SUBJECT CODE : CCS335

Strictly as per Revised Syllabus of

ANNA UNIVERSITY

Choice Based Credit System (CBCS)

Vertical - 3 (Cloud Computing and Data Centre Technologies) (CSE/IT/AI&DS)

Vertical - 2 (Full Stack Development for IT) (IT/AI&DS)

Vertical - 2 (Cloud Computing and Data Centre Technologies) (CS&BS)

CLOUD COMPUTING

Iresh A. Dhotre

M.E. (Information Technology)

Ex-Faculty, Sinhgad College of Engineering,

Pune.

TECHNICAL PUBLICATIONS

SINCE 1993

An Up-Thrust for Knowledge

CLOUD COMPUTING

Subject Code : CCS335

Vertical - 3 (Cloud Computing and Data Centre Technologies) (CSE/IT/AI&DS)

Vertical - 2 (Full Stack Development for IT) (IT/AI&DS)

Vertical - 2 (Cloud Computing and Data Centre Technologies) (CS&BS)

Published by :

TECHNICAL PUBLICATIONS
An Up-Thrust for Knowledge

Amit Residency, Office No.1, 412, Shaniwar Peth,
Pune - 411030, M.S. INDIA, Ph.: +91-020-24495496/97
Email : info@technicalpublications.in Website : www.technicalpublications.in

Printer :
Yogiraj Printers & Binders
Sr.No. 10/1A,
Ghule Industrial Estate, Nanded Village Road,
Tal. - Haveli, Dist. - Pune - 411041.

ISBN 978-93-5585-436-0

9 789355 854360

AU 21

PREFACE

The importance of **Cloud Computing** is well known in various engineering fields. Overwhelming response to my books on various subjects inspired me to write this book. The book is structured to cover the key aspects of the subject **Cloud Computing**.

The book uses plain, lucid language to explain fundamentals of this subject. The book provides logical method of explaining various complicated concepts and stepwise methods to explain the important topics. Each chapter is well supported with necessary illustrations, practical examples and solved problems. All the chapters in the book are arranged in a proper sequence that permits each topic to build upon earlier studies. All care has been taken to make students comfortable in understanding the basic concepts of the subject.

Representative questions have been added at the end of section to help the students in picking important points from that section.

The book not only covers the entire scope of the subject but explains the philosophy of the subject. This makes the understanding of this subject more clear and makes it more interesting. The book will be very useful not only to the students but also to the subject teachers. The students have to omit nothing and possibly have to cover nothing more.

I wish to express my profound thanks to all those who helped in making this book a reality. Much needed moral support and encouragement is provided on numerous occasions by my whole family. I wish to thank the **Publisher** and the entire team of **Technical Publications** who have taken immense pain to get this book in time with quality printing.

Any suggestion for the improvement of the book will be acknowledged and well appreciated.

Author
D. A. Dhotre

Dedicated to Readers

SYLLABUS

Cloud Computing - [CCS335]

UNIT I CLOUD ARCHITECTURE MODELS AND INFRASTRUCTURE

Cloud Architecture : System Models for Distributed and Cloud Computing - NIST Cloud Computing Reference Architecture - Cloud deployment models - Cloud service models; Cloud Infrastructure : Architectural Design of Compute and Storage Clouds - Design Challenges. **(Chapter - 1)**

UNIT II VIRTUALIZATION BASICS

Virtual Machine Basics - Taxonomy of Virtual Machines - Hypervisor - Key Concepts - Virtualization structure - Implementation levels of virtualization - Virtualization Types : Full Virtualization - Para Virtualization - Hardware Virtualization - Virtualization of CPU, Memory and I/O devices. **(Chapter - 2)**

UNIT III VIRTUALIZATION INFRASTRUCTURE AND DOCKER

Desktop Virtualization - Network Virtualization - Storage Virtualization - System-level of Operating Virtualization - Application Virtualization - Virtual clusters and Resource Management - Containers vs. Virtual Machines - Introduction to Docker - Docker Components - Docker Container - Docker Images and Repositories. **(Chapter - 3)**

UNIT IV CLOUD DEPLOYMENT ENVIRONMENT

Google App Engine - Amazon AWS - Microsoft Azure; Cloud Software Environments - Eucalyptus - OpenStack. **(Chapter - 4)**

UNIT V CLOUD SECURITY

Virtualization System - Specific Attacks : Guest hopping - VM migration attack - hyperjacking. Data Security and Storage; Identity and Access Management (IAM) - IAM Challenges - IAM Architecture and Practice. **(Chapter - 5)**

TABLE OF CONTENTS

UNIT I

UNIT II

UNIT III

UNIT IV

UNIT V

UNIT I

1 Cloud Architecture Models and Infrastructure

Syllabus

Cloud Architecture : System Models for Distributed and Cloud Computing - NIST Cloud Computing Reference Architecture - Cloud deployment models - Cloud service models; Cloud Infrastructure: Architectural Design of Compute and Storage Clouds - Design Challenges

Contents

1.1 History of Cloud Computing

AU : Dec.-20,21,22

- Idea of cloud computing was introduced by computer scientist John McCarthy publicly in 1961. Then in 1969, Leonard Kleinrock, a chief scientist of the ARPANET project comments about Internet.
- The general public has been leveraging forms of Internet-based computer utilities since the mid-1990s through various incarnations of search engines, e-mail services, open publishing platforms and other types of social media.
- Though consumer-centric, these services popularized and validated core concepts that form the basis of modern-day cloud computing. The Salesforce.com provides remote service from 1990 to organizations.
- Amazon launched its web services in 2002 and it provides services to organizations for storage and remote computing. Cloud computing definition as per Gartner "a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies".
- In 2008, Gartner's original definition of cloud was changed. In the definition, "massively scalable" was used instead of "scalable and elastic."
- **NIST definition of cloud :** Cloud computing is a pay-per-use model for enabling available, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service-provider interaction.
- The above cloud definition was published by NIST in 2009, followed by a revised version after further review and industry input that was published in September of 2011.
- Cloud computing refers to a variety of services available over the Internet that deliver compute functionality on the service provider's infrastructure.
- Its environment (infrastructure) may actually be hosted on either a grid or utility computing environment, but that doesn't matter to a service user.

1.1.1 Introduction to Cloud Computing

- Cloud computing refer to a variety of services available over the Internet that deliver compute functionality on the service provider's infrastructure.
- Its environment (infrastructure) may actually be hosted on either a grid or utility computing environment, but that doesn't matter to a service user.

- Cloud computing is a general term used to describe a new class of network based computing that takes place over the Internet, basically a step on from Utility Computing.
- In other words, this is a collection/group of integrated and networked hardware, software and Internet infrastructure (called a platform).
- Cloud computing refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards.
- Fig. 1.1.1 shows cloud symbol. It denotes cloud boundary.
- Using the Internet for communication and transport provides hardware, software and networking services to clients.
- These platforms hide the complexity and details of the underlying infrastructure from users and applications by providing very simple graphical interface or API.

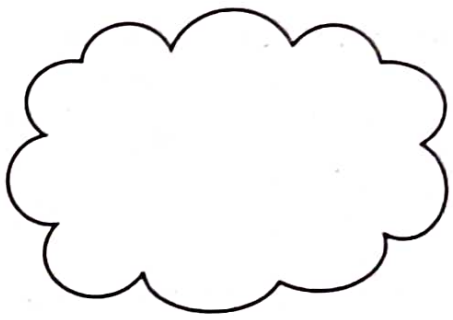


Fig. 1.1.1 Cloud symbol

- In addition, the platform provides on demand services that are always on anywhere, anytime and anyplace. Pay for use and as needed.
- The hardware and software services are available to the general public, enterprises, corporations and business markets.

IT resources :

- IT resources are of two types : Software based and hardware based.
- Software based resources are virtual server, custom software program and hardware based means physical server and networking devices.
- IT resources include server, virtual server, storage device, networking device, services and software programs.
- An on-premise IT resource can access and interact with a cloud-based IT resource.
- An on-premise IT resource can be moved to a cloud, thereby changing it to a cloud-based IT resource.
- **Cloud provider :** A person, organization, or entity responsible for making a service available to interested parties. When assuming the role of cloud provider, an organization is responsible for making cloud services available to cloud

consumers, as per agreed upon Service Level Agreement (SLA) guarantees. Cloud provider have their own IT resources.

- **Cloud consumer :** A person or organization that maintains a business relationship with, and uses service from, Cloud Providers. The cloud consumer uses a cloud service consumer to access a cloud service.
- **Cloud service owner :** The person or organization that legally owns a cloud service is called a cloud service owner. The cloud service owner can be the cloud consumer, or the cloud provider that owns the cloud within which the cloud service resides.
- **Resource administrator :** Cloud resource administrator is the person or organization responsible for administering a cloud-based IT resource. The cloud consumer or cloud provider, or even third-party organization could be a cloud resource administrator

Cloud types :

- Most people separate cloud computing into two distinct sets of models :
  1. **Deployment models :** This refers to the location and management of the cloud's infrastructure.
  2. **Service models :** This consists of the particular types of services that you can access on a cloud computing platform.
- Fig. 1.1.2 shows NIST cloud computing definitions.

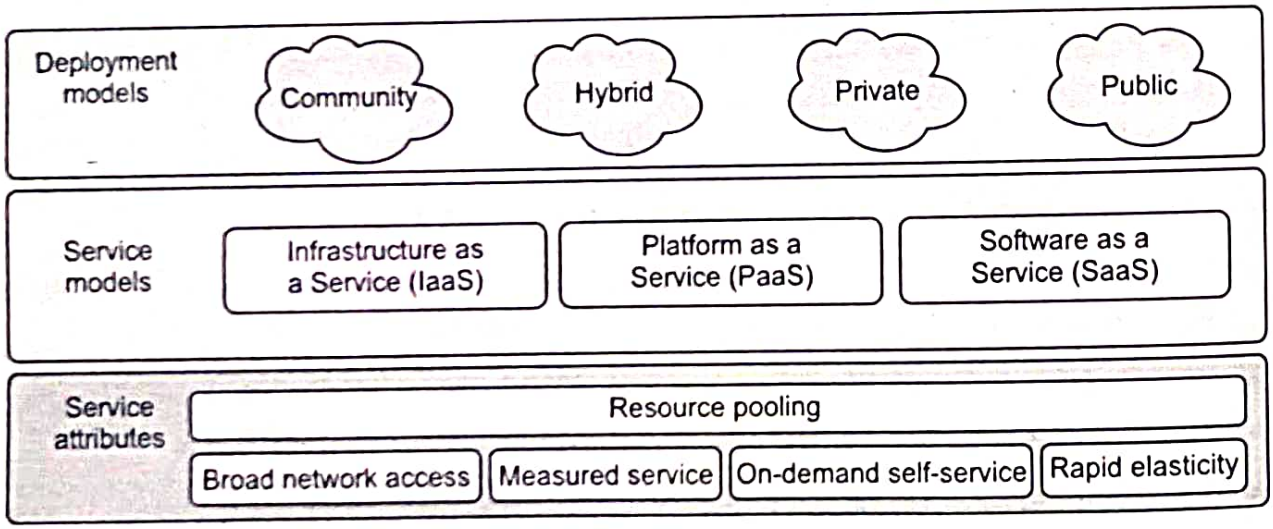


Fig. 1.1.2

- **On-demand self-service :** A client can provision computer resources without the need for interaction with cloud service provider personnel.
- **Broad network access :** Access to resources in the cloud is available over the network using standard methods in a manner that provides platform - independent

access to clients of all types. This includes a mixture of heterogeneous operating systems, and thick and thin platforms such as laptops, mobile phones, and PDA.

- **Resource pooling :** A cloud service provider creates resources that are pooled together in a system that supports multi-tenant usage. Physical and virtual systems are dynamically allocated or reallocated as needed.
- **Rapid elasticity :** Resources can be rapidly and elastically provisioned
- **Measured service :** The use of cloud system resources is measured, audited, and reported to the customer based on a metered system.

1.1.2 Cloud Components

- Cloud computing solutions are made up of several elements. Fig. 1.1.3 shows cloud components.

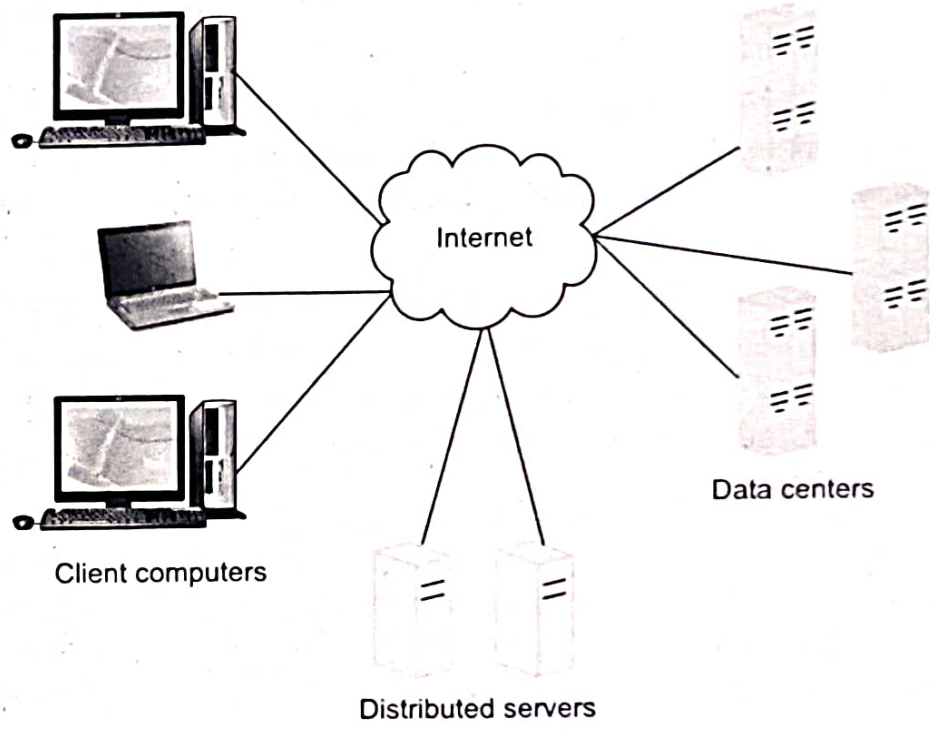


Fig. 1.1.3 Cloud components

1. **Clients :** Mobile, terminals or regular computers.
2. **Benefits :** Lower hardware costs, lower IT costs, security, data security, less power consumption, ease of repair or replacement, less noise.
3. **Data centers :** Collection of servers where the application to subscribe is housed. It could be a large room in the basement of your building or a room full of servers on the other side of the world

4. **Virtualizing servers :** Software can be installed allowing multiple instances of virtual servers to be used and a dozen virtual servers can run on one physical server.
5. **Distributed servers :** Servers don't all have to be housed in the same location. It can be in geographically disparate locations. If something were to happen at one site, causing a failure, the service would still be accessed through another site. If the cloud needs more hardware, they can add them at another site.

1.1.3 Characteristics of Cloud Computing

1. **On-demand self-service :** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed without requiring human interaction with each service's provider.
2. **Ubiquitous network access :** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms.
3. **Location-independent resource pooling :** The provider's computing resources are pooled to serve all consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.
4. **Rapid elasticity :** Capabilities can be rapidly and elastically provisioned to quickly scale up, and rapidly released to quickly scale down.
5. **Pay per use :** Capabilities are charged using a metered, fee-for-service, or advertising-based billing model to promote optimization of resource use.

1.1.4 Role and Boundaries in Cloud Computing

- Organizations and humans can assume different types of predefined roles depending on how they relate to and/or interact with a cloud and its hosted IT resources. The cloud computing defines these roles and identifies their main interactions.

1. Cloud provider :

- A person, organization or entity responsible for making a service available to interested parties. When assuming the role of cloud provider, an organization is responsible for making cloud services available to cloud consumers, as per agreed upon Service Level Agreement (SLA) guarantees. Cloud providers have their own IT resources.
- Fig. 1.1.4 shows cloud provider.

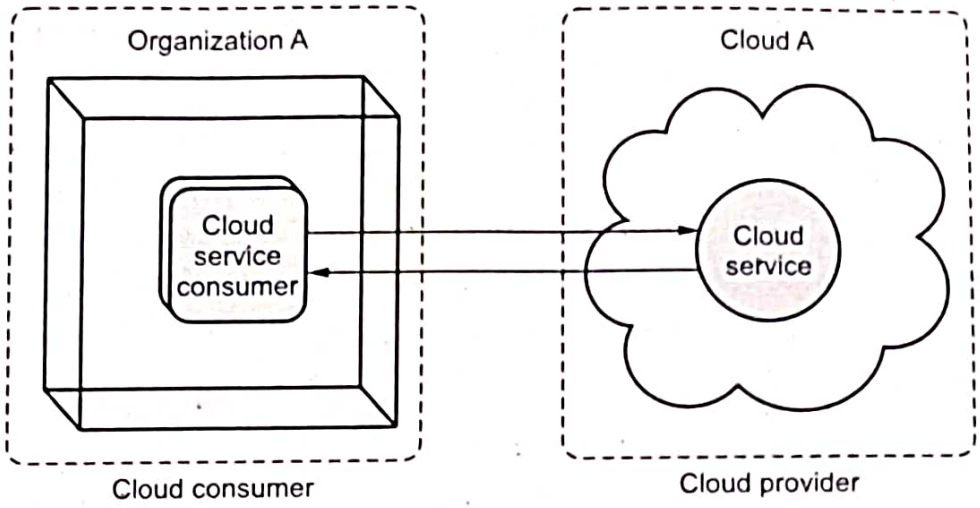


Fig. 1.1.4 Cloud service and cloud service consumer

- A cloud provider would have a significant number of roles responsible for the management of its cloud resources including those responsible for selling, onboarding, configuring and supporting cloud services for its consumers.

2. Cloud consumer :

- A person or organization that maintains a business relationship with and uses service from, cloud providers. The cloud consumer uses a cloud service consumer to access a cloud service.
- Anyone who purchases a cloud service is a consumer and within the consumer there could be an array of roles responsible for configuring and managing the resources from the cloud provider depending on the services obtained.

3. Cloud service owner :

- The person or organization that legally owns a cloud service is called a cloud service owner. The cloud service owner can be the cloud consumer or the cloud provider that owns the cloud within which the cloud service resides.
- Fig. 1.1.5 shows cloud service owner.

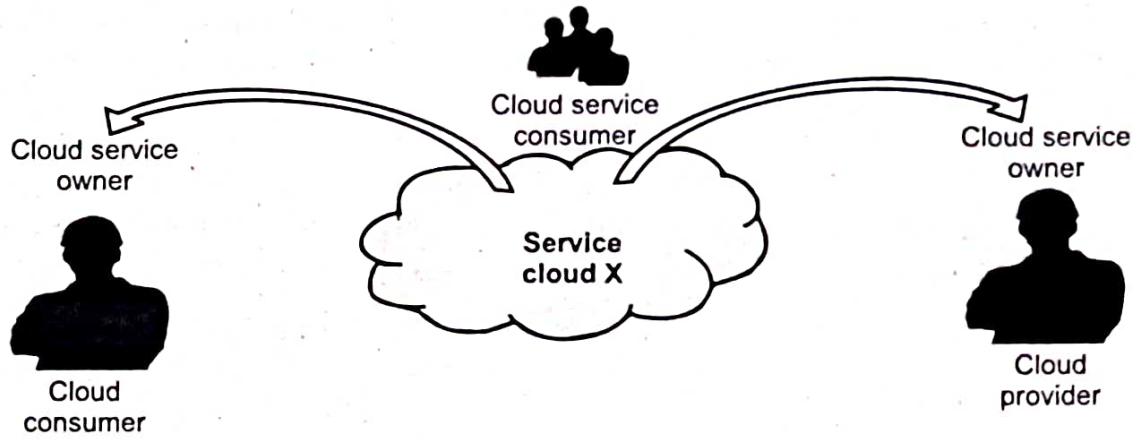


Fig. 1.1.5 Cloud service owner

- The reason a cloud service owner is not called a cloud resource owner is because the cloud service owner role only applies to cloud services.

4. Resource administrator :

- Cloud resource administrator is the person or organization responsible for administering a cloud-based IT resource. The cloud consumer or cloud provider or even third-party organization could be a cloud resource administrator.
- For example, a cloud service owner can contract a cloud resource administrator to administer a cloud service.

5. Cloud auditor :

- Cloud auditor is a party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation. Generally, cloud auditors are categorized based on intent.
- For the most part, their focus is on risk and compliance, especially around information security. Other auditors can provide advisory services especially to consumers looking to cut down their bills or raise the level of efficiency in the resources consumed.

6. Cloud broker :

- Cloud broker is any entity that manages the use, performance, and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers.
- Cloud brokers support consumers to get value for money by playing the advisory role especially for consumers who have a hybrid mix of resources from multiple providers.

7. Cloud carrier :

- Cloud carrier is an intermediary that provides connectivity and transport of cloud services from cloud providers to cloud consumers.
- Most ISPs have taken the role of cloud carriers as they provide the requisite bandwidth needed to connect consumers with providers as well as capabilities that support the connectivity.

8. Trust boundary :

- Logical perimeter that typically spans beyond physical boundaries to represent the extent to which IT resources are trusted. Fig. 1.1.6 shows trust boundary.
- When analysing cloud environments, the trust boundary is most frequently associated with the trust issued by the organization acting as the cloud consumer.

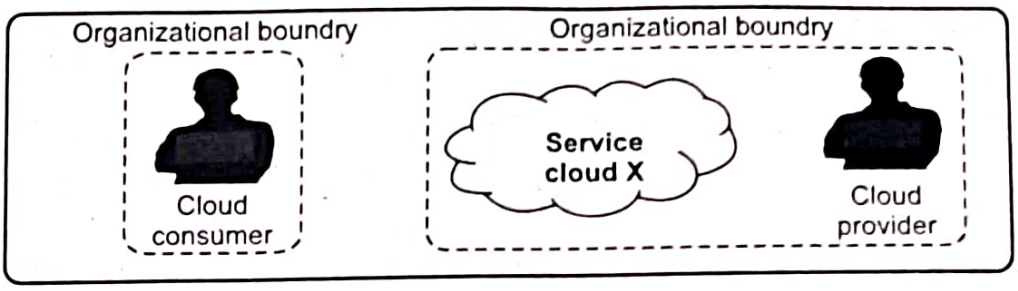


Fig. 1.1.6 Trust boundary

1.1.5 Cloud Applications

1. Through cloud cost flexibility, online marketplace gains access to more powerful analytics online. Cloud takes away the need to fund the building of hardware, installing software or paying dedicated software license fees.
2. Greater business scalability enables online video retailer to meet spikes in demand : Cloud enables businesses not just IT operations to add or provision computing resources just at the time they're needed.
3. Greater market adaptability provides online entertainment platform the ability to reach any type of customer device. A third of the executives we surveyed believe cloud can help them adapt to diverse user groups with a diverse assortment of devices.
4. Masked complexity enables access to services, no matter how intricate the technology they're built on.
5. With context-driven variability, "intelligent assistants" are possible. "Because of its expanded computing power and capacity, cloud can store information about user preferences, which can enable product or service customization," the report states.
6. Ecosystem connectivity enables information exchange across business partners.

1.1.6 Pros and Cons of Cloud Computing

Pros of cloud computing :

1. **Lower computer costs :** Since applications run in the cloud, not on the desktop PC, your desktop PC does not need the processing power or hard disk space demanded by traditional desktop software.
2. **Improved performance :** Computers in a cloud computing system boot and run faster because they have fewer programs and processes loaded into memory.
3. **Reduced software costs :** Instead of purchasing expensive software applications, you can get most of what you need for free.

4. **Instant software updates :** When you access a web-based application, you get the latest version - without needing to pay for or download an upgrade.
5. **Improved document format compatibility :** You do not have to worry about the documents you create on your machine being compatible with other user's applications or operating systems.
6. **Unlimited storage capacity :** Cloud computing offers virtually limitless storage.
7. **Increased data reliability :** Unlike desktop computing, in which if a hard disk crashes and destroy all your valuable data, a computer crashing in the cloud should not affect the storage of your data.
8. **Universal document access :** All your documents are instantly available from wherever you are.
9. **Latest version availability :** The cloud always hosts the latest version of your documents; as long as you are connected, you are not in danger of having an outdated version.
10. **Easier group collaboration :** Sharing documents leads directly to better collaboration.
11. **Device independence :** Move to a portable device and your applications and documents are still available.

Cons of cloud computing :

1. It requires a constant Internet connection : Cloud computing is impossible if you cannot connect to the Internet.
2. Features might be limited.
3. Stored data might not be secure : With cloud computing, all your data is stored on the cloud.
4. Does not work well with low-speed connections.

1.1.7 Difference between Cloud and Traditional Data Centers

Cloud	Traditional data centers
Cloud is a virtual resource that helps businesses to store, organize and operate data efficiently.	Data center is a physical resource that helps businesses to store, organize and operate data efficiently.
Infrastructure (hardware) cost is less.	Infrastructure (hardware) cost is more.

the maintenance cost of the cloud is less because service providers maintain it	The maintenance cost of the data center is quite high because developers are required for the maintenance.
Scaling of cloud requires less amount of investment	Scaling a data center requires a huge amount of investment
Cloud is more affordable but offers limited powers in comparison to data centers.	Data centers are less affordable but offers more power in comparison to cloud.

1.1.8 Multitenant Technology

- A multi - tenant cloud is a cloud computing architecture that allows customers to share computing resources in a public or private cloud. Each tenant's data is isolated and remains invisible to other tenants.
- It allows multiple users to work in a software environment at the same time, each with their own separate user interface, resources and services. The multitenant application design was created to enable multiple users (tenants) to access the same application logic simultaneously.
- Multitenancy can describe hardware or software architectures in which multiple systems, applications, or data from different enterprises are hosted on the same physical hardware.
- Multitenant applications typically include a level of customization for tenants, such as customizing the look and feel of the application or allowing the tenant to decide on specific access control permissions and restrictions for users.
- "Tenants" is a term for a group of users or software applications that all share access to the hardware through the underlying software. Multiple tenants on a server all share the memory, which is dynamically allocated and cleaned up as needed. They also share access to system resources, such as the network controller.
- Fig. 1.1.7 shows multi-tenant technology.

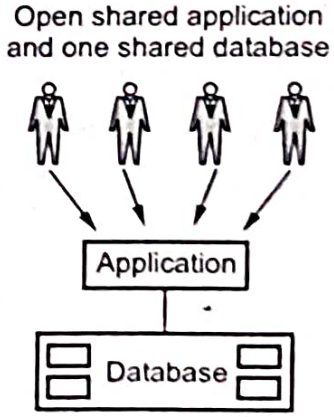


Fig. 1.1.7 Multi-tenant technology

- Multi-tenant architecture is to offer shared tenancy on public cloud providers like Amazon Web Services, Microsoft Azure and Google Cloud.
- Tenants can individually customize features of the application, such as :
  1. **User interface** : Tenants can define a specialized look for their application interface.
  2. **Business process** : Tenants can customize the rules, logic, and workflows of the business processes that are implemented in the application.
  3. **Data model** : Tenants can extend the data schema of the application to include, exclude, or rename fields in the application data structures.
  4. **Access control** : Tenants can independently control the access rights for users and groups.
- Common characteristics of multitenant applications are as follows :
  1. **Usage isolation** - The usage behaviour of one tenant does not affect the application availability and performance of other tenants.
  2. **Data security** - Tenants cannot access data that belongs to other tenants.
  3. **Recovery** - Backup and restore procedures are separately executed for the data of each tenant.
  4. **Application upgrade** - Tenants are not negatively affected by the synchronous upgrading of shared software artifacts.
  5. **Scalability** - The application can scale to accommodate increases in usage by existing tenants and/or increases in the number of tenants.
  6. **Metered usage** - Tenants are charged only for the application processing and features that are actually consumed.
  7. **Data tier isolation** - Tenants can have individual databases, tables and schemas isolated from other tenants.

Benefits of a multitenancy technology :

1. **Costs savings** : It yields tremendous economy of scale for the provider so he can offer the service at a lower cost to customers.
2. **Improved quality, user satisfaction, and customer retention** : A multitenant application is one large community hosted by the provider which can gather operational information from the collective user population and make frequent, incremental improvements to the service that benefit the entire user community at once.
3. **Improved security** : Most current enterprise security models are perimeter-based, making them vulnerable to inside attacks.

University Questions

1. *Explain the following challenges in cloud.*
   i) Security **Marks 5**
   ii) Data lock-in and standardization. **Marks 5**
   iii) Fault tolerance and disaster recovery. **AU : Dec.-20, Marks 3**
2. *Formulate stage-by-stage evolution of cloud with neat sketch and formulate any three benefits, drawbacks achieved by it in the banking and insurance sectors.* **AU : Dec.-21, Marks 13**
3. *Explain about evolution of cloud computing in detail.* **AU : Dec.-22, Marks 13**

1.2 System Models for Distributed and Cloud Computing

- Large number of autonomous computer nodes are used for building distributed system and cloud computing. These nodes are interconnected by LANs, WANs or SANs in a hierarchical manner. Now a days, new technology is applied for networking. Few LAN switches can easily connect hundreds of machines as a working cluster. A WAN can connect many local clusters to form a very large cluster of clusters. tech
- Massive system are formed using LAN and WAN system with cluster. Massive systems are considered highly scalable and can reach web scale connectivity, either physically or logically.
- Massive systems are classified into four groups : clusters, P2P networks, computing grids and Internet clouds over huge data centers.

Parameters	Cloud	Computer cluster	Peer to Peer Network	Grid Computing
Architecture, network size and connectivity	Virtualized cluster of servers over data centers via SLA	N/W of compute nodes interconnected by SAN, LAN, or WAN hierarchically	Flexible network of client machines logically connected by an overlay network	Heterogeneous Clusters interconnected by high - speed network links over selected resource sites
Control and resources management	Dynamic resource provisioning of servers, storage, and networks	Homogeneous nodes with distributed control, running UNIX or Linux	Autonomous client nodes, free in and out, with self - organization	Centralized control, server oriented with authenticated security

Applications and network-centric services	Upgraded web search, utility computing and outsourced computing services	High - performance computing search engines and web services	Most appealing to business file sharing, content delivery and social networking	Distributed supercomputing, global problem solving and data center services
Example	Google App Engine, Bluecloud, AWS, Microsoft Azure	Google search engine, SunBlade, IBM Road Runner, Cray XT4.	Gnutella, eMule, BitTorrent, Napster, KaZaA, Skype, JXTA	TeraGrid, UK EGEE, D-Grid, ChinaGrid

- A **distributed system** is a collection of independent computers that appears to its users a single coherent system. A distributed system can consist of any number of possible configurations, such as mainframes, personal computers, workstations, minicomputers and so on.
- The major milestones have led to cloud computing are mainframes computing, cluster computing and grid computing.

1.2.1 Clusters of Cooperative Computers

- Computing cluster consists of interconnected stand - alone computers which work cooperatively as a single integrated computing resource.

1.2.1.1 Cluster Architecture

- Fig. 1.2.1 shows architecture of a typical server cluster. To build a larger cluster with more nodes, the interconnection network can be built with multiple levels of Gigabit Ethernet. Using hierarchical construction with SAN, LAN, or WAN, we can build scalable clusters with an increasing number of nodes.

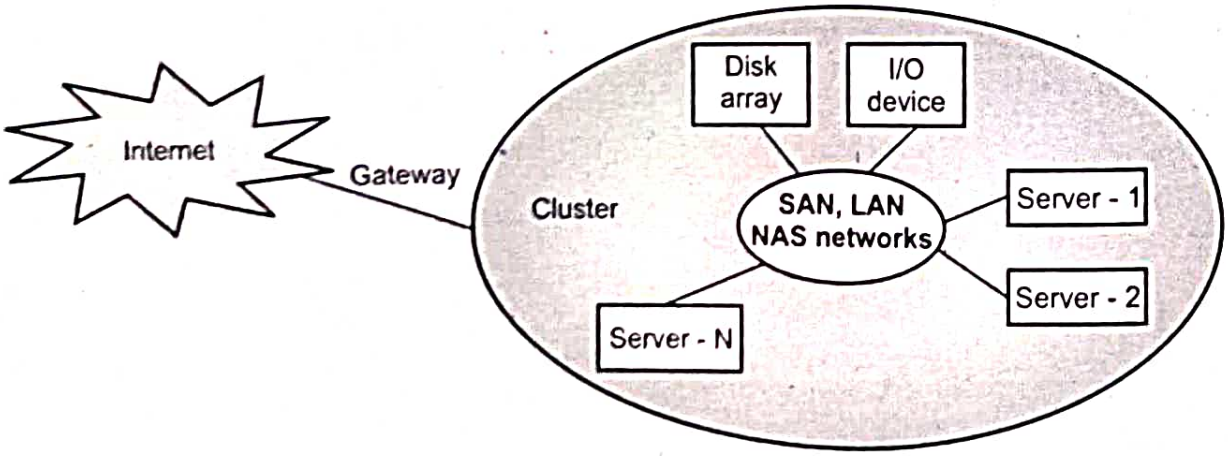


Fig. 1.2.1 Architecture of a typical server cluster

- Cluster is connected to the Internet via a Virtual Private Network (VPN) gateway. The gateway IP address locates the cluster. Most clusters have loosely coupled node computers. All resources of a server node are managed by their own OS. Thus, most clusters have multiple system images as a result of having many autonomous nodes under different OS control.

1.2.1.2 Single System Image

- Single System Image (SSI) is an abstraction that provides the illusion that a multicomputer or cluster is a single machine. There are individual instances of the Operating Systems (OSs) running on each node of a multicomputer, processes working together are spread across multiple nodes and files may reside on multiple disks.
- An SSI provides a unified view of this collection to users, programmers and system administrators. This unification makes a system easier to use and more efficient to manage.
- Multicomputers consist of nodes, each with its own memory, CPUs and a network interface. In the case of clusters, each node is a stand - alone computer made of commodity, off-the-shelf parts. Instead of viewing this collection of computers as individual systems, it is easier and more economical if users, programmers and system administrators can treat the collection as a single machine.
- Single System Image (SSI) consisting of single entry point, single file hierarchy, single I/O space, single networking scheme, single control point, single job management system, single memory space and single process space. The ultimate goal of SSI is for a cluster to be as easy to use as a desktop computer.
- Single job management system: All cluster jobs can be submitted from any node to a single job management system.
- Single user interface : The users use the cluster through a single graphical interface. Such an interface is available for workstations and PCs.

1.2.2 Grid Computing Infrastructures

- Grid computing is a distributed computing system where a group of computers are connected to create and work as one large virtual computing power, storage, database, application and service.
- Computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high - end computational capabilities. A computational grid is a loose network of computers linked to perform grid computing.

- Fig. 1.2.2 shows grid computing.

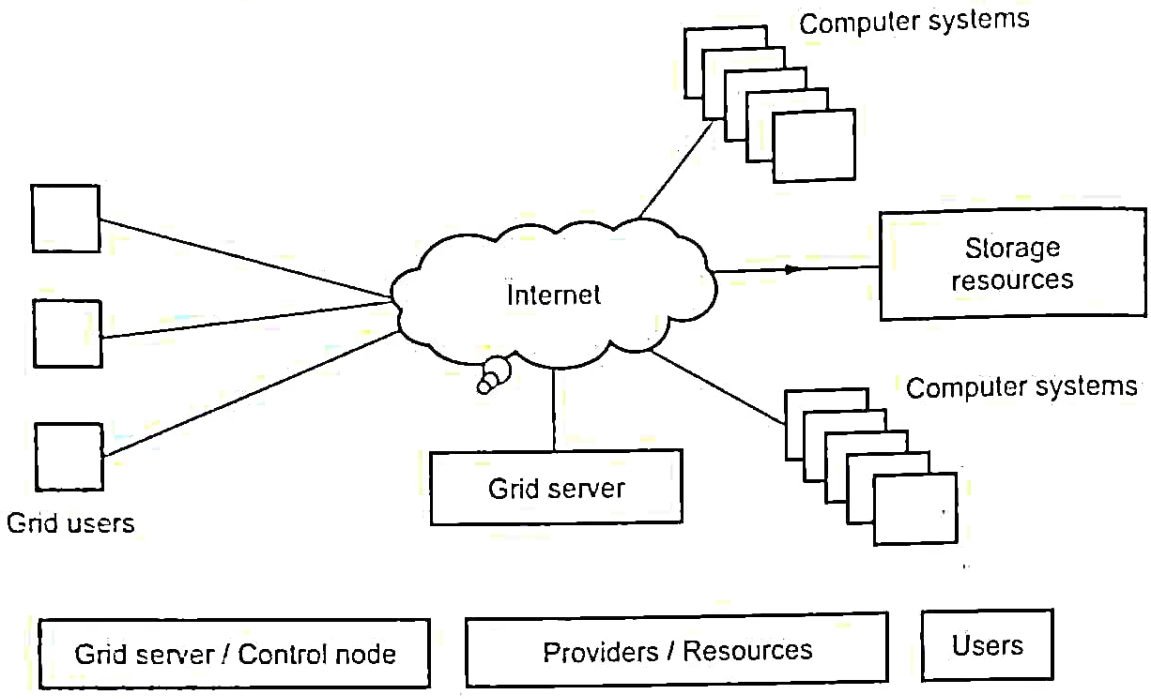


Fig. 1.2.2 Grid computing

- In a computational grid, a large computational task is divided up among individual machines, which run calculations in parallel and then return results to the original computer.
- These individual machines are nodes in a network, which may span multiple administrative domains and may be geographically distant.
- Grid systems are classified into two categories : Computational or data grids and P2P grids.

1.2.3 Peer-to-Peer Network Families

- The P2P architecture offers a distributed model of networked systems. First, a P2P network is client - oriented instead of server - oriented.
- In a P2P system, every node acts as both a client and a server, providing part of the system resources. Peer machines are simply client computers connected to the Internet. All client machines act autonomously to join or leave the system freely. No central coordination or central database is needed.
- Initially, the peers are totally unrelated. Each peer machine joins or leaves the P2P network voluntarily. Only the participating peers form the physical network at any time.
- P2P network does not use a dedicated interconnection network. The physical network is simply an Ad hoc network formed at various Internet domains

randomly using the TCP/IP protocols. Thus, the physical network varies in size and topology dynamically due to the free membership in the P2P network.

- P2P overlay networks are built at application layer which is on the top of the network topology. These overlays are used for indexing and peer discovery which makes the P2P system independent from the physical network topology. Contents are directly exchanged between the underlying Internet Protocol (IP) networks.

1.2.3.1 Overlay Networks

- Files are distributed in the participating peers. Based on communication or file - sharing needs, the peer IDs form an overlay network at the logical level. This overlay is a virtual network formed by mapping each physical machine with its ID, logically, through a virtual mapping.
- There are two types of overlay networks : Unstructured and structured.
- **Structured P2P overlay network** topology is a tightly controlled network. The contents are placed only at specified locations but not at random peers. Structured peer-to-peer overlay networks are sometimes referred as Distributed Hash Table (DHT), are scalable networks which supports Internet - scale applications.
- The applications of structured P2P overlays are construction of large - scale networks, decentralized applications, distributed storage, group communication and content distribution. The advantage of this overlay is messages correctly reach the destination even if large number of nodes crashes.
- An **Unstructured P2P overlay network** has no prior knowledge about the topology of the network. Here the peers join the network without any specific rules. A resource may take a long time for the search operation because most of the time there is no relation between the name of resources and their locations.
- The advantages of this overlay are : Easy implementation, simplicity, keyword search and dynamic environments. The major drawback of this overlay is the scalability problem.

1.2.3.2 P2P Application Families

- P2P networks are classified into four groups : Distributed File Sharing, Collaborative Platform, Distributed P2P and computing P2P Platform.

System parameters	Distributed file sharing	Collaborative platform	Distributed P2P	Computing P2P platform
Application	Content distribution of MP3 music, video, open software	Instant messaging, collaborative design and gaming	Scientific exploration and social networking	Open networks for public resources
Problems	Loose security and serious online copyright violations	Lack of trust, disturbed by spam, privacy and peer collusion	Security holes, selfish partners and peer collusion	Lack of standards or protection protocols
Example	Gnutella, Napster, eMule, BitTorrent, Aimster	ICQ, AIM, Groove, Magi, Multiplayer Games, Skype	SETI@home, Geonome@home	JXTA, .NET, FightingAid@ home

1.2.3.3 P2P Computing Challenges

- Hardware, software and network requirements are three problems face by P2P computing. There are too many hardware models and architectures to select from; incompatibility exists between software and the OS; and different network connections and protocols make it too complex to apply in real applications.
- Data locality, network proximity and interoperability are three design objectives in distributed P2P applications.
- P2P performance is affected by routing efficiency and self - organization by participating peers. Fault tolerance, failure management and load balancing are other important issues in using overlay networks.
- Security, privacy and copyright violations are major worries by those in the industry in terms of applying P2P technology in business applications. In a P2P network, all clients provide resources including computing power, storage space and I/O bandwidth.
- The distributed nature of P2P networks also increases robustness, because limited peer failures do not form a single point of failure.
- P2P networks are reliable for a small number of peer nodes. They are only useful for applications that require a low level of security and have no concern for data sensitivity.

1.2.4 Cloud Computing over the Internet

- Definition of cloud computing by IBM : "A cloud is a pool of virtualized computer resources. A cloud can host a variety of different workloads, including batch - style backend jobs and interactive and user - facing applications."
- Cloud allows workloads to be deployed and scaled out quickly through rapid provisioning of virtual or physical machines. The cloud supports redundant, self - recovering, highly scalable programming models that allow workloads to recover from many unavoidable hardware/software failures.

1.2.4.1 Internet Clouds

- Cloud computing applies a virtualized platform with elastic resources on demand by provisioning hardware, software and data sets dynamically. Fig. 1.2.3 shows virtualized resources from data centers to form an Internet cloud.

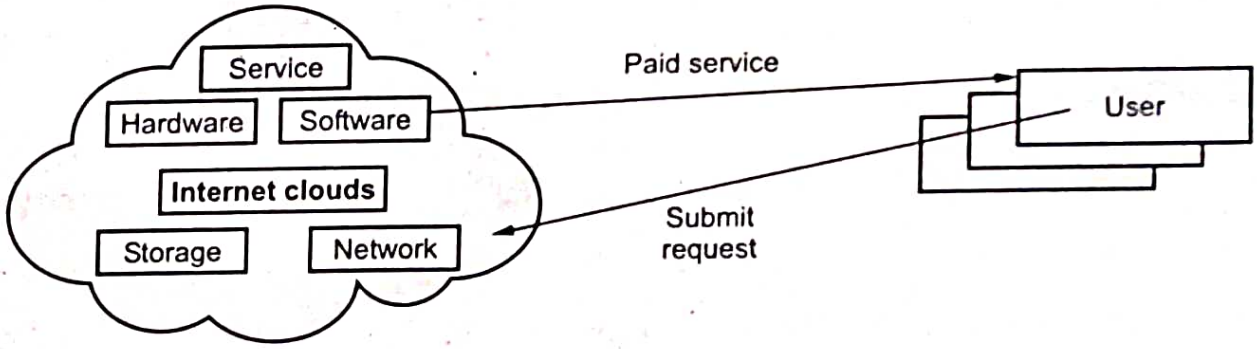


Fig. 1.2.3 Virtualized resources from data centers to form an Internet cloud

- The idea is to move desktop computing to a service - oriented platform using server clusters and huge databases at data centers. Cloud computing leverages its low cost and simplicity to benefit both users and providers.
- Machine virtualization has enabled such cost - effectiveness. The cloud ecosystem must be designed to be secure, trustworthy and dependable.

1.2.4.2 Cloud Landscape

- Cloud service models are as follows :
  1. Software as a Service : SaaS provider dispose the applied software unified on their server, the user can subscribe applied software service from the manufacturer through Internet.
  2. Platform as a Service (PaaS) : PaaS takes develop environment as a service to supply. This layer provides a platform for creating applications.
  3. Infrastructure as a Service (Iaas) : In this layer, servers, network devices and storage disks are made available to organizations as services on a need-to basis.

- Reasons to adapt the cloud for upgraded Internet applications and web services are as follows :
  1. Desired location in areas with protected space and higher energy efficiency
  2. Sharing of peak-load capacity among a large pool of users, improving overall utilization.
  3. Separation of infrastructure maintenance duties from domain - specific application development
  4. Significant reduction in cloud computing cost, compared with traditional computing paradigms
  5. Cloud computing programming and application development
  6. Service and data discovery and content/service distribution
  7. Privacy, security, copyright and reliability issues
  8. Service agreements, business models and pricing policies

1.2.5 Difference between Distributed, Grid and Cloud Computing

Distributed computing	Grid computing	Cloud computing
Small to medium size	Large size	Small to large size
Low security requirement	High security requirement	Low security requirement
It is homogeneous	It is heterogeneous	It is heterogeneous
Network type is private	Network type is private	Network type is public
It is based on Ethernet	It is based on Ethernet	It is based on Ethernet
SLA requirement is strict	SLA requirement is high	SLA requirement is low

1.3 NIST Cloud Computing Reference Architecture

- Fig. 1.3.1 shows NIST cloud computing reference architecture. It defines five major actors : *cloud consumer, cloud provider, cloud carrier, cloud auditor and cloud broker.*
- Each actor is an entity (organization) that participates in a transaction or process and/or performs tasks in cloud computing.
- **Cloud consumer :** A person or organization that maintains a business relationship with and uses service from, cloud providers.
- **Cloud provider :** A person, organization, or entity responsible for making a service available to interested parties.

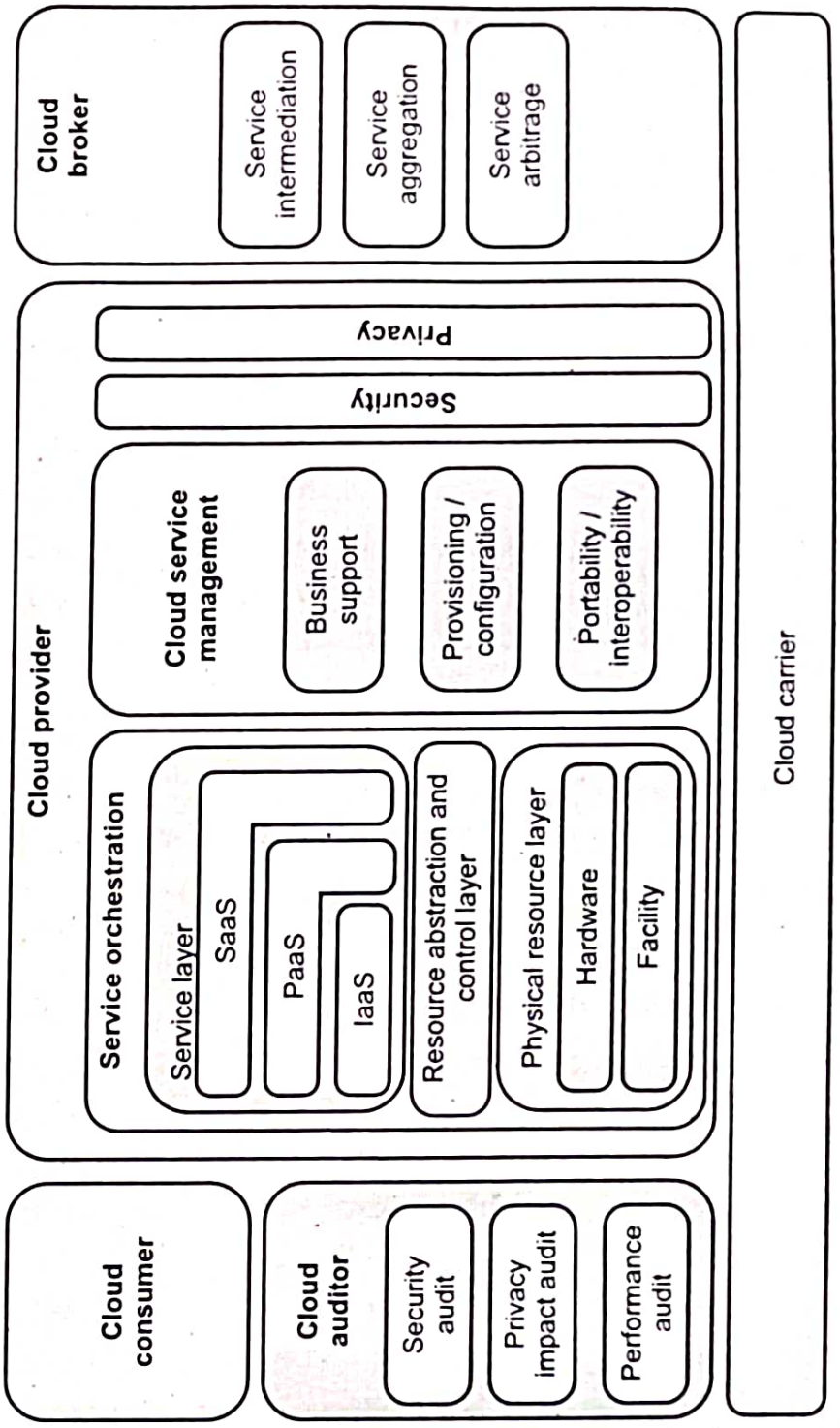


Fig. 1.3.1 NIST cloud computing reference architecture

- **Cloud auditor :** A party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
- **Cloud broker :** An entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers.

- **Cloud carrier :** An intermediary that provides connectivity and transport of cloud services from cloud providers to cloud consumers.
- **Cloud Services Broker (CSB) :** The CSB is typically a third - party entity or company that looks to extend value to multiple customers of cloud - based services through relationships with multiple cloud service providers. It acts as a liaison between cloud services customers and cloud service providers, selecting the best provider for each customer and monitoring the services. A CSB provides :
  1. **Service intermediation :** A CSB enhances a given service by improving some specific capability and providing value - added services to cloud consumers. The improvement can be managing access to cloud services, identity management, performance reporting, enhanced security, etc.
  2. **Service aggregation :** A CSB combines and integrates multiple services into one or more new services. The broker provides data integration and ensures the secure data movement between the cloud consumer and multiple cloud providers.
  3. **Service arbitrage :** Service arbitrage is similar to service aggregation except that the services being aggregated are not fixed. Service arbitrage means a broker has the flexibility to choose services from multiple agencies. The cloud broker, for example, can use a credit - scoring service to measure and select an agency with the best score.

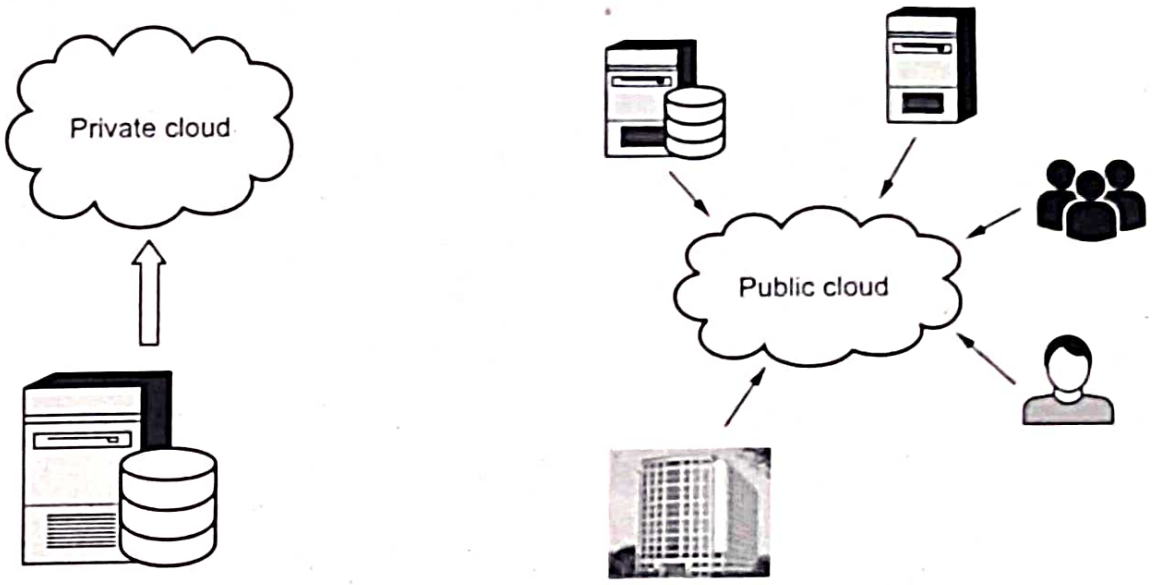
1.4 Cloud Deployment Models

AU : Dec.-21,22

- Cloud deployment models are refers to the location and management of the cloud's infrastructure.
- Deployment models are defined by the ownership and control of architectural design and the degree of available customization. Cloud deployment models are private public and community clouds.
- Fig. 1.4.1 shows cloud deployment model. (See Fig. 1.4.1 on next page)

1. Public cloud :

- The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.
- Public cloud is a huge data centre that offers the same services to all its users. The services are accessible for everyone and much used for the consumer segment.
- Examples of public services are Facebook, Google and LinkedIn.
- **Public cloud benefits :**
  a) Low investment hurdle : Pay for what user use.
  b) Good test/development environment for applications that scale to many servers.



(a) Private cloud (b) Public cloud

Fig. 1.4.1 Cloud deployment model

- **Public cloud risks :**
  a) Security concerns : Multi-tenancy and transfers over the Internet.
  b) IT organization may react negatively to loss of control over data center function.

2. Private cloud :

- The cloud infrastructure is operated solely for a single organization. It may be managed by the organization or a third party and may exist on-premises or off-premises.
- **Private cloud benefits :**
  a) Fewer security concerns as existing data center security stays in place.
  b) IT organization retains control over data center.
- **Private cloud risks :**
  a) High investment hurdle in private cloud implementation, along with purchases of new hardware and software.
  b) New operational processes are required; old processes not all suitable for private cloud.

3. Community cloud :

- The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements, policy or compliance considerations). It may be managed by the organizations or a third party and may exist on-premises or off-premises.

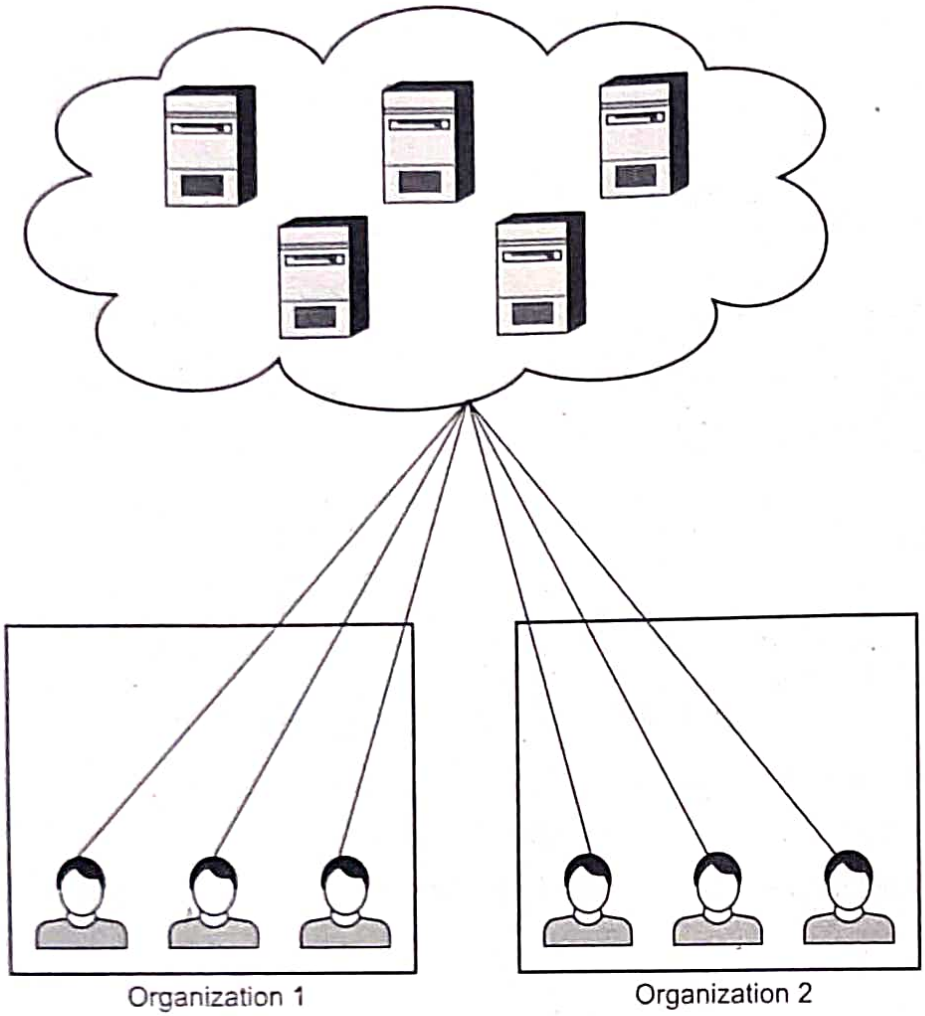


Fig. 1.4.2 Community cloud

4. Hybrid cloud :

- The cloud infrastructure is a composition of two or more clouds (private, community or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).
- **Hybrid cloud benefits :**
  a) Operational flexibility : Run mission critical on private cloud, dev/test on public cloud.
  b) Scalability : Run peak and bursty workloads on the public cloud.
- **Hybrid cloud risks :**
  a) Hybrid clouds are still being developed; not many in real use.
  b) Control of security between private and public clouds, some of same concerns as in public cloud.

1.4.1 Difference between Public and Private Cloud

Public cloud	Private cloud
Public cloud infrastructure is offered via web applications and also as web services over Internet to the public.	Private cloud infrastructure is dedicated to a single organization.
Support multiple customer.	Support dedicated customer.
Full utilized of infrastructure.	Does not utilize shared infrastructure.
Security is low as compared to private cloud.	High level of security.
Low cost	High cost
Azure, Amazon Web Services, Google App Engine and Force.com are a few examples of public clouds.	An example of the Private Cloud is NIRIX's one Server with dedicated servers.

University Questions

1. *Outline the various deployment models of cloud with neat sketch and identify which among them could be applied to formulate cloud structure for a small firm.* **AU : Dec.-21, Marks 13**
2. *Compare and contrast : Public, Private and Hybrid clouds.* **AU : Dec.-22, Marks 8**

1.5 Cloud Service Models

- Service models describe the type of service that the service provider is offering. The best-known service models are software as a service, platform as a service, and Infrastructure as a service.
- The service models build on one another and define what a vendor must manage and what the client's responsibility is.
- Service models : This consists of the particular types of services that you can access on a cloud computing platform.
- Cloud service is any service made available to users on demand via the Internet from a cloud computing provider's servers as opposed to being provided from a company's own on-premises servers.
- Cloud services are designed to provide easy, scalable access to applications, resources and services and are fully managed by a cloud services provider.

- A cloud service can exist as a simple web-based software program with a technical interface invoked via the use of a messaging protocol or as a remote access point for administrative tools or larger environments and other IT resources.
- The organization that provides cloud-based IT resources is the cloud provider. Cloud providers normally own the IT resources for lease by cloud consumers and could also resell IT resources leased from other providers.

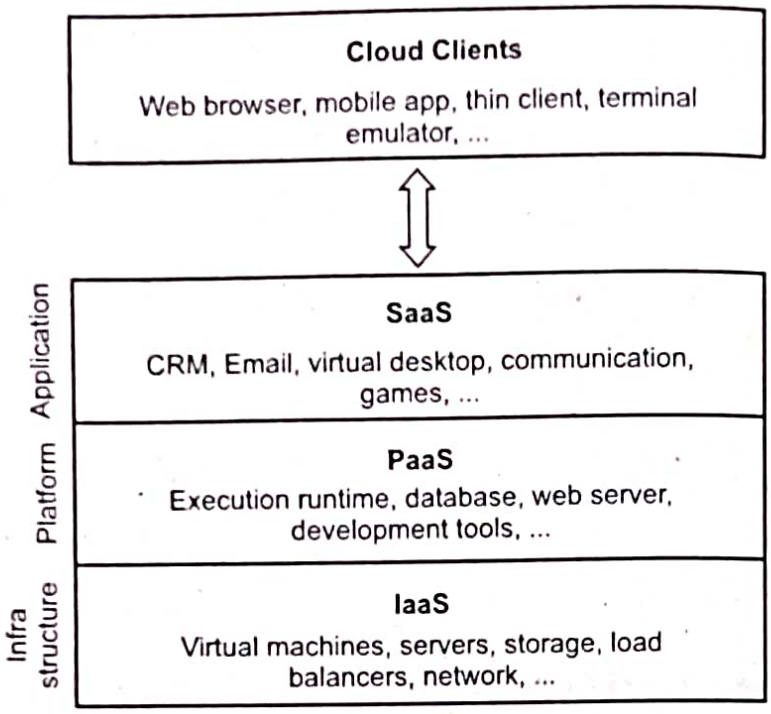


Fig. 1.5.1 Cloud computing stack

- Cloud computing, often described as a stack, has a broad range of services built on top of one another under the name cloud.
- Fig. 1.5.1 shows cloud computing stack.
- Flavors of cloud computing is as follows;
  1. SaaS applications are designed for end-users, delivered over the web.
  2. PaaS is the set of tools and services designed to make coding and deploying those applications quick and efficient.
  3. IaaS is the hardware and software that powers it all - servers, storage, networks, operating systems.

1.6 Software as a Service (SaaS)

AU : Dec.-22

- Software-as-a-Service (SaaS), is a cloud based software delivery model that allows end users to access software applications over the internet. With a SaaS model, the software is hosted on remote servers, maintained and updated by the service provider and made available to customers via web browsers, mobile apps and APIs.
- The public cloud provider manages all the hardware and traditional software, including middleware, application software and security. So SaaS customers can dramatically lower costs; deploy, scale and upgrade business solutions more

quickly than maintaining on - premises systems and software; and predict total cost of ownership with greater accuracy.

- Fig. 1.6.1 shows SaaS model.

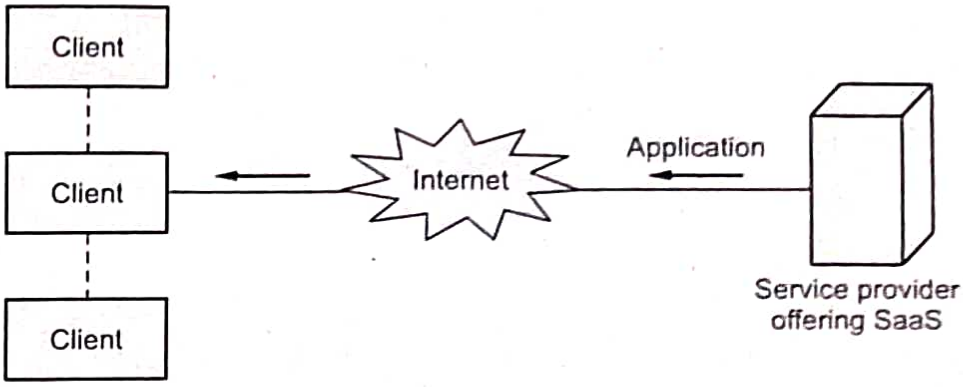


Fig 1.6.1 SaaS

- Model in which an application is hosted as a service to customers who access it via the Internet.
- The provider does all the patching and upgrades as well as keeping the infrastructure running.
- The traditional model of software distribution, in which software is purchased for and installed on personal computers, is referred to as product.
- In this model, the user, client or consumer runs an application from a cloud infrastructure. Through an interface such as a web browser, the client or user may access this application from a variety of devices.
- The complete application is offered as on demand service. This saves the client from having to invest in any software licenses or servers up front and can save the provider money since they are maintaining and providing only a single application.
- In this model, the client does not manage cloud infrastructure, networks or servers, storage or operating systems. Even, Microsoft, Google and Zoho offer SaaS.
- The SaaS concept can be defined as providing robust "web - based, on - demand software, storage and various applications" to organizations.
- The SaaS model has emerged as an alternative to traditional one - time licensing for providing and maintaining the software needed by knowledge workers within organizations.
- Popular software as a service examples include Office 365, Google G Suite (Apps), Dropbox, Salesforce, SAP Concur and Zoom.

1.6.1 Challenges of SaaS

- Lack of the following features prevents the massive adoption of clouds :
  1. Controllability
  2. Visibility and flexibility
  3. Security and Privacy
  4. High Performance and Availability
  5. Integration and Composition
  6. Standards
- Challenges :
  1. **Integration conundrum :** Organization without a method of synchronizing data between multiple lines of businesses are at a serious disadvantage in terms of maintaining accurate data, forecasting, and automating key business processes. Real-time data and functionality sharing is an essential ingredient for clouds.
  2. Application programming interfaces are not proper and insufficient.
  3. **Security for transmission of data :** Data integrity, confidentiality, quality and value have to be preserved as services and applications are interlinked and saddled to work together.

1.6.2 Characteristics of SaaS

1. Software applications or services are stored remotely.
2. A user can then access these services or software applications via the Internet.
3. In most cases, a user does not have to install anything onto their host machine, all they require is a web browser to access these services and in some cases, a browser may require additional plug-in/add-on for certain services.
4. Network - based management and access to commercially available software from central locations rather than at each customer's site, enabling customers to access applications remotely via the Internet.
5. Application delivery from a one-to-many model, as opposed to a traditional one-to-one model.

1.6.3 Merits and Demerits of SaaS

Merits :

- Accessibility : Ability to run via an internet browser 24/7 from any device
- Operational management : No installation, equipment updates or traditional licensing management

- Cost effective : No upfront hardware costs and flexible payment methods such as pay-as-you-go models
- Scalability : Easily scale a solution to accommodate changing needs
- Data storage : Data is routinely saved in the cloud
- Analytics : Access to data reporting and intelligence tools

Demerits :

- Stability : SaaS needs a strong internet connection for usage. Network problems and slow internet speed leads to delays and reduced productivity.
- Loss of control : In-house software applications give business owners a high degree of control.
- Limited customization : Most SaaS applications offer little in the way of customization from the vendor
- Slower speed : SaaS solutions can have more latency than client/server apps
- Security risks : While the SaaS provider secures the application itself, strict measures should be taken with sensitive data

University Question

1. *Discuss the features of software as a service and explain in detail about SaaS with example.*

AU : Dec.-22, Marks 13

1.7 Platform as a Service (PaaS)

AU : Dec.-20

- Platform as a service is another application delivery model and also known as cloud - ware. Supplies all the resources required to build applications and services completely from the Internet, without having to download or install software.
- PaaS offers a complete environment for developers to run their applications. The environment includes hardware, routers, operating system, runtime environment, middleware, database, web server and more. PaaS users, like developers, can deploy their applications on the PaaS provider's infrastructure and platform.
- Services include : Application design, development, testing, deployment and hosting, team collaboration, web service integration, database integration, security, scalability, storage, state management and versioning.
- PaaS is closely related to SaaS but delivers a platform from which to work rather that an application to work with.
- This model involves software encapsulated and offered as a service, from which higher levels of service may then be built. The user, customer or client in this

model is the one building applications which then run on the provider's infrastructure.

- This in turn provides customers and clients with the capability to deploy applications onto the cloud infrastructure using programming tools and languages, which the provider supports.
- The customer still does not manage the framework, network, servers or operating system, but has control over deployed applications and sometimes over the hosting environment itself.
- The PaaS provider hosts everything, servers, networks, storage, operating system software, databases, development tools at their data center. Typically customers can pay a fixed fee to provide a specified amount of resources for a specified number of users, or they can choose 'pay-as-you-go' pricing to pay only for the resources they use.
- A Platform-as-a-Service solution works by combining three **principle components :** cloud infrastructure, software and a Graphic User Interface (GUI).
- Most PaaS models include :
  a) The physical infrastructure required to support development
  b) Software solutions, like tools for building applications
  c) A graphic user interface, or GUI, that provides the avenues through which users can work
- Cloud infrastructure includes operating system software, virtual machines, firewalls, storage and networking. In all examples of PaaS, these serve as the technological foundation of system a safe, interconnected computing environment where work can be done.
- The software component is used for the development of applications, including building, deploying and managing them. In a PaaS setup, it is the software that enables the creation of products.
- The GUI forms the connection between the PaaS system and the people that use it. Therefore, the GUI has to link developers with the tools they need to design solutions.
- Examples of PaaS providers are SAP, Heroku, Microsoft Azure, VMWare, Google App Engine and Swisscom
- Organizations typically use PaaS for these scenarios :
  1. Development framework : PaaS provides a framework that developers can build upon to develop or customize cloud - based applications. PaaS developers create applications using built-in software components. Cloud features such as

scalability, high - availability and multi - tenant capability are included, reducing the amount of coding that developers must do.

2. Analytics or business intelligence : Tools provided as a service with PaaS allow organizations to analyze and mine their data, finding insights and patterns and predicting outcomes to improve forecasting, product design decisions, investment returns and other business decisions.
3. Additional services : PaaS providers may offer other services that enhance applications, such as workflow, directory, security and scheduling.

1.7.1 Characteristics of PaaS

1. It support multi - tenant architecture.
2. It support for development of group collaboration.
3. PaaS systems can be deployed as public cloud services or as private cloud services.
4. Provision of runtime environments. Typically each runtime environment supports either one or a small set of programming languages and frameworks.
5. Support for custom applications. Support for the development, deployment and operation of custom applications.
6. Preconfigured capabilities Many PaaS systems are characterized by capabilities that are preconfigured by the provider, with a minimum of configuration available to developers and customer operations staff.
7. Support for porting existing applications. While many PaaS systems are primarily designed to support "born on the cloud" applications.
8. Security is an important characteristic in PaaS. It needs to provide authentication and authorization to differentiate the access rights of different users.

1.7.2 Benefits and Disadvantages of PaaS

Benefits of PaaS

1. Scalability including rapid allocation and deallocation of resources with a pay-as-you-use model
2. Reduced capital expenditure
3. Reduced lead times with on-demand availability of resources
4. Self-services with reduced administration costs
5. Reduced skill requirements
6. Support of team collaboration
7. Ability to add new users quickly

Disadvantages of PaaS :

1. Security and compliance risks : PaaS software is generally offered in a public cloud where it is shared by multiple users, which creates a higher risk of security vulnerabilities.
2. Vendor lock-in : PaaS solutions for each business requirement might differ and the chosen vendor may not be able to provide convenient options for frameworks, customization, or migration.
3. Loss of operational control : Developers may have to trade off the abstraction for more granular control over application components.
4. Integration and migration : If all the components of a system are not built on cloud, integration and migration between both could pose challenges.

University Questions

1. *Explain the various layered cloud architectural development design for effective cloud computing environment.* **AU : Dec.-20, Marks 13**
2. *Demonstrate the architectural design of compute and storage clouds.* **AU : Dec.-20, Marks 13**

1.8 Infrastructure as a Service

- IaaS gives the storage room likeness to the in-house datacenter stood out from various organizations sorts.
- Center datacenter framework segments are capacity, servers (registering units), the system itself, and administration apparatuses for foundation upkeep and checking.
- Each of these parts has made a different market specialty. While some little organizations have practical experience in just a single of these IaaS cloud specialties, vast cloud suppliers like Amazon or Right Scale have offerings over all IaaS territories.
- Fig. 1.8.1 shows IaaS. (See Fig. 1.8.1 on next page)
- It offers the hardware so that your organization can put whatever they want onto it. Rather than purchase servers, software, racks, and having to pay for the datacenter space for them, the service provider rents those resources :
  1. Server space 2. Network equipment
  3. Memory 4. CPU cycles 5. Storage space
- Again, the customer is not managing cloud infrastructure, but in this case, the customer does control operating systems, deployed applications, storage, and sometimes-certain networking components

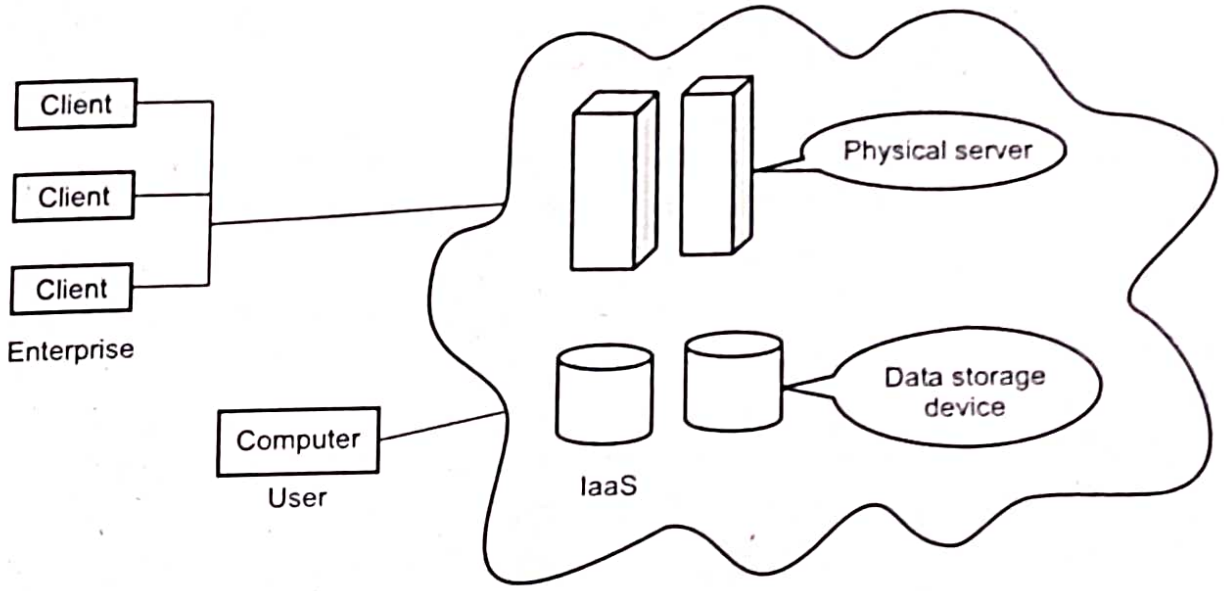


Fig. 1.8.1 IaaS

- Examples : Amazon EC2, Rackspace Mosso, GoGrid
- IaaS server types :
  1. **Physical server :** Actual hardware is allocated for the customer's dedicated use.
  2. **Dedicated virtual server :** The customer is allocated a virtual server, which runs on a physical server that may or may not have other virtual servers.
  3. **Shared virtual server :** The customer can access a virtual server on a device that may be shared with other customers.

How does IaaS work ?

- In an IaaS environment, the cloud provider acts as a host to the resources needed by the cloud consumers. Cloud consumers can access the resources virtually with an internet connection to run their applications and workloads.
- IaaS provider performs the following :
  1. Maintains network resources, compute resources, storage resources and data center infrastructure.
  2. Provides maintenance-free virtualized access to the hardware or infrastructure resources (mentioned above) on a pay-as-you-go basis.
  3. Creates a continuous virtual environment for cloud consumers.
  4. Provides easy access and control over individual IaaS components to the cloud consumers.
- Instead of using a physical data center, or hardware, IaaS provides these as a service on-demand. IaaS can be implemented on public, private and hybrid cloud models.

- In a public cloud setup, the customer's workloads run on data centers that are owned and maintained by the public cloud provider. The infrastructure is pooled across multiple organizations and institutions. The provider offers infrastructure over the internet as a service through dedicated connections and takes care of the virtualization software. The provider may also give access to physical resources (not virtualized) as per the organization's needs.
- In a private cloud setup, the infrastructure is available to only one organization, which is similar to having an on-premise data center, but managed by a cloud provider.
- A hybrid model offers a mix of virtual machines and container - based applications, deployed on public cloud or data centers.
- ***How IaaS relates to virtualization, automation and containerization ?***
- All the physical resources are virtualized using a hypervisor before they can be accessed by the cloud consumer (subscriber). Consumers can access the infrastructure from anywhere using an internet connection, or a Virtual Private Network (VPN) for additional security.
- The cloud provider has virtual machines that the subscriber (organization) can use to install their choice of operating system, software, database and other components. Fig. 1.8.2 shows IaaS architecture with virtualization.

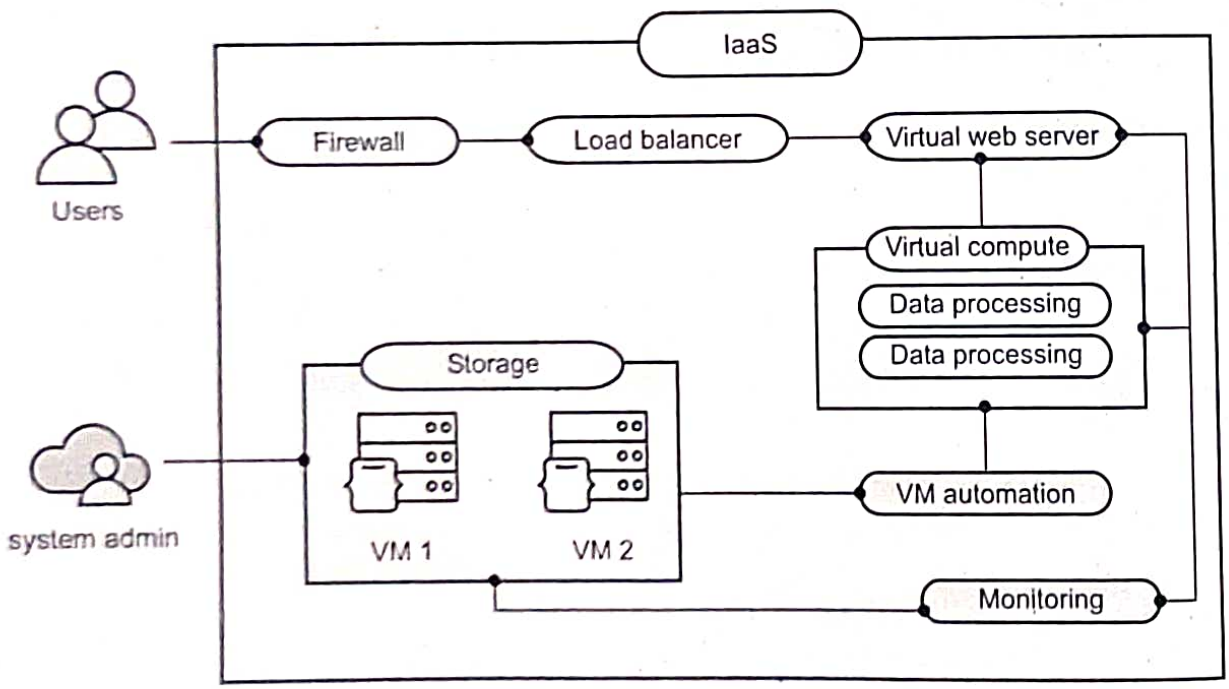


Fig. 1.8.2 IaaS architecture with virtualization

- Through virtualization, virtual machines provide complete environments that function as a virtual computer system with their own CPUs, memory, network interface and storage. In IaaS, these are created on a physical hardware system in a data center. Software called a hypervisor separates the machine's resources from the hardware and provisions them appropriately so they can be used by the VM.
- VM automation helps in reducing the time taken for maintenance of the logical infrastructure (IaaS). Load balancing can be done for better performance and higher availability. Firewall provides the necessary security for clients to access the application.
- A cloud provider uses network protocols like fiber channel, iSCSI and NFS to split a physical network into logical partitions (or views). For example, the hypervisor can provide networking as a service to the virtual machines using routing, bridging, or network address translation.

1.8.1 Advantages and Disadvantages of IaaS

Advantages of IaaS :

1. Elimination of an expensive and staff - intensive data center
2. Ease of hardware scalability
3. Reduced hardware cost
4. On - demand, pay as you go scalability
5. Reduction of IT staff
6. Suitability for ad hoc test environments
7. Allows complete system administration and management
8. Support multiple tenants

Disadvantages of IaaS :

1. Because of the multi - tenant nature of IaaS, resources like bandwidth and disk space may be unevenly shared or used up by a certain consumer, which may affect the overall network performance.
2. There are network outages from the cloud vendor's end.
3. There is dependency on vendors for infrastructure upgrades and maintenance.
4. Although providers and subscribers follow security guidelines, the organization is dependent on the provider for data security.

1.8.2 Difference between IaaS, PaaS and SaaS

IaaS	PaaS	SaaS
IaaS gives users automated and scalable environments	PaaS provides a framework for quickly developing and deploying applications	SaaS makes applications available through the internet.
Amazon Web Services, for example, offers IaaS through the Elastic Compute Cloud, or EC2	Google Cloud Platform provides another PaaS option in App Engine	SaaS applications such as Gmail, Dropbox, Salesforce, or Netflix
In Iaas, infrastructure as a service.	In Paas, platform as a service	In SaaS, software as a service
Virtual platform on which required operating environment and application deployed	Operating environment included	Operating environment largely irrelevant, fully functional application provided
IaaS is a cloud service that provides basic computing infrastructure: servers, storage, and networking resources. In other words, IaaS is a virtual data center	PaaS refers to cloud platforms that provide runtime environments for developing, testing, and managing applications	SaaS allows people to use cloud-based web applications.
Major IaaS providers include Amazon Web Services, Microsoft Azure, and Google Compute Engine.	Examples of PaaS services are Heroku and Google App Engine.	email services such as Gmail and Hotmail are examples of cloud-based SaaS services.
IaaS services are available on a pay-for-what-you-use model	PaaS solutions are available with a pay-as-you-go pricing model.	SaaS services are usually available with a pay-as-you-go pricing model
Used by IT administrator	Used by software developers	Used by end user

1.9 Identity as a Service

- Identity as a Service (IDaaS) is cloud-based authentication operated by a third-party provider.
- Identity as a service (IDaaS) are SaaS-based Identity And Access Management (IAM) offerings that allow organizations to use Single Sign-on (SSO using SAML or OIDC), authentication and access controls to provide secure access to their growing number of software and SaaS applications.

- Five key capabilities are required to make enterprise IDaaS solutions possible :
  1. **Single Sign-on (SSO) :** With single sign-on employees, partners and customers obtain easy, fast and secure access to all SaaS, mobile and enterprise applications with a single authentication using corporate credentials.
  2. **Multi-factor Authentication (MFA) :** MFA typically includes adaptive authentication methods-options to step up as risk increases based on situational changes, user behavior or application sensitivity.
  3. **Access security :** Access security is policy-based access management for applications and APIs to enhance security beyond SSO.
  4. **Directory :** While most enterprises prefer to integrate IDaaS with their existing user stores, they may use a cloud directory, especially to support customers and/or partners.
  5. **Provisioning :** Through SCIM support and integration with on-premises provisioning, user data is synced with web and enterprise applications.
- IDaaS supplies cloud-based authentication or identity management to enterprises who subscribe. The goal is to ensure users are who they claim to be, and to give them the right kinds of access to software applications, files, or other resources at the right times. If the infrastructure to make this happen is built on site, then the company has to figure out what to do every time a problem comes up.

Advantages of IDaaS :

1. Deliver access services efficiently and cost-effectively.
2. Protect against internal and external security threats
3. With IDaaS, costs drop to the subscription fee and the administration work
4. Your team has to keep up servers; purchase, upgrade, and install software; back up data regularly; pay hosting fees

1.10 Cloud Infrastructure : Architectural Design of Compute and Storage Clouds

AU : Dec.-22

- Major design goals of a cloud computing platform is scalability, virtualization, efficiency, and reliability. Clouds support Web 2.0 applications.
- The cloud management receives the user request and then finds the correct resources, and then calls the provisioning services which invoke resources in the cloud. The cloud management software need to support both physical and virtual machines.

- The platform needs to establish a very large-scale HPC infrastructure. The hardware and software systems are combined together to make it easy and efficient to operate. The system scalability can benefit from cluster architecture.
- A cloud platform should be built to serve many users simultaneously. Therefore, multitasking is a necessity to assess distributed system performance.
- Five basic performance metrics are shown in Fig. 1.10.1. Refined performance models could be extended form basic attributes to include program behavior, environmental demand, QoS and cost-effectiveness.

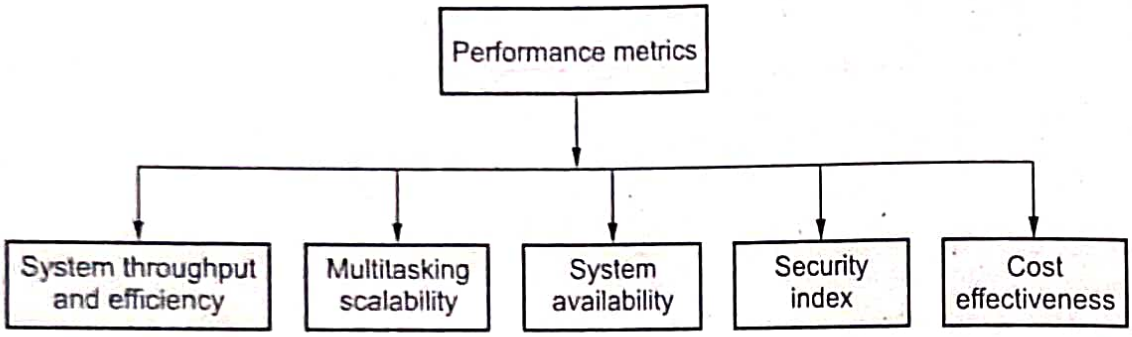


Fig. 1.10.1 Performance metrics

- Enabling technologies for clouds : The key driving forces behind cloud computing are the ubiquity of broadband and wireless networking, falling storage costs, and progressive improvements in Internet computing software.
- Cloud users are able to demand more capacity at peak demand, reduce costs, experiment with new services, and remove unneeded capacity, whereas service providers can increase the system utilization via multiplexing, virtualization and dynamic resource provisioning.
- Resource virtualization enables rapid cloud deployment-faster and fast disaster recovery. Service-oriented architecture (SOA) also plays a vital role. The progress in providing Software as a Service, Wed.2.0 standards and Internet performance have all contributed to the emergence of cloud services.
- The cloud computing resources are built in data centers, which are typically owned and operated by a third-party provider. Consumers do not need to know the underlying technologies.
- Web service providers offer special APIs that enable developers to exploit Internet clouds. Monitoring and metering units are used to track the usage and performance of resources provisioned. The software infrastructure of a cloud platform must handle all resource management and do most of the maintenance, automatically

1.10.1 Layered Cloud Architecture Development

- Fig. 1.10.2 shows layered architectural development of the cloud platform for IaaS, PaaS, and SaaS applications over the Internet and intranet.

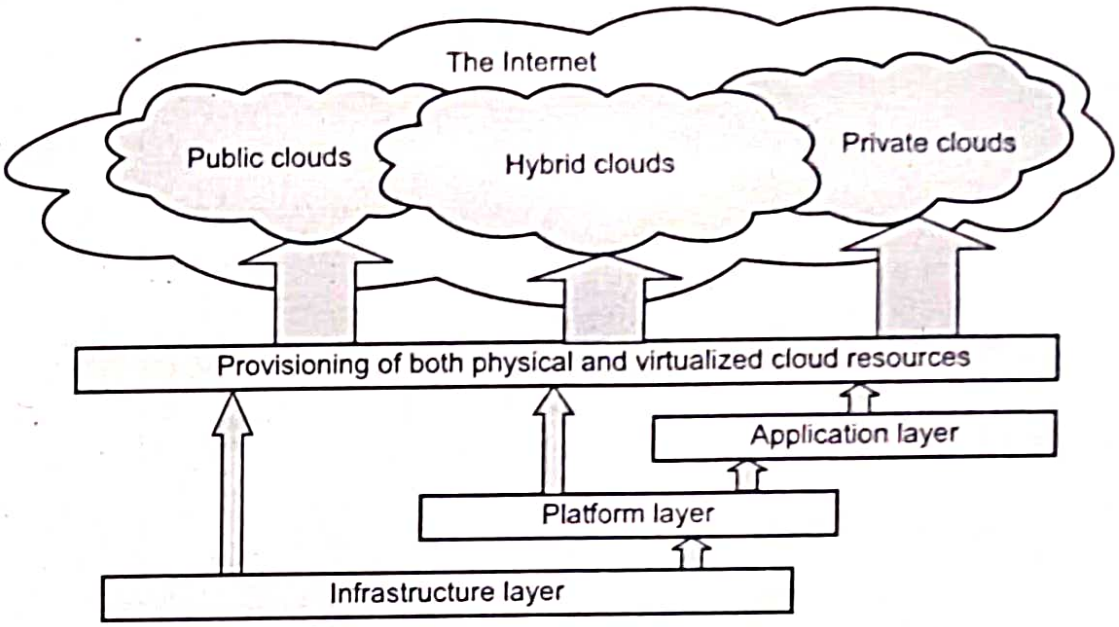


Fig. 1.10.2 Layered architectural development of the cloud platform

- The architecture of a cloud is developed at three layers : Infrastructure, platform, and application. These three development layers are implemented with virtualization and standardization of hardware and software resources provisioned in the cloud.
- The services to public, private, and hybrid clouds are conveyed to users through the networking support over the Internet and intranets involved. It is clear that the infrastructure layer is deployed first to support IaaS type of services.
- This infrastructure layer serves as the foundation to build the platform layer of the cloud for supporting PaaS services. The infrastructure layer is built with virtualized compute, storage and network resource.
- The platform layer is for general-purpose and repeated usage of the collection of software resources. The application layer is formed with a collection of all needed software modules for SaaS application.

1.10.2 Design Challenges

1. Service availability and data lock-in problem
2. Data privacy and security concerns
3. Unpredictable performance and bottlenecks
4. Distributed storage and wide-spread software bug

5. Cloud scalability, interoperability and standardization
6. Software licensing and reputation sharing

1.10.3 Generic Cloud Architecture

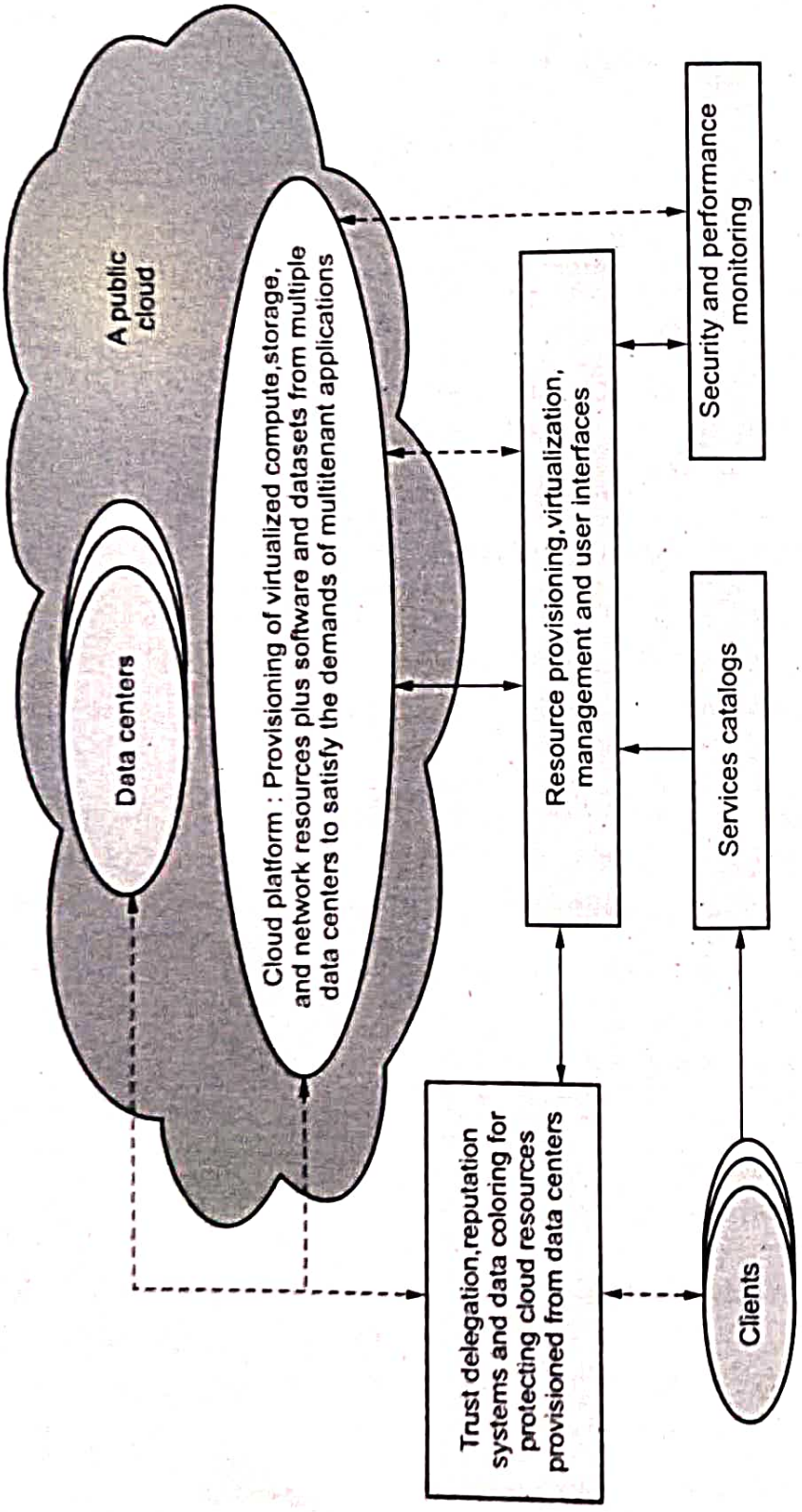


Fig. 1.10.3 Security - aware cloud architecture

- Fig. 1.10.3 shows a security - aware cloud architecture. The Internet cloud contains large number of cluster of servers. These servers are provisioned on demand to perform collective web services or distributed applications using data - center resources.
- The cloud platform is formed dynamically by provisioning servers, software and database resources. Servers in the cloud can be physical machines or Virtual machines. User interfaces are applied to request services.
- The cloud platform demands distributed storage and accompanying services. The cloud computing resources are built into the data centers, which are typically owned and operated by a third - party provider.
- Consumers do not need to know the underlying technologies. In a cloud, software becomes a service. The cloud demands a high degree of trust of massive amounts of data retrieved from large data centers. We need to build a framework to process large - scale data stored in the storage system. This demands a distributed file system over the database system.
- Other cloud resources are added into a cloud platform, including Storage Area Networks (SANs), database systems, firewalls and security devices. Web service providers offer special APIs that enable developers to exploit Internet clouds. Monitoring and metering units are used to track the usage and performance of provisioned resources.

1.10.4 Market - Oriented Cloud Architecture

- Fig. 1.10.4 shows architectural framework of cloud computing. (See Fig. 1.10.4 on next page)
  1. **Users/Brokers :** They submit their service requests from anywhere in the world to the cloud.
  2. **SLA resource allocator :** It is a kind of interface between users and cloud service provider which enable the SLA - oriented resource management.
  3. **Service request examiner and admission control :** It interprets the submitted request for QoS requirements before determining whether to accept or reject the request. Based on resource availability in the cloud and other parameters decide.
  4. **Pricing :** It is in charge of billing based on the resource utilization and some factors. Some factors are request time, type etc.
  5. **Accounting :** Maintains the actual usage of resources by request so that the final cost can be charged to the users.
  6. **VM monitor :** Keeps tracks on the availability of VMs and their resources.

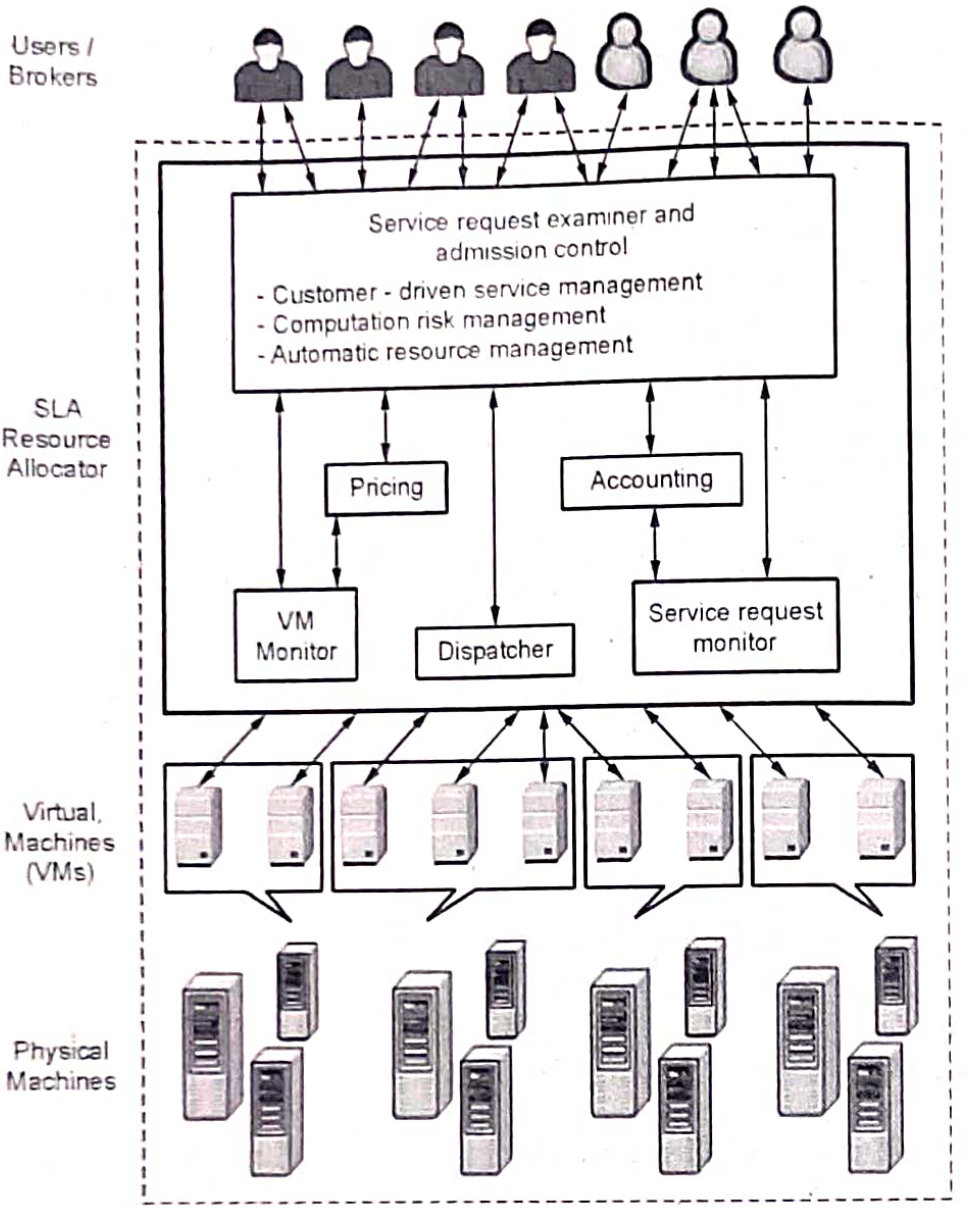


Fig. 1.10.4 Architectural framework

7. **Dispatcher :** The dispatcher mechanism starts the execution of admitted requests on allocated VMs.
8. **Service request monitor :** The request monitor mechanism keeps track of execution of request in order to be in tune with SLA.

Cloud computing service layers :

Parameters	Services	Description
Application Focused	Services	Services - Complete business services such as PayPal, OpenID, OAuth, Google Maps, Alexa
	Application	Application - Cloud based software that eliminates the need for local installation such as Google Apps, Microsoft Online

	Development	Development - Software development platforms used to build custom cloud based applications (PAAS and SAAS) such as SalesForce
Infrastructure Focused	Platform	Platform - Cloud based platforms, typically provided using virtualization, such as Amazon ECC, Sun Grid
	Storing	Storage - Data storage or cloud based NAS such as CTERA, iDisk, CloudNAS
	Hosting	Hosting - Physical data centers such as those run by IBM, HP, NaviSite, etc.

Cloud components :

- Cloud computing solutions are made up of several elements.
  1. **Clients :** Mobile, terminals or regular computers.
  2. **Benefits :** Lower hardware costs, lower IT costs, security, data security, less power consumption, ease of repair or replacement, less noise.
  3. **Data centers :** Collection of servers where the application to subscribe is housed. It could be a large room in the basement of your building or a room full of servers on the other side of the world
  4. **Virtualizing servers :** Software can be installed allowing multiple instances of virtual servers to be used and a dozen virtual servers can run on one physical server.
  5. **Distributed servers :** Servers don't all have to be housed in the same location. It can be in geographically disparate locations. If something were to happen at one site, causing a failure, the service would still be accessed through another site. If the cloud needs more hardware, they can add them at another site.

University Questions

1. *List the architectural design challenges in cloud environment.* **AU : Dec.-22, Marks 5**
2. *Elaborate on any one cloud environment with details architecture.* **AU : Dec.-22, Marks 5**

1.11 Migrating into the Cloud

- Cloud migration is the process of transferring data, application code and other technology-related business processes from an on-premise or legacy infrastructure to the cloud environment.
- Cloud migration is a phenomenal transformation in the business information system domain as it provides adequate services for the growing needs of

businesses. However, moving data to the cloud requires preparation and planning in deciding on an approach.

- Technology companies providing cloud and managed services should work closely with businesses and support them on three main aspects as part of their cloud migration journey :
  a) Developing the right approach by ensuring clarity on the desired business outcome or end result.
  b) Taking advantage of relevant talent and expertise to address companie's core business issues.
  c) Using appropriate toolsets that are user-friendly, secure and viable in the long term.
- Your cloud computing migration strategy must deliver elasticity, agility, and scalability to help your organization take advantage of emerging opportunities and pivot to address business and industry changes.
- Achieving these goals will broadly employ a mix of Software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) cloud architecture layers. You may employ these across hybrid and multi-cloud environments based on business needs and best practices.
- Cloud migration strategy should focus on four broad areas :
  a) Security that determines application and data access controls, data security at rest and in transit and backup and data recovery.
  b) Governance and compliance that makes sure you meet regulatory compliance and all aspects of data security to go beyond the limits of shared security models of cloud services providers.
  c) Cost management that looks at immediate and long-term cost savings that guide where and how you migrate, manage, and monitor applications and workloads.
  d) Accessibility, resilience, and scalability that are driven by customer and workforce UX needs, service enhancements and innovation.
- Benefits of migrating to the cloud include :
  a) Increased agility and flexibility
  b) Ability to innovate faster
  c) Easing of increasing resource demands
  d) Better managing of increased customer expectations
  e) Reduction in costs
  f) Deliver immediate business results

g) Simplify IT

h) Shift to everything as-a-service

i) Better consumption management

j) Cloud scalability

k) Improved performance

Cloud migration process :

- **Cloud migration process is divided into three phases :** ***plan, execute (run) and monitor***

1. **Plan :** Cloud migration requires solid planning to be successful. It define following parameters :

   a. Identify business objective

   b. Identify key business drivers

   c. Get executive sponsorship

   d. Providing full visibility into your on-premise environment, including all system dependencies.

   f. Evaluating performance, server, and security requirements.

2. **Execute or run :** Once your environment has been assessed and a plan has been mapped out, it's necessary to execute your migration. The main challenge here is carrying out your migration with minimal disruption to normal operation, at the lowest cost, and over the shortest period of time.

3. **Monitor :** After a successful migration, tools and processes should be implemented to monitor the new cloud environment.

- A formal issue tracking process should be created to ensure that everyone impacted by the migration has a way to report problems, and IT has a simple way to manage all requests.

Challenges of cloud migration :

1. **Interoperability :** It is no easy feat to get existing applications to communicate with newer cloud environments.

2. **Resource availability :** The migration process might require taking in-house servers temporarily offline. But downtime could be disastrous to application performance, thus customer loyalty, if not supported by a proper plan for disaster recovery.

3. **Data integrity :** How will user keep data secure while moving it to the cloud where have less control ?

4. **Resource management :** Not all IT professionals trust the cloud yet. If team was used to managing physical servers, they might need educating on the new infrastructure or even reconfiguring to introduce new roles.

1.11.1 Seven Step Model of Migrating into the Cloud

- Seven-step model is as follows :
  1. Different variances of cloud services
  2. Cloud as a tool
  3. Cloud compatible
  4. Current cost
  5. To manage rather than operate
  6. To simplify
  7. Gain more knowledge.
- **Step 1 :** Cloud migration assessments comprise assessments to understand the issues involved in the specific case of migration at the application level or the code, the design, the architecture, or usage levels.
  These assessments are about the cost of migration as well as about the ROI that can be achieved in the case of production version.
- **Step 2 :** An isolating all systemic and environmental dependencies of the enterprise application components within the captive data center.
- **Step 3 :** Generating the mapping constructs between what shall possibly remain in the local captive data center and what goes onto the cloud.
- **Step 4 :** Substantial part of the enterprise application needs to be rearchitected, redesigned, and reimplemented on the cloud.
- **Step 5 :** We leverage the intrinsic features of the cloud computing service to augment our enterprise application in its own small ways.
- **Step 6 :** We validate and test the new form of the enterprise application with an extensive test suite that comprises testing the components of the enterprise application on the cloud as well.
- **Step 7 :** Test results could be positive or mixed. In the latter case, we iterate and optimize as appropriate. After several such optimizing iterations, the migration is deemed successful.

1.12 Two Marks Questions with Answers

Q.1 Define cloud computing. AU : Dec.-21

Ans. : NIST definition of cloud : Cloud computing is a pay-per-use model for enabling available, convenient, on - demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service - provider interaction

Q.2 What is the use of elasticity in cloud ? AU : Dec-22

Ans. : The elastic nature of cloud services has provided enterprises with incredible flexibility in consuming resources for computing, storage, infrastructure and more. With access to a rapidly growing ecosphere of cloud products on - demand, enterprises have been able to achieve the agility, scalability and cost savings required to increase competitiveness and fuel digital transformation.

Q.3 Define on-demand provisioning. AU : Dec.-22

OR Depict the importance of on - demand provisioning in e-commerce applications. AU : Dec.-21

Ans. : On-Demand Computing (ODC) is a delivery model in which computing resources are made available to the user as needed. The resources may be maintained within the user's enterprise or made available by a cloud service provider.

Q.4 Why do we need a hybrid cloud ? Justify. AU : Dec.-22

Ans. : Hybrid cloud solutions enable us to migrate and manage workloads between these various cloud environments, allowing to create more versatile setups based on specific business needs. Many organizations choose to adopt hybrid cloud platforms to reduce costs, minimize risk and extend their existing capabilities to support digital transformation efforts.

Q.5 Give the advantages of storage as a service. AU : Dec.-22

Ans. :

- Scalability: STaaS provides a high level of scalability.
- High - level backup and recovery
- Optimizes storage to be cost-effective

Q.6 What is a hybrid cloud ? AU : Dec.-21

Ans. : Hybrid cloud integrates public cloud services, private cloud services and on-premises infrastructure and provides orchestration, management and application portability across all three.

Q.7 List the main characteristics of cloud computing. AU : Dec-20

Ans. : Cloud computing's characteristics include on - demand self - service, broad network access and being very elastic and scalable.

Q.8 Illustrate the virtual appliances in cloud computing. AU : Dec-20

Ans. : Virtual appliance is a pre - installed and pre - configured software solution on one or more virtual machines that is optimized for a specific function. A virtual appliance does not require locally installed hardware and can be remotely accessed by users. Its purpose is to simplify the delivery and operation of an application, so only the OS components required to support the application's functions are included.

Q.9 Differentiate public cloud and private cloud. AU : Dec.-20

Ans. :

Public cloud	Private cloud
Public cloud infrastructure is offered via web applications and also as web services over Internet to the public.	Private cloud infrastructure is dedicated to a single organization.
Support multiple customer	Support dedicated customer
Full utilized of infrastructure.	Does not utilize shared infrastructure
Security is low as compared to private cloud	High level of security
Low cost	High cost
Azure, Amazon Web Services, Google App Engine and Force.com are a few examples of public clouds	An example of the private cloud is NIRIX's one server with dedicated servers

Q.10 Summarize the benefits and drawbacks of using "Platform as a Service". AU : Dec.-20

Ans. : Benefits :

1. Reduced capital expenditure.
2. Reduced skill requirements.
3. Ability to add new uses quickly.

Drawbacks of PaaS :

1. Security and compliance risks : PaaS software is generally offered in a public cloud where it is shared by multiple users, which creates a higher risk of security vulnerabilities.
2. Vendor lock-in : PaaS solutions for each business requirement might differ and the chosen vendor may not be able to provide convenient options for frameworks, customization, or migration.
3. Loss of operational control : Developers may have to trade off the abstraction for more granular control over application components.

Q.11 List cloud enabling technology.

Ans. : Enabling technologies are as follows :

1. Broadband networks and internet architecture
2. Data center technology
3. Virtualization technology
4. Web technology
5. Multitenant technology

Q.12 What is cloud service ?

Ans. :

- Cloud service is any service made available to users on demand via the Internet from a cloud computing provider's servers as opposed to being provided from a company's own on - premises servers.
- Cloud services are designed to provide easy, scalable access to applications, resources and services, and are fully managed by a cloud services provider.

Q.13 What Is a dynamic infrastructure platform ?

Ans. : A dynamic infrastructure platform is a system that provides computing resources, particularly servers, storage and networking, in a way that they can be programmatically allocated and managed.

Q.14 What is cloud adoption ?

Ans. : Cloud adoption is a strategic move by organizations of reducing cost, mitigating risk and achieving scalability of data base capabilities. Cloud adoption may be up to various degrees in an organization, depending on the depth of adoption. In fact the depth of adoption yields insight into the maturity of best practices, enterprise-ready cloud services availability.

Q.15 What is cloud reference model ? What are the applications of this models ?

AU : Dec.-15

Ans. : The cloud computing reference model is an abstract model that characterizes and standardizes the functions of a cloud computing environment by partitioning it into abstraction layers and cross - layer functions. The three cross - layer functions are business continuity, security and service management

Q.16 List and explain cloud deployment models.

Ans. : Cloud deployment models are refers to the location and management of the cloud's infrastructure. Deployment models are defined by the ownership and control of architectural design and the degree of available customization. Cloud deployment models are private, public and community clouds

Q.17 List and explain cloud deployment models.

Ans. : Cloud deployment models are refers to the location and management of the cloud's infrastructure. Deployment models are defined by the ownership and control of architectural design and the degree of available customization. Cloud deployment models are private, public and community clouds

Q.18 What is public cloud ?

Ans. : Public cloud is built over the Internet and can be accessed by any user who has paid for the service. Public clouds are owned by service providers and are accessible through a subscription

Q.19 What is private clouds ?

Ans. : A private cloud is built within the domain of an intranet owned by a single organization. Therefore, it is client owned and managed and its access is limited to the owning clients and their partners.

Q.20 What is community cloud ?

Ans. : The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements, policy, or compliance considerations). It may be managed by the organizations or a third party and may exist on - premises or off - premises

Q.21 What do you understand by SaaS ? **AU : June-16**

Ans. : Software-as-a-Service (SaaS) is a software delivery model that provides access to applications through the Internet as a Web-based service. It provides a means to free users from complex hard - ware and software management by offloading such tasks to third parties, which build applications accessible to multiple users through a Web browser

Q.22 What do you mean by term scalability in terms of cloud ? **AU : June-16**

Ans. : The ability to scale on demand constitutes one of the most attractive features of cloud computing. Scalability refers to the idea of a system in which every application or piece of infrastructure can be expanded to handle increased load.

Q.23 How Cloud computing provides scalability and fault tolerance ? **AU : Dec.-15**

Ans. : Fault tolerance is the process of finding faults and failures in a system. If a fault occurs or there is a hardware failure or software failure then also the system should work properly

Q.24 What is cloud ecosystem ? **AU : Dec.-16**

Ans. :

- Cloud computing ecosystem are business process, application services, platform services and Infrastructure services.

- A cloud ecosystem is a complex system of interdependent components that all work together to enable cloud services.
- In cloud computing, the ecosystem consists of hardware and software as well as cloud customers, cloud engineers, consultants, integrators and partners.

Q.25 What is meant by IaaS in cloud computing ?

Ans. : IaaS is a cloud service model where the physical resources or hardware like storage, compute and network are provisioned and maintained by a cloud provider. Organizations can access these resources virtually via the internet for as long as they want.

Q.26 What is the use of IaaS in cloud computing ?

Ans. : IaaS provides the basic infrastructure like storage, network and compute resources and other additional offerings like load balancing, clustering and security to organizations in a virtualized manner over the internet. This way, organizations can focus on building their applications, without worrying about maintaining the infrastructure or installing the hardware and software, reducing costs and speeding up the application development.

❑❑❑

Notes

UNIT II

2 Virtualization Basics

Syllabus

Virtual Machine Basics - Taxonomy of Virtual Machines - Hypervisor - Key Concepts - Virtualization structure - Implementation levels of virtualization - Virtualization Types : Full Virtualization - Para Virtualization - Hardware Virtualization - Virtualization of CPU, Memory and I/O devices.

Contents

2.1 Virtual Machine Basics

- Virtual Machine (VM) is a virtual environment that functions as a virtual computer system with its own CPU, memory, network interface, and storage, created on a physical hardware system.
- A Virtual machine is a software construct that mimics the characteristics of a physical server.
- A Virtual Machine (VM) is a software program or operating system that not only exhibits the behavior of a separate computer, but is also capable of performing tasks such as running applications and programs like a separate computer.
- In a pure virtual machine architecture the operating system gives each process the illusion that it is the only process on the machine. The user writes an application as if only its code were running on the system.
- Each user interacts with the computer by typing commands to the virtual machine on a virtual system console and receiving results back from the machine as soon as they are computed.
- Each user directs the virtual machine to perform different commands. These commands are then executed on the physical machine in a multiprogramming environments.
- Virtualization is an abstraction layer that decouples the physical hardware from the operating system to deliver greater IT resource utilization and flexibility.
- It allows multiple virtual machines, with heterogeneous operating systems to run in isolation, side-by-side on the same physical machine.
- Fig. 2.1.1 shows virtual machine.

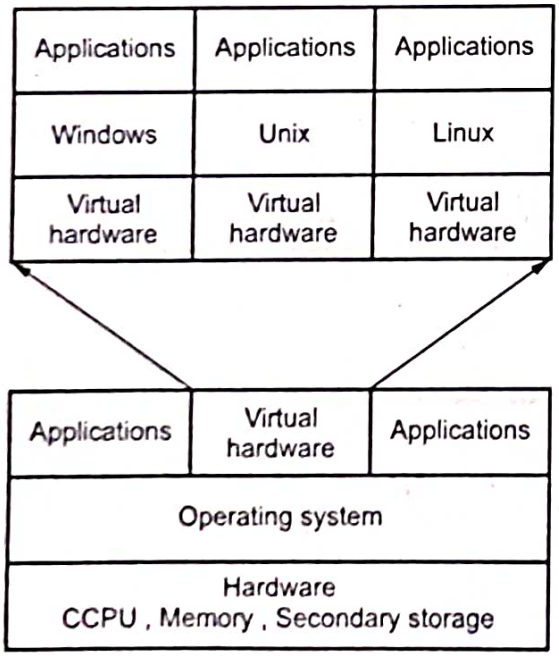


Fig. 2.1.1 Virtual machine

Benefits :

1. There is no overlap amongst memory as each Virtual Memory has its own memory space.
2. Virtual machines are completely isolated from the host machine and other virtual machines.
3. Data does not leak across virtual machines.
4. Can use multiple operating system environments on the same computer
5. The cost reduction is possible using small virtual servers on a more powerful single server.

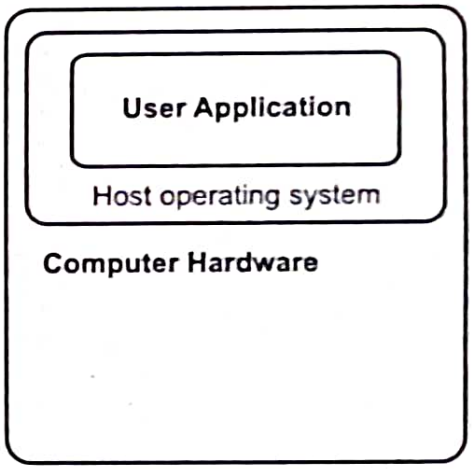
Disadvantages :

1. Virtual machines are less efficient than real machines because they access the hardware indirectly.
2. A virtual machine can be infected with the weaknesses of the host machine
3. Difficulty in direct access to hardware, for example, specific cards or USB devices
4. Great use of disk space, since it takes all the files for each operating system installed on each virtual machine.

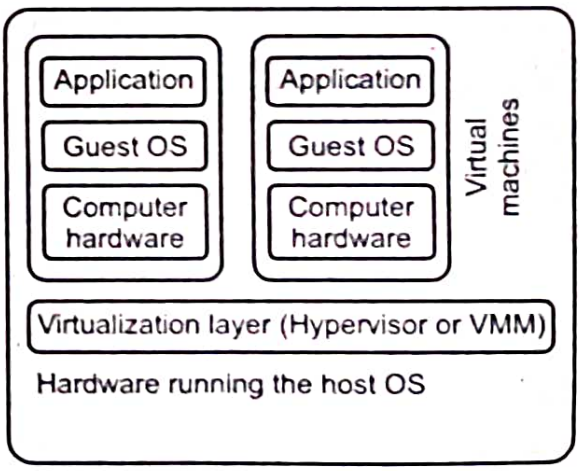
2.2 Taxonomy of Virtual Machines

AU : Dec.-21,22

- Virtualization is a broad term that refers to the abstraction of resources across many aspects of computing. For our purposes : One physical machine to support multiple virtual machines that run in parallel.
- Virtualization is a frame work or methodology of dividing the resources of computer into multiple execution environments.
- Virtualization is an abstraction layer that decouples the physical hardware from the operating system to deliver greater IT resource utilization and flexibility.
- It allows multiple virtual machines, with heterogeneous operating systems to run in isolation, side-by-side on the same physical machine.
- Fig. 2.2.1 shows before and after virtualization.



(a) Before virtualization



(b) After virtualization

Fig. 2.2.1

- Virtualization means running multiple machines on a single hardware. The "Real" hardware invisible to operating system. OS only sees an abstracted out picture. Only Virtual Machine Monitor (VMM) talks to hardware.
- It is "a technique for hiding the physical characteristics of computing resources from the way in which other systems, applications, or end users interact with those resources.
- This includes making a single physical resource appear to function as multiple logical resources; or it can include making multiple physical resources appear as a single logical resource."
- It is divided into two main categories :
  1. Platform virtualization involves the simulation of virtual machines.
  2. Resource virtualization involves the simulation of combined, fragmented or simplified resources.
- Fig. 2.2.2 shows taxonomy of virtualization.

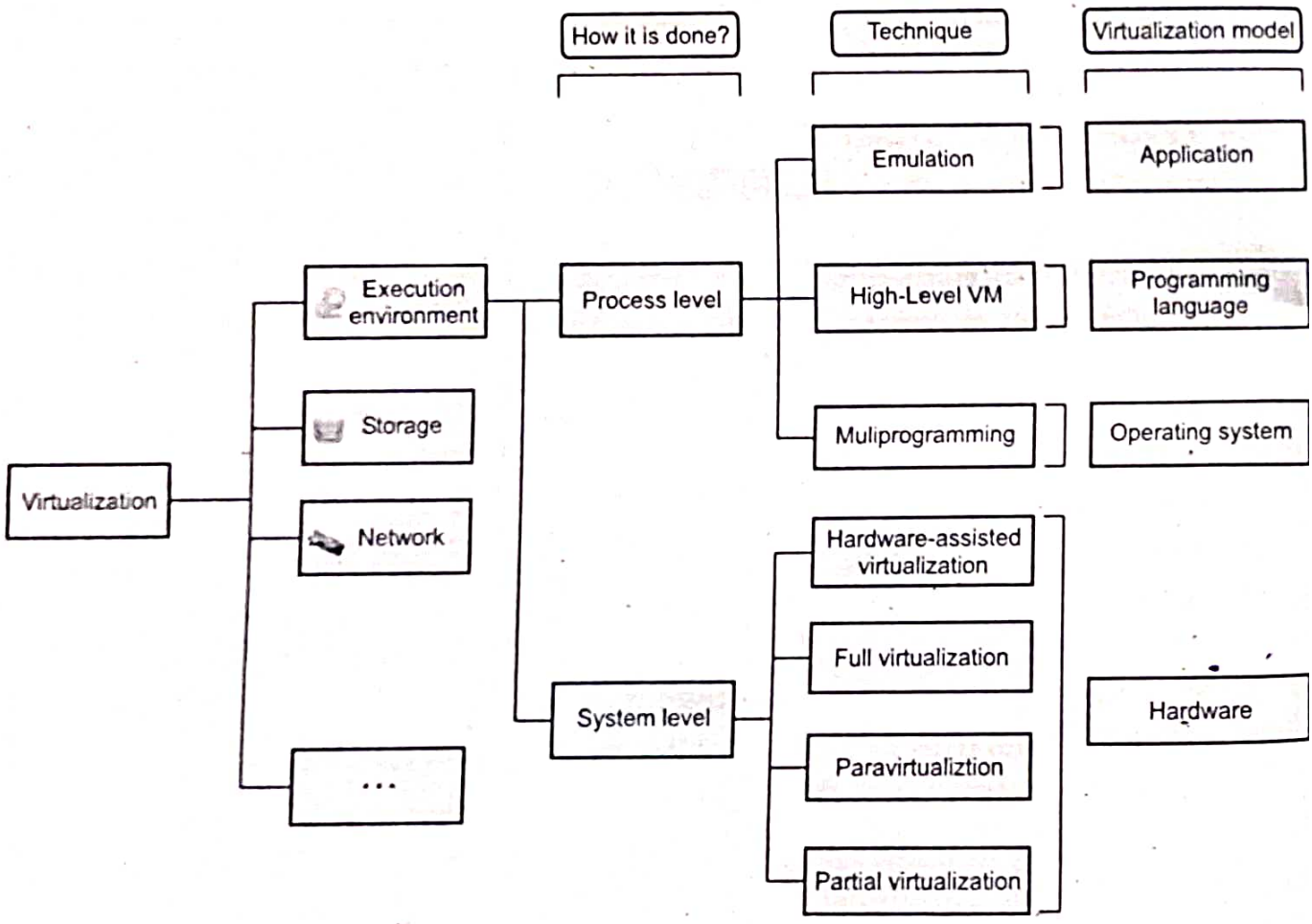


Fig. 2.2.2 Taxonomy of virtualization

- Virtualization is mainly used to emulate execution environment, storage and network. Execution environment classified into two types : process level and system level.

- Process level is implemented on top of an existing operating system.
- System level is implemented directly on hardware and do not or minimum requirement of existing operating system.

2.2.1 Difference between Virtualization and Cloud Computing

Sr. No.	Virtualization	Cloud Computing
1.	Virtualization is the process of creating a virtual environment on an existing server to run your desired program, without interfering with any of the other services provided by the server or host platform to other users.	Cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive.
2.	Location of virtual machine is on a specific host.	Location of virtual machine is on any host.
3.	Instance storage is persistent.	Instance storage is shortly lived.
4.	Virtualization uses customizable VM resource like CPU and RAM.	Cloud computing uses standard VM resource like CPU and RAM
5.	Recovery from failures: attempt to recover failed VM.	Recovery from failures : Discard instance spin up new one.

2.2.2 Pros and Cons of Virtualization

a) Pros

1. Data center and energy-efficiency savings : As companies reduce the size of their hardware and server footprint, they lower their energy consumption
2. Operational expenditure savings : Once servers are virtualized, your IT staff can greatly reduce the ongoing administration and management of manual work.
3. Reduced costs : It reduced cost of IT infrastructure.
4. Data does not leak across virtual machine.
5. Virtual machine is completed isolated from host machine and other virtual machine.
6. Simplifies resource management by pooling and sharing resources.
7. Significantly reduce downtime.
8. Improved performance of IT resources.

b) Cons

1. Not all hardware or software can be virtualized.
2. Not all servers are applications are specifically designed to be virtualization-friendly.

University Questions

1. *Outline the various levels of virtualization with an example for each category.*
AU : Dec.-21, Marks 13

2. *What is virtualization ? List the various levels of virtualization ? Explain.*
AU : Dec.-22, Marks 10

2.3 Hypervisor

- In computing, a hypervisor is a virtualization platform that allows multiple operating systems to run on a host computer at the same time. The term usually refers to an implementation using full virtualization.
- A hypervisor is a software layer installed on the physical hardware, which allows splitting the physical machine into many virtual machines. This allows multiple operating systems to be run simultaneously on the same physical hardware.
- The operating system installed on the virtual machine is called a guest OS, and is sometimes also called an instance. The hardware the hypervisor runs on is called the host machine.
- A hypervisor management console, which is also called a virtual machine manager (VMM), is computer software that enables easy management of virtual machines.
- Hypervisors are currently classified in two types : type 1 and type 2

2.3.1 Type 1

- Type 1 hypervisor is software that runs directly on a given hardware platform. A "guest" operating system thus runs at the second level above the hardware.
- Fig. 2.3.1 shows Type 1 hypervisor.

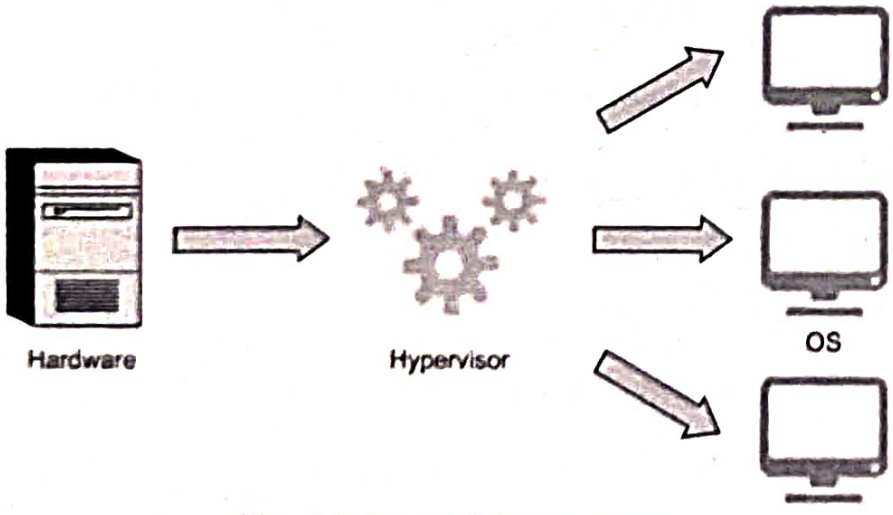


Fig. 2.3.1 Type 1 hypervisor

- Type 1 VMs have no host operating system because they are installed on a bare system. An operating system running on a Type 1 VM is a full virtualization because it is a complete simulation of the hardware that it is running on.
- Type 1 hypervisor is also called a native or bare-metal hypervisor that is installed directly on the hardware, which splits the hardware into several virtual machines where we can install guest operating systems.
- Virtual machine management software helps to manage this hypervisor, which allows guest OSes to be moved automatically between physical servers based on current resources requirements.
- It is completely independent from the Operating System.
- The hypervisor is small as its main task is sharing and managing hardware resources between different operating systems.
- A major advantage is that any problems in one virtual machine or guest operating system do not affect the other guest operating systems running on the hypervisor.

2.3.2 Type 2 Hypervisor

- This is also known as Hosted Hypervisor.
- In this case, the hypervisor is installed on an operating system and then supports other operating systems above it.
- It is completely dependent on host Operating System for its operations. Fig. 2.3.2 shows type 2 hypervisor.

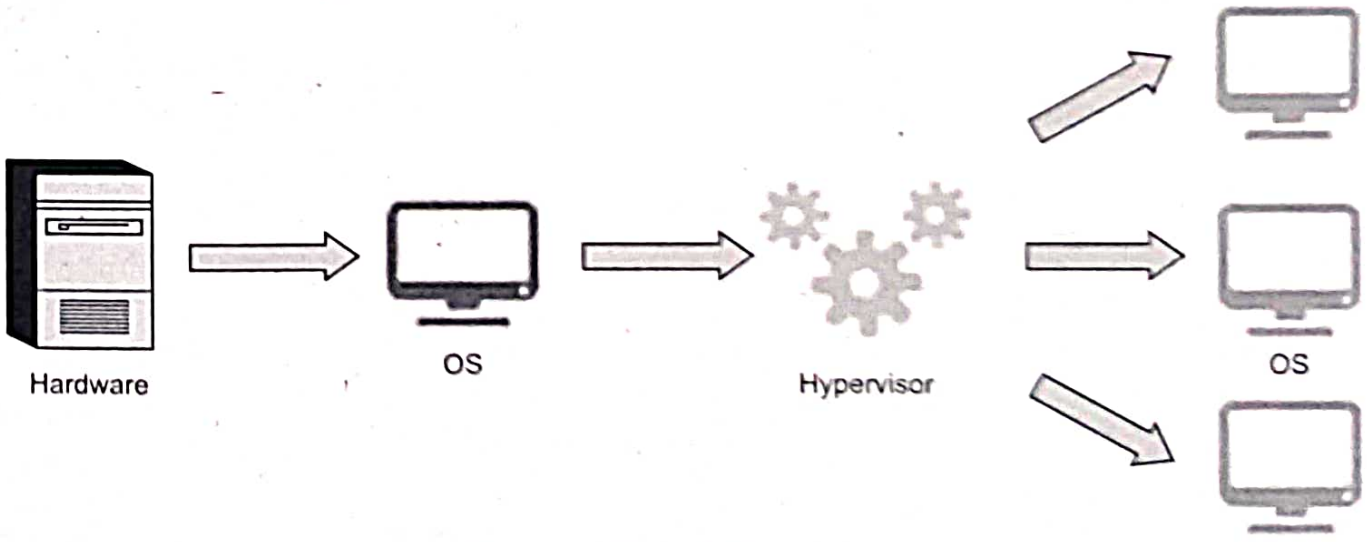


Fig. 2.3.2 Type 2 Hypervisor

- While having a base operating system allows better specification of policies, any problems in the base operating system affects the entire system as well even if the hypervisor running above the base OS is secure.

- Type 2 hypervisors don't support over/dynamic allocation of RAM, so care is required when allocating resources to virtual machines
- This is why we call type 2 hypervisors hosted hypervisors. As opposed to type 1 hypervisors that run directly on the hardware, hosted hypervisors have one software layer underneath. What we have in this case is :
  1. A physical machine.
  2. An operating system installed on the hardware (Windows, Linux, MacOS).
  3. A type 2 hypervisor software within that operating system.
  4. The actual instances of guest virtual machines.
- Type 2 hypervisors are typically found in environments with a small number of servers. Type 2 hypervisors are convenient for testing new software and research projects.

2.3.3 Paravirtualization

- Paravirtualization is a type of virtualization in which a guest operating system (OS) is recompiled, installed inside a virtual machine (VM), and operated on top of a hypervisor program running on the host OS
- Para-virtualization refers to communication between the guest OS and the hypervisor to improve performance and efficiency.
- Para-virtualization involves modifying the OS kernel to replace non-virtualizable instructions with hyper-calls that communicate directly with the virtualization layer hypervisor.
- The hypervisor also provides hyper-call interfaces for other critical kernel operations such as memory management, interrupt handling and time keeping.
- Fig. 2.3.3 shows para-virtualization architecture.

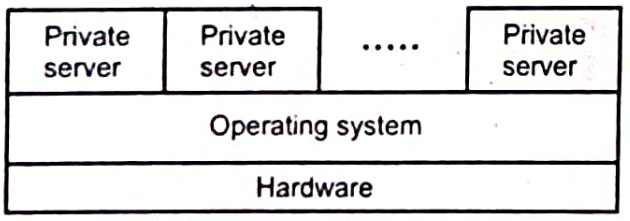


Fig. 2.3.3 Para-virtualization architecture

- In Para-virtualization, the virtual machine does not necessarily simulate hardware, but instead offers a special API that can only be used by modifying the "guest" OS. This system call to the hypervisor is called a "hypercall" in Xen.
- Xen is an open source para-virtualization solution that requires modifications to the guest operating systems but achieves near native performance by collaborating with the hypervisor.

- Microsoft Virtual PC is a para-virtualization virtual machine approach. User-mode Linux (UML) is another para-virtualization solution that is open source.
- Each guest operating system executes as a process of the host operating system. Cooperative Linux, is a virtualization solution that allows two operating systems to cooperatively share the underlying hardware.
- Linux-V server is an operating system-level virtualization solution for GNU/Linux systems with secure isolation of independent guest servers.
- The Linux KVM is virtualization technology that has been integrated into the mainline Linux kernel . Runs as a single kernel loadable module, a Linux kernel running on virtualization-capable hardware is able to act as a hypervisor and support unmodified Linux and Windows guest operating systems.
- Para-virtualization shares the process with the guest operating system.

Problems with para-virtualization

1. Para-virtualized systems won't run on native hardware
2. There are many different para-virtualization systems that use different commands, etc.

- The main difference between full virtualization and paravirtualization in Cloud is that full virtualization allows multiple guest operating systems to execute on a host operating system independently while paravirtualization allows multiple guest operating systems to run on host operating systems while communicating.

2.3.4 Difference between Type 1 and Type 2 Hypervisor

Type 1 Hypervisor	Type 2 Hypervisor
This is also known as Bare Metal or Embedded or Native Hypervisor	This is also known as Hosted Hypervisor
It is completely independent from the Operating System	It is completely dependent on host Operating System for its operations
It works directly on the hardware of the host and can monitor operating systems that run above the hypervisor	In this case, the hypervisor is installed on an operating system and then supports other operating systems above it
It support hardware virtualization	It support OS virtualization
Examples : VMware ESXi Server and Microsoft Hyper-V	Examples : Vmware Workstation, Microsoft Virtual PC, Oracle Virtual Box
Higher performance and scalability because of being bare metal type	Low performance as a result of host operating system overhead

2.4 Implementation Levels of Virtualization

- Virtualization is implemented at various levels :
  1. Instruction set architecture level
  2. Hardware abstraction level
  3. Operating system level
  4. Library support level
  5. User application level

2.4.1 Instruction Set Architecture Level

- The definition of the storage resources and the instructions that manipulate data are documented in what is referred to as Instruction Set Architecture (ISA).
- ISA view of a machine corresponds to the machine and assembly language levels. For example, MIPS binary code can run on an x86-based host machine with the help of ISA emulation.
- Instruction set emulation leads to virtual ISAs created on any hardware machine. The basic emulation method is through code interpretation. An interpreter program interprets the source instructions to target instructions one by one.
- The key to virtualize a CPU lies in the execution of the guest instruction, including both system-level and user-level instructions virtualizing a CPU can achieved in one two ways :
  1. **Emulation :** The only processor vitalization mechanism available when the ISA of the guest is different from the ISA of the host.
  2. **Direct native execution :** Possible only if the ISA of the host is identical to the ISA of the guest.
- Fig. 2.4.1 shows ISA emulation.

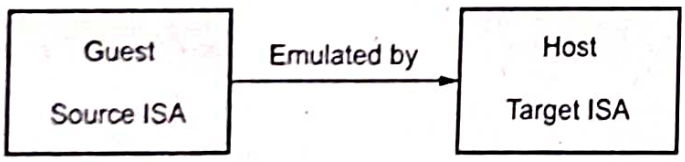


Fig. 2.4.1 ISA emulation

- Emulation is the process of implementing the interface and functionality of one system (or subsystem) on a system (or subsystem) having different interface and functionality.

- In other words, emulation allows a machine implementing one ISA (the target), to reproduce the behavior of a software compiled for another ISA (the source). Emulation can be carried out using :
  1. Interpretation
  2. Binary traslation

2.4.2 Hardware Abstraction Level

- This type of virtualization is performed right on top of the bare hardware. On the hand, this approach generates a virtual hardware environment for a VM. On the other hand, the process manages the underlying hardware through virtualization.
- The idea is to virtualizes a computer's resources, such as processors memory, and I/O devices. The intention is to upgrade the hardware utilization rate by multiple users concurrently.
- The Xen hypervisor has been applied to virtualize x86-based machines to run Linux or other guest OS applications.

2.4.3 Operating System Level Virtualization

- Operating-system-level virtualization is a server-virtualization method where the kernel of an operating system allows for multiple isolated user-space instances, instead of just one. Such instances, which are sometimes called containers and software containers.
- This refers to an abstraction layer between traditional OS and user applications.
- This type of virtualization creates isolated containers on a single physical server and the OS instances to utilize the hard-ware and software in data centers.
- Containers behave like real servers. With containers you can create a portable, consistent operating environment for development, testing, and deployment.
- This virtualization creates virtual hosting environments to allocates hardware resources among a large number of mutually distrusting users.
- Operating-system-level virtualization usually imposes little to no overhead, because programs in virtual partitions use the operating system's normal system call interface and do not need to be subjected to emulation or be run in an intermediate virtual machine.
- Operating system-level virtualization is not as flexible as other virtualization approaches since it cannot host a guest operating system different from the host one, or a different guest kernel.

- Instead of trying to run an entire guest OS, container virtualization isolates the guests, but doesn't try to virtualize the hardware. Instead, you have containers for each virtual environment.
- With container-based technologies, you'll need a patched kernel and user tools to run the virtual environments. The kernel provides process isolation and performs resource management.

Why operating system level virtualization is required ?

- Operating system level virtualization provides feasible solution for hardware level virtualization issue. It inserts a virtualization layer inside an operating system to partition a machine's physical resources.
- It enables multiple isolated VMs within a single operating system kernel. This kind of VM is often called a virtual execution environment (VE), Virtual Private System (VPS), or simply container.
- From the user's point of view, virtual execution environment look like real servers.
- This means a virtual execution environment has its own set of processes, file system, user accounts, network interfaces with IP addresses, routing tables, firewall rules etc.
- Although VEs can be customized for different people, they share the same operating system kernel. Therefore, OS-level virtualization is also called single-OS image virtualization.

Challenges to cloud computing in OS level virtualization ?

- Cloud computing is transforming the computing landscape by shifting the hardware and staffing costs of managing a computational center to third parties.
- Cloud computing has at least two challenges :
  1. The ability to use a variable number of physical machines and virtual machine instances depending on the needs of a problem. For example, a task may need only a single CPU during some phases of execution but may need hundreds of CPUs at other times.
  2. It is related to slow operation of instantiating new virtual machine. Currently, new virtual machines originate either as fresh boots or as replicates of a template VM, unaware of the current application state. Therefore, to better support cloud computing, a large amount of research and development should be done.

Advantages of OS virtualization :

1. OS virtualization provide least overhead among all types of virtualization solution.
2. They offer highest performance and highest density of virtual environment.
3. Low resource requirements.
4. High Scalability.

Disadvantage of OS virtualization :

1. They support only one operating system as base and guest OS in a single server.
2. It supports library level virtualization.

2.4.4 Library Support Level

- Library-level virtualization is also known as user-level Application Binary Interface (ABI).
- This type of virtualization can create execution environments for running alien programs on a platform rather than creating a VM to run the entire operating system.
- It is done by API call interception and remapping.
- Virtualization with library interfaces is possible by controlling the communication link between applications and the rest of a system through API hooks.
- **Example :** Wine, WAB, LxRun, Visual MainWin
- **Advantage :** It has very low implementation effort
- **Shortcoming and limitation :** Poor application flexibility and isolation.

2.4.5 User Application Level

- Virtualization at the application level virtualizes an application as a VM. On a traditional OS, an application often runs as a process. Therefore, application-level virtualization is also known as process-level virtualization.
- A fully virtualized application is not installed in the traditional sense, although it is still executed as if it were. The application behaves at runtime like it is directly interfacing with the original operating system and all the resources managed by it, but can be isolated to varying degrees.
- Full application virtualization requires a virtualization layer. Application virtualization layers replace part of the runtime environment normally provided by the operating system.
- The layer intercepts all disk operations of virtualized applications and transparently redirects them to a virtualized location, often a single file.

- The application remains unaware that it accesses a virtual resource instead of a physical one. Since the application is now working with one file instead of many files spread throughout the system, it becomes easy to run the application on a different computer and previously incompatible applications can be run side-by-side.
- The most popular approach is to deploy High Level Language (HLL) VMs. Here the virtualization layer sits as an application program on top of the operating system, and the layer exports an abstraction of a VM that can run programs written and compiled to a particular abstract machine definition. Any program written in the HLL and compiled for this VM will be able to run on it.
- **Benefits :**
  1. Application virtualization uses fewer resources than a separate virtual machine.
  2. Application virtualization also enables simplified operating system migrations.
  3. Applications can be transferred to removable media or between computers without the need of installing them, becoming portable software.
- **Limitations :**
  1. Not all computer programs can be virtualized
  2. Lower performance

2.5 Virtualization Types : Full Virtualization

AU : Dec.-21

- Full Virtualization doesn't need to modify the host OS; it relies upon binary translation to trap and to virtualize certain sensitive instructions.
- Fig. 2.5.1 shows full virtualization.

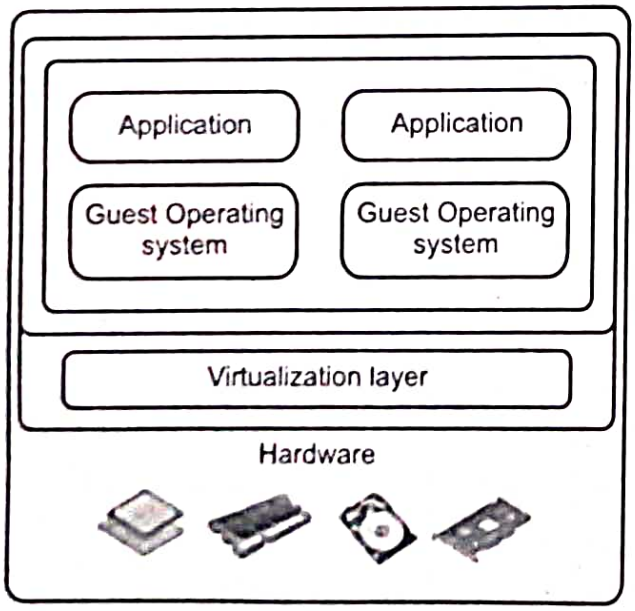


Fig. 2.5.1 Full virtualization

- VMware Workstation applies full virtualization, which uses binary translation to automatically modify x86 software on-the-fly to replace critical instructions
- Normal instructions can run directly on the host OS. This is done to increase the performance overhead - normal instructions are carried out in the normal manner, but the difficult and precise executions are first discovered using a trap and executed in a virtual manner.
- This is done to improve the security of the system and also to increase the performance.

Host based virtualization :

- Virtualization implemented in a host computer rather than in a storage subsystem or storage appliance.
- Virtualization can be implemented either in host computers, in storage subsystems or storage appliances, or in specific virtualization appliances in the storage interconnect fabric.
- The guest OS are installed and run on top of the virtualization layer. Dedicated applications may run on the VMs. Certainly, some other applications can also run with the host OS directly.
- **Advantages** of host-based architecture :
  1. The user can install this VM architecture without modifying the host OS.
  2. The host-based approach appeals to many host machine configurations.

2.5.1 Memory Virtualization

- Memory virtualization features allow abstraction isolation and monitoring of memory on a per Virtual Machine (VM) basis. These features may also make live migration of VMs possible, add to fault tolerance, and enhance security.
- Example features include Direct Memory Access (DMA) remapping and Extended Page Tables (EPT), including their extensions: accessed and dirty bits, and fast switching of EPT contexts.
- The VMkernel manages all machine memory. The VMkernel dedicates part of this managed machine memory for its own use. The rest is available for use by virtual machines.
- Virtual machines use machine memory for two purposes : each virtual machine requires its own memory and the VMM requires some memory and a dynamic overhead memory for its code and data.
- The virtual memory space is divided into blocks, typically 4 kB, called pages. The physical memory is also divided into blocks, also typically 4 kB.

- When physical memory is full, the data for virtual pages that are not present in physical memory are stored on disk. ESX/ESXi also provides support for large pages.
- The VMM is responsible for mapping the guest physical memory to the actual machine memory.
- Each page table of a guest OS has a page table allocated for it in the VMM. The page table in the VMM which handles all these is called a shadow page table.
- As it can be seen all this process is nested and inter-connected at different levels through the concerned address.
- If any change occurs in the virtual memory page table or TLB, the shadow page table in the VMM is updated accordingly.

2.5.2 I/O Virtualization

- I/O Virtualization involves managing of the routing of I/O requests between virtual devices and shared physical hardware.
- The there are three ways to implement this are full device emulation, para-VZ and direct I/O.
- I/O virtualization features facilitate offloading of multi-core packet processing to network adapters as well as direct assignment of virtual machines to virtual functions, including disk I/O.

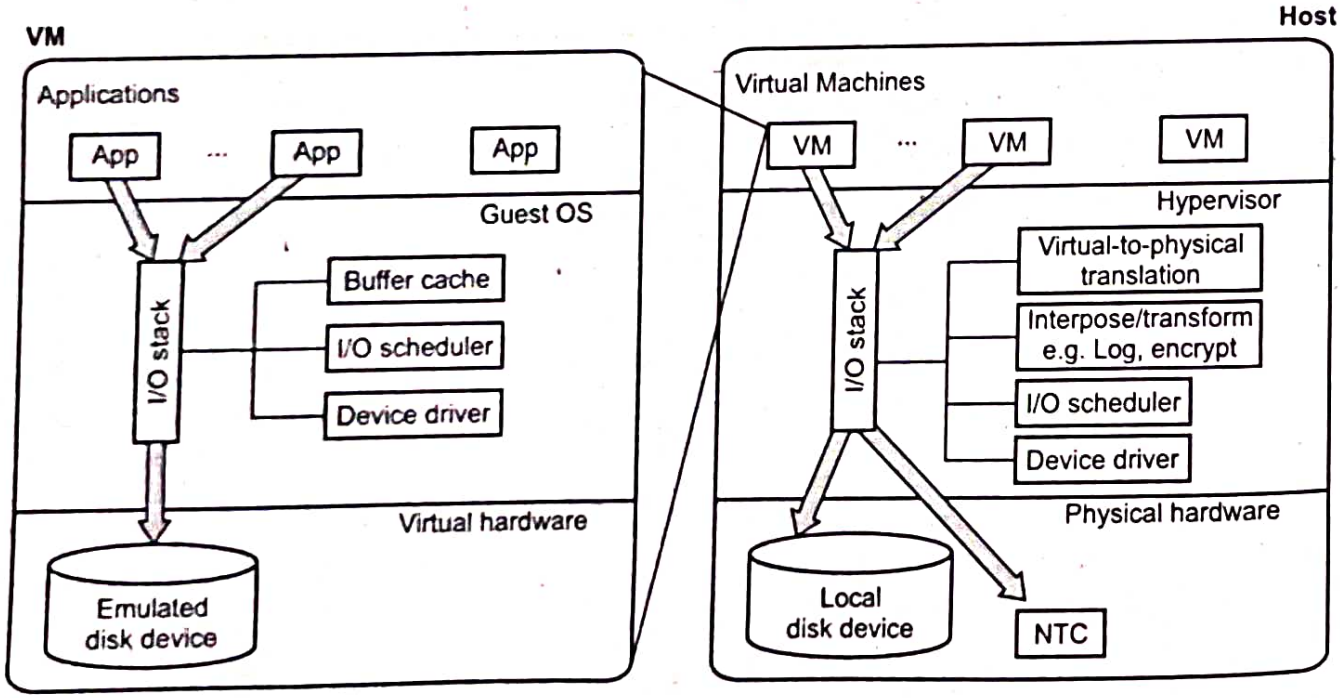


Fig. 2.5.2 I/O virtualization

- Examples include Virtual Machine Device Queues (VMDQ), Single Root I/O Virtualization.
- Fig. 2.5.2 shows I/O virtualization. (See Fig. 2.5.2 on previous page)
  1. **Full Device Emulation :** This process emulates well-known and real-world devices. All the functions of a device or bus infrastructure such as device enumeration, identification, interrupts etc. are replicated in the software, which itself is located in the VMM and acts as a virtual device. The I/O requests are trapped in the VMM accordingly.
  2. **Para-virtualization :** This method of I/O VZ is taken up since software emulation runs slower than the hardware it emulates. In para-VZ, the frontend driver runs in Domain-U; it manages the requests of the guest OS. The backend driver runs in Domain-0 and is responsible for managing the real I/O devices. This methodology (para) gives more performance but has a higher CPU overhead.
  3. **Direct I/O virtualization :** This lets the VM access devices directly; achieves high performance with lower costs. Currently, it is used only for the mainframes.

2.5.3 Difference between Full and Para Virtualization

Sr. No.	Full Virtualization	Para Virtualization
1.	Full Virtualization relies upon binary translation to trap and to virtualize certain sensitive instructions.	Para-Virtualization refers to communication between the guest OS and the hypervisor to improve performance and efficiency.
	Example : VMware	Example : Xen architecture
2.	Full Virtualization doesn't need to modify the host OS.	Para-Virtualization involves modification of OS kernel.
3.	Normal instructions can run directly on the host OS.	Para-virtualized systems won't run on native hardware.
4.	Full Virtualization uses binary translation and direct execution.	Para-Virtualization uses hyper - calls.
5.	Performance is good.	Performance is better in certain cases.
6.	Guest software does not require any modification since the undelying hardware is fully simulated.	Hardware is not simulated and the guest software run their own isolated domains.

2.5.4 Virtualization of CPU

- Certain processors such as Intel VT provide hardware assistance for CPU virtualization.

- When using this assistance, the guest can use a separate mode of execution called guest mode. The guest code, whether application code or privileged code, runs in the guest mode.
- On certain events, the processor exits out of guest mode and enters root mode. The hypervisor executes in the root mode, determines the reason for the exit, takes any required actions, and restarts the guest in guest mode.
- When you use hardware assistance for virtualization, there is no need to translate the code. As a result, system calls or trap-intensive workloads run very close to native speed.
- Some workloads, such as those involving updates to page tables, lead to a large number of exits from guest mode to root mode. Depending on the number of such exits and total time spent in exits, this can slow down execution significantly.
- CPU virtualization features enable faithful abstraction of the full prowess of Intel CPU to a virtual machine.
- All software in the VM can run without any performance, as if it was running natively on a dedicated CPU. Live migration from one Intel CPU generation to another, as well as nested virtualization, is possible.

2.5.5 Binary Translation with Full Virtualization

- This approach relies on binary translation to trap and to virtualize certain sensitive and non-virtualizable instructions with new sequences of instructions that have the intended effect on the virtual hardware. Meanwhile, user level code is directly executed on the processor for high performance virtualization.
- Fig. 2.5.3 shows full virtualization with binary translation.

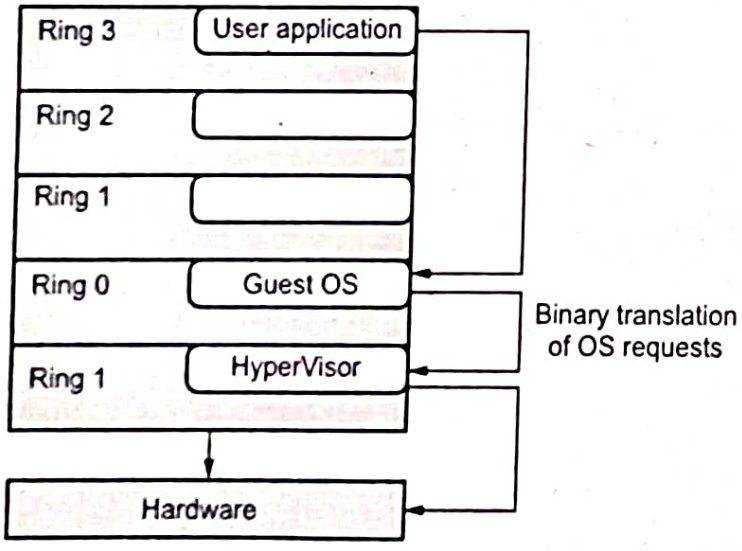


Fig. 2.5.3 Full virtualization with binary translation

- This combination of binary translation and direct execution provides full virtualization as the guest OS is completely decoupled from the underlying hardware by the virtualization layer.
- The guest OS is not aware it is being virtualized and requires no modification.
- The hypervisor translates all operating system instructions at run-time on the fly and caches the results for future use, while user level instructions run unmodified at native speed.
- VMware's virtualization products such as VMWare ESXi and Microsoft Virtual Server are examples of full virtualization.
- The performance of full virtualization may not be ideal because it involves binary translation at run-time which is time consuming and can incur a large performance overhead.

University Question

1. *Outline the problems in virtualizing in CPU, I/O and memory devices and suggest how it could be overridden for efficient utilization of cloud services.* **AU : Dec.-21, Marks 13**

2.6 Two Marks Questions with Answers

Q.1 What is virtualization ?

Ans. : Virtualization is an abstraction layer that decouples the physical hardware from the operating system to deliver greater IT resource utilization and flexibility. It allows multiple virtual machines, with heterogeneous operating systems to run in isolation, side-by-side on the same physical machine. Virtualization means running multiple machines on a single hardware. The "Real" hardware invisible to operating system. OS only sees an abstracted out picture. Only Virtual Machine Monitor (VMM) talks to hardware.

Q.2 What are the benefits of virtualization in the context of cloud computing ?

Ans. :

1. It is possible to achieve a more efficient use of resources.
2. Portability and self-containment also contribute to reducing the costs of maintenance.
3. A virtual execution environment can be configured as a sandbox, thus preventing any harmful operation.

Q.3 List disadvantages of virtualization.

Ans. :

1. Performance degradation.
2. Virtualization can some time lead to an inefficient use of the host.
3. Virtualization opens the door to a new and unexpected form of phishing.

Q.4 What is operating system level virtualization ?

Ans. : Operating-system-level virtualization is a server-virtualization method where the kernel of an operating system allows for multiple isolated user-space instances, instead of just one. Such instances, which are sometimes called containers and software containers.

Q.5 What are hardware virtualization techniques ?

Ans. : This technology allows simulating the hardware interface expected by an operating system. Hardware virtualization allows the coexistence of different software stacks on top of the same hardware. These stacks are contained inside virtual machine instances, which operate in complete isolation from each other.

Q.6 What is application server virtualization ?

Ans. : Application server virtualization abstracts 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.

Q.7 Why operating system level virtualization is required ?

Ans. :

- Operating system level virtualization provides feasible solution for hardware level virtualization issue. It inserts a virtualization layer inside an operating system to partition a machine's physical resources.
- It enables multiple isolated VMs within a single operating system kernel. This kind of VM is often called a virtual execution environment (VE), Virtual Private System (VPS) or simply container .
- From the user's point of view, virtual execution environment look like real servers.
- This means a virtual execution environment has its own set of processes, file system, user accounts, network interfaces with IP addresses, routing tables, firewall rules etc.

Q.8 Define emulation.

Ans. : Emulation is the process of implementing the interface and functionality of one system (or subsystem) on a system (or subsystem) having different interface and functionality.

Q.9 List advantages of OS virtualization.

Ans. :

1. OS virtualization provide least overhead among all types of virtualization solution.
2. They offer highest performance and highest density of virtual environment.
3. Low resource requirements.
4. High Scalability

Q.10 Define I/O virtualization.

Ans. : I/O virtualization involves managing of the routing of I/O requests between virtual devices and shared physical hardware. There are three ways to implement this are full device emulation, para-VZ and direct I/O.

Q.11 What is Type 2 hypervisor ?

Ans. : Type 2 hypervisor is software that runs within an operating system environment. A "guest" operating system thus runs at the third level above the hardware.

Notes

UNIT III

3 Virtualization Infrastructure and Docker

Syllabus

Desktop Virtualization - Network Virtualization - Storage Virtualization - System-level of Operating Virtualization - Application Virtualization - Virtual clusters and Resource Management - Containers vs. Virtual Machines - Introduction to Docker - Docker Components - Docker Container - Docker Images and Repositories.

Contents

3.1 Desktop Virtualization

- Desktop virtualization is a technology that allows the creation and storage of multiple user desktop instances on a single host, residing in a data center or the cloud. It is achieved by using a hypervisor, which resides on top of the host server hardware to manage and allow virtual desktops to utilize the computing power of the underlying server hardware.
- Fig. 3.1.1 shows desktop virtualization.

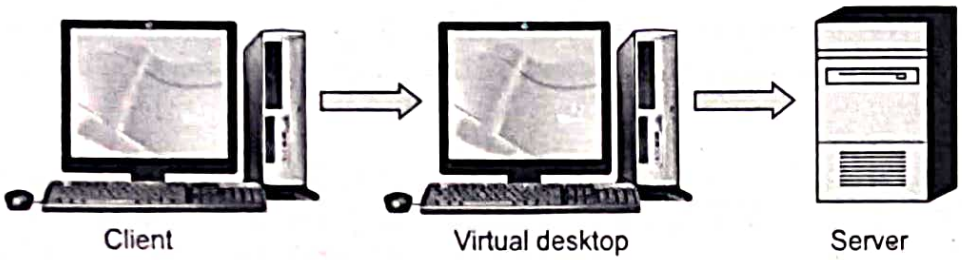


Fig. 3.1.1 Desktop virtualization

- The three most popular types of desktop virtualization are Virtual Desktop Infrastructure (VDI), Remote Desktop Services (RDS) and Desktop-as-a-Service (DaaS).

3.1.1 Types of Desktop Virtualization

1. Virtual desktop infrastructure

- A popular type of desktop virtualization is Virtual Desktop Infrastructure (VDI). VDI uses a VM to deliver persistent or non-persistent virtual desktops to many types of connected devices.
- With a persistent virtual desktop, each user has a unique, dedicated desktop image they can customize with apps and data, knowing the desktop will be saved for future use.
- A non-persistent VDI allows users to access a virtual desktop from an identical pool when they need it. Once the user logs out of a non-persistent VDI, the VDI reverts to its unaltered state.
- Characteristics of VDI :
  i) Virtual desktops live within virtual machines on a centralized server.
  ii) Each virtual desktop includes an operating system image, typically Microsoft Windows.
  iii) The virtual machines are host-based, meaning multiple instances of them can be housed on the same server within the datacenter.

iv) End clients, such as PCs, tablets or thin client terminals, must be constantly connected to the centrally managed server so they can maintain access to the virtualized desktops they're hosting.

v) The connection broker is a software layer that acts as an intermediary between users and virtual resources, which finds a virtual desktop within the resource pool for each client upon successful access of the VDI environment.

- Here are some reasons why VDI is beneficial :
  a) Save money on licensing and individual Workstations/PCs by using thin clients.
  b) Fully secured virtual environment that is fully monitored and managed.
  c) Centralized management and backups.
  d) Secure remote access from anywhere in the world.
  e) Cost reduction for multiple software licenses.
- **Disadvantages :**
  a) If an individual requires different applications from the other users, they will require a completely different image, without changing the applications for other users.
  b) A substantial initial outlay is required for the main server hardware, storage and network infrastructure. This might not be feasible for some smaller businesses.
  c) Administrators, savvy to the limitations, problem solving and installation of VDIs will either have to be brought in or existing IT staff given the relevant training.
  d) If a problem occurs, this will generally affect all users, rather than being able to isolate problems if operating systems run off individual PCs.

2. Remote Desktop Services

- Remote Desktop Services (RDS) or Remote Desktop Session Host (RDSH) are beneficial where only limited applications require virtualization. They allow users to remotely access Windows applications and desktops using the Microsoft Windows server operating system.
- RDS is a more cost-effective solution, since one Windows server can support multiple users.

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

- Desktop-as-a-service (DaaS) is a flexible desktop virtualization solution that uses cloud-based virtual machines backed by a third-party provider. Using DaaS,

organizations can outsource desktop virtualization solutions that help a user to access computer applications and desktops from any endpoint platform or device.

3.1.2 Benefits of Desktop Virtualization

1. Resource utilization : Since IT resources for desktop virtualization are concentrated in a data center, resources are pooled for efficiency.
2. Remote workforce enablement : Since each virtual desktop resides in central servers, new user desktops can be provisioned in minutes and become instantly available for new users to access.
3. VDI offers security improvements compared with running everything locally.

3.2 Network Virtualization

- Network virtualization refers to the technology that enables partitioning or aggregating a collection of network resources and presenting them to various users in a way that each user experiences an isolated and unique view of the physical network.
- Network virtualization creates virtual networks whereby each application sees its own logical network independent of the physical network.
- A virtual LAN (VLAN) is an example of network virtualization that provides an easy, flexible, and less expensive way to manage networks.
- VLANs make large networks more manageable by enabling a centralized configuration of devices located in physically diverse locations.
- Fig. 3.2.1 shows network virtualization.
- Consider a company in which the users of a department are separated over a metropolitan area with their resources centrally located at one office.
- In a typical network, each location has its own network

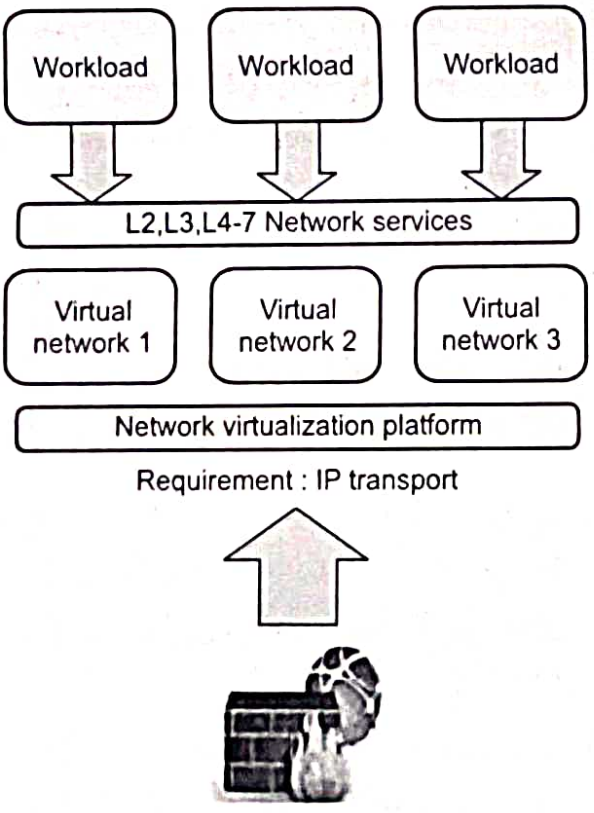


Fig. 3.2.1 Network virtualization

connected to the others through routers. When network packets cross routers, latency influences network performance.

- With VLANs, users with similar access requirements can be grouped together into the same virtual network. This setup eliminates the need for network routing.
- As a result, although users are physically located at disparate locations, they appear to be at the same location accessing resources locally.
- In addition to improving network performance, VLANs also provide enhanced security by isolating sensitive data from the other networks and by restricting access to the resources located within the networks.
- Network virtualization decouples the roles of the traditional Internet Service Providers (ISPs) into Infrastructure Providers (InPs) and Service Providers (SPs)
- Benefits :
  1. Reduces the number of physical devices needed.
  2. Easily segment networks.
  3. Permits rapid change / scalability and agile deployment.
  4. Security from destruction of physical devices.

3.3 Storage Virtualization

- Storage virtualization is a major component for storage servers, in the form of functional RAID levels and controllers. Operating systems and applications with device can access the disks directly by themselves for writing.
- Storage virtualization in cloud computing pools multiple physical storage arrays from Storage Area Networks (SANs) and makes them appear as a single virtual storage device. Virtualization storage separates the storage management software from the underlying hardware infrastructure to provide more flexibility and scalable pools of storage resources.
- Fig. 3.3.1 shows storage virtualization.
- Storage virtualization refers to the abstraction of storage systems from applications or computers. It is a foundation for the implementation of other technologies, such as thin provisioning and data protection, which are transparent to the server.
- Storage virtualization provides the ability to pool storage systems into a consolidated, shared capacity that can be managed from a central point of control.
- Example of storage virtualizations are host-based volume management, LUN creation, tape storage virtualization and disk addressing.
- Storage virtualization has the following characteristics :
  1. The availability of logical volumes separate from physical hard disk constraints.

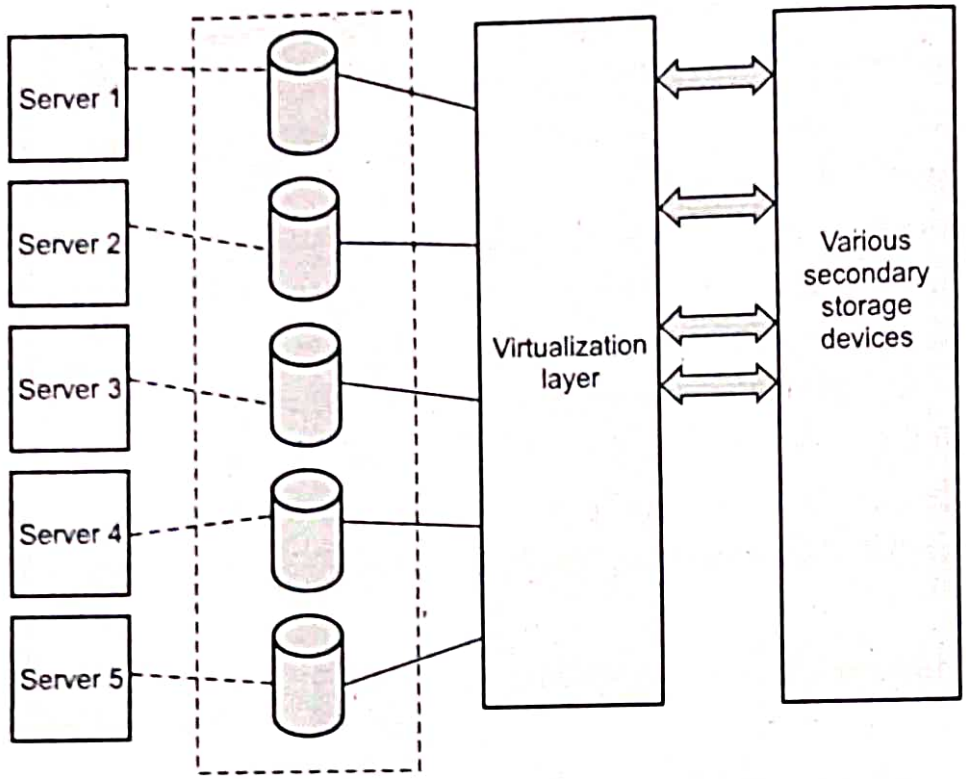


Fig. 3.3.1 Storage virtualization

2. The capability of abstracting multivendor storage devices into one group and reallocating storage space independently of size or physical location.
3. The capability of having automated storage optimization and management.

- Top level servers assigned one virtual volume, which is currently in use by an application. These virtual volumes are mapped to the actual storage in the arrays. When an I/O is sent to a virtual volume, it is redirected through the virtualization at the storage network layer to the mapped physical array.
- Primary types of storage virtualizations are block level virtualization and file virtualization.
- Currently there are three methods of storage virtualization :
  1. Server-based virtualization : This method places a management program on the host system and has the benefit of leveraging the SAN asset as it is.
  2. Fabric-based virtualization : This can be done via network switches or appliance servers. In both instances, independent appliances, such as switches, routers and dedicated servers are placed between servers and storage and have a storage virtualization function. The purpose behind this is to reduce the impact on the existing SAN and servers.
  3. Storage array-based virtualization : This is a virtualization implemented at the storage-system level.

3.3.1 Storage Virtualization Challenges

- Storage virtualization has evolved at a time when data explosion threatened to throw enterprise storage management totally out of gear.
- Traditionally, managing disk storage was once simple : If enterprises needed more space, they got a bigger disk drive. However, as data storage needs grew, multiple disk drives had to be added. Over time technologies such as RAID, network-attached storage and storage-area networks evolved to tackle these storage challenges.
- But managing and maintaining thousands of disk drives presented an even more serious challenge and storage virtualization emerged to tackle these.
  1. Scalability : Ensure storage devices perform appropriate requirements. Each array is managed independently.
  2. Functionality : Virtualized environment must provide same or better functionality. It must be continue to leverage existing functionality on arrays.
  3. Manageability : Virtualization device breaks end-to-end view of storage infrastructure and must integrate existing management tools.
  4. Support : Interoperability in multi-vendor environment.
- A good storage virtualization solution should :
  1. Enhance the storage resources it is virtualizing through the aggregation of services to increase the return of existing assets.
  2. Not add another level of complexity in configuration and management.
  3. Improve performance rather than act as a bottleneck in order for it to be scalable. Scalability is the capability of a system to maintain performance linearly as new resources are added.
  4. Provide secure multi-tenancy so that users and data can share virtual resources without exposure to other users' bad behavior or mistakes.
  5. Not be proprietary, but virtualize other vendor storage in the same way as its own storage to make the management seamless.

3.3.2 Types of Storage Virtualization

- Storage virtualization provides the ability to pool storage systems into a consolidated, shared capacity that can be managed from a central point of control. Virtualization can be implemented in both storage area network and network attached storage.
- Storage virtualization are of two types : Block level and File level.

3.3.3 Block Level Virtualization

- Block level virtualization is used in storage area network. The act of applying virtualization to one or more block-based storage services for the purpose of providing a new block service to clients. Some examples of block virtualization are disk aggregation.
- Block which is used for data storage is progression of bytes and bits and is made up of a proposed length. Data which is aligned in these blocks is called as blocked and inserting data into the block is called blocking.
- Block level storage virtualization provides storage to operating systems and applications in the form of virtual disks. Fig. 3.3.2 Shows block level virtualization.

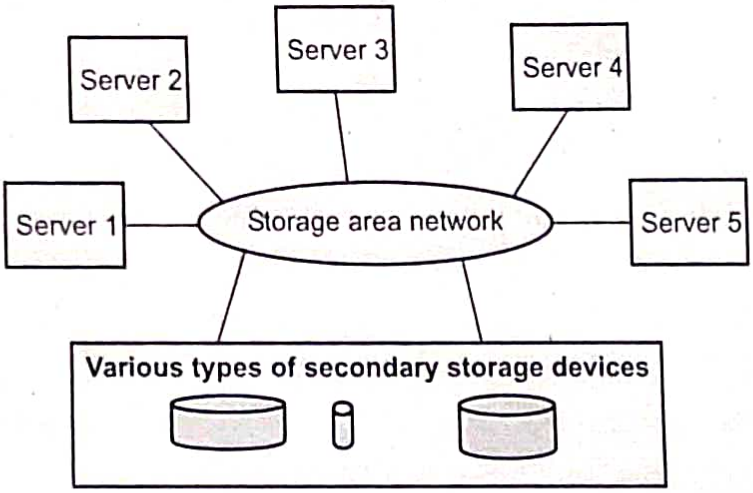


Fig. 3.3.2 Block level virtualization

- There are two types of block level virtualization. One is disk level virtualization, whereby an abstraction process moves data from a physical disk level to a LUN level and is presented as though it were a physical device.
- Another method is storage level virtualization, which, unlike disk level virtualization, hides the physical layer of RAID controllers and disks and hides and virtualizes the entire storage system.
- SCSI commands are transmitted in between the initiator and target. There is no overhead file system like an ext3.
- Block level file system utilizes FC, iSCSI and FCOE protocol.
- Block level file storage is pretty expensive but is very much reliable. It is highly customizable storage and is versatile and speedy.
- Block-level virtualization is usually just called storage virtualization and serves applications such as database software that need block-level access to data. The disks will typically (but not always) reside in Storage Area Network arrays (SANs).

3.3.4 File Level Virtualization

- Network attached storage uses file level virtualization.
- File level storage virtualization provides storage volumes to operating systems and applications in the form of files and directories. Access to storage is through network protocols, such as Common Internet File System and Network File Systems.
- Storage resources and capacity is may be underutilized because files are bound to a specific file server. It is necessary move the file from one server to another server.
- File-level storage is the predominant storage technology used on hard drives, Network-Attached Storage (NAS) systems and similar storage systems. Fig. 3.3.3 shows file level virtualization.

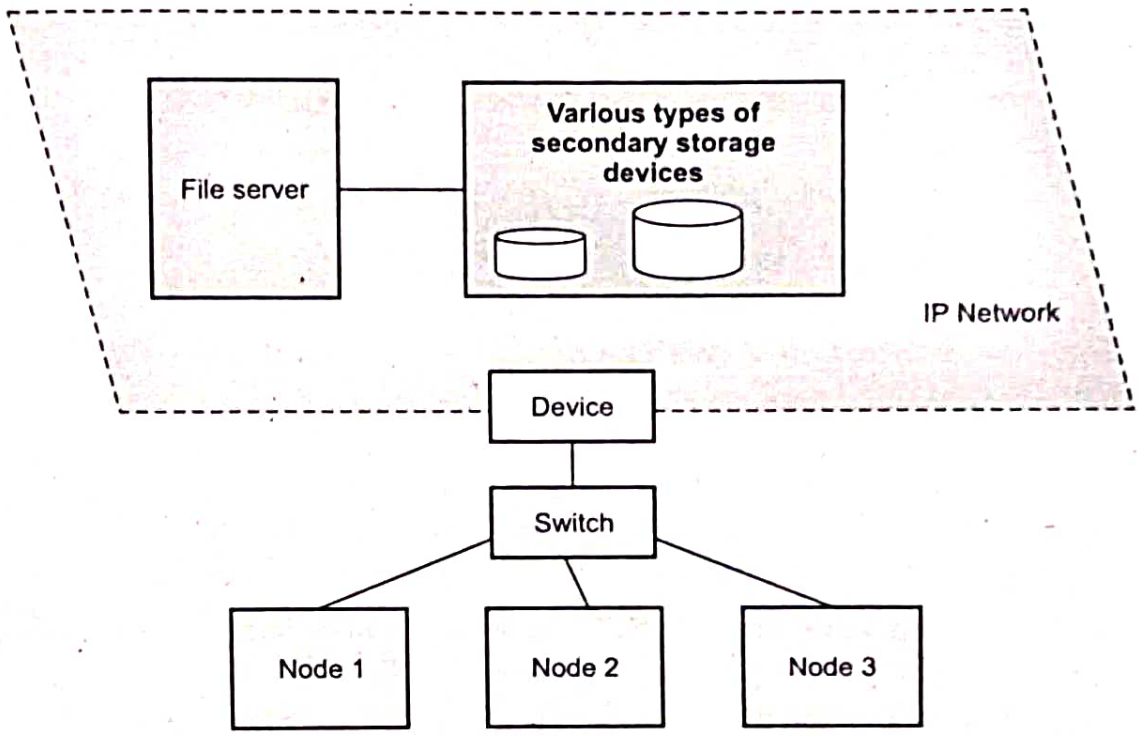


Fig. 3.3.3 File level virtualization

- Moving large number of files in not possible because it requires the server is to be down. Server and some applications need to be reconfigured with the new path. It creates the problem for network administrators for improving the storage efficiency while maintaining the required service level.
- This file of virtualization only simplifies the file mobility. It provides location transparency to user. File level storage works with an ext3 file system. Data is written and read into files, which have variable lengths.

- File level storage will not support of virtual machine file system. It support external boot up, which is essential for ESX and ESXi host servers.
- This type storage cannot handle heavy traffic on the network. Recovery of files is much faster in this level of data storage system. Storage resources and capacity is may be underutilized because files are bound to a specific file server. It is necessary move the file from one server to another server.

3.3.5 Difference between Block Level and File Level Virtualization

Block level	File level
Block-Level virtualization works before the file system exists. It replaces controllers and takes over at the disk level.	The server that uses the storage must have software installed on it in order to enable file-level usage.
It is based on SAN.	It is based on NAS.
Block addresses are used to Read/Write data to the storage media.	Files are accessed by "semantics" instructions. Data inside files is accessed by byte-ranges within the file.
Storage is accessible using fibre channel or iSCSI.	File level storage is usually accessible using common file level protocols such as CIFS and NFS.

3.3.6 Benefits of Storage Virtualization

- Benefits of storage virtualization :
  1. Data is stored in more convenient locations away from the specific host.
  2. The storage devices are able to perform advanced functions like de-duplication, replication, thin provisioning and disaster recovery functionality.
  3. By abstracting the storage level, IT operations can become more flexible in how storage is partitioned, provided and protected.
  4. Improved physical resource utilization.
  5. Lower total cost of ownership : Virtualized storage allows more to be done with the same or less storage.

3.4 System - Level of Operating Virtualization

- Operating - system - level virtualization is a server-virtualization method where the kernel of an operating system allows for multiple isolated user-space instances, instead of just one. Such instances, which are sometimes called containers and software containers.
- This refers to an abstraction layer between traditional OS and user applications.

- This type of virtualization creates isolated containers on a single physical server and the OS instances to utilize the hard-ware and software in data centers.
- Containers behave like real servers. With containers you can create a portable, consistent operating environment for development, testing and deployment.
- This virtualization creates virtual hosting environments to allocates hardware resources among a large number of mutually distrusting users.
- Operating - system - level virtualization usually imposes little to no overhead, because programs in virtual partitions use the operating system's normal system call interface and do not need to be subjected to emulation or be run in an intermediate virtual machine.
- Operating system-level virtualization is not as flexible as other virtualization approaches since it cannot host a guest operating system different from the host one, or a different guest kernel.
- Instead of trying to run an entire guest OS, container virtualization isolates the guests, but doesn't try to virtualize the hardware. Instead, you have containers for each virtual environment.
- With container-based technologies, you'll need a patched kernel and user tools to run the virtual environments. The kernel provides process isolation and performs resource management.

Why operating system level virtualization is required ?

- Operating system level virtualization provides feasible solution for hardware level virtualization issue. It inserts a virtualization layer inside an operating system to partition a machine's physical resources.
- It enables multiple isolated VMs within a single operating system kernel. This kind of VM is often called a virtual execution environment (VE), Virtual Private System (VPS) or simply container.
- From the user's point of view, virtual execution environment look like real servers.
- This means a virtual execution environment has its own set of processes, file system, user accounts, network interfaces with IP addresses, routing tables, firewall rules etc.
- Although VEs can be customized for different people, they share the same operating system kernel. Therefore, OS-level virtualization is also called single-OS image virtualization.

Challenges to cloud computing in OS level virtualization ?

- Cloud computing is transforming the computing landscape by shifting the hardware and staffing costs of managing a computational center to third parties.
- Cloud computing has at least two challenges :
  1. The ability to use a variable number of physical machines and virtual machine instances depending on the needs of a problem. For example, a task may need only a single CPU during some phases of execution but may need hundreds of CPUs at other times.
  2. It is related to slow operation of instantiating new virtual machine. Currently, new virtual machines originate either as fresh boots or as replicates of a template VM, unaware of the current application state. Therefore, to better support cloud computing, a large amount of research and development should be done.

Advantages of OS virtualization :

1. OS virtualization provide least overhead among all types of virtualization solution.
2. They offer highest performance and highest density of virtual environment.
3. Low resource requirements.
4. High Scalability.

Disadvantage of OS virtualization :

1. They support only one operating system as base and guest OS in a single server.
2. It supports library level virtualization.

3.5 Application Virtualization

- Virtualization at the application level virtualizes an application as a VM. On a traditional OS, an application often runs as a process. Therefore, application-level virtualization is also known as process-level virtualization.
- A fully virtualized application is not installed in the traditional sense, although it is still executed as if it were. The application behaves at runtime like it is directly interfacing with the original operating system and all the resources managed by it, but can be isolated to varying degrees.
- Full application virtualization requires a virtualization layer. Application virtualization layers replace part of the runtime environment normally provided by the operating system.

- The layer intercepts all disk operations of virtualized applications and transparently redirects them to a virtualized location, often a single file.
- The application remains unaware that it accesses a virtual resource instead of a physical one. Since the application is now working with one file instead of many files spread throughout the system, it becomes easy to run the application on a different computer and previously incompatible applications can be run side-by-side.
- The most popular approach is to deploy High Level Language (HLL) VMs. Here the virtualization layer sits as an application program on top of the operating system, and the layer exports an abstraction of a VM that can run programs written and compiled to a particular abstract machine definition. Any program written in the HLL and compiled for this VM will be able to run on it.
- Benefits :
  1. Application virtualization uses fewer resources than a separate virtual machine.
  2. Application virtualization also enables simplified operating system migrations.
  3. Applications can be transferred to removable media or between computers without the need of installing them, becoming portable software.
- **Limitations :**
  1. Not all computer programs can be virtualized.
  2. Lower performance.

3.6 Virtual Clusters and Resource Management

- As with traditional physical servers, Virtual Machines (VMs) can also be clustered. A VM cluster starts with two or more physical servers.
- Most virtualization platforms, including XenServer and VMware ESX Server, support a bridging mode which allows all domains to appear on the network as individual hosts. By using this mode, VMs can communicate with one another freely through the virtual network interface card and configure the network automatically.
- Virtual clusters enable admins to deploy, track and manage containers across various systems to ensure performance, security and governance and low costs.
- With many VMs, an inefficient configuration always causes problems with overloading or underutilization.
- Amazon's EC2 provides elastic computing power in a cloud. EC2 permits customers to create VMs and to manage user accounts over the time of their use. Xen Server and VMware ESXi Server support a bridging mode which allows all domains to appear on the network as individual hosts. With this mode VMs can

communicate with one another freely through the virtual network interface card and configure the network automatically.

Physical versus virtual clusters :

- Virtual Clusters are built with VMs installed at one or more physical clusters. The VMs in a virtual cluster are interconnected by a virtual network across several physical networks.

Virtual cluster features :

a) Virtual machines can be restarted on other hosts if the host where the virtual machine running fails.

b) Distributed Resource Scheduler : Virtual machines can be load balanced so that none of the hosts is too overloaded or too much empty in the cluster.

c) Live migration : Of virtual machines from one host to other.

- Fig. 3.6.1 shows cloud platform with virtual cluster.

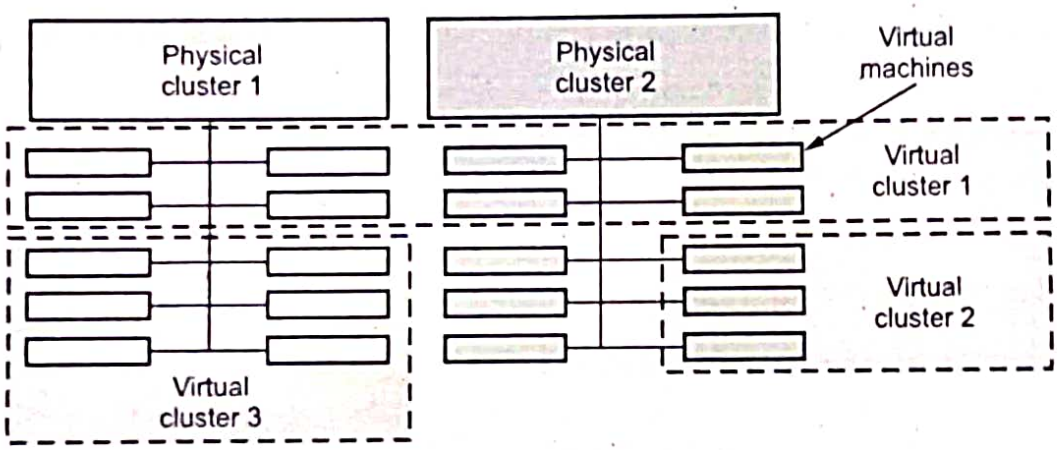


Fig. 3.6.1 Cloud platform example with three virtual clusters over two physical clusters

- The provisioning of VMs to a virtual cluster is done dynamically and they have the following properties :

a) Virtual cluster nodes can be either physical or virtual with different operating systems.

b) VM runs with a guest OS that manages the resources in the physical machine.

c) The purpose of using VMs is to consolidate multiple functionalities on the same server.

d) VMs can be replicated in multiple servers to promote parallelism, fault tolerance and disaster discovery.

e) The no. of nodes in a virtual cluster can grow or shrink dynamically.

f) The failure of some physical nodes will slow the work but the failure of VMs will cause no harm.

Characteristics virtual cluster :

1. Virtual machine or physical machine is used as virtual cluster nodes. Multiple VM running with different types of OS can be deployed on the same physical node.
2. Virtual machine runs with guest operating system. Host OS and VM OS are different but it manages the resources in the physical machine.
3. Virtual machine can be replicated in multiple servers and it support distributed parallelism, fault tolerance and disaster recovery.
4. Number of nodes of a virtual cluster may change accordingly.
5. If Virtual machine failes, it can not affect the host machine.

- **Virtual cluster is managed by four ways :**
  1. We can use a guest-based manager, by which the cluster manager resides inside a guest OS. Ex. : A Linux cluster can run different guest operating systems on top of the Xen hypervisor.
  2. We can bring out a host-based manager which itself is a cluster manager on the host systems. Ex. : VMware HA (High Availability) system that can restart a guest system after failure.
  3. An independent cluster manager, which can be used on both the host and the guest - making the infrastructure complex.
  4. Finally, we might also use an integrated cluster (manager), on the guest and host operating systems; here the manager must clearly distinguish between physical and virtual resources.

3.6.1 Virtualization in Disaster Recovery

- Data is a prime asset for all business organizations and it needs to be protected from getting lost, hacking, phishing and identity theft. Virtualization is the process of producing a virtual version of a system, software or even a working environment rather than a physical counterpart, as defined by the definition.
- Disaster Recovery (DR) relies upon the replication of data and computer processing in an off-premises location not affected by the disaster. When servers go down because of a natural disaster, equipment failure or cyber attack, a business needs to recover lost data from a second location where the data is backed up.
- With a disaster recovery plan, you can organize the actions to take in case of any disaster or incident. This will speed up the response time and minimize downtime.
- **Reducing downtime :** Virtualization software allows businesses to create image-based backups of their virtual machines. This means that in the event of a disaster, businesses can restore their systems quickly, rather than hours or days

needed to rebuild systems from scratch. Therefore, we can reduce and even eliminate downtime. Just access the data from another device to keep working.

- **Create off-site backups :** When we create business backups, we must consider having at least one copy of files on an off-site backup. This will allow us to rapidly recover files if anything helps with our local business data and storage device. When server can have issues, but power outages, fires and other natural disasters can also affect the place. Virtualization can automatically send backup files to off-site backup device.
- **Recover data from failed drives :** If hard drive dies or RAID fails, data virtualization system can help us. Since it is cloud storage, virtualization can keep copies of files even if storage devices stop working.
- **Test disaster recovery plans :** Virtualization can create a test environment on the system for testing the disaster recovery plan whenever required. This allows businesses to ensure that their disaster recovery plan is effective and can be executed efficiently when required.
- **Duplicate data for remote access :** With virtualization, businesses can duplicate their data in real-time or at specific intervals to a remote site. This allows remote users to access their data, applications and systems in the event of a disaster.

3.7 Introduction to Docker

- Docker is quickly changing the way that organizations are deploying software at scale.
- Docker is a tool that promises to easily encapsulate the process of creating a distributable artifact for any application, deploying it at scale into any environment, and streamlining the workflow and responsiveness of agile software organizations.
- **Benefits :**
  1. Packaging software in a way that leverages the skills developers already have.
  2. Bundling application software and required OS file systems together in a single standardized image format.
  3. Abstracting software applications from the hardware without sacrificing resources.

3.7.1 Process Simplification

- Docker can simplify both workflows and communication and that usually starts with the deployment story.
- Fig. 3.7.1 shows workflow with and without docker.

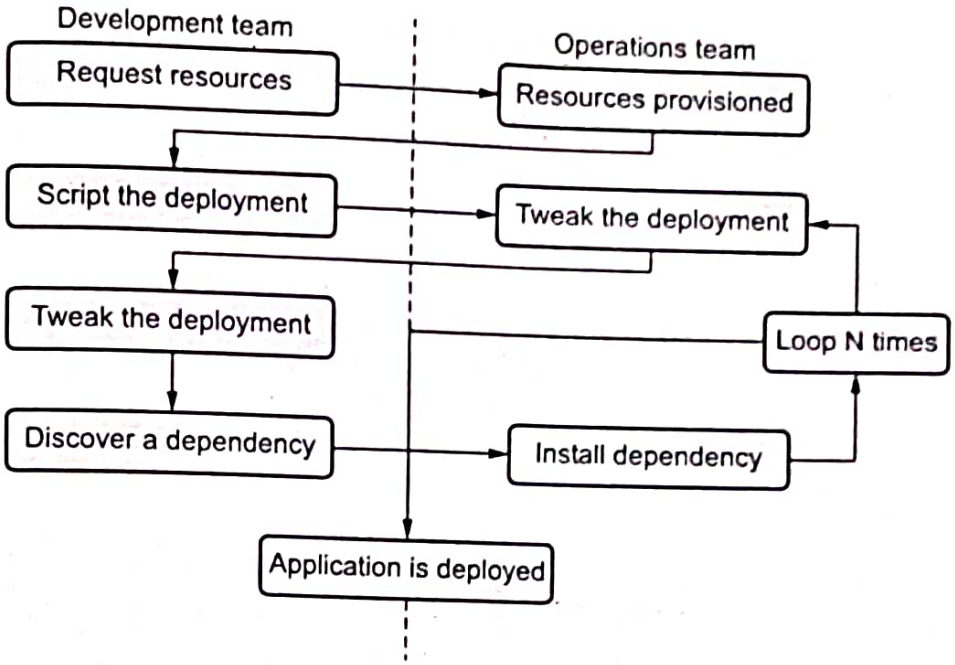


Fig. 3.7.1 Traditional deployment workflow (without docker)

1. Application developers request resources from operations engineers.
2. Resources are provisioned and handed over to developers.
3. Developers script and tool their deployment.
4. Operations engineers and developers tweak the deployment repeatedly.
5. Additional application dependencies are discovered by developers.
6. Operations engineers work to install the additional requirements.
7. Go to step 5 and 6.
8. The application is deployed.

- Fig. 3.7.2 shows Docker deployment workflow.

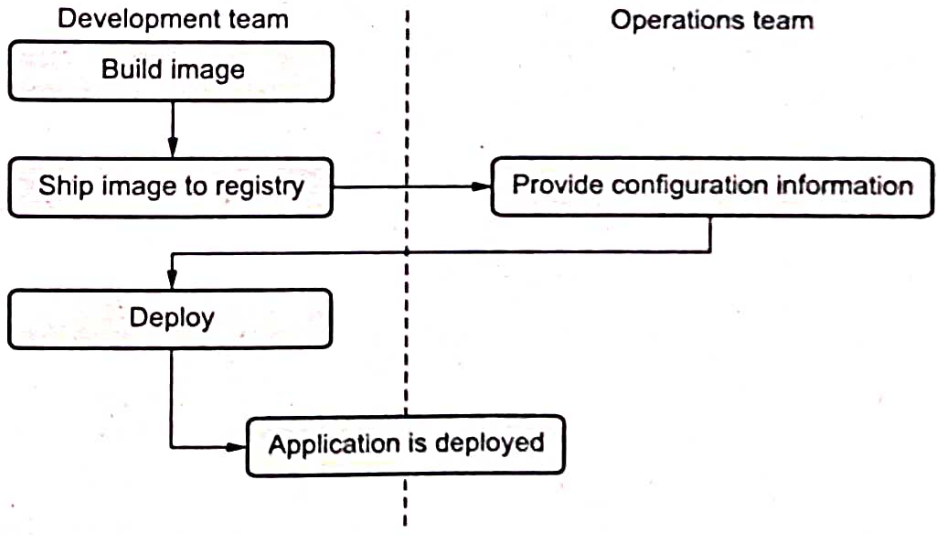


Fig. 3.7.2 Docker deployment workflow

1. Developers build the Docker image and ship it to the registry.
2. Operations engineers provide configuration details to the container and provision resources.
3. Developers trigger deployment.

3.7.2 Broad Support and Adoption

- Docker is increasingly well supported, with the majority of the large public clouds. For example, Docker runs on AWS Elastic Beanstalk, Google AppEngine, IBM Cloud, Microsoft Azure, etc.
- Google's Eric Brewer announced that Google would be supporting Docker as its primary internal container format. Rather than just being good PR for these companies, what this means for the Docker community is that there is starting to be a lot of money backing the stability and success of the Docker platform.
- When docker released their libswarm development library at docker-Con 2014, an engineer from Orchard demonstrated deploying a docker container to a heterogeneous mix of cloud providers at the same time.
- The Docker-client runs directly on most major operating systems, but because the Docker server uses Linux containers, it does not run on non-Linux systems.
- Docker has traditionally been developed on the Ubuntu Linux distribution, but today most Linux distributions and other major operating systems are now supported where possible.

3.7.3 Architecture

- The fundamental architecture of Docker is a simple client - server model, with only one executable that acts as both components, depending on how you invoke the docker command.
- Underneath those simple exteriors, Docker heavily leverages kernel mechanisms such as IPTABLES, virtual bridging, cgroups, namespaces and various filesystem drivers.
- Fig. 3.7.3 shows docker architecture.

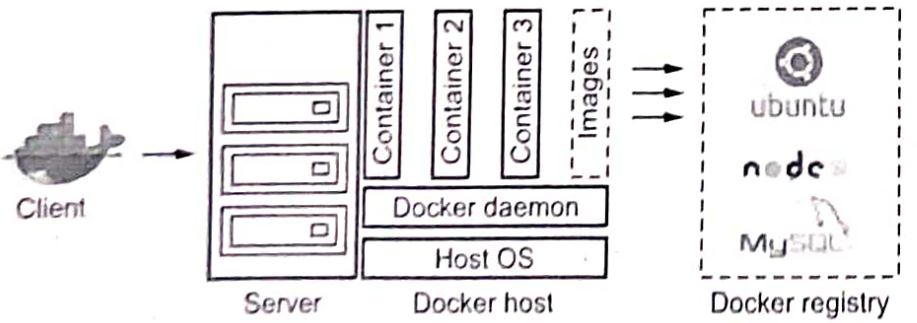


Fig. 3.7.3 Docker architecture

- It consists of two parts : The client and the server. Registry is one more components which stores docker images and metadata about those images.
- Docker Engine is a client-server based application with following components -
  1. A server which is a continuously running service called a **daemon process.**
  2. A REST API which interfaces the programs to use talk with the daemon and give instruct it what to do.
  3. A command line interface client.
- Docker client is the primary service using which docker users communicate with the docker. When we use commands "docker run" the client sends these commands to dockerd, which execute them out.

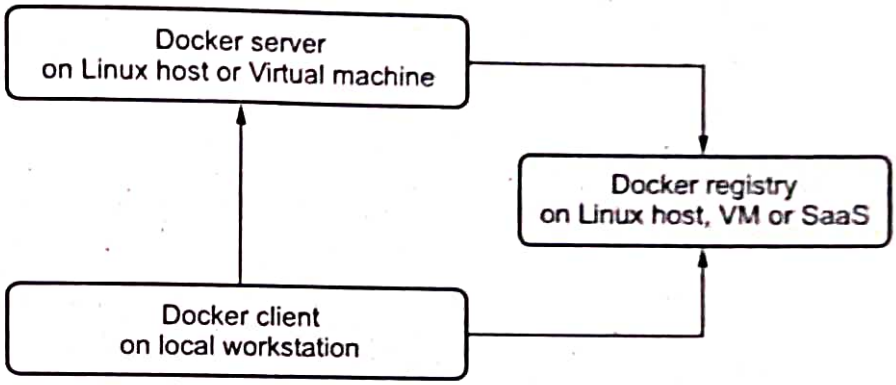


Fig. 3.7.4 Data flow

- The command used by docker depend on docker AP. In docker client can interact more than one daemon process.
- The docker images are building the block of docker or docker image is a read-only template with instructions to create a docker container. Docker images are the most build part of docker life cycle.
- The server does the ongoing work of running and managing your containers, and you use the client to tell the server what to do.
- The docker daemon can run on any number of servers in the infrastructure and a single client can address any number of servers.
- Clients drive all of the communication, but docker servers can talk directly to image registries when told to do so by the client.
- Clients are responsible for directing servers what to do and servers focus on hosting containerized applications.
- Docker registry keeps docker images. We can run our private registry.
- When we run the docker pull and docker run commands, the required images are pulled from our configured registry directory.

- Using docker push command, the image can be uploaded to our configured registry directory.

3.7.4 Container and Kubernetes

- A container image is a ready-to-run software package that includes everything a program needs to execute, including the code and any run-times it needs, application and system libraries and default values for any important settings.
- Container orchestration is concerned with the management of container lifecycles, particularly in large, dynamic environments. Container orchestration is used by software teams to control and automate a variety of tasks on container management.
- Container orchestration works in any context where containers are employed. It can assist you in deploying the same program across several environments without having to rewrite it.
- Kubernetes is an open-source container management platform that unifies a cluster of machines into a single pool of compute resources. With kubernetes, you organize your applications in groups of containers, which it runs using the Docker engine, taking care of keeping your application running as you request.
- Kubernetes is an open source container orchestration platform that automates many of the manual processes involved in deploying, managing, and scaling containerized applications.
- Kubernetes was originally developed and designed by engineers at Google.
- The primary responsibility of kubernetes is container orchestration. That means making sure that all the containers that execute various workloads are scheduled to run physical or virtual machines.
- The containers must be packed efficiently following the constraints of the deployment environment and the cluster configuration. In addition, kubernetes must keep an eye on all running containers and replace dead, unresponsive or otherwise unhealthy containers.
- Kubernetes uses docker to run images and manage containers.
- Kubernetes allows several containers to work in harmony, reducing operational burden. Interestingly, this includes docker containers. Kubernetes can be integrated with the docker engine and uses "Kubelets" to coordinate the scheduling of docker containers.
- The docker engine runs the container image, which is created by running docker build. The higher-level concepts (load balancing, service discovery and network policies) are controlled by kubernetes. When combined, both docker and

kubernetes can develop a modern cloud architecture. However, it should be remembered the two systems, at their core, are fundamentally different.

- Fig. 3.7.5 shows kubernetes architecture.

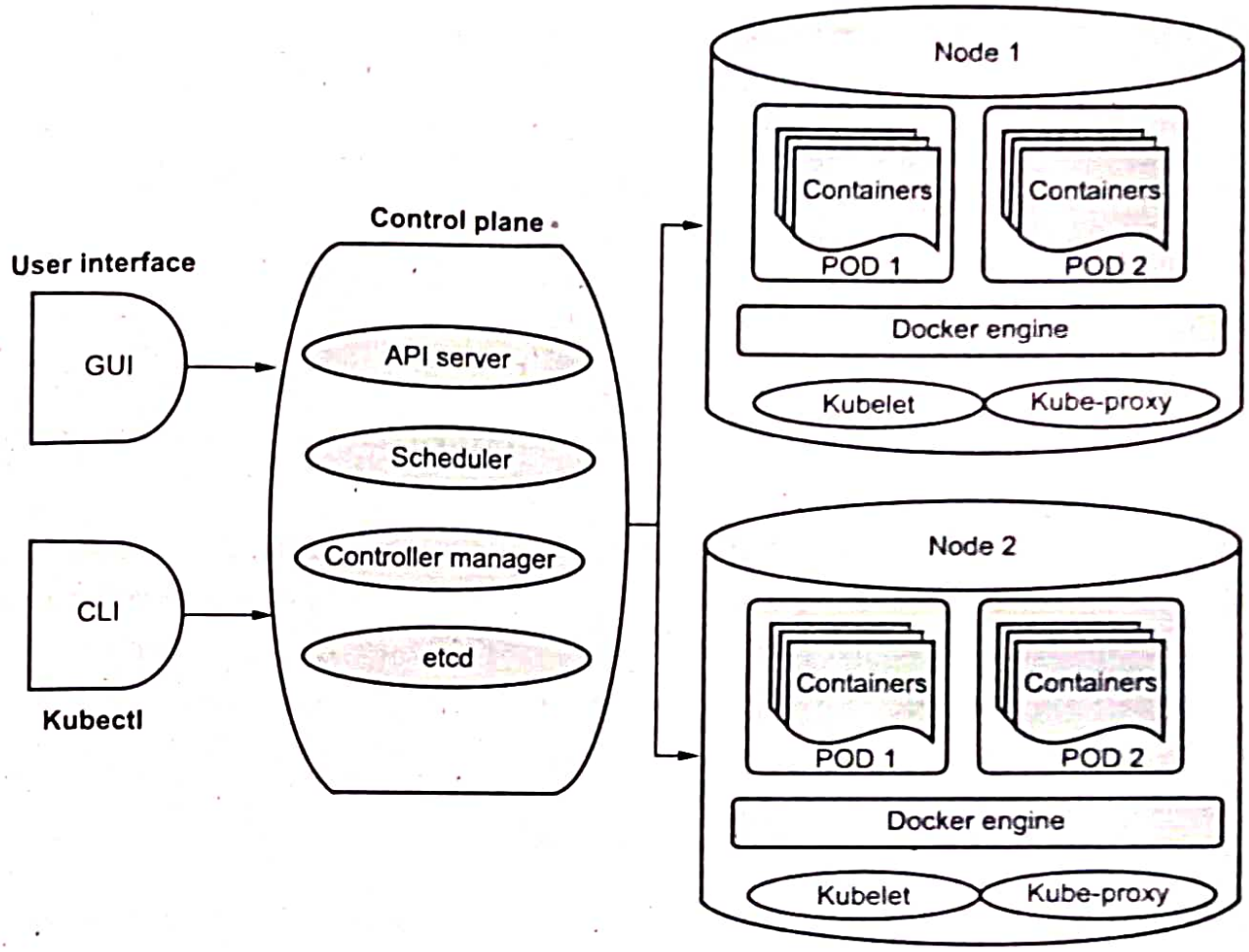


Fig. 3.7.5 Kubernetes architecture

- **Kubelet :** This function runs on nodes, reads container manifests, and assures defined containers have started and are running.
- **Node :** These perform the assigned tasks, with the kubernetes master controlling them.
- **Master :** This controls the kubernetes nodes and is the source of all task assignments.
- **Pod :** When one or more containers are deployed to one node. Containers in a pod will share a host name, an IP address, IPC and other resources.
- **Replication controller :** Controls the number of "identical" copies in a pod that should be running in different locations on the cluster.
- **Service :** This will decouple the work definitions from the pods. Service requests are automatically sent to the right pod, regardless of location.

- **Kubectl :** The primary configuration tool for kubernetes.
- **Kubernetes objects :** These are persistent entities within the Kubernetes system. They are used to represent the state of the cluster.

3.8 Two Marks Questions with Answers

Q.1 What is storage virtualization in cloud computing ?

Ans. : Storage virtualization in cloud computing is sharing physical storage into multiple storage devices that appear as a single virtual device.

Q.2 What is networking virtualization ?

Ans. : Network virtualization refers to the technology that enables partitioning or aggregating a collection of network resources and presenting them to various users in a way that each user experiences an isolated and unique view of the physical network. Network virtualization creates virtual networks whereby each application sees its own logical network independent of the physical network.

Q.3 What is virtual desktop infrastructure ?

Ans. : Virtual desktop infrastructure is a term that refers to using a virtualized desktop that is hosted on a virtual machine that lives on a server.

Q.4 What are the three key components of virtual desktop infrastructure ?

Ans. : Three key components of virtual desktop infrastructure are host, connection broker and end points.

Q.5 What is cloud analytics ?

Ans. : Cloud analytics is a type of cloud service model where data analysis and related services are performed on a public or private cloud. Cloud analytics can refer to any data analytics or business intelligence process that is carried out in collaboration with a cloud service provider.

Q.6 What is file level storage virtualization ?

Ans. : File level storage virtualization provides storage volumes to operating systems and applications in the form of files and directories. Access to storage is through network protocols, such as common Internet file system and network file systems. storage resources and capacity is may be underutilized because files are bound to a specific file server. It is necessary move the file from one server to another server.

❑❑❑

UNIT IV

4 Cloud Deployment Environment

Syllabus

Google App Engine - Amazon AWS - Microsoft Azure; Cloud Software Environments - Eucalyptus - OpenStack.

Contents

4.1 Google App Engine

AU : Dec.-21

- Google App Engine (GAE) is a Platform as a Service cloud computing platform for developing and hosting web applications in Google-managed data centers.
- Google App Engine is a way to write your own web applications and have them hosted on Google servers. It enables developers to build their web applications on the same scalable system that power Google applications.
- An app is a piece of software which can run on the computer, internet, phone or any other electronic device. Google refers to their online services as Apps. They also sell a specific suite of services known as Google Apps.
- Google's providing both SaaS and PaaS solutions in cloud computing. Some of the examples for SaaS solutions including Google Apps which including Gmail, Doc, etc. and PaaS includes Google App engine.
- Services provided by App engine includes :
  a) Platform as a Service (PaaS) to build and deploy scalable applications.
  b) Hosting facility in fully-managed data centers.
  c) A fully-managed, flexible environment platform for managing application server and infrastructure.
  d) Support in the form of popular development languages and developer tools.
- **Major feature of Google App Engine :**
  1. Automatic scaling and load balancing.
  2. Authentication using Google Accounts API.
  3. Provides dynamic web services based on common standards.
  4. Integration with other Google Cloud Services and API.
  5. Support persistent storage, with query access sorting and transaction management features.
- Google App engine offers users the ability to build and host web applications on Google's infrastructure.

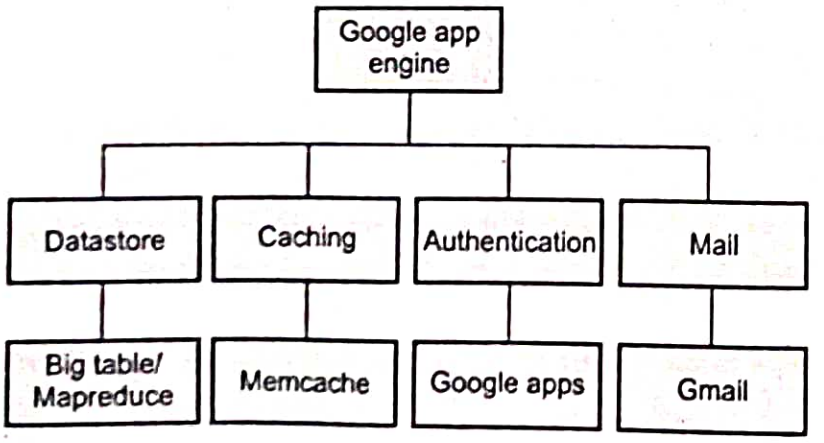


Fig. 4.1.1

- The App Engine offers a number of services that enable you to perform several common operations when managing your application. The following APIs are available to access these services :
  1. **Mail :** Using the mail API, the developers can send email messages.
  2. **Memcache :** The Memcache service gives the users the benefit of working efficiently by providing high retrieval speed, even when multiple users access the same application at the same instance of time.
  3. **Image manipulation :** The Image service allows you to manipulate images of your application. With the use of this API, you can resize, crop, rotate and flip images in JPEG and PNG formats.
- In the PaaS space Google is a key player. App Engine is a platform to create, store and run applications on Google's servers using development languages as java and python.
- App Engine includes tools for managing the data store, monitoring the site and its resource consumption and debugging and logging. A user can serve the app from his own domain name using Google Apps.
- **Key features of GAE programming mode using java and python.**
- The Google App engine Software Development Kit (SDK) provides Java and Python programming languages.
- The languages have their own web server application that contains all Google App Engine services on a local computer. The web server also simulates a secure sandbox environment.
- The Google App engine SDK has APIs and libraries including the tools to upload applications. The architecture defines the structure of applications that run on the Google App engine.

1. Python :

- The Google App engine allows implementation of applications using python programming language and running them on its interpreter.
- The Google App engine provides rich APIs and tools for designing web applications, data modeling, managing, accessing apps data, support for mature libraries and frameworks like Django.
- The main characteristics of Google App engine are its DataStore, configuration file app.yaml and how it serves an application.

2. Java :

- The Google App engine provides tools and APIs required for the development of web applications that run on the Google App engine Java run time.

- The application interacts with the environment using servlets and web technologies like Java Server Pages (JSPs) which can be developed using Java6.
- The GAE environment uses Java SE Runtime JRE platform 6 and libraries which the applications can access using APIs.
- Java SDK has implementations for Java Data Objects (JDO) and Java Persistence (JPA) interface.
- To exchange email messages with Google App engine, it provides the Google App Engine mail service through the Java Mail API.
- Support for other languages like JavaScript, Ruby or Scalar is also provided by Google App engine with the use of JVM compatible compilers and interpreters.
- When Google App engine gets a web request that corresponds to the URL mentioned in the applications deployment descriptor it invokes a servlet corresponding to the request and uses Java Servlets API to provide requested data and accepts response data.
- Google App engine makes it easy to build an applications that runs reliably, even under heavy load and with large amounts of data.
- App engine includes the below features :
  a) Dynamic web serving, with full support for common web technologies.
  b) Persistent storage with queries, sorting and transactions.
  c) Automatic scaling and load balancing.
  d) APIs for authenticating users and sending email using Google accounts.
  e) Scheduled tasks for triggering events at specified times and regular intervals.

University Question

1. *Write detailed steps to set the google app engine environment for executing any program of your choice.* **AU : Dec.-21, Marks 13**

4.2 Amazon AWS

- Amazon Web Services (AWS) is a cloud computing platform from Amazon that provides customers with a wide array of cloud services.
- Amazon first debuted its Amazon Web Services in 2006 as a way to enable the use of online services by client-side applications or other web sites via HTTP, REST or SOAP protocols.
- Amazon bills customers for Amazon AWS based on their usage of the various Amazon Web Services.

- In 2012, Amazon launched the AWS Marketplace to accommodate and grow the emerging ecosystem of AWS offerings from third-party providers that have built their own solutions on top of the Amazon Web Services platform.
- The AWS Marketplace is an online store for Amazon Web Services customers to find, compare and begin using AWS software and technical services.
- Amazon Web Services is a secure cloud services platform, offering compute power, database storage, content delivery and other functionality to help businesses scale and grow.
- In 2017, AWS comprised more than 90 services spanning a wide range including computing, storage, networking, database, analytics, application services, deployment, management, mobile, developer tools and tools for the Internet of Things.
- Today, Amazon Web Services provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world.
- In 2016 AWS partnered with Digital Currency Group to create a laboratory environment allowing companies to experiment with block chain technologies.
- In January 2018, Amazon launched an autoscaling service on AWS.

What is Amazon Web Services ?

- Amazon Web Services (AWS) is a collection of remote computing services (web services) that together make up a cloud computing platform, offered over the Internet by Amazon.com.
- The AWS Cloud infrastructure is built around Regions and Availability Zones (AZs). A Region is a physical location in the world where we have multiple AZs. AZs consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.
- These AZs offer you the ability to operate production applications and databases that are more highly available, fault tolerant, and scalable than would be possible from a single data center.
- The AWS cloud operates 42 AZs within 16 geographic regions around the world, with five more availability zones and two more regions coming online in 2017.
- Each availability zone is designed as an independent failure zone. This means that availability zones are physically separated within a typical metropolitan region and are located in lower risk flood plains.

4.2.1 Components

- AWS consists of many cloud services that you can use in combinations tailored to your business or organizational needs.
- With Amazon Web Services you will find a complete cloud platform ready to use for virtually any workload.
- The user requests to the server by the method such as e-mail either to register or to transfer the domain.
- Your request which includes all information will be sent to Amazon API Gateway restful service.
- API Gateway will transfer the collected user information to an AWS Lambda function.
- AWS Lambda function will generate an e-mail and forward it to the 3rd party mail server using Amazon SES.
- Components of Amazon Web Service architecture are Amazon API Gateway, AWS Lambda, Amazon Simple Email Service.
- API Gateway is a front-door to access data, business logic and functionality. API Gateway will provide a restful API endpoint for our AWS Lambda function.
- API works at small as well as large-scale and helps developers to manage, spectator, create and provide security to the API's.

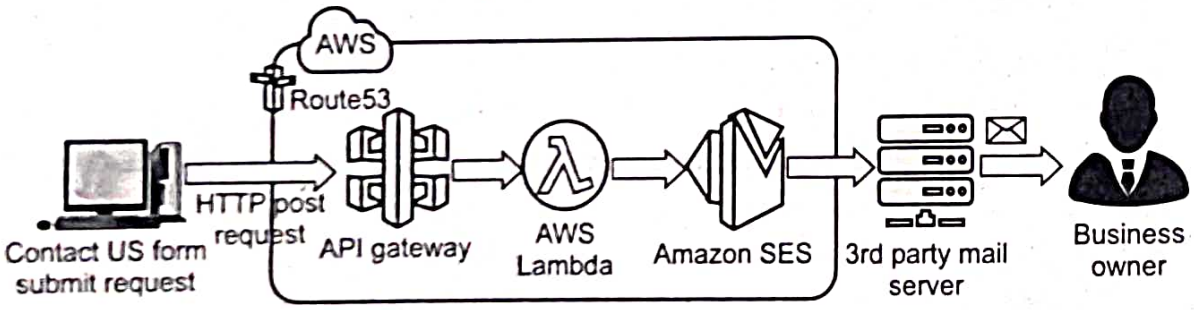


Fig. 4.2.1 AWS

- AWS Lambda is a compute service that runs your back-end code and responds to events such as object uploads to Amazon S3 bucket, Dynamo DB or in-app activity. The Lambda function will get all the information from a user through API Gateway.
- Amazon Simple email service helps us to send e-mail with minimal setup and maximum deliverability. It is integrated with AWS management console so that you can monitor your sending activity. Amazon Simple Email Service helps us by monitoring insecurity.

4.2.2 Advantages and Disadvantages of AWS

Advantages :

1. Easy to use.
2. No capacity limits : Organizations launch different projects and the guess what capacity they will need.
3. Provides speed and agility.
4. Secure and reliable : AWS provides security and also helps to protect the privacy as it is stored in AWS data centers.

Disadvantages :

1. Limitations of Amazon EC2 : AWS sets default limits on resources which vary from region to region. These resources consist of images, volumes and snapshots.
2. Technical support fee : AWS charges you for immediate support.
3. Security limitations.

4.2.3 Compute Service

- Compute services contains the fundamental element of cloud computing systems. Example of compute service is Amazon EC2.
- Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers and system administrators.
- Amazon EC2 reduces the time required to obtain and boot new server instances (called Amazon EC2 instances) to minutes, allowing user to quickly scale capacity, both up and down, as your computing requirements change.
- EC2 allows creating Virtual Machines (VM) on-demand. Pre-configured template Amazon Machine Image (AMI) can be used get running immediately. Creating and sharing your own AMI is also possible via the AWS Marketplace.

4.2.3.1 Amazon Machine Image

- Amazon Machine Image (AMI) is a template for software configuration (Operating System, Application Server and Applications).
- Machine imaging is a process that is used to provide system portability and provision and deploy systems in the cloud through capturing the state of systems using a system image.

- A system image makes a copy or a clone of the entire computer system inside a single file. The image is made by using a program called system imaging program and can be used later to restore a system image
- An AMI typically contains three things : Template, permission to launch, block device mapping.
  1. **Template :** For the root volume for the instances (An application server, an OS, and applications)
  2. **Permissions to launch :** Which account can use this AMI to launch instances.
  3. **Block device mapping :** That specifies the volumes to attach to the instance during its launch time.
- AMIs provide a template for the root volume required to launch a particular instance. This will typically include the operating systems, an application server and applications.
- It also includes in the AMI are launch permissions that restrict the ability to launch instances from that AMI to defined AWS accounts. Finally, a block device mapping specifies the volumes to attach to the instance once it is launched.
- Once an AMI has been created and registered, it can be used to launch new instances. An AMI can be copied to different regions, and it can also be deregistered. Fig. 4.2.2 shows an AMI lifecycle.

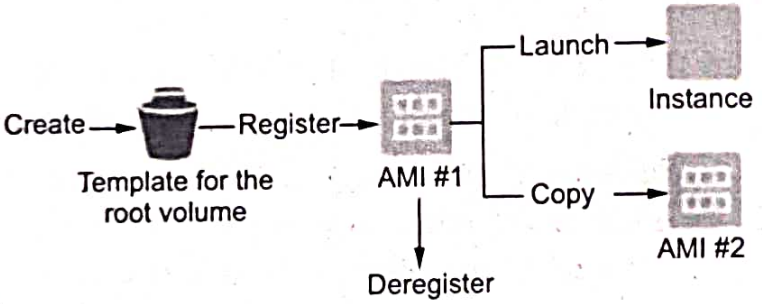


Fig. 4.2.2 AMI lifecycle

- Once an AMI is created, it is stored in an S3 bucket and the user can decide whether to make it available to other users or keep it for personal use.
- Instance is AMI running on virtual servers in the cloud. Each instance type offers different compute and memory facilities. Create an Amazon Machine Image containing your applications, libraries, data and associated configuration settings. Or use pre-configured, templated images to get up and running immediately.
- Auto scaling allows automatically scale of the capacity up seamlessly during demand spikes to maintain performance and scales down during demand lulls to minimize costs.
- Elastic load balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. It provides tools to build failure resilient applications by launching application instances in separate availability zones.

- AMIs can be attained directly from AWS, can be created and shared through communities or can be purchased from vendors via the AWS Marketplace.
- Pay only for resources actually consume, instance-hours. VM Import/Export enables user to easily import virtual machine images from existing environment to Amazon EC2 instances and export them back at any time.
- Boto is a Python package that provides programmatic connectivity to Amazon Web Services.
- The AMI files are encrypted and compressed for security purpose and stored in Amazon S3 (Simple Storage System) buckets as a set of 10 MB chunks.
- Machine imaging is mostly run on virtualization platform due to this it is also called as virtual appliances and running virtual machines are called instances.
- The AMI file system is not a standard bit-for-bit image of a system that is common to many disk imaging programs. AMI omits the kernel image and stores a pointer to a particular kernel that is part of the AWS kernel library.
- Among the choices are Red Hat Linux, Ubuntu, Microsoft Windows, Solaris and others. Files in AMI are compressed and encrypted and an XML file is written that describes the AMI archive.
- Machine images are sometimes referred to as "virtual appliances", systems that are meant to run on virtualization platforms.

4.2.3.2 EC2 Instances

- Amazon Elastic Compute Cloud (Amazon EC2) instances represent virtual machines. EC2 instances are launched by created by an Amazon Machine Image (AMI). An AWS template that describes and defines the OS and operating environment for one or more EC2 instances of one or more EC2 instance types.
- Each instance type delivers a mix of CPU, memory, storage and networking capacity, across one or more size options and should be carefully matched to your workload's unique demands.
- EC2 functions :
  1. Load variety of operating system.
  2. Install custom applications.
  3. Manage network access permission.
  4. Run image using as many/few systems as customer desire.
- Currently available configurations for EC2 instances are as follows :
  1. **Standard instances :** Among the most popular and widely used EC2 instance types. Standard instances have memory to CPU ratios suitable for most

general-purpose applications. General-purpose instances include A1, M5, M5a, M4, T3, T3a and T2.

2. **Micro instances :** Micro instances can be used for small Web applications with limited traffic.
3. **High-memory instances :** EC2 high memory instances offer 6, 9, 12, 18, and 24 TB of memory in an instance. These instances are purpose-built to run large in-memory databases.
4. **High-CPU instances :** This types of instances are used in compute-intensive applications.
5. **Cluster compute instances :** Cluster compute instances provide a high-performance network interconnect along with a high-performance CPU.
6. **Cluster GPU instances :** This class provides instances featuring graphic processing units (GPUs) and high compute power, large memory and extremely high I/O and network performance.

- EC2 instances can be run either by using the command-line tools provided by Amazon, which connects the Amazon Web Service that provides remote access to the EC2 infrastructure.
- EC2 advantages :
  1. Amazon EC2 enables you to increase or decrease capacity within minutes.
  2. User have complete control of your Amazon EC2 instances.
  3. Support flexible cloud hosting services.
  4. Secure : Amazon EC2 works in conjunction with Amazon VPC to provide security and robust networking functionality.
  5. Reliable : Amazon EC2 offers a highly reliable environment where replacement instances can be rapidly and predictably commissioned.

4.2.3.3 Configuring Amazon EC2 Linux Instances

- Let's get started with Amazon Elastic Compute Cloud (Amazon EC2) by launching, connecting to and using a Linux instance. An instance is a virtual server in the AWS cloud. With Amazon EC2, you can setup and configure the operating system and applications that run on your instance.
- When you sign up for AWS, you can get started with Amazon EC2 using the AWS Free Tier.
- The instance is an Amazon EBS-backed instance (meaning that the root volume is an EBS volume). You can either specify the availability zone in which your instance runs or let Amazon EC2 select an availability zone for you. When you launch your instance, you secure it by specifying a key pair and security group.

When you connect to your instance, you must specify the private key of the key pair that you specified when launching your instance.

- Various steps to configure Amazon EC2 Linux instance is shown in Fig. 4.2.3.

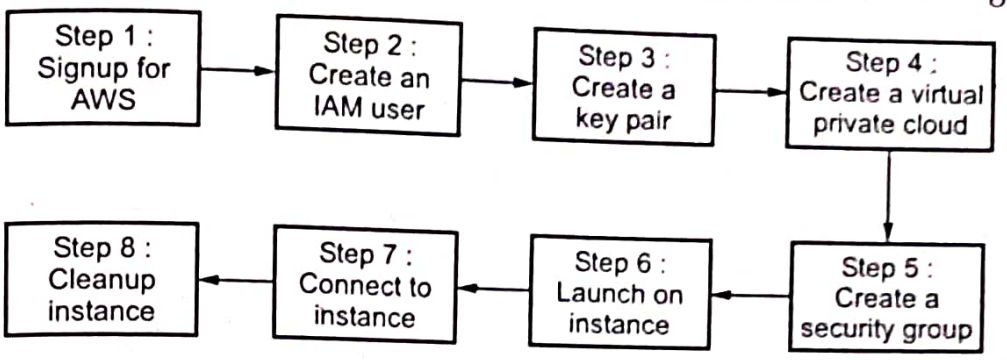


Fig. 4.2.3 Steps to signup for EC2

Step 1 : SignUp for AWS

- When you signup for Amazon Web Services (AWS), your AWS account is automatically signed up for all services in AWS, including Amazon EC2. You are charged only for the services that you use.
- With Amazon EC2, you pay only for what you use. If you are a new AWS customer, you can get started with Amazon EC2 for free.

Step 2 : Create an IAM user

- Services in AWS, such as Amazon EC2, require that you provide credentials when you access them, so that the service can determine whether you have permission to access its resources. The console requires your password.
- You can create access keys for your AWS account to access the command line interface or API. However, we don't recommend that you access AWS using the credentials for your AWS account; we recommend that you use AWS Identity and Access Management (IAM) instead.
- Create an IAM user and then add the user to an IAM group with administrative permissions or grant this user administrative permissions. You can then access AWS using a special URL and the credentials for the IAM user. If you signed up for AWS but have not created an IAM user for yourself, you can create one using the IAM console.

Step 3 : Create a key pair

- AWS uses public-key cryptography to secure the login information for your instance. A Linux instance has no password; you use a key pair to log in to your instance securely. You specify the name of the key pair when you launch your instance, then provide the private key when you log in using SSH.

- If you haven't created a key pair already, you can create one using the Amazon EC2 console. Note that if you plan to launch instances in multiple regions, you'll need to create a key pair in each region.

Step 4 : Create a Virtual Private Cloud (VPC)

- Amazon VPC enables you to launch AWS resources into a virtual network that you've defined, known as a Virtual Private Cloud (VPC). The newer EC2 instance types require that you launch your instances in a VPC. If you have a default VPC, you can skip this section and move to the next task, create a security group. To determine whether you have a default VPC, open the Amazon EC2 console and look for default VPC under account attributes on the dashboard.

Step 5 : Create a security group

- Security groups act as a firewall for associated instances, controlling both inbound and outbound traffic at the instance level. You must add rules to a security group that enable you to connect to your instance from your IP address using SSH. You can also add rules that allow inbound and outbound HTTP and HTTPS access from anywhere. Note that if you plan to launch instances in multiple regions, you'll need to create a security group in each region.

Step 6 : Launch an instance

- You can launch a Linux instance using the AWS management console as described in the following procedure.
  1. Open the Amazon EC2 console at **https://console.aws.amazon.com/ec2/.**
  2. From the console dashboard, choose **Launch Instance.**
  3. The **Choose an Amazon Machine Image (AMI)** page displays a list of basic configurations, called Amazon Machine Images (AMIs), that serve as templates for your instance. Select an HVM version of Amazon Linux 2. Notice that these AMIs are marked "Free tier eligible."
  4. On the **Choose an Instance Type** page, you can select the hardware configuration of your instance. Select the t2.micro type, which is selected by default. Notice that this instance type is eligible for the free tier.
  5. Choose **Review and Launch** to let the wizard complete the other configuration settings for you.
  6. On the **Review Instance Launch** page, under security groups, you'll see that the wizard created and selected a security group for you. You can use this security group or alternatively you can select the security group that you created when getting setup using the following steps.
     a) Choose **Edit security groups.**

b) On the **Configure Security Group** page, ensure that **Select an existing security** group is selected.

c) Select your security group from the list of existing security groups and then choose **Review and Launch.**

7. On the **Review Instance Launch** page, choose **Launch.**
8. When prompted for a key pair, select **Choose an existing key pair,** then select the key pair that you created when getting setup. When you are ready, select the acknowledgement check box and then choose launch instances.
9. A confirmation page lets you know that your instance is launching. Choose **View Instances** to close the confirmation page and return to the console.
10. On the **Instances** screen, you can view the status of the launch. It takes a short time for an instance to launch. When you launch an instance, its initial state is pending. After the instance starts, its state changes to running and it receives a public DNS name.
11. It can take a few minutes for the instance to be ready so that you can connect to it. Check that your instance has passed its status checks; you can view this information in the status checks column.

Step 7 : Connect to your Instance

Several ways to connect to your Linux instance is shown in Table 4.2.1.

Your computer OS	Topic
Linux	Connecting to your Linux instance using SSH.
Windows	Connecting to your Linux instance from Windows using PuTTY.
	Connecting to your Linux instance from Windows using Windows Subsystem for Linux.
Other	Connecting to your Linux instance using MindTerm.

Table 4.2.1 Ways to connect to Linux instance

Step 8 : Cleanup your instance

- After you've finished with the instance, you should cleanup by terminating the instance.
- Terminating an instance effectively deletes it; you can't reconnect to an instance after you've terminated it.
- If you launched an instance that is not within the AWS free tier, you'll stop incurring charges for that instance as soon as the instance status changes to shutting down or terminated. If you'd like to keep your instance for later, but not incur charges, you can stop the instance now and then start it again later.

- To terminate your instance following steps can be used :
  1) In the navigation pane, choose instances. In the list of instances, select the instance.
  2) Choose actions, instance state, terminate.
  3) Choose yes, terminate when prompted for confirmation.
- Amazon EC2 shuts down and terminates your instance. After your instance is terminated, it remains visible on the console for a short while and then the entry is deleted.

4.2.4 Storage Service

- AWS provides a collection of services for data storage and information management. It is represented by Amazon Simple Storage Service (S3).
- Amazon S3 has a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web. S3 can serve as a raw data store for IoT systems for storing raw data, such as sensor data, log data, audio and video data.

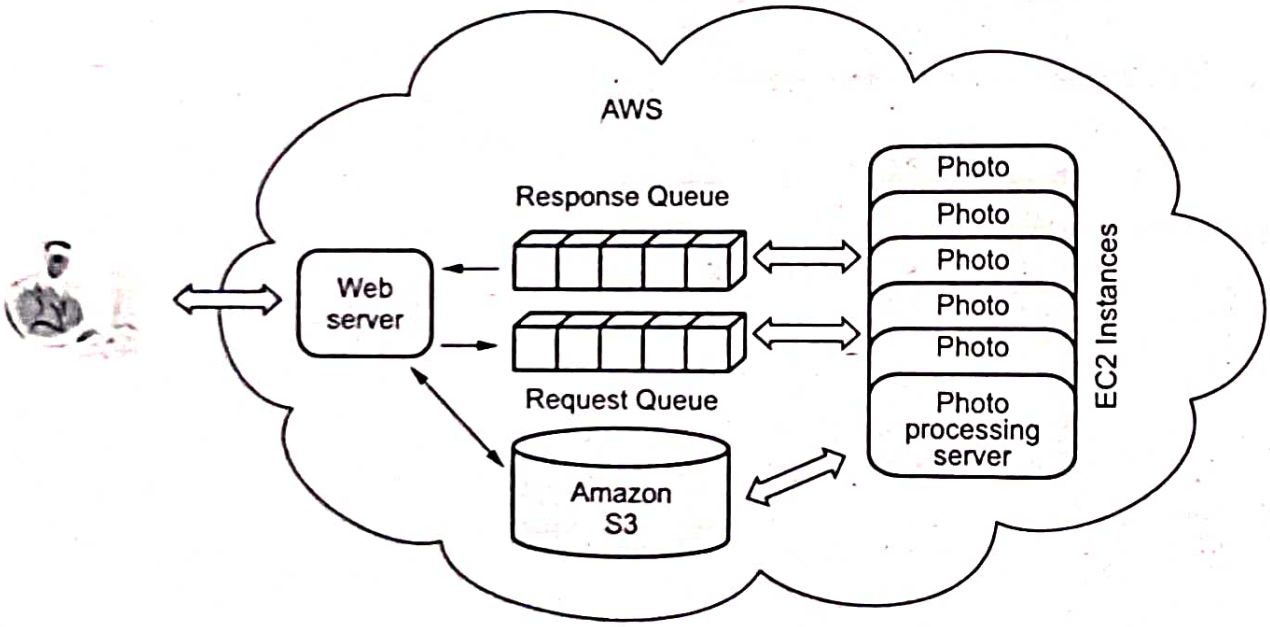


Fig. 4.2.4 Amazon S3 working

Features :

1. Unlimited storage.
2. Highly scalable : In terms of storage, request rate and concurrent users.
3. Reliable : Store redundant data in multiple facilities and on multiple devices.
4. Secure : Flexibility to control who / how / when / where to access the data.
5. Performance : Choose region to optimize for latency / minimize costs.

- Example : Online photo processing service.

Procedure :

1. Web server receive request.
2. Put request message in the queue.
3. Pictures stored in S3.
4. Multiple EC2 instances run photo processing.
5. Put back in the queue.
6. Return.

- Store data on Amazon's distributed system containing multiple servers within Amazon's data center locations. Amazon doesn't offer you a GUI based tool to access your data. You can use one of the several tools online or build one through APIs.
- Amazon EC2 provides three type of storage option : Amazon EBS, Amazon S3 and Instance Storage. Amazon EBS (Elastic Block Store) provides with persistent, block-level storage. Basically additional hard disk that you can attach to instance. It suitable for apps which require database, filesystem, block level storage.
- A **bucket** is a container for objects stored in Amazon S3. Every object is contained in a bucket. For example, if the object named "photos/puppy.jpg" is stored in the rakshita bucket, then it is addressable using the URL http://rakshita.s3.amazonaws.com/photos/puppy.jpg
- Buckets serve several purposes : They organize the Amazon S3 namespace at the highest level, they identify the account responsible for storage and data transfer charges, they play a role in access control and they serve as the unit of aggregation for usage reporting.
- **Objects** are the fundamental entities stored in Amazon S3. Objects consist of object data and metadata. The data portion is opaque to Amazon S3. The metadata is a set of name-value pairs that describe the object. These include some default metadata, such as the date last modified and standard HTTP metadata, such as content-type. You can also specify custom metadata at the time the object is stored.
- A **key** is the unique identifier for an object within a bucket. Every object in a bucket has exactly one key. Because the combination of a bucket, key and version ID uniquely identify each object, Amazon S3 can be thought of as a basic data map between "bucket + key + version" and the object itself. Every object in Amazon S3 can be uniquely addressed through the combination of the web service endpoint, bucket name, key and optionally, a version.

- **Regions :** You can choose the geographical region where Amazon S3 will store the buckets you create. Objects stored in a region never leave the region unless you explicitly transfer them to another region.

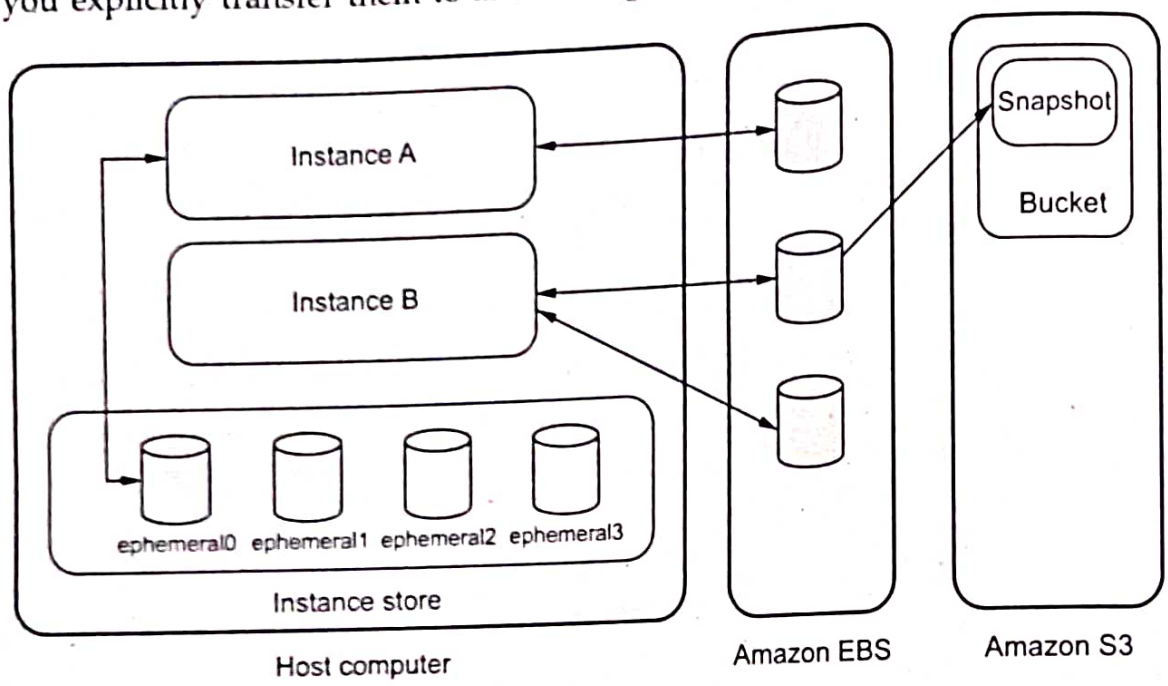


Fig. 4.2.5 Amazon EBS and S3

4.2.4.1 Bucket

- Amazon S3 defines a bucket name as a series of one or more labels, separated by periods, that adhere to the following rules : The bucket name can be between 3 and 63 characters long, and can contain only lower-case characters, numbers, periods and dashes
- Amazon S3 defines a bucket name as a series of one or more labels, separated by periods, that adhere to the following rules :
  1. The bucket name can be between 3 and 63 characters long and can contain only lower-case characters, numbers, periods and dashes.
  2. Each label in the bucket name must start with a lowercase letter or number.
  3. The bucket name cannot contain underscores, end with a dash, have consecutive periods or use dashes adjacent to periods.
  4. The bucket name cannot be formatted as an IP address (198.51.100.24).
- A bucket is owned by the AWS account that created it. By default, you can create up to 100 buckets in each of your AWS accounts. If you need additional buckets, you can increase your bucket limit by submitting a service limit increase.
- The following are the rules for naming S3 buckets in all AWS Regions :
  1. Bucket names must be unique across all existing bucket names in Amazon S3.

2. Bucket names must comply with DNS naming conventions.
3. Bucket names must be at least 3 and no more than 63 characters long.
4. Bucket names must not contain uppercase characters or underscores.
5. Bucket names must start with a lowercase letter or number.
6. Bucket names must be a series of one or more labels. Adjacent labels are separated by a single period (.). Bucket names can contain lowercase letters, numbers, and hyphens. Each label must start and end with a lowercase letter or a number.
7. Bucket names must not be formatted as an IP address (for example, 192.168.5.4).
8. When you use virtual hosted-style buckets with Secure Sockets Layer (SSL), the SSL wildcard certificate only matches buckets that don't contain periods. To work around this, use HTTP or write your own certificate verification logic. We recommend that you do not use periods (".") in bucket names when using virtual hosted-style buckets.

4.2.4.2 Amazon Elastic Block Store

- Amazon Elastic Block Store (Amazon EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud.
- Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability.
- EBS volumes are highly available and reliable storage volumes that can be attached to any running instance that is in the same Availability Zone.
- EBS volumes are particularly well-suited for use as the primary storage for file systems, databases, or for any applications that require fine granular updates and access to raw, unformatted, block-level storage.
- The size of an EBS volume can be configured by the user and can range from 1 GB to 1 TB.
- The network-based EBS storage service is delivered in volumes, which can be attached to an EC2 instance and used just like a disk drive. Because a volume can become unformatted, it must have a file system installed (formatted) on it before it can be used.
- Amazon EBS is well suited to both database-style applications that rely on random reads and writes, and to throughput-intensive applications that perform long, continuous reads and writes.

- Amazon EBS encryption offers you a simple encryption solution for your EBS volumes without the need for you to build, manage and secure your own key management infrastructure.
- When you create an encrypted EBS volume and attach it to a supported instance type, data stored at rest on the volume, disk I/O and snapshots created from the volume are all encrypted.
- Amazon EBS encryption uses AWS Key Management Service (AWS KMS) master keys when creating encrypted volumes and any snapshots created from your encrypted volumes.
- EBS can make your applications more reliable, because the storage is separate from any specific instance.
- A large repository of public data set snapshots can be restored to EBS volumes and seamlessly integrated into AWS cloud-based applications.
- Performance metrics, such as bandwidth, throughput, latency and average queue length, are available through the AWS Management Console.
- These metrics, provided by Amazon CloudWatch, allow you to monitor the performance of your volumes to make sure that you are providing enough performance for your applications without paying for resources you don't need.
- Amazon EBS storage costs depend on how much EBS storage, in terms of gigabyte-per-month, is provisioned in a particular account.
- While EC2 instances only accrue charges while they're running, the EBS volumes attached to instances continue to retain information and hence accrue charges, even when the instance is stopped.
- An EBS snapshot is a point-in-time backup of an EBS volume. It is a "copy" of the data on in EBS volume. EBS snapshots are billed at a lower rate than active EBS volumes.
- If an EBS block has low access volume, the active volume of this EBS block can be deleted after the information stored in EBS is copied to an EBS Snapshot.

EBS Snapshots

- Amazon EBS provides the ability to create snapshots (backups) of any EBS volume and write a copy of the data in the volume to Amazon S3, where it is stored redundantly in multiple Availability Zones.
- The volume does not need to be attached to a running instance in order to take a snapshot.
- As you continue to write data to a volume, you can periodically create a snapshot of the volume to use as a baseline for new volumes.

- These snapshots can be used to create multiple new EBS volumes or move volumes across Availability Zones. Snapshots of encrypted EBS volumes are automatically encrypted.
- When you create a new volume from a snapshot, it's an exact copy of the original volume at the time the snapshot was taken.
- EBS volumes that are restored from encrypted snapshots are automatically encrypted. The snapshots can be shared with specific AWS accounts or made public.
- When you create snapshots, you incur charges in Amazon S3 based on the volume's total size. For a successive snapshot of the volume, you are only charged for any additional data beyond the volume's original size.
- Snapshots are incremental backups, meaning that only the blocks on the volume that have changed after your most recent snapshot are saved.

4.2.4.3 Amazon ElastiCache

- It is a fully managed caching service.
- ElastiCache is protocol-compliant with Memcached, an open source, high-performance, distributed memory object caching system for speeding up dynamic web applications by alleviating database load.
- According to the Amazon website, ElastiCache makes it easy to deploy, operate, and scale an in-memory cache in the cloud.
- The service improves the performance of web applications by enabling information retrieval from a fast, managed, in-memory caching system, instead of relying entirely on slower disk-based databases.
- ElastiCache is a managed, in-memory data store service. It has two engines AWS Redis and Memcached which is used to power real-time applications.
- Memcached is a general-purpose distributed memory caching system. It is often used to speed up dynamic database-driven websites by caching data and objects in RAM to reduce the number of times an external data source must be read.
- Memcached is free and open-source software, licensed under the Revised BSD license.
- ElastiCache offloads the administrative overhead of running a caching service by :
  1. Creating the server pool based on commands issued via the AWS Management Console or API.
  2. Managing the pool to ensure caching server availability.

3. Automatically patching servers with necessary software changes and migrating your data from an un-patched version to a new, patched version.
4. Allowing you to grow or shrink the pool with a simple command.

- ElastiCache runs in the Amazon Virtual Private Cloud environment, giving you complete control over network access to your cache cluster.
- Amazon ElastiCache automatically detects and replaces failed nodes, reducing the overhead associated with self-managed infrastructures and provides a resilient system that mitigates the risk of overloaded databases, which slow website and application load times.
- Node is the smallest building block of an ElastiCache deployment. It is a fixed-size chunk of secure, network-attached RAM. Each cache node runs an instance of either Memcached or Redis.
- Memcached cluster can have up to 20 nodes.

4.2.4.4 Amazon SimpleDB

- SimpleDB provides a simplified data model based on the relational database data model. SimpleDB provides support for semi structured data, the model for which is based on the concept of domains, items, and attributes.
- This service works in close conjunction with Amazon Simple Storage Service (Amazon S3) and Amazon Elastic Compute Cloud, collectively providing the ability to store, process and query data sets in the cloud. These services are designed to make web-scale computing easier and more cost-effective for developers.
- SimpleDB differs from relational databases where user must define a schema for each database table before user can use it and where user must explicitly change that schema before user can store data differently.
- In SimpleDB, there is no schema requirement. Although user still have to consider the format of data, this approach has the benefit of freeing from the time it takes to manage schema modifications.
- The lack of schema means that there are no data types; all data values are treated as variable length character data. As a result, there is literally nothing extra to do if user want to add a new field to an existing database. Just add the new field to whichever data items require it. There is no rule that forces every data item to have the same fields.
- The drawbacks of a schema-less database include the lack of automatic integrity checking in the database and an increased burden on the application to handle formatting and type conversions.

4.2.4.5 Amazon CloudFront

- Amazon CloudFront is a content delivery web service (CDN). It integrates with other AWS Cloud services to give developers and businesses an easy way to distribute content to users across the world with low latency, high data transfer speeds and no minimum usage commitments.
- Amazon CloudFront uses RTMP protocol for video streaming and HTTP or HTTPS for web content. Content delivery networks are suited for delivery of bulky data, like video streaming, downloading larger files and software and to make website access faster.
- Amazon CloudFront is a pay-as-you-go model that can easily be integrated with all Amazon Web Services.
- Amazon CloudFront operates by caching the instance of each object on its different CDN locations, therefore reducing the time it takes to deliver content.
- Amazon CloudFront accesses the data from Amazon S3 through supported application programming interfaces and places it in regional data buckets.

Advantages :

1. No server hardware infrastructure to set up or maintain.
2. No up-front investment in software licenses.
3. No long -term commitment.
4. Global delivery using CloudFront.
5. Pay for what you use.
6. Easy to get started with self service management console.

4.3 Microsoft Azure

- Windows Azure is a cloud computing platform and infrastructure, created by Microsoft, for building, deploying and managing applications and services through a global network of Microsoft - managed data centers.
- Azure queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS. A single queue message can be up to 64 KB in size and a queue can contain millions of messages, up to the total capacity limit of a storage account.
- Azure is a virtualized infrastructure to which a set of additional enterprise services has been layered on top, including, a virtualization service called Azure AppFabric that creates an application hosting environment. AppFabric is a cloud-enabled version of the .NET framework.

- Windows Azure is Microsoft's application platform for the public Cloud. Applications can be deployed on to Azure in various models.
- Windows Azure is used to :
  1. Build a web application that runs and stores its data in Microsoft data centers.
  2. Store data while the applications that consume this data run on premise (outside the public Cloud).
  3. Create virtual machines to develop and test, or run SharePoint and other out-of-the-box applications.
  4. Develop massively scalable applications with many users.
  5. Offer a wide range of services.
- Azure has three components : compute, storage and fabric.
  1. **Compute :** Windows Azure provides a hosting environment for managed code. It provides a computation service through roles. Windows Azure supports three types of roles :
     a) Web roles used for web application programming and supported by IIS7.
     b) Worker roles are also used for background processing of web roles.
     c) Virtual Machine (VM) roles are generally used for migrating windows server applications to Windows Azure in an easy way.
  2. **Storage :** Windows Azure provides storage in the cloud. It provides four different types of storage services :
     a) Queues for messaging between web roles and worker roles.
     b) Tables for storing structural data.
     c) BLOBs (Binary Large Objects) to store text, files or large data.
     d) Windows Azure Drives (VHD) to mount a page blob. They can easily be downloaded and uploaded via blobs.
  3. AppFabric provides infrastructure services for developing, deploying and managing Windows Azure application. It provides five services : Service bus, Access, Caching, Integration and Composite.
- Fig. 4.3.1 shows Windows Azure platform architecture.
- Microsoft Azure is a cloud computing service created by Microsoft for building, testing, deploying and managing applications and services through a global network of Microsoft-managed data centers.
- It provides software as a service (SaaS), platform as a service and infrastructure as a service and supports many different programming languages, tools and

frameworks, including both Microsoft-specific and third-party software and systems.

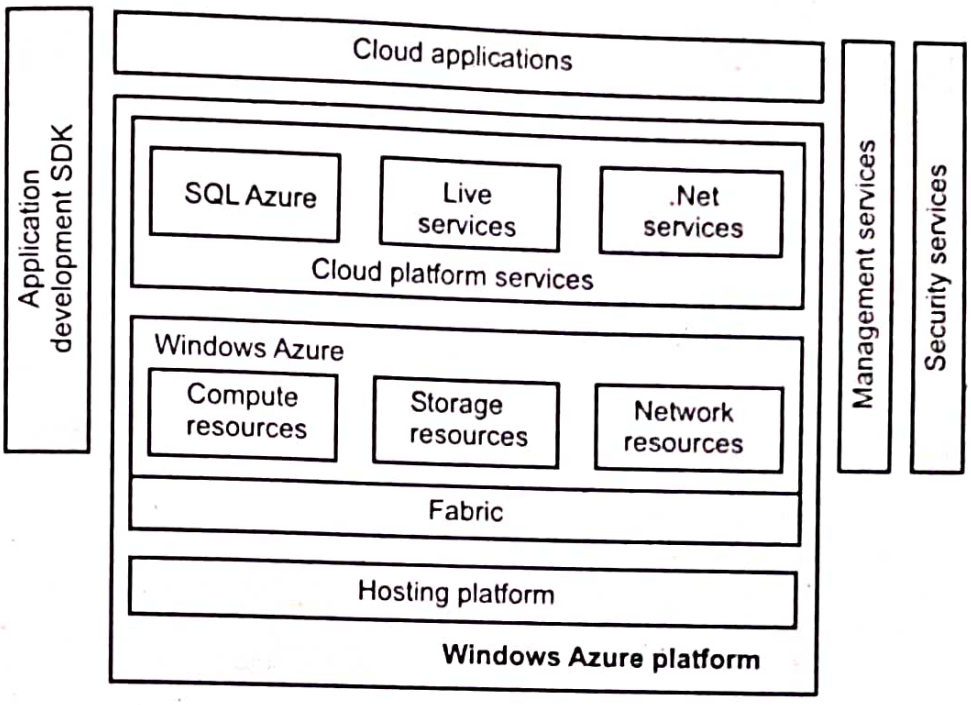


Fig. 4.3.1 Windows Azure platform architecture

- Windows Azure provides resources and services for consumers. For example, hardware is abstracted and exposed as compute resources.
- Physical storage is abstracted as storage resources and exposed through very well-defined interfaces.
- A common windows fabric abstracts the hardware and the software and exposes virtual compute and storage resources.
- Each instance of an application is automatically managed and monitored for availability and scalability.
- If an application goes down, the Fabric is notified and a new instance of the application is created. Because virtualization is a key element in cloud computing, no assumption must be made on the state of the underlying hardware hosting the application.
- Advantages of Microsoft Azure
  1. Microsoft Azure offers high availability.
  2. It offers you a strong security profile.
  3. It is a cost-effective solution for an IT budget.
  4. Azure allows you to use any framework, language or tool.
  5. Azure allows businesses to build a hybrid infrastructure.

4.4 Cloud Software Environments : Eucalyptus

- Eucalyptus stands for "Elastic Utility Computing Architecture for Linking Your programs to Useful Systems". It is used to build private, public and hybrid clouds. It can also produce your own data center into a private cloud and allow you to extend the functionality to many other organizations.
- Eucalyptus in cloud computing is an open-source software platform for carrying out Infrastructure-as-a-Service in a hybrid cloud computing or private cloud computing environment.
- Eucalyptus is open-source software for building AWS-compatible private and hybrid clouds. As an Infrastructure as a Service (IaaS) product, Eucalyptus allows your users to provision your compute and storage resources on-demand.
- Eucalyptus has the following key features :
  a) Support for multiple users with the help of a single cloud.
  b) Support for linux and windows virtual machines.
  c) Accounting reports.
  d) Use of WS-security to ensure secure communication between internal resources and processes.
  e) The option to configure policies and service level agreements based on users and the environment.
  f) Provisions for group, user management and security groups.
- Challenges
  a) **Extensibility :** Simple architecture and open internal APIs.
  b) **Client-side interface :** Amazon's EC2 interface and functionality (familiar and testable).
  c) **Networking :** Virtual private network per cloud and must function as an overlay.
  d) **Security :** Must be compatible with local security policies.
  e) Packaging, installation, maintenance : system administration staff is an important constituency for uptake.
- Fig. 4.4.1 shows Eucalyptus architecture.

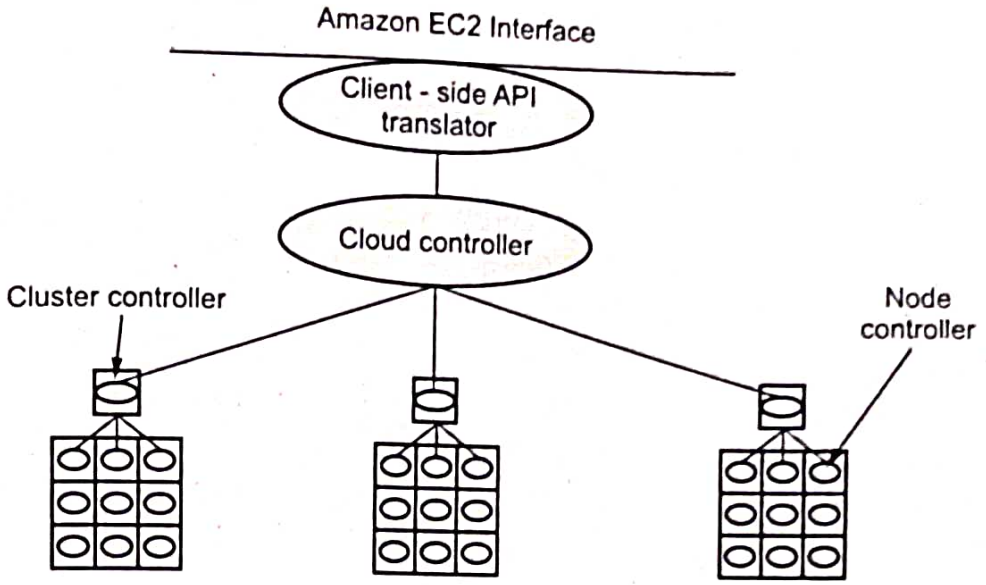


Fig. 4.4.1 Eucalyptus architecture

- **Components of eucalyptus in cloud computing :**
  1. **Node controller :** The Node Controller (NC) is the component that executes on the physical resources that host VM instances and is responsible for instance start up, inspection, shutdown and clean-up.
  2. **Cluster controller :** A collection of NCs that logically belong together report to a single Cluster Controller (CC) that typically executes on a cluster head node or server that has access to both private and public networks. The CC is responsible for gathering state information from its collection of NCs, scheduling incoming VM instance execution requests to individual NCs and managing the configuration of public and private instance networks.
  3. **Cloud controller :** Each Eucalyptus installation includes a single Cloud Controller (CLC) that is the user-visible entry point and global decision-making component of a Eucalyptus installation. The CLC is responsible for processing incoming user-initiated or administrative requests, making high-level VM instance scheduling decisions, processing Service-Level Agreements (SLAs) and maintaining persistent system and user metadata.
- The CLC itself is composed of a collection of services that handle user requests and authentication, persistent system and user metadata, and the management and monitoring of VM instances.
  4. **Client interface :** The CLC's client interface service essentially acts as a translator between the internal Eucalyptus system interfaces and some defined external client interface.

- For example, Amazon provides a WSDL document that describes a Web-service SOAP based client interface to their service as well as a document describing an HTTP Query-based interface, both of which can be translated by the CLC user interface service into Eucalyptus internal objects.

5. **Administrative interface :** In addition to supporting primary tasks, such as starting and stopping instances, a cloud infrastructure must support administrative tasks, such as adding and removing users and disk images.

- Eucalyptus supports such tasks though a Web based interface, implemented by the cloud controller and command line tools. Unlike the client interface, however, the administrative interface is unique to Eucalyptus.

6. **Instance control :** Creation of virtual machine instance metadata in Eucalyptus is managed by a component of the CLC named the VmControl service.

7. **SLA implementation and management :** Service-level agreements (SLAs) are implemented as extensions to the message handling service which can inspect, modify, and reject the message, as well as the state stored by VmControl.

- Eucalyptus does not assume that all worker nodes will have publicly routable IP addresses. Each cloud allocation will have one or more public IP addresses. All cloud images have access to a private network interface. Two types of networks internal to a cloud allocation.

4.4.1 Eucalyptus Installation

- To install Eucalyptus, perform the following tasks :

1. **Plan your installation :** In order to get the most out of a Eucalyptus deployment.
2. **Configure dependencies :** Before you install Eucalyptus, ensure you have the appropriate dependencies installed and configured.
3. **Install repositories :** Downloads RPM packages.
4. Configure eucalyptus.
5. Start eucalyptus.
6. Register eucalyptus services.
7. Configure the runtime environment.

- Features of eucalyptus in cloud computing are :

a) Supports both Windows and Linux virtual machines.

b) API is viable with the Amazon EC2 platform.

c) Viable with Simple Storage Service (S3) and Amazon Web Services (AWS).

Installing the node controller :

- There are two main ways of going about installing Eucalyptus.
- The first way is to download the required RPMs onto your machine, install each of them and then manually configure the cloud as per your needs.
- The second way is much faster and will get your Eucalyptus cloud up and running in a matter of minutes.
- Installing the node controller is a very simple process. Once your machine boots from the Eucalyptus Faststart DVD, select the option 'Install CentOS 6 with Eucalyptus Node Controller' from the boot screen.

Installing the cloud controller :

- Installation of the cloud controller is very similar to the nodes, with a few exceptions. Once your machine boots from the Eucalyptus Faststart DVD, select the option 'Install CentOS 6 with Eucalyptus Frontend' from the boot screen.
- Again, select the appropriate 'Language' and 'Keyboard settings' according to your needs.
- Provide a 'Static IP' and a suitable 'Host Name' to your cloud controller in the 'Network Configuration' wizard.
- Once done, you will be provided with an interface to supply a 'Public IP Range/ List' for your Eucalyptus cloud. You need to enter a valid IP address range here. These public IPs will be mapped to individual Eucalyptus instances (virtual machines) once they are launched in the cloud.

4.4.2 Advantages of Eucalyptus

- Eucalyptus can be utilised to benefit both the eucalyptus private cloud and the eucalyptus public cloud.
- Clients can run Amazon or Eucalyptus machine pictures as examples on both clouds.
- It isn't extremely mainstream on the lookout yet is a solid contender to CloudStack and OpenStack.
- It support application programming interface similarity with all the Amazon Web Services.
- Eucalyptus can be utilised with DevOps apparatuses like chef and puppet.

4.5 OpenStack

AU : Dec.-21

- OpenStack is a recently open-sourced, IaaS cloud-computing platform founded by Rackspace Hosting and NASA and is used widely in industry.

- OpenStack is an open-source cloud platform. OpenStack software controls large pools of compute, storage, and networking resources throughout a data center, all managed by a dashboard that gives administrators control while empowering their users to provision resources through a web interface.
- To produce the ubiquitous Open-Source cloud computing platform that will meet the needs of public and private cloud providers regardless of size, by being simple to implement and massively scalable.
- Components of OpenStack are as follows :
  1. Horizon - Dashboard : It provides a modular web-based user interface for all the OpenStack services. With this web GUI, user can perform most operations on your cloud like launching an instance, assigning IP addresses and setting access controls.
  2. Keystone is a framework for authentication and authorization for all the OpenStack services. It handles API requests as well as providing configurable catalog, policy, token and identity services. Keystone is a framework for authentication and authorization for all the OpenStack services.
  3. Nova : It provides virtual servers upon demand. Nova is the most complicated and distributed component of OpenStack. A large number of processes cooperate to turn end user API requests into running virtual machines.
  4. Glance - Image Store : It provides discovery, registration and delivery services for disk and server images.
  5. Quantum - Network : It provides " network connectivity as a service " between interface devices managed by other OpenStack services. The service works by allowing users to create their own networks and then attach interfaces to them. Quantum has a pluggable architecture to support many popular networking vendors and technologies.
  6. Cinder allows block devices to be exposed and connected to compute instances for expanded storage and better performance.
  7. Object store allows you to store or retrieve files. It provides a fully distributed, API-accessible storage platform that can be integrated directly into applications or used for backup, archiving and data retention

University Question

1. *Detail the structure of OpenStack and explain each of its components.*

AU : Dec.-21, Marks 13

4.6 Two Marks Questions with Answers

Q.1 What is Amazon Web Services ?

Ans. : Amazon Web Services (AWS) is a collection of remote computing services (web services) that together make up a cloud computing platform, offered over the Internet by Amazon.com. Amazon Web Services (AWS) is a cloud computing platform from Amazon that provides customers with a wide array of cloud services.

Q.2 What is AWS ecosystem ?

Ans. : • AWS ecosystem is made up of three subsystems :

1. AWS computing services provided by Amazon.
2. Computing services provided by third parties that operate on AWS.
3. Complete applications offered by third parties that run on AWS.

Q.3 What do you understand by third party cloud services ?

Ans. : Composing service that belongs to different vendors or integrating them into existing software systems. The service-oriented model, which is the basis of cloud computing, facilitates such an approach and provides the opportunity for developing a new class of services that can be called third-party cloud services.

Q.4 What is eucalyptus ?

Ans. :

- Eucalyptus stands for Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems.
- It is an open-source software framework that provides the platform for private cloud computing implementation on computer clusters.
- Eucalyptus implements Infrastructure as a Service (IaaS) methodology for solutions in private and hybrid clouds.
- Eucalyptus provides a platform for a single interface so that users can calculate the resources available in private clouds and the resources available externally in public cloud services.

Q.5 List the features of eucalyptus.

Ans. : Features include :

1. Supports both Linux and Windows Virtual Machines (VMs).
2. Application program interface - (API)compatible with Amazon EC2.
3. Compatible with Amazon Web Services (AWS) and Simple Storage Service (S3).
4. Works with multiple hypervisors including VMware, Xen and KVM.
5. Can be installed and deployed from source code or DEB and RPM.

Q.6 What is azure queues ?

Ans. : Azure queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS. A single queue message can be upto 64 KB in size and a queue can contain millions of messages, upto the total capacity limit of a storage account.

Q.7 How virtualization employed in azure ?

Ans. : Azure is a virtualized infrastructure to which a set of additional enterprise services has been layered on top, including, a virtualization service called azure AppFabric that creates an application hosting environment. AppFabric is a cloud-enabled version of the .NET Framework.

Q.8 List the major feature of Google app engine. Which kind of problems can be solved using Google app engine ?

Ans. : Major feature of Google app engine :

1. Automatic scaling and load balancing.
2. Authentication using Google Accounts API.
3. Provides dynamic web services based on common standards.
4. Integration with other Google cloud services and API.
5. Support persistent storage, with query access sorting and transaction management features.
6. Google app engine offers users the ability to build and host web applications on Google's infrastructure.

❑❑❑

UNIT V

5 Cloud Security

Syllabus

Virtualization System - specific Attacks : Guest hopping - VM migration attack - hyperjacking. Data Security and Storage; Identity and Access Management (IAM) - IAM Challenges - IAM Architecture and Practice.

Contents

5.1 Overview of Cloud Security

- Cloud security is the protection of data stored online via cloud computing platforms from theft, leakage, and deletion. Methods of providing cloud security include firewalls, penetration testing, tokenization, Virtual Private Networks (VPN), and avoiding public internet connections.
- Cloud security refers to an array of policies, technological procedures, services, and solutions designed to support safe functionality when building, deploying, and managing cloud-based applications and associated data.
- Cloud security is designed to protect the following, regardless of your responsibilities :
  a) **Physical networks** - Routers, electrical power, cabling, climate controls, etc.
  b) **Data storage** - Hard drives, etc.
  c) **Data servers** - Core network computing hardware and software
  d) **Computer virtualization frameworks** - Virtual machine software, host machines and guest machines
  e) **Operating systems (OS)** - Software that houses
  f) **Middleware** - Application Programming Interface (API) management
  g) **Runtime environments** - Execution and upkeep of a running program
  h) **Data** - All the information stored, modified and accessed
  i) **Applications** - Traditional software services (email, tax software, productivity suites, etc.)
  j) **End-user hardware** - Computers, mobile devices, Internet of Things (IoT) devices, etc.
- Cloud computing security addresses both physical and logical security issues across all the different service models of software, platform and infrastructure. It also addresses how these services are delivered in the public, private, hybrid and community delivery models.

5.1.1 Cloud Security Challenges and Risks

- Cloud computing security challenges fall into three broad categories :
  1. **Data protection :** Securing your data both at rest and in transit.
  2. **User authentication :** Limiting access to data and monitoring who accesses the data.
  3. **Disaster and data breach :** Contingency planning.

- **Data protection :** Data needs to be encrypted at all times, with clearly defined roles when it comes to who will be managing the encryption keys.
- **User authentication :** Data resting in the cloud needs to be accessible only by those authorized to do so, making it critical to both restrict and monitor who will be accessing the company's data through the cloud. In order to ensure the integrity of user authentication, companies need to be able to view data access logs and audit trails to verify that only authorized users are accessing the data.
- **Contingency planning :** With the cloud serving as a single centralized repository for a company's mission-critical data, the risks of having that data compromised due to a data breach or temporarily made unavailable due to a natural disaster are real concerns.
- If information is encrypted while passing through the cloud, who controls the encryption/decryption keys ? Is it the customer or the cloud vendor ? Most customers probably want their data encrypted both ways across the Internet using secure sockets layer protocol.
- They also most likely want their data encrypted while it is at rest in the cloud vendor's storage pool. Be sure that you, the customer, control the encryption/decryption keys, just as if the data were still resident on your own servers.
- Data integrity means ensuring that data is identically maintained during any operation.
- Cloud-based services will result in many mobile IT users accessing business data and services without traversing the corporate network. This will increase the need for enterprises to place security controls between mobile users and cloud-based services.
- Placing large amounts of sensitive data in a globally accessible cloud leaves organizations open to large distributed threats, attackers no longer have to come onto the premises to steal data, and they can find it all in the one "virtual" location.
- Virtualization efficiencies in the cloud require virtual machines from multiple organizations to be co-located on the same physical resources. Although traditional data center security still applies in the cloud environment, physical segregation and hardware-based security cannot protect against attacks between virtual machines on the same server.
- Operating system and application files are on a shared physical infrastructure in a virtualized cloud environment and require system, file, and activity monitoring to provide confidence and auditable proof to enterprise customers that their resources have not been compromised or tampered with.

- In the cloud computing environment, the enterprise subscribes to cloud computing resources and the responsibility for patching is the subscriber's rather than the cloud computing vendor's.
- The need for patch maintenance vigilance is imperative. Lack of due diligence in this regard could rapidly make the task unmanageable or impossible, leaving you with "virtual patching" as the only alternative.
- **Confidentiality :** Confidentiality refers to limiting information access. Sensitive information should be kept secret from individuals who are not authorized to see the information.
- In cloud environments, confidentiality primarily pertains to restricting access to data in transit and storage.
- Integrity can extend to how data is stored, processed, and retrieved by cloud services and cloud-based IT resources.
- Some common cloud security threats include :
  a) Risks of cloud-based infrastructure including incompatible legacy IT frameworks, and third-party data storage service disruptions.
  b) Internal threats due to human error such as misconfiguration of user access controls.
  c) External threats caused almost exclusively by malicious actors, such as malware, phishing, and DDoS attacks.

5.1.2 Cloud Security Architecture

- Cloud security architecture describes all the hardware and technologies designed to protect data, workloads, and systems within cloud platforms.
- Fig. 5.1.1 shows NIST cloud computing security reference architecture approach. The reference architecture identifies the five major cloud actors; consumer, provider, broker, carrier, and auditor.
- Secure cloud computing architecture encompasses three core capabilities: confidentiality, integrity, and availability.
  1. Confidentiality is the ability to keep information secret and unreadable to the people who shouldn't have access to that data.
  2. Integrity is the idea that the systems and applications are exactly what you expect them to be and function exactly as you expect them to function.
  3. Availability speaks to Denial-of-Service (DoS) attacks. Perhaps an attacker can't see or change your data. But if an attacker can make systems unavailable to you or your customers, then you can't carry out tasks that are essential to maintain your business.

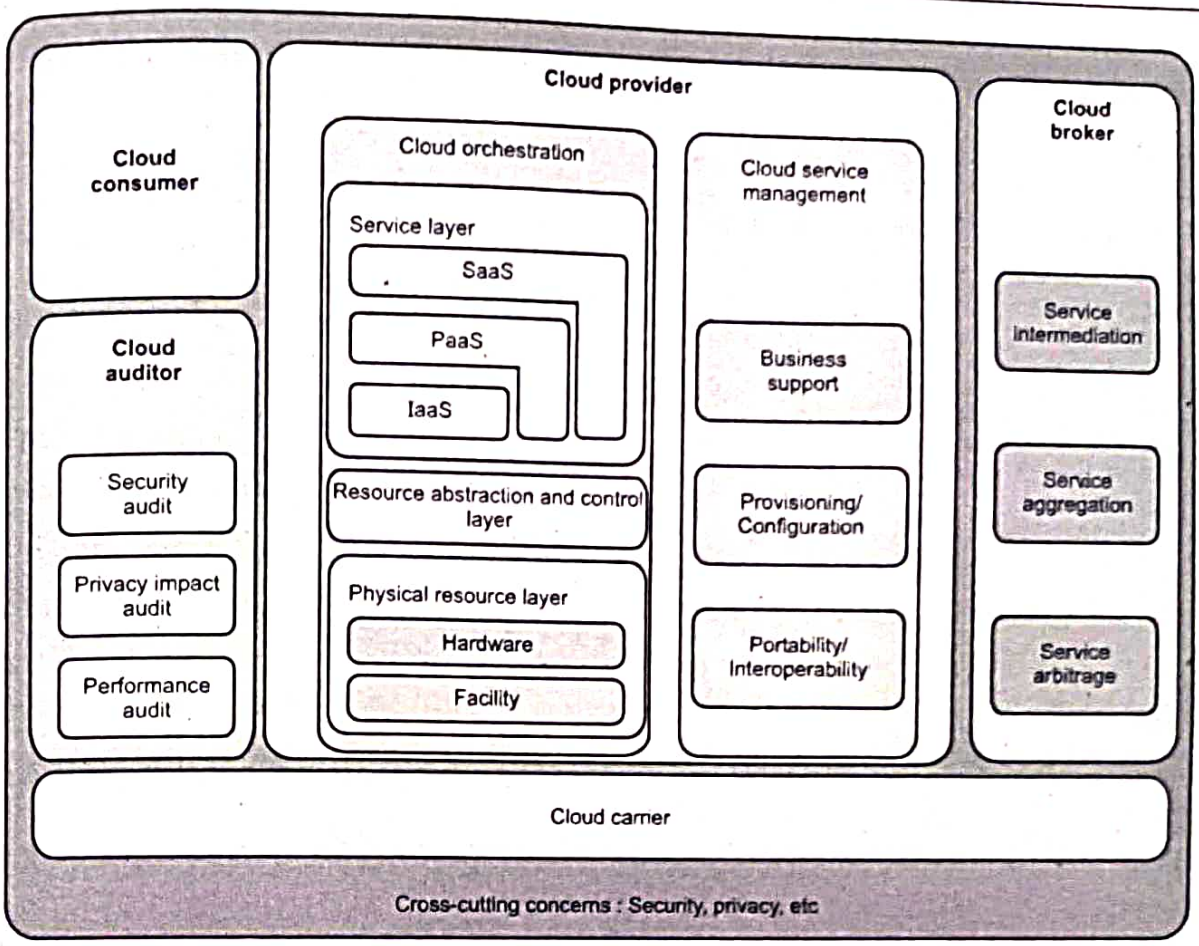


Fig. 5.1.1 Cloud computing security reference architecture

Actor	Definition
Cloud consumer	A person or organization that maintains a business relationship with and uses service form, *Cloud Providers.*
Cloud provider	A person, organization or entity responsible for making a service available to interested parties.
Cloud auditor	A party that can conduct an independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
Cloud broker	An entity that manages the use, performance and delivery of cloud services and negotiates relationships between *Cloud Providers* and *Cloud Consumers.*
Cloud carrier	An intermediary that provides connectivity and transport of cloud services from *Cloud Providers* to *Cloud Consumers.*

5.1.3 Cloud Security Services

- The basic security services for information security include assurance of data confidentiality, integrity and availability.
- Fig. 5.1.2 shows organization of data security and privacy in cloud computing.

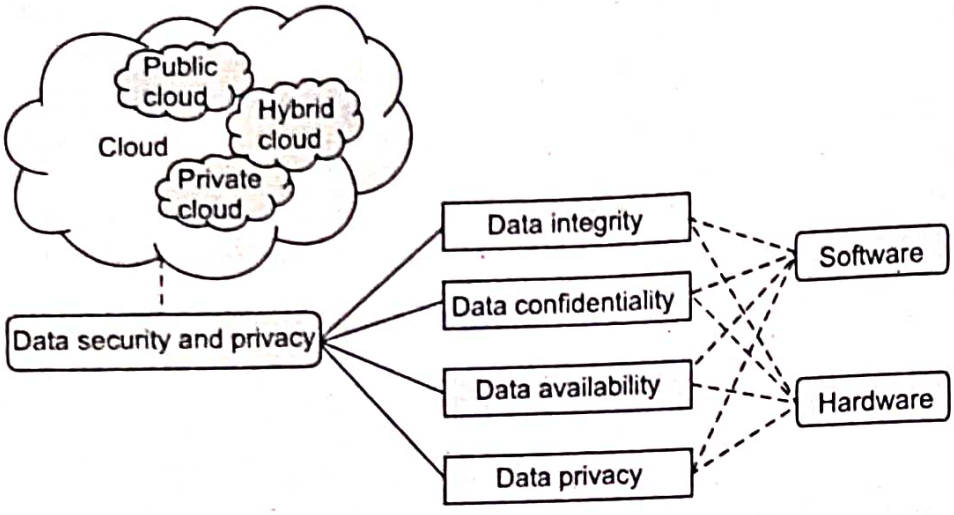


Fig. 5.1.2

1. Confidentiality :

- Confidentiality refers to limiting information access. Sensitive information should be kept secret from individuals who are not authorized to see the information. In cloud environments, confidentiality primarily pertains to restricting access to data in transit and storage.
- Data confidentiality is important for users to store their private or confidential data in the cloud. Authentication and access control strategies are used to ensure data confidentiality.
- The data confidentiality, authentication and access control issues in cloud computing could be addressed by increasing the cloud reliability and trustworthiness.
- Because the users do not trust the cloud providers and cloud storage service providers are virtually impossible to eliminate potential insider threat, it is very dangerous for users to store their sensitive data in cloud storage directly.
- Simple encryption is faced with the key management problem and cannot support complex requirements such as query, parallel modification and fine-grained authorization.

2. Integrity :

- This service protects data from malicious modification. When having outsource their data to remote cloud servers, cloud users must have a way to check whether or not their data at rest or in transit are intact. Such a security service would be of the core value to cloud users.
- Integrity can extend to how data is stored, processed and retrieved by cloud services and cloud-based IT resources.

- Data integrity in the cloud system means preserving information integrity. The data should not be lost or modified by unauthorized users.
- Data integrity in the cloud system means preserving information integrity. The data should not be lost or modified by unauthorized users.
- Data integrity is the basis to provide cloud computing service such as SaaS, PaaS and IaaS.
- Besides data storage of large-scaled data, cloud computing environment usually provides data processing service. Data integrity can be obtained by techniques such as RAID-like strategies and digital signature.

3. Availability :

- This service assures that data stored in the cloud are available on each user retrieval request. This service is particularly important for data at rest in cloud servers and related to the fulfillment of service level agreement.
- Data availability means the following : When accidents such as hard disk damage, IDC fire, and network failures occur, the extent that user's data can be used or recovered and how the users verify their data by techniques rather than depending on the credit guarantee by the cloud service provider alone.
- The cloud service provider should ensure the data security, particularly data confidentiality and integrity. The cloud provider should share all such concerns with the client and build trust relationship in this connection. The cloud vendor should provide guarantees of data safety and explain jurisdiction of local laws to the clients.
- Disaster recovery plan is a plan designed to recover all the vital business processes during a disaster with in a limited amount of time. This plan has all the procedures required to handle the emergency situations.
- A disaster recovery process should have provable recovery capability, and hence it provides the most efficient method to be adopted immediately after a disaster occurs.

5.1.4 Security Authorization Challenges in Cloud

- Authorization is the function of specifying access rights/privileges to resources related to information security and computer security in general and to access control in particular.
- Authorization determines what the user can access and what he cannot access

1. Auditing :

- Cloud security audit can help by assessing and prioritizing risks, evaluating current controls, identifying the gaps in existing cloud security strategy and programs and making recommendations tied to business priorities.
- Functions performed by IT auditors :
  a. Backup controls
  b. Data center security
  c. System development standards
  d. System and transaction controls
  e. Contingency plan.

2. Accountability :

- This is the process that keeps track of a user's activity while attached to a system; the trail included the amount of time attached, the resources accessed, and how much data transferred.
- Accounting data is used for trending, detecting breaches and forensic investigating. Keeping track of users and their activities serves many purposes.
- For example, tracing back to events leading up to a cyber security incident can prove very valuable to a forensics analysis and investigation case.

5.1.5 Cloud Security Threats

1. Traffic eavesdropping

- Data being passively intercepted by a malicious service agent for illegitimate information gathering purpose while being transferred to or within a cloud

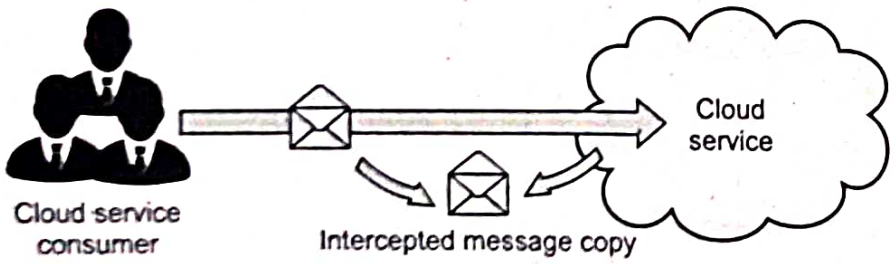


Fig. 5.1.3

- Aim to discredit the confidentiality of data and the relationship between the cloud consumer and cloud provider.
- It is hard to detect for a long period of time because of passive nature of the attack.

2. Malicious intermediary

- Messages intercepted and altered by a malicious service agent discrediting the message's confidentiality and/or integrity.
- Possible malicious contents insertion before forwarding it to its destination.

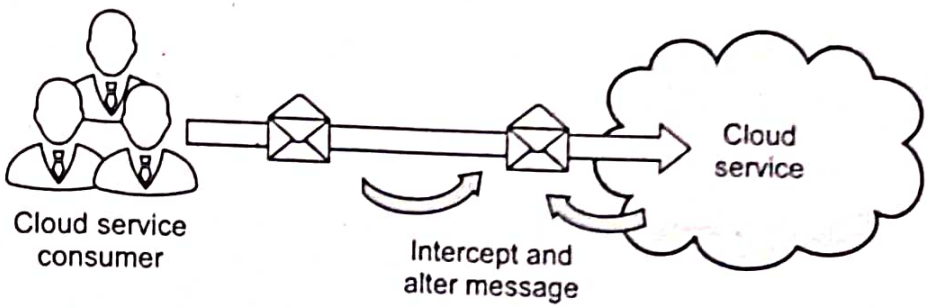


Fig. 5.1.4

3. Denial of Service (DoS)

- Intentional sabotage on shard physical IT resource by overloading it so that the IT resource can hardly be allocated to other consumers sharing the same IT resource.
- Typically intentional overloading shared IT resource by generating excessive messages, consuming full network bandwidth, or sending multiple requests that consume excessive CPU time and memory.

4. Insufficient authorization

- A case when access is granted to an attacker erroneously or too broadly, resulting in the attacker getting access to IT resources that are normally protected.
- Another case **(Weak Authentication)** when weak passwords or shared accounts are used to protect IT resources.

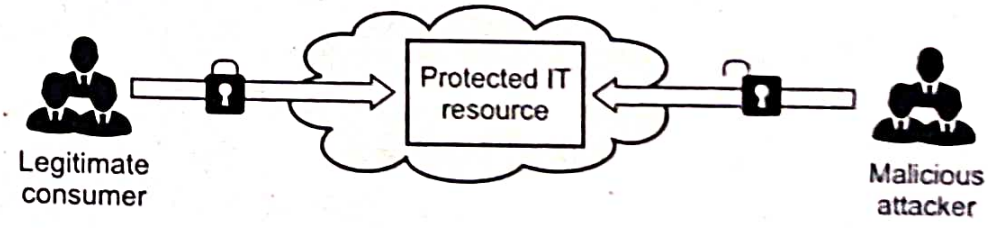


Fig. 5.1.5

5. Virtualization attack (Overlapping Trust Boundaries)

- Physical resources shared by multiple virtual users in virtualized environment by the nature of resource virtualization.
- Possible inherent risk that some cloud consumers could abuse their access right to attack the underlying physical IT resources.

5.1.6 Secure Cloud Software Requirement

- Requirements of secure cloud software are as follows :
  1. **Secure development practices :** It includes data handling, code practices, language options, input validation and content injection, physical security of the system.
  2. **Approaches to cloud software requirements engineering :** A resource perspective on cloud software security requirements, goal-oriented software security requirements and monitoring internal and external requirements.
  3. **Cloud security policy implementation and decomposition :** Includes implementation issues, decomposing critical security issues into secure cloud software requirements (Confidentiality, integrity, availability, authentication and identification, authorization, auditing).

5.2 Virtualization System - specific Attacks

- Cloud computing security challenges fall into three broad categories :
  1. **Data protection :** Securing your data both at rest and in transit.
  2. **User authentication :** Limiting access to data and monitoring who accesses the data.
  3. **Disaster and data breach :** Contingency planning.
- Data protection : Data needs to be encrypted at all times, with clearly defined roles when it comes to who will be managing the encryption keys.
- User authentication : Data resting in the cloud needs to be accessible only by those authorized to do so, making it critical to both restrict and monitor who will be accessing the company's data through the cloud. In order to ensure the integrity of user authentication, companies need to be able to view data access logs and audit trails to verify that only authorized users are accessing the data.
- Contingency planning : With the cloud serving as a single centralized repository for a company's mission-critical data, the risks of having that data compromised due to a data breach or temporarily made unavailable due to a natural disaster are real concerns.
- If information is encrypted while passing through the cloud, who controls the encryption/decryption keys ? Is it the customer or the cloud vendor ? Most customers probably want their data encrypted both ways across the Internet using secure sockets layer protocol.
- They also most likely want their data encrypted while it is at rest in the cloud vendor's storage pool. Be sure that you, the customer, control the

encryption/decryption keys, just as if the data were still resident on your own servers.

- Data integrity means ensuring that data is identically maintained during any operation.
- Cloud-based services will result in many mobile IT users accessing business data and services without traversing the corporate network. This will increase the need for enterprises to place security controls between mobile users and cloud-based services.
- Placing large amounts of sensitive data in a globally accessible cloud leaves organizations open to large distributed threats, attackers no longer have to come onto the premises to steal data, and they can find it all in the one "virtual" location.
- Virtualization efficiencies in the cloud require virtual machines from multiple organizations to be co-located on the same physical resources. Although traditional data center security still applies in the cloud environment, physical segregation and hardware-based security cannot protect against attacks between virtual machines on the same server.
- Operating system and application files are on a shared physical infrastructure in a virtualized cloud environment and require system, file, and activity monitoring to provide confidence and auditable proof to enterprise customers that their resources have not been compromised or tampered with.
- In the cloud computing environment, the enterprise subscribes to cloud computing resources, and the responsibility for patching is the subscriber's rather than the cloud computing vendor's.
- The need for patch maintenance vigilance is imperative. Lack of due diligence in this regard could rapidly make the task unmanageable or impossible, leaving you with "virtual patching" as the only alternative.
- Confidentiality : Confidentiality refers to limiting information access. Sensitive information should be kept secret from individuals who are not authorized to see the information.
- In cloud environments, confidentiality primarily pertains to restricting access to data in transit and storage.
- Integrity can extend to how data is stored, processed, and retrieved by cloud services and cloud-based IT resources.
- Some common cloud security threats include :
  a) Risks of cloud-based infrastructure including incompatible legacy IT frameworks, and third-party data storage service disruptions.

b) Internal threats due to human error such as misconfiguration of user access controls.

c) External threats caused almost exclusively by malicious actors, such as malware, phishing, and DDoS attacks.

5.2.1 Guest - hopping Attack

- In guest-hopping attacks, due to the separation failure between shared infrastructures, an attacker gets access to a virtual machine by penetrating another virtual machine hosted in the same hardware.
- One possible mitigation of guest-hopping attack is the Forensics and VM debugging tools to observe any attempt to compromise the virtual machine.
- Another solution is to use the High Assurance Platform (HAP), which provides a high degree of isolation between virtual machines.
- **Guest to host attack/guest escape :** Once the attacker has found a vulnerability in the virtualization layer in combine with improper configurations of both the host and the guest, attacker can bypass the virtualization layer and access the host machine.
- Since the host machine contains multiple guests, the attacker can control all the guest machines and monitor any interaction between the guests and the host. In addition, the attacker can lunch various attacks, like, corrupting resources, memory, CPU and launch arbitrary code.
- **Guest to guest attack / guest hopping :** In this attack, the attacker can inject a malware in one guest, and once attacker gets a control over the virtual machine, they can spread this malware to other virtual machines or attacking the virtualization layer itself. Thus, controlling all the virtual machines that exist on the host machine. The attacker then can monitor the usage of various resources, like, CPU, memory, etc. which affects the confidentiality of the guest machine.
- In addition, the attacker has the ability to manipulate existing data in the virtual machines, modifying their configurations, injecting malicious code, etc. Thus, affecting the integrity and the availability of the data.
- **Guest mobility :** Guest machine contents are stored as files in the host machine's hard desk drive, thus, easing the process of transferring or copying the contents of one guest to another host through the network.
- With this usability, security problems arise, if the guest is infected with malicious malware, the other host will be contaminated with the same malware. Thus, the

attacker will have control over multiple virtual machines on multiple hosts and possibly use the same technique to affect multiple virtual machines.

- **Guest denial of service attack :** In virtualization, the host machine allocates resources such as RAM, CPU, storage and network bandwidth for each guest machine. DOS attack occurs when one guest machine occupies all the resources resulting in denying other guest machines from utilizing host's resources.
- **Virtual machine overflow :** In this attack, the attacker runs a malicious script on the guest machine and fills the allocated memory region with meaningless characters, exceeding the allowed boundaries for the guest machine and as a result the machine crashes.
- After that, the attacker can access the host's memory pointer's and directing them to run the attacker's malicious script. By that, the attacker can gain root access over the host machine and thus having access over all the guest machines that resides in the host machine.
- **Virtualization memory leak :** Each guest machine has a specific space in host's memory and if the host did not properly free the allocated memory, a virtual memory leak can occur.
- The attacker can exploit this vulnerability by using this allocated space to execute several attacks, like DOS and buffer overflow attack.

5.2.2 VM Migration Attack : Hyperjacking

- Hyperjacking is another illicit method that can be used to spy on victims, control devices and steal valuable information. Hyperjacking involves the compromise and unauthorized control of a virtual machine.
- Hypervisors form the backbone of virtual machines. These are software programs that are responsible for creating, running and managing VMs. A single hypervisor can host multiple virtual machines, or multiple guest operating systems, at one time, which also gives it the alternative name of Virtual Machine Manager (VMM).
- There are two kinds of hypervisors. The first is known as a "**bare metal**" or "native" hypervisor, with the second being a "host" hypervisor.
- Hyperjacking involves installing a rogue hypervisor that can take complete control of a server. Regular security measures are ineffective because the OS will not even be aware that the machine has been compromised.
- Hypervisors are the key target of hyperjacking attacks. In a typical attack, the original hypervisor will be replaced via the installation of a rogue, malicious hypervisor that the threat actor has control of. By installing a rogue hypervisor

under the original, the attacker can therefore gain control of the legitimate hypervisor and exploit the VM.

- By having control over the hypervisor of a virtual machine, the attacker can, in turn, gain control of the entire VM server. This means that they can manipulate anything in the virtual machine.
- This mechanism is due to a lack of separation between control flows and data flows, guest OS access to the hypervisor (e.g., via a management tool on the guest OS), or an unpatched system. The exploitation of this mechanism can result in the attacker gaining unlimited access to the entire virtualization server and the guest VMs. This attack mechanism can result from poorly managed control and data flows as well as poorly managed shared access to resources.

Virtual machine migration services :

- VM migration are of two types : Hot migration and cold migration.

i) Hot migration :

- A hot migration is referred to as a live migration. It is a staged migration where the virtual machine stays powered on during the initial full synchronization and the subsequent delta sync, using the vSphere vMotion feature.
- There are two types of hot migration :
  1. **Compute resource -** A migration of a virtual server from one compute resource to another.
  2. **Full migrate -** A migration of a virtual server with or without disks and NICs between compute resources, data stores and networks.
- The live migration process transfers the VM memory, network connectivity and storage as the OS continues to run. The obvious advantages of a live migration are that we do not have to interrupt operations.
- The best time to do a live migration on VMware is when server needs maintenance or an update, or when we need to switch a VM to a different host.
- The process allows for :
  1. A clean separation between hardware and software, including the separation of concerns between the users and operator of a data center or cluster.
  2. Consolidation of clustered hardware into a single management domain. This means that if we need to remove a certain physical machine from service for maintenance, we can migrate OS instances to one or more alternative machines to relieve the load on congested host machines.
- Live migration can also be used for load balancing in which work is shared among computers in order to optimize the utilization of available CPU resources

ii) Cold migration

- Cold migration involves moving a powered-off or suspended virtual machine to a new host. It also usually means relocating configuration and disk files for these powered-off or suspended virtual machines to new storage locations.
- Cold migration includes moving virtual machines from one virtual switch to another or from one data center to another. Fig. 5.2.1 shows cold migration.

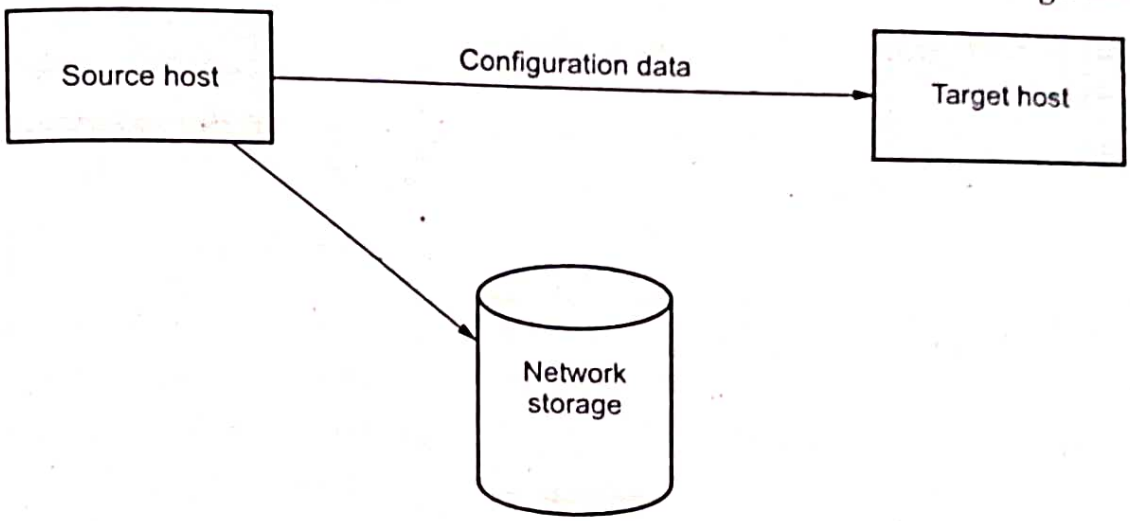


Fig. 5.2.1 Cold migration

- Cold migration is easy to implement and is summarized as follows :
  a) The configuration files, including NVRAM file, log files and the disks of the virtual machines, are moved from the source host to the destination host's associated storage area.
  b) The virtual machine is registered with the new host.
  c) After the migration is completed, the old version of the virtual machine is deleted from the source host.

5.3 Data Security and Storage

- Cloud computing security challenges fall into three broad categories :
  1. **Data protection :** Securing your data both at rest and in transit
  2. **User authentication :** Limiting access to data and monitoring who accesses the data
  3. **Disaster and data breach :** Contingency planning.
- Data protection : Data needs to be encrypted at all times, with clearly defined roles when it comes to who will be managing the encryption keys.
- User authentication : Data resting in the cloud needs to be accessible only by those authorized to do so, making it critical to both restrict and monitor who will be accessing the company's data through the cloud. In order to ensure the integrity

of user authentication, companies need to be able to view data access logs and audit trails to verify that only authorized users are accessing the data.

- Contingency planning : With the cloud serving as a single centralized repository for a company's mission-critical data, the risks of having that data compromised due to a data breach or temporarily made unavailable due to a natural disaster are real concerns.
- **Security challenges for cloud service customers :**
  1. **Ambiguity in responsibility :** A CSC uses services based on different service categories as well as different deployment models. If the responsibilities are not clearly defined in any of these cases then it may result in inconsistency or may leave an open gate for attacks.
  2. **Loss of trust :** Because of the abstraction of the security implementation details between a CSC and a CSP, it is difficult for a CSC to get details of the security mechanisms that the CSP has implemented to keep the cloud data secure.
  3. **Loss of governance :** When the CSC uses cloud services, it has to move its data onto the cloud and has to provide certain privileges to the CSP for handling the data in the cloud. This may result in misconfiguration or an attack due to the abstraction of the CSP's cloud practices and due to the privileges that need to be given to the CSP.
  4. **Loss of privacy :** CSC's privacy may be violated due to leakage of private information while the CSP is processing CSC's private data or using the private information for a purpose that the CSP and CSC haven't agreed upon.
  5. **Cloud service provider lock-in :** This issue arises if a CSP doesn't abide by the standard functions or frameworks of cloud computing and hence makes it difficult for a CSC using its services to migrate to any other CSP. The use of non-standard functions and cloud framework makes the CSP non-inter-operable with other CSPs and also leaves CSC open to security attacks.
  6. **Misappropriation of intellectual property :** A CSC may face this challenge due to the possibility that a CSC's data on the cloud might leak to third parties that are using the same CSP for their cloud services. This leakage may violate the CSC's copyrights and may result in the disclosure of CSC's private data.
  7. **Loss of software integrity :** A CSC encounters this challenge due to the fact that its software is running in the cloud once it is given to the CSP. It is possible that this software might be tampered with or might be affected while the software is running in the CSP and is not in CSC's control, resulting in CSC's loss over its software.

5.3.1 Advantages

- **Data centralization :** service provider takes responsibility of storage and small organization need not spend more money for personal storage device.
- **Incident response :** IaaS providers contribute dedicated legal server which can be used on demand.
- Forensic image verification time.
- **Logging :** storage requirement for benchmark logs is mechanically solved.

5.3.2 Disadvantages

- Loss of control : The enterprise's loss of control in enhancing the network's security is the most significant disadvantage of cloud computing security. The responsibility of securing the network is shared between the Cloud Service Provider (CSP) and the enterprise.
- Reduced visibility and control : when migrating to a cloud based computing model, organizations will lose a degree of visibility and control, with some responsibility for policies and infrastructure moving to the cloud provider.
- Unsecure API and interfaces.
- Data segregation.

5.4 Identity and Access Management (IAM) AU : Dec.-21

- Identity and Access Management (IAM) can help a user to manage to compute, store, manage and application services in the AWS cloud. It uses access control techniques through which a user is familiar with which includes users, groups and permission.
- With the help of a single AWS IAM, the user can manage the customer and their needs. It provides Amazon AWS building blocks which help the user to build the applications for the security purpose.
- AWS identity and access management help the user to focus on the features and functionality which includes the security on the other side of the things. AWS IAM can also rotate access keys on the virtual machine instances.
- Functions :
  1. To manage AWS IAM users and their access.
  2. To manage Amazon IAM roles and their permissions.
  3. To manage to federate users and their permissions.

5.4.1 Identity Management and Access Control

- AWS Identity and Access Management (IAM) is a web service that helps you securely control access to AWS resources.
- When you first create an AWS account, you begin with a single sign-in identity that has complete access to all AWS services and resources in the account. This identity is called the AWS account root user and is accessed by signing in with the email address and password that you used to create the account.
- For each AWS account, you can create multiple users with different credentials. For each user, you can give different rights.
- **IAM Users** are account objects that allow an individual user to access your AWS environment with a set of credentials. You can issue user accounts to anyone you want to view or administer objects and resources within your AWS environment. Permissions can be applied individually to a user, but the best practice for permission assignments is to assign them via the use of groups.
- **IAM groups** are objects that have permissions assigned to them via policies allowing the members of the group access to specific resources. Having users assigned to these groups allows for a uniform approach to access management and control.
- **IAM roles** are again objects created within IAM which have policy permissions associated to them. However, instead of being associated with users as groups are, roles are assigned to instances at the time of launch. This allows the instance to adopt the permissions given by the role without the need to have access keys stored locally on the instance.
- Security groups are used to control access to EC2 instances. Because AWS uses flat Layer 3 networking, any instance within a user account can communicate with any other instance.
- AWS Identity Access Management allows to establish access rules and permissions to specific users and applications.
  1. Set up permissions for users and applications.
  2. Create user groups for common rules assignment.
  3. Cloud Trail allows to monitor the access.
  4. Identity federation : allow users to log in with their company credentials.
  5. Temporary security credentials, obtained by calling AWS STS APIs like AssumeRole or GetFederationToken.
- **IAM policy** - A document that defines the effect, actions, resources, and optional conditions.
- **IAM role** - An identity with permission policies, to which users can be assigned.

- **IAM group** - A group of users to which common policies can be attached.
- Best practices regarding security groups are as follows :
  1. Avoid using the default security group.
  2. Use meaningful names.
  3. Open only the ports you need to open.
  4. Partition applications.
  5. Restrict system administrator access.

5.4.2 Security Policies

- User can manage access in AWS by creating policies and attaching them to IAM identities or AWS resources.
- A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. AWS evaluates these policies when a principal entity (user or role) makes a request.
- IAM policies define permissions for an action regardless of the method that you use to perform the operation.

Types of Policy :

1. **Identity-based policies :** Attach managed and inline policies to IAM identities (users, groups to which users belong, or roles). Identity-based policies grant permissions to an identity.
2. **Resource-based policies :** Attach inline policies to resources. For example : resource-based policies are Amazon S3 bucket policies and IAM role trust policies. Resource-based policies grant permissions to a principal entity that is specified in the policy. Principals can be in the same account as the resource or in other accounts.
3. **Permissions boundaries :** Use a managed policy as the permissions boundary for an IAM entity (user or role). That policy defines the maximum permissions that the identity-based policies can grant to an entity, but does not grant permissions.
4. **Organizations SCPs :** Use an AWS Organizations Service Control Policy (SCP) to define the maximum permissions for account members of an organization or Organizational Unit (OU). SCPs limit permissions that identity-based policies or resource-based policies grant to entities (users or roles) within the account, but do not grant permissions.
5. **Access Control Lists (ACLs) :** Use ACLs to control which principals in other accounts can access the resource to which the ACL is attached. ACLs are similar to resource-based policies, although they are the only policy type that does not use the JSON policy document structure. ACLs are cross-account permissions

policies that grant permissions to the specified principal entity. ACLs cannot grant permissions to entities within the same account.

6. **Session policies :** Pass an advanced session policy when you use the AWS CLI or AWS API to assume a role or a federated user. Session policies limit the permissions that the role or user's identity-based policies grant to the session. Session policies limit permissions for a created session, but do not grant permissions.

5.4.3 IAM Abilities and Limitation

- Path names must begin and end with a forward slash (/).
- Names of users, groups, roles, policies, instance profiles and server certificates must be alphanumeric, including the following common characters : plus (+), equal (=), comma (,), period (.), at (@), underscore (_), and hyphen (-).
- Names of users, groups and roles must be unique within the account.
- User passwords (login profiles) can contain any Basic Latin (ASCII) characters.

5.4.4 Machine Imaging

- Machine imaging is a process that is used to provide system portability and provision and deploy systems in the cloud through capturing the state of systems using a system image.
- A system image makes a copy or a clone of the entire computer system inside a single file. The image is made by using a program called **system imaging program** and can be used later to restore a system image.
- For example : Amazon Machine Image (AMI) is a system image that is used in the cloud computing. The Amazon Web Services uses AMI to store copies of a virtual machine.
- An AMI is a file system image that contains an operating system, all device drivers and any applications and state information that the working virtual machine would have.
- The AMI files are encrypted and compressed for security purpose and stored in Amazon S3 (Simple Storage System) buckets as a set of 10 MB chunks.
- Machine imaging is mostly run on virtualization perform due to this it is also called as **virtual appliances** and running virtual machines are called **instances.**
- The AMI file system is not a standard bit-for-bit image of a system that is common to many disk imaging programs. AMI omits the kernel image and stores a pointer to a particular kernel that is part of the AWS kernel library.

- Among the choices are Red Hat Linux, Ubuntu, Microsoft Windows, Solaris and others. Files in AMI are compressed and encrypted and an XML file is written that describe the AMI archive.
- Machine Images are sometimes referred to as "virtual appliances", systems that are meant to run on virtualization platforms.

5.4.5 IAM Challenges

- The major challenges faced by the IAM in the cloud are as follows :

1. Identity provisioning / de-provisioning

- This concerns with providing a secure and timely management of on-boarding (provisioning) and off-boarding (de-provisioning) of users in the cloud.
- When a user has successfully authenticated to the cloud, a portion of the system resources in terms of CPU cycles, memory, storage and network bandwidth is allocated. Depending on the capacity identified for the system, these resources are made available on the system even if no users have been logged on.

2. Maintaining a single ID across multiple platforms and organizations

- It is tough for the organizations to keep track of the various logins and ID that the employees maintain throughout their tenure. The centralised federated identity management is the answer for this issue. Here users of cloud services are authenticated using a company chosen identity provider.

3. Security when using 3rd party or vendor network

- A lot of services and applications used in the cloud are from 3rd party or vendor networks. You may have secured your network, but can not guarantee that their security is adequate.

4. Compliance visibility : Who has access to what ?

- When it comes to cloud services, it's important to know who has access to applications and data, where they are accessing it and what they are doing with it.
- IAM should be able to provide a centralised compliance reports across access rights, provisioning/de-provisioning and end-user and administrator activity. There should be a central visibility and control across all your systems for auditing purposes.
- Identity and access management is an important aspect of any business. It's a process that allows organizations to manage user access to data and resources and ensures the security of that data. While the process is not easy, it is important to get it right so it does not become a roadblock to your business. This can be achieved by having the right tools in place and following best practices.

5.4.6 IAM Architecture and Practice

- Fig. 5.4.1 shows architecture of IAM.

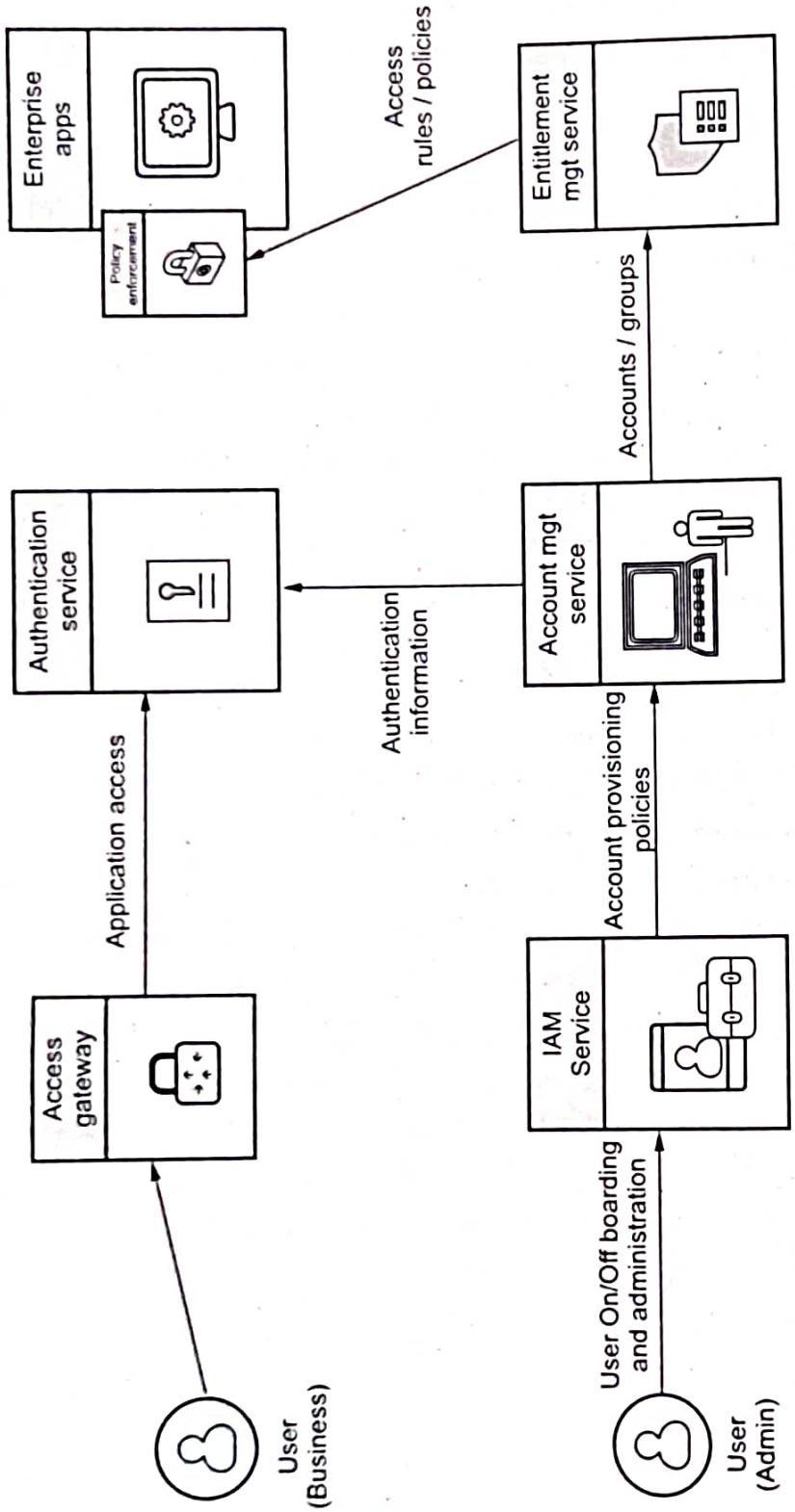


Fig. 5.4.1 Architecture of IAM

- **User management :** It consists of activities for the control and management over the identity life cycles.
- **Authentication management :** It consists of activities for effectively controlling and managing the processes for determining which user is trying to access the services and whether those services are relevant to him or not.
- **Authorization management :** It consists of activities for effectively controlling and managing the processes for determining which services are allowed to access according to the policies made by the administrator of the organization.
- **Access management :** It is used in response to a request made by the user wanting to access the resources with the organization.
- **Data management and provisioning :** The authorization of data and identity are carried towards the IT resource through automated or manual processes.
- **Monitoring and auditing :** Based on the defined policies the monitoring, auditing and reporting are done by the users regarding their access to resources within the organization.
- **Operational activities of IAM :** In this process, we onboard the new usérs on the organization's system and application and provide them with necessary access to the services and data.
- **Credential and attribute management :** Credentials are bound to an individual user and are verified during the authentication process. These processes generally include allotment of username, static or dynamic password, handling the password expiration, encryption management and access policies of the user.
- **Entitlement management :** These are also known as **authorization policies** in which we address the provisioning and de-provisioning of the privileges provided to the user for accessing the databases, applications and systems.
- **Identity federation management :** In this process, we manage the relationships beyond the internal networks of the organization that is among the different organizations. The federations are the associate of the organization that came together for exchanging information about the user's resources to enable collaboration and transactions.
- **Centralization of authentication and authorization :** It needs to be developed in order to build custom authentication and authorization features into their application, it also promotes the loose coupling architecture.

5.4.7 Single Sign - On

- A mechanism enabling one cloud service consumer to be authenticated by a **security broker** which establishes a security context that is persisted while the

cloud service consumer accesses other cloud services or cloud-based IT resources in order for the cloud service consumer not to **re-authenticate** itself with every subsequent request.

Implementation mechanisms

- Not a trivial job at all to propagate the authentication and authorization information for a cloud service consumer across multiple cloud services, especially with a numerous cloud services or cloud-based IT resources to be invoked as part of **the same overall runtime activity.**
- SSO (or security broker) mechanism to enable mutually independent cloud services and IT resources to generate and circulate **runtime authentication** and **authorization credentials (security token)** in order to allow the credentials provided by the cloud service consumer at its login time to be valid through out the duration of the same session.
- Security brokerage mechanism is especially useful when a cloud service consumer needs to access cloud services residing on different clouds.
- Not to counter security threats directly, but **to enhance** the usability of cloud-based environments for access and management of distributed IT resources and solutions without violating security policies.

University Question

1. *What is IAM and detail the segregation roles carried out by IAM when services of multiple organizations are maintained within the same geographical location ?*

AU : Dec.-21, Marks 13

5.5 Two Marks Questions with Answers

Q.1 Define cloud security.

Ans. : Cloud computing security consists of a set of policies, controls, procedures and technologies that work together to protect cloud-based systems, data and infrastructure. These security measures are configured to protect data, support regulatory compliance and protect customer's privacy as well as setting authentication rules for individual users and devices.

Q.2 Discuss the different cloud security services.

Ans. : Cloud security services are authentication, authorization, auditing and accountability.

Q.3 How security policies are implemented on cloud computing ?

Ans. : Cloud security is a shared responsibility of the cloud provider and customer.

- **Step 1 :** Perform data classification (Statement of sensitivity);
- **Step 2 :** Perform threat risk assessment on the solution;
- **Step 3 :** Address threats/risks identified by implementing the proper controls;
- **Step 4 :** Continuously monitor and periodically audit systems and services.

Q.4 What is multitenancy issue in cloud computing ?

Ans. : A multi-tenant cloud is a cloud computing architecture that allows customers to share computing resources in a public or private cloud. Each tenant's data is isolated and remains invisible to other tenants.

Q.5 Discuss the problem associated with cloud computing.

Ans. :

- Problem associated with cloud computing are security, integration and interoperability, governance and regulatory compliance.
- Some governments or enterprises may need to enforce strict limits on the spatial and temporal existence of data. For example, a government might want to keep the data of its citizens within the country and for an exact duration.

Q.6 What do you understand by virtualization security management ?

Ans. : Virtualization security is the collective measures, procedures and processes that ensure the protection of a virtualization infrastructure / environment. It addresses the security issues faced by the components of a virtualization environment and methods through which it can be mitigated or prevented.

Q.7 What is the difference between identity management and access management ?

Ans. : Identity management confirms that user are user and stores information about us. An identity management database holds information about user identity. Access management uses the information about our identity to determine which software suites we are allowed access to and what we are allowed to do when we access them. For example, access management will ensure that every manager with direct reports has access to an app for timesheet approval, but not so much access that they can approve their own timesheets.

Q.8 What is AWS identity and access management ?

Ans. : Amazon Web Services (AWS) identity and access management is simply the IAM system that is built into AWS. By using AWS IAM, we can create AWS users and groups and grant or deny them access to AWS services and resources. AWS IAM is available free of charge.

❑❑❑

Notes

SOLVED MODEL QUESTION PAPER

[As Per New Syllabus]

Cloud Computing

Vertical - 3 (Cloud Computing and Data Centre Technologies) (CSE/IT/AI&DS)
Vertical - 2 (Full Stack Development for IT) (IT/AI&DS)
Vertical - 2 (Cloud Computing and Data Centre Technologies) (CS&BS)

Time : Three Hours] **[Maximum Marks : 100**

Answer ALL Questions

PART A - (10 × 2 = 20 Marks)

Q.1 *What is meant by IaaS in cloud computing ?* **(Refer Two Marks Q.25 of Chapter - 1)**

Q.2 *What is private clouds ?* **(Refer Two Marks Q.19 of Chapter - 1)**

Q.3 *Define I/O virtualization.* **(Refer Two Marks Q.10 of Chapter - 2)**

Q.4 *What is application server virtualization ?* **(Refer Two Marks Q.6 of Chapter - 2)**

Q.5 *What is networking virtualization ?* **(Refer Two Marks Q.2 of Chapter - 3)**

Q.6 *What is cloud analytics ?* **(Refer Two Marks Q.5 of Chapter - 3)**

Q.7 *How virtualization employed in azure ?* **(Refer Two Marks Q.7 of Chapter - 4)**

Q.8 *What is AWS ecosystem ?* **(Refer Two Marks Q.2 of Chapter - 4)**

Q.9 *Define cloud security.* **(Refer Two Marks Q.1 of Chapter - 5)**

Q.10 *What is multitenancy issue in cloud computing ?* **(Refer Two Marks Q.4 of Chapter - 5)**

PART B - (5 × 13 = 65 Marks)

Q.11 a) **i)** *What is cloud computing ? Explain characteristics, pros and cons of cloud computing.* **(Refer section 1.1)** **[7]**

ii) *Explain cloud service model.* **(Refer section 1.5)** **[6]**

OR

b) **i)** *Discuss briefly infrastructure as a service. What are advantages and disadvantages of IaaS.* **(Refer section 1.8)** **[7]**

ii) *Explain migrating into the cloud.* **(Refer section 1.11)** **[6]**

Q.12 a) *What is hypervisor ? Explain type 1 and type 2 hypervisor. Write difference between type 1 and type 2 hypervisor.* **(Refer section 2.3)** **[13]**

OR

b) **i)** *Explain following : Memory virtualization, I/O virtualization* **(Refer sections 2.5.1 and 2.5.2)** **[6]**

ii) *What is virtual machine ? Explain advantages and disadvantages of virtual machine.* **(Refer section 2.1)** **[7]**

Q.13 a) **i)** *What is docker ? Explain docker architecture.* **(Refer section 3.7)** **[7]**

ii) *Discuss briefly virtual clusters and resource management.* **(Refer section 3.6)** **[6]**

OR

b) **i)** *What do you mean block level virtualization ? Explain difference between block level and file level virtualization.* **(Refer section 3.3)** **[7]**

ii) *Explain about desktop virtualization. Explain types of desktop virtualization.* **(Refer section 3.1)** **[6]**

Q.14 a) *Write short note on following : i) Microsoft azure ii) Eucalyptus* **(Refer sections 4.3 and 4.4)** **[13]**

OR

b) *Write short note on following : i) OpenStack ii) Google App Engine* **(Refer sections 4.5 and 4.1)** **[13]**

Q.15 a) **i)** *Discuss about identity and access management.* **(Refer section 5.4)** **[7]**

ii) *Explain in detail cloud security services.* **(Refer section 5.1.3)** **[6]**

OR

b) **i)** *Explain guest - hopping attack and hyperjacking.* **(Refer section 5.2)** **[7]**

ii) *Discuss cloud security challenges and risks.* **(Refer section 5.1.1)** **[6]**

PART C - (1 × 15 = 15 Marks)

Q.16 a) *What is EC2 instances ? Explain configuring Amazon EC2 Linux instances.* **(Refer section 4.2.3)** **[15]**

OR

b) *What is public and private cloud ? Explain difference between public and private cloud.* **(Refer section 1.4)** **[15]**

❑❑❑