as to the rights that third parties have to your data.

1.2 Historical Developments:

The idea of renting computing services by leveraging large distributed computing facilities has been around for long time. It dates back to the days of mainframe computers.

Five core technologies that played an important role in realisation of cloud computing are: distributed systems, virtualisation, Web 2.0, service orientation, and utility computing.

→ Distributed systems:

It is a collection of independent computers that appears to its users as a single coherent system.

Three major milestones have led to cloud computing

*Mainframe computing: Mainframes were powerful, highly reliable computers specialised for large data movement and massive I/O operations. They were mostly used by large organisations for bulk data processing tasks such as online transactions, enterprise resource planning, and other operations involving the processing of large amount of data

*Clusters computing: It started as low-cost alternative to the use of mainframes and super computers. Cluster technology contributed considerably to the evolution of tools and frameworks for distributed computing, including Condor, Parallel Virtual Machine (PVM), and Message Passing Interface (MPI)

* Grid Computing: It appeared in 1990s as an evolution of cluster computing. In an analogy to the power grid, it proposed a new approach to access large computational power, huge storage facilities and variety of services.

Several developments made possible the diffusion of computing grids

a) Clusters became quite common resources

b) they were often underutilized

c) new problems were requiring computational power

d) the improvements in networking made possible long-distance, high-bandwidth connectivity

→ Virtualization

It encompasses a collection of solutions allowing the abstraction of some of the fundamental elements for computing such as hardware, runtime environments, storage and networking. Application of virtualization has always been limited by technologies that didnot allow an efficient use of virtualization solutions.

Virtualization is essentially a technology that allows creation of different computing environments. Virtualization technologies are also used to replicate runtime environment for programs. This technique allows isolating the execution of applications and providing a finer control on the resources they access.

→ Web 2.0

Web is the primary interface through which cloud computing delivers its services. At present, Web encompasses a set of technologies and services that facilitate interactive information sharing, collaboration, user-centered design and application composition. This evolution has transformed the Web into a rich platform for application development and is known as Web 2.0

It brings interactivity and flexibility into Web pages, providing enhanced user experience by gaining Web based access to all the functions that are normally found in desktop applications.

Eg: Google Docs, Google Maps, Flickr, Facebook, Twitter

→ Service oriented Computing (SOC)

SOC is the core reference model for cloud computing systems. This approach adopts the concept of services as the main building blocks of application and system development

A service is supposed to be loosely coupled, reusable, programable language independent and location transparent.

SOC introduces two important concepts.

Quality of service (QoS)

Software-as-a Service (SaaS)

* QoS identifies a set of functional and non functional attributes that can be used to evaluate the behaviour of a service from different perspectives.
* SaaS introduces a new delivery model for applications. The term has been inherited from the world of Application Service Providers (ASPs)

→ Utility Oriented Computing

Utility Computing is a vision of computing that defines a service-provisioning model for compute services in which resources such as storage, compute power, applications and infrastructure are packaged and offered on a pay per use basis. The idea of providing computing as a utility has a long history but has become a reality today with advent of cloud computing

1.3 Building Cloud Computing Environments

The creation of cloud computing environments encompasses both the development of applications and systems that leverage cloud computing solutions and the creation of frameworks, platforms, and infrastructures delivering cloud computing services

Application Development

Applications that leverage cloud computing benifit from its capability to dynamically scale on demand

* One class of applications that take the biggest advantage of this feature is that of web applications. With the diffusion of Web 2.0 technologies, the web has become a platform for developing rich and complex applications, including enterprise applications that now leverage the Internet for service delivery and user interaction.

* Another class of applications that can potentially gain considerable advantage by leveraging cloud computing is represented by resource-intensive applications. These can be either data-intensive or compute-intensive applications
* Cloud computing provides a solution for on-demand and dynamic scaling across the entire stack of computing. This is achieved by
  - Providing methods for renting compute power, storage, and networking
  - Offering run time environments designed for scalability and dynamic sizing
  - providing application services that mimic the behaviour of desktop applications

Infrastructure and system development

* Distributed computing is a foundational model for cloud computing because cloud systems are distributed systems
* Web 2.0 technologies constitute the interface through which cloud computing services are delivered, managed and provisioned
* Cloud computing is often summarized with the acronym XaaS → Everything-as-a-service that clearly underlines the central role of service orientation.
* Virtualization is another element that plays a fundamental role in cloud computing. This technology is a core feature of the infrastructure used by cloud providers.

V. Imp

**COMPUTING PLATFORMS AND TECHNOLOGIES

Development of a cloud computing application happens by leveraging platforms and frameworks that provide different types of services, from the bare metal infrastructure to customizable applications serving specific purposes

*1 1. Amazon web services (AWS)

* AWS offers comprehensive cloud IaaS services ranging from virtual compute, storage, and networking to complete computing ~~stak~~ stacks.
* AWS is mostly known for its compute storage-on-demand services, namely Elastic Compute Cloud (EC2) and Simple Storage Services (S3)
* EC2 provides users with customizable virtual hardware and also provides the capability to save a specific ~~rum~~ running instance as an image.
* S3 is organized into buckets; these are containers of objects that are stored in binary form and can be enriched with attributes

2. Google App Engine

* GAE is a scalable runtime environment mostly devoted to executing Web applications
* It provides both a secure execution environment and a collection of services that simplify the development of scalable and high performance Web applications. These services include in-memory caching scalable data store, job queues and messaging
* Developers can build and test applications on their own machines using AppEngine Software Development Kit (SDK)
* The languages supported are Python, Java and Go

*3. Microsoft Azure

* It is a cloud operating system and a platform for developing applications in the cloud.
* Applications in Azure are organised around the concept of roles, which identify a distribution unit for applications and embody the application's logic
* There are three types of role.
  - Web role : designed to host a web application
  - Worker role: used to perform workload processing
  - Virtual machine role : provides virtual environment

4. Hadoop

Apache Hadoop is an open source framework that is suited for processing large data sets on commodity hardware. Hadoop is an implementation of MapReduce, an application programming model developed by google, which ~~wt~~ provides two fundamental operations for data processing : map and reduce. It provides the runtime environment.

5. Force.com and Salesforce.com

* Force.com is a cloud computing platform for developing social enterprise applications.
* The platform is the basis for SalesForce.com, a Software-as-a-Solution for a customer relationship management.
* Force.com allows developers to create applications by composing ready to use blocks
* The platform provides complete support for developing applications

6. Manjrasoft Aneka

* It is a cloud application platform for rapid creation of scalable applications and their development on various types of clouds in a seamless and elastic manner.
* It supports a collection of programing abstractions for developing applications and a distributed runtime environment that can be deployed on heterogeneous hardware.
* Developers can choose different abstractions to design their applications.

2. VIRTUALIZATION

*** Virtualization technology is one of the fundamental components of cloud computing. It allows the creation of a secure, customizable, and isolated execution environment for running applications, even if they are untrusted, without affecting other users' applications. The basis of this technology is the ability of a computer program to emulate an executing environment separate from the one that hosts such programs.

2.1 Introduction

Virtualization is meant to provide an abstract environment - whether virtual hardware or OS - to run applications. It plays a fundamental role in efficiently delivering Infrastructure-as-a-service (IaaS) solutions for cloud computing.

Virtualization technologies have gained renewed interested recently due to the confluence of several phenomena.

* Increased performance and computing capacity

* Underutilized hardware and software resources
  Due to increased performance and computing capacity
  • the effect of limited or sporadic use of resource
* Lack of space
* Greening initiatives
* Rise of administrative costs

V.Imp
2.2 Characteristics of virtualized environments

In a virtualized environment there are three major components: guest, host and virtualization layer.

- Guest → represents the system component that interacts with virtualization layer rather than with a host
- Host → represents the original environment where the guest is suppose to be managed.
- Virtualization layer → it is responsible for recreating the same or a different environment where the guest will operate

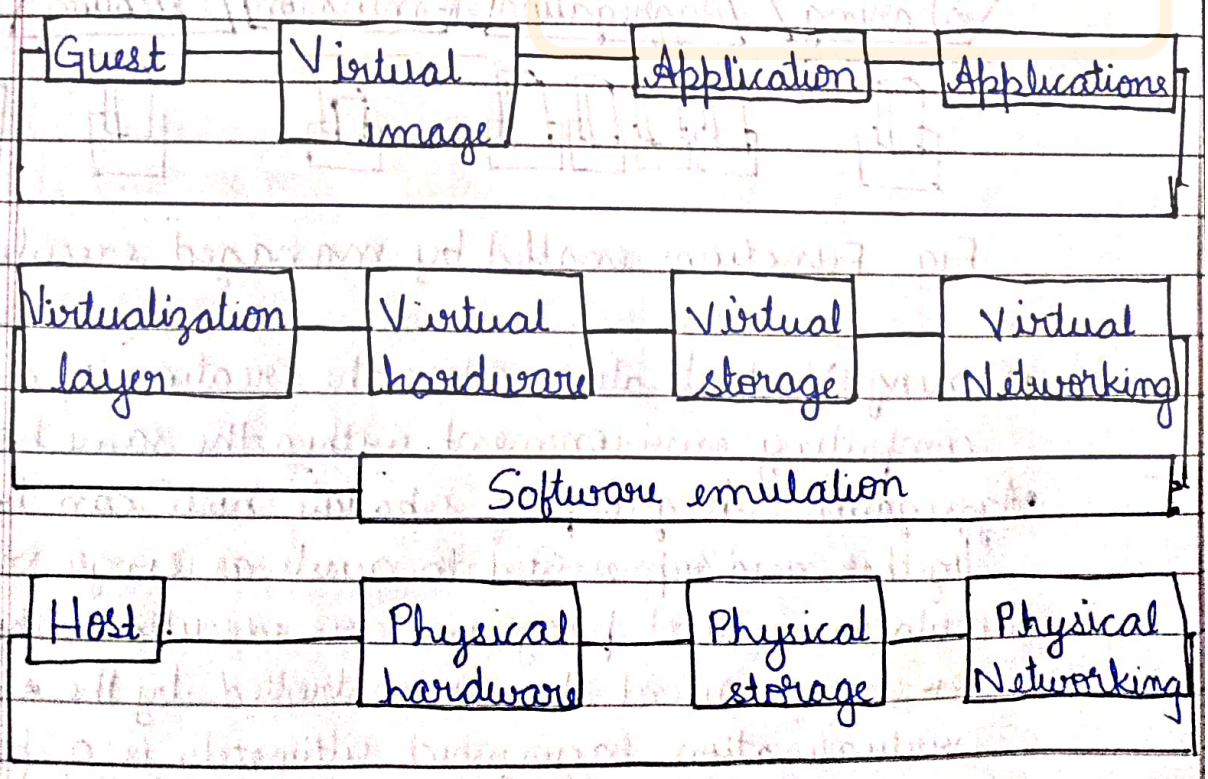


Fig: Virtualization reference model

The characteristics of virtualized solutions are,

* Increased security

The virtual machine represents an emulated environment in which the guest is executed. All the operations of the guest are generally performed against the virtual machine, which translates and applies them to host. This level of indirection allows the virtual machine manager to control and filter the activity of the guest, thus preventing some harmful operations from being performed

* Managed execution

Virtualization of the execution environment not only allows increased security but also features like sharing, aggregation, emulation and isolation

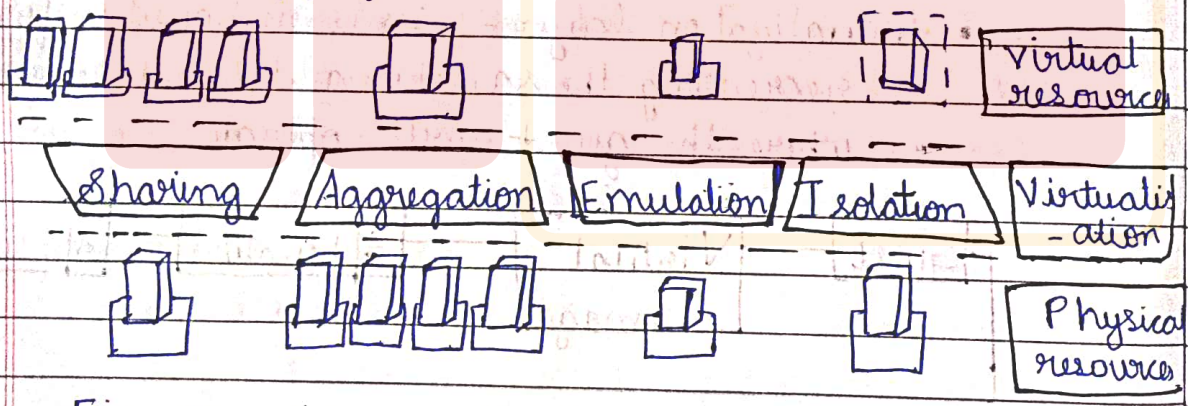


Fig: Functions enabled by managed execution

- Sharing: Virtualization allows the creation of a separate computing environment within the same host
- Aggregation: A group of separate hosts can be tied together and represented to guests as single virtual host
- Emulation: Guest programs are executed within an environment that is controlled by the virtualization layer, which ultimately is a program
- Isolation: The guest program performs its activity

by interacting with an abstraction layer, which provides access to the underlying resources.

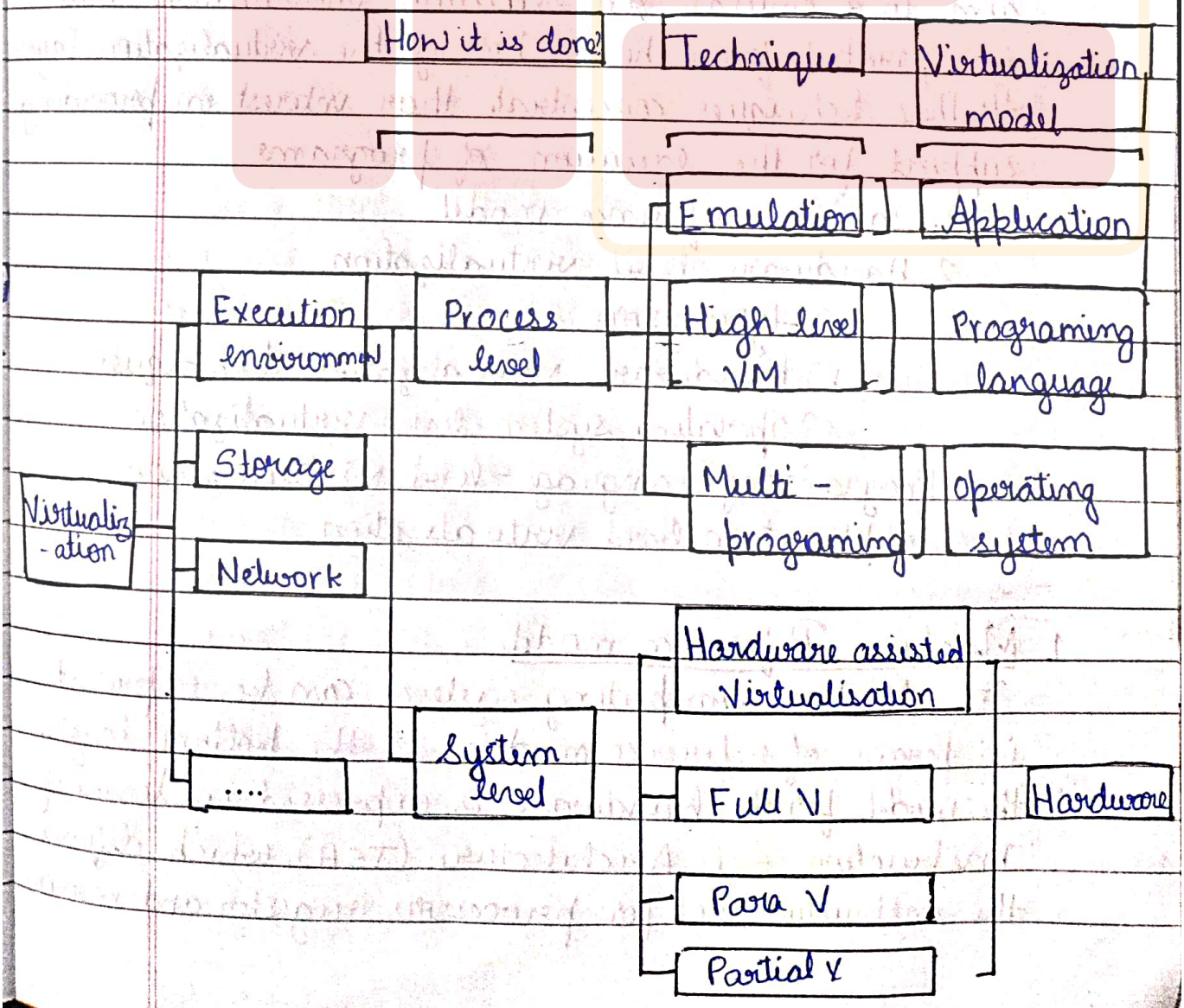
*Portability

The concept of portability applies in ~~d~~ different ways according to the specific type of virtualization

- In hardware virtualization solution, the guest is packed into a virtual image that can be safely moved and executed on top of virtual machines
- In programing level virtualization, the binary code can be run without any recompilation on any implementation of the corresponding virtual machine

Imp

2.3 Taxonomy of virtualization techniques



* Virtualization covers a wide range of emulation techniques that are applied to different areas of computing.
* The first classification discriminates against the service or entity that is being emulated
* It is mainly used to emulate execution environment, storage and networks.
* Among these categories, execution virtualization constitutes the oldest, most popular and most developed area. Therefore, it deserves major investigation and further categorization

Execution Virtualization

Execution virtualization includes all techniques that aim to emulate an execution environment that is separated from the one hosting the virtualization layer. All these techniques concentrate their interest on providing support for the execution of programs.

1. Machine reference model
2. Hardware level virtualization
   a) Hypervisors
   b) Hardware virtualization techniques
   c) Operating system-level virtualization
3. Programing language-level virtualization
4. Application level virtualization

1 Machine Reference model

Modern computing systems can be expressed in terms of reference model. At the bottom layer, the model for the hardware is expressed in terms of Instruction Set Architecture (ISA), which defines the instruction set for processors, register and memory

* ISA is the interface between hardware and software and it is important to the operating system (OS) developer (System ISA) and to the User ISA.
* Application Binary Interface (ABI) separates the OS ~~cover system~~ layer from the applications and libraries, which are managed by the OS.
* The highest level of abstraction is represented by the Application Programming Interface (API), which interfaces applications to libraries and/or the underlying OS.

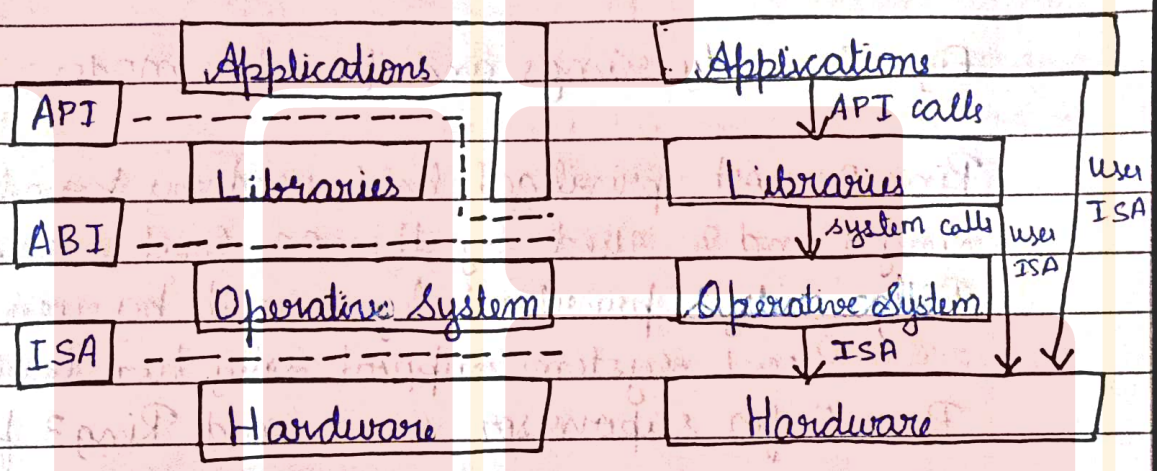


Fig: Machine reference model

* The instruction set exposed by the hardware has been divided into different security classes that define who can operate with them. The first distinction can be made between privileged and non privileged instructions.
* Non privileged instructions are those which can be used without interfering with other tasks because they donot access shared resources.

Privileged are those that are executed under specific ~~instructions~~ restrictions and are mostly used for sensitive operations.

For instance, a possible implementation features a hierarchy of privileges in the form of ring based security

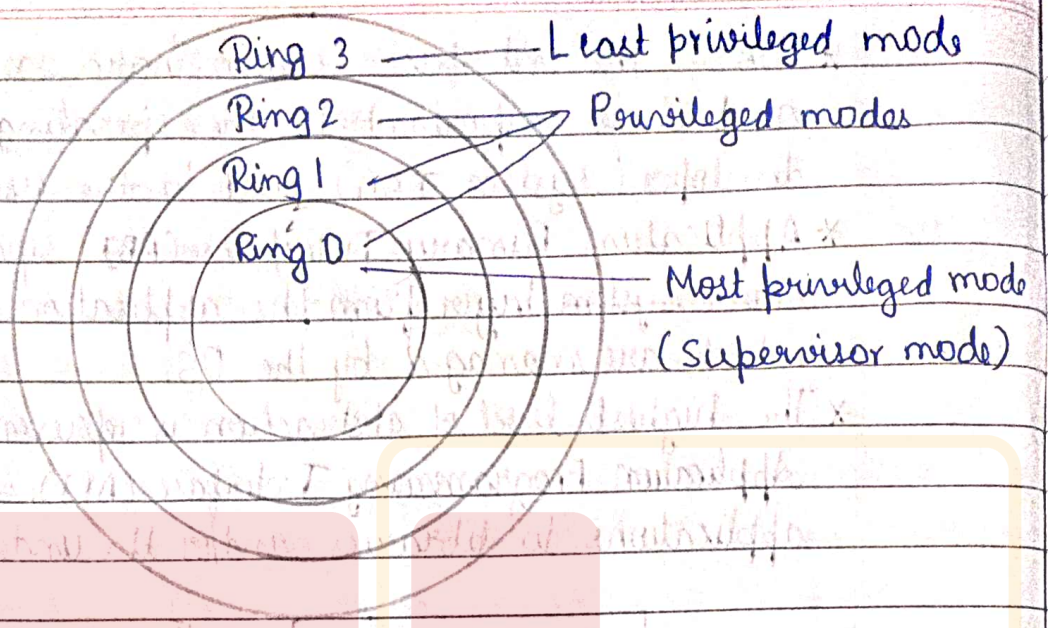


Fig: Security rings and privilege modes

Ring 0: most privileged level, used by kernel of OS
Ring 1 and 2: used by the OS level services
Ring 3: least privileged level, used by user.
Recent systems support only two levels, with Ring 0 for supervisor mode and Ring 3 for user mode.

2. <u>Hardware level Virtualization</u>

It is a virtualization technique that provides an abstract execution environment in terms of computer hardware on top of which a guest OS can be run.

In this model, guest is represented by the OS, the host by the physical computer hardware, the virtual machine by its emulation, and the virtual machine manager by the hypervisor.

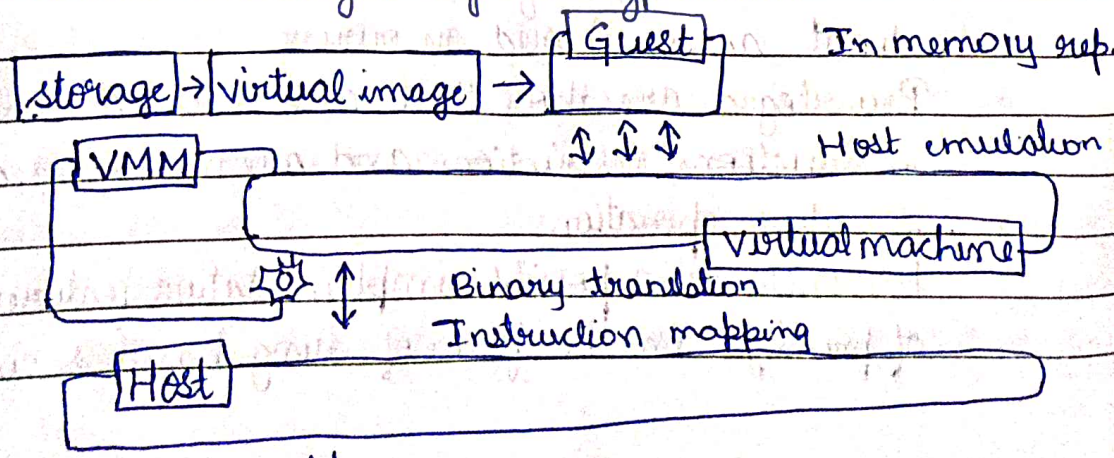


Fig: Reference model

a) Hypervisors

* A fundamental element of hardware virtualization is the hypervisor, or virtual machine manager (VMM).
* It recreates the hardware environment in which guest OS are installed.

imp ** There are two major types of hypervisors: Type I and Type II

* Type I hypervisors: run directly on top of the hardware. Therefore, they take place of the operating systems and interact directly with the ISA interface exposed by the underlying hardware. It is also called native virtual machine since it runs natively on hardware.
* Type II hypervisors: require the support of an OS to provide virtualization services. This means that they are programs managed by the OS, which interact with it through the ABI. It is also called a hosted virtual machine, since it is hosted with in an OS.

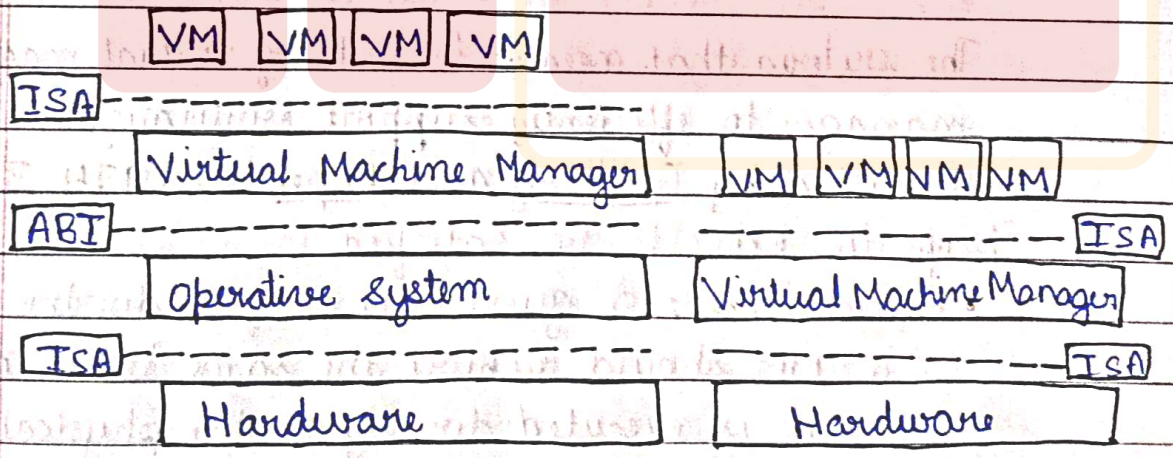


Fig: Hosted (left) and native (right) virtual machines

* Virtual machine manager is internally organised as shown below. Three main modules - dispatcher, allocator and interpreter, coordinate their activity in order to emulate the underlying hardware.
* Dispatcher - constitutes the entry point of the monitor

Allocator - responsible for deciding the system resources to be provided to the VM.

Interpreter - module consists of ~~instructions~~ interpreter routines. These are executed whenever a VM executes a privileged instruction.

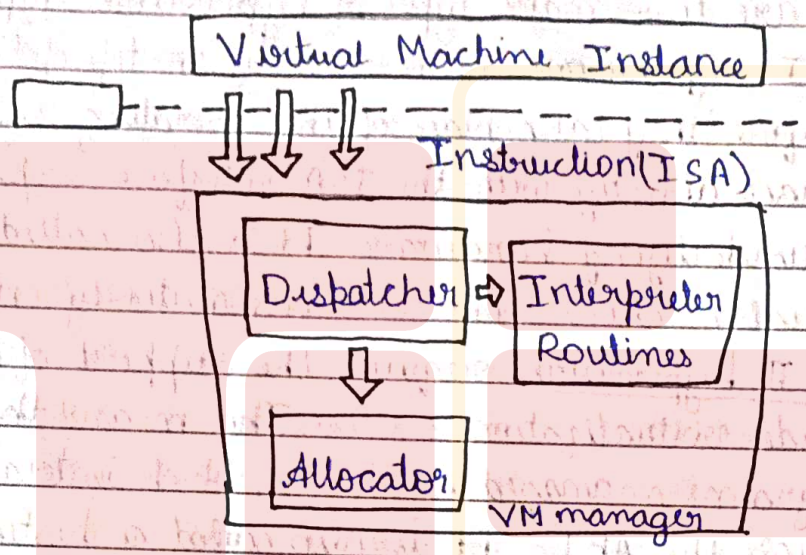


Fig: Hypervisor reference architecture

The criteria that need to be met by virtual machine (VMM) manager to efficiently support virtualization were established by Goldberg and Popek in 1974. Three properties have to be satisfied

- Equivalence : A guest running under the control of a VMM should exhibit the same behaviour as when it is executed directly on the physical host
- Resource control ; The VMM should be in complete control of virtualized resources.
- Efficiency : A statistically dominant fraction of the machine instructions should be executed without intervention from the VMM.

Imp ***

Popek and Goldberg provided a classification of the instruction set and proposed three theorems that define the properties that hardware instructions need to satisfy inorder to efficiently support virtualization

Theorem 3.1: For any conventional third-generation computer, a VMM may be constructed if the set of sensitive instructions for that computer is a subset of the set of privileged instructions.

This theorem establishes that all the instruction that change the configuration of the system resources should generate a trap in user mode and be executed under the control of VMM.

Theorem 3.2: A conventional third generation computer is recursively virtualizable if;

- It is virtualizable and
- A VMM without any timing dependencies can be constructed for it

Recursive virtualization is the ability to run a VMM on top of another VMM. This allows nesting hypervisors as long as the capacity of the underlying resources can accomodate that. Virtualizable hardware is a prerequisite to recursive virtualization

Theorem 3.3: A hybrid VMM may be constructed for any conventional third generation machine in which the set of user sensitive instructions is a subset of the privileged instructions

There is another term, hybrid virtual Machine (HVM), which is less efficient than virtual machine

system. In the case of an HVM, more instructions are interpreted rather than being executed directly. All instructions in virtual supervisor mode are interpreted. Whenever there is an attempt to execute a behaviour sensitive or control sensitive instruction, HVM controls the execution directly or gains the control via trap. Here all instructions are caught by HVM that are simulated.

***

b) Hardware virtualization techniques:

- Hardware assisted virtualization: This refers to a scenario in which the hardware provides architectural support for building a virtual machine manager able to run a guest OS in complete isolation. This technique was originally introduced in the IBM System /370.
- Full Virtualization: It refers to the ability to run a program, most likely an OS, directly on top of a virtual machine and without any modification as though it was run on the raw hardware. The advantage of full virtualization is complete isolation, which leads to enhanced security, easy of emulation of different architecture and coexistence of different systems on the same platform.
- Para Virtualization: It expose a software interface to the VM that is slightly modified from the host and, as a consequence, guests need to be modified. Its aim is to provide the capability to demand the execution of performance-critical operations directly on the host, thus preventing performance losses that would otherwise be ~~expe~~ experienced in managed execution.
- Partial Virtualization: It provides a partial

emulation of the underlying hardware, thus not allowing the complete execution of the guest OS in complete isolation. It allows many applications to run transparently, but not all the features of the OS can be supported

c) Operating system level virtualization
* It offers the opportunity to create different and separate execution environments for applications that are managed concurrently.
* The virtualization is done within a single OS, where the OS kernel allows for multiple isolated user space instances
* ex: Free BSD jails, IBM Logical partition (LPAR) Open VZ, iCore Virtual accounts, Free VPS

3. Programming Language-level virtualization
* It is mostly used to achieve ease of deployment of applications, managed execution and portability across different platforms and OS.
* It consists of VM executing the byte code of a program, which is the result of the compilation process
* Also called as process virtual machines
* Advantage: has the ability to provide a uniform execution environment across different platforms
* Security is another advantage of it

4. Application level virtualization
* It is a technique allowing applications to be in run ~~on~~time environments that donot natively support all the features requires by such applications

* Emulation can also be used to execute program binaries compiled for different hardware architecture. In this case, one of these strategies can be implemented
- Interpretation - Every source instruction is ~~converted to native instructions~~ interpreted by an emulator for executing native ISA instructions
- Binary translation - Every source instruction is converted into native instructions with equivalent functions.

Other types of Virtualization

1. <u>Storage Virtualization</u> :
It is a system administration practice that allows decoupling the physical organisation of the hardware from its logical representation. Using this, users donot have to be worried about the specific location of the data, which can identified using a logical path

2. <u>Network Virtualization</u> :
It combines hardware appliances and specific software for the creation and management of a virtual network. It can aggregate different physical networks into external or internal network virtualization

3. <u>Desktop Virtualization</u>
It abstracts the desktop environment available on a personal computer in order to provide access to it using client/server approach. It makes accessible a different system as though

it were natively installed on the host, but this system is remotely stored on a different host and accessed through a network connection

4. Application server virtualization

A collection of application servers that provide the same services as a single virtual application server by using load balancing strategies and providing a high availability infrastructure for the services hosted in the application server.

3.4. Virtualization and cloud computing

Virtualization plays an important role in cloud computing since it allows for the appropriate degree of customization, security, isolation, and manageability that are fundamental for delivering IT services on demand.

Besides being an enabler for computation on demand, virtualization also gives the opportunity to design more efficient computing systems by means of consolidation, which is performed transparently to cloud computing service users.

Since virtualization allows us to create isolated and controllable environments, it is possible to serve these environments with the same resources without them interfering with each other. This opportunity is particularly attractive when resources are underutilized, because it allows reducing the number of active resources by aggregating virtual machines over small number of resource that become fully utilized. This service is also known as server consolidation, while the movement of VM instances is called VM Migration.

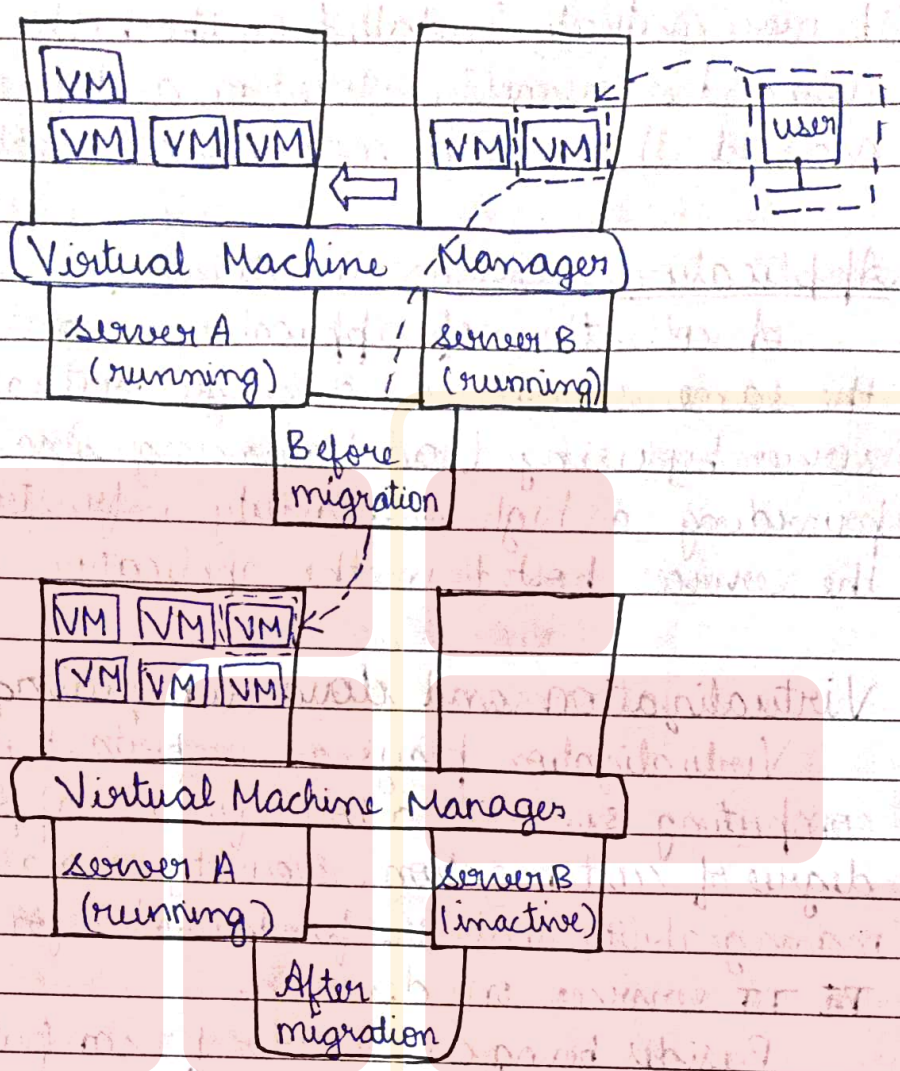


Fig: Live migration and server consolidation

imp

2.5 Pros and cons of virtualization

The capillary diffusion of the internet connection and the advancements in computing technology have made virtualization an interesting opportunity to deliver on-demand IT infrastructure and services

Advantages of virtualization

* Gain better performance and efficiency from resources in the existing computing component using CPU virtualization
* Boost VM security
* Portability

* Faster provisioning of applications and resources
* Save money on hardware virtualization software involve less cost
* Gain peace of mind and provide better reliability interms of disaster recovery as well as better backup and retrieval capabilities.

Disadvantages of Virtualization

* Performance degradation:
  Performance is definitely one of the major concern in virtualization. Since virtualization interposes an abstraction layer between the guest and the host, the guest can experience increased latencies
* Inefficiency and degraded user experience
  Virtualization can sometime lead to an inefficient use of the host. Some of the specific features of the host cannot be exposed by the abstraction layer and then become inaccessible.
* Security holes and new threats
  The capability of emulating a host in a completely transparent manner led the way to malicious programs that are designed to extract sensitive information from the guest.

2.6 Technology Examples

A wide range of virtualization technology is available especially for virtualizing computing environments.

Imp* Xen: Para Virtualization

* Xen is an open source initiative implementing a virtualization platform based on para-virtualization.

* Xen based technology are used for either desktop virtualization or server virtualization and recently it has also been to provide cloud computing solutions by means of Xen cloud platform

Imp

VMWare : Full Virtualization

* VM Ware's technology is based on the concept of full virtualization, where the underlying hardware is replicated and made available to the guest OS which runs unaware of such abstraction layers and doesnot need to be modified.
* VM ware implements full virtualization either in desktop environment, by means of type II hypervisors or in the server environment by means of Type I hypervisor.
* In both cases, full virtualization is made possible by means of direct execution and binary translation, thus allowing the virtualization of architecture such as X86

Microsoft Hyper-V

Hyper-V is a Microsoft's hardware virtualisation product. It lets the user create and run a software version of a computer called virtual machine

Hyper-V runs each VM in its own isolated space, so that user can avoid problems such as crash affecting the other workloads, or to give different people or services access to different systems

Advantages: • Improve business continuity
• Use hardware more effectively
• Establish or expand a private cloud environment.