MCA 401A: Cloud Computing

UNIT I

Cloud Architecture and Model: Technologies for Network-Based System – System Models for Distributed and Cloud Computing – NIST Cloud Computing Reference Architecture, Cloud Models: Characteristics – Cloud Services – Cloud models (IaaS, PaaS, SaaS) – Public Vs Private Cloud –Cloud Solutions - Cloud ecosystem – Service management – Computing on demand.

UNIT II

Virtualization: Basics of Virtualization - Types of Virtualization - Implementation Levels of Virtualization - Virtualization Structures - Tools and Mechanisms - Virtualization of CPU, Memory, I/O Devices - Virtual Clusters and Resource management – Virtualization for Data-center Automation. VMWare, Virtual Box Virtualization software.

UNIT III

Cloud Infrastructure: Architectural Design of Compute and Storage Clouds – Layered Cloud Architecture Development – Design Challenges - Inter Cloud Resource Management – Resource Provisioning and Platform Deployment – Global Exchange of Cloud Resources. Federation in the Cloud – Four Levels of Federation – Federated Services and Applications – Future of Federation

UNIT IV

Programming Model: Parallel and Distributed Programming Paradigms – MapReduce , Twister and Iterative MapReduce – Hadoop Library from Apache – Mapping Applications - Programming Support – Software environments for service development; Amazon, Azure, GoogleApp Engine, AWS - Cloud Environments -Eucalyptus, Open Nebula, OpenStack, Aneka, CloudSim. Cloud Storage – Storage-as-a-Service – Advantages of Cloud Storage – Cloud Storage Providers – S3.

UNIT V

Security In The Cloud : Security Overview – Cloud Security Challenges and Risks – Software-as-a-Service Security – Security Governance – Risk Management – Security Monitoring – Security Architecture Design – Data Security – Application Security – Virtual Machine Security - Identity Management and Access Control.

Text Books:

1. Kai Hwang, Geoffrey C Fox, Jack G Dongarra, "Distributed and Cloud Computing, From Parallel
   Processing to the Internet of Things", Morgan Kaufmann Publishers, 2012.
2. Rittinghouse, John W., and James F. Ransome, ―Cloud Computing: Implementation, Management and
   Security‖, CRC Press, 2017.

Reference Books

1. John W.Rittinghouse and James F.Ransome, "Cloud Computing: Implementation, Management, and Security", CRC Press, 2010.

2. Kumar Saurabh, “Cloud Computing – insights into New-Era Infrastructure”, Wiley India,2011.

Lecture notes

Unit-1

Cloud computing:

Cloud computing[1] is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user.[2] Large clouds often have functions distributed over multiple locations, each of which is a data center. Cloud computing relies on sharing of resources to achieve coherence and typically uses a "pay as you go" model, which can help in reducing capital expenses but may also lead to unexpected operating expenses for users.

Technologies for network:

A list of cloud computing technologies are given below -

- Virtualization
- Service-Oriented Architecture (SOA)
- Grid Computing
- Utility Computing

Virtualization

Virtualization is the process of creating a virtual environment to run multiple applications and operating systems on the same server. The virtual environment can be anything, such as a single instance or a combination of many operating systems, storage devices, network application servers, and other environments.

The concept of Virtualization in cloud computing increases the use of virtual machines. A virtual machine is a software computer or software program that not only works as a physical computer but can also function as a physical machine and perform tasks such as running applications or programs as per the user's demand.

Types of Virtualization

A list of types of Virtualization is given below -

i. Hardware virtualization
ii. Server virtualization

iii. Storage virtualization

iv. Operating system virtualization

v. Data Virtualization

Service-Oriented Architecture (SOA)

Service-Oriented Architecture (SOA) allows organizations to access **on-demand** cloud-based computing solutions according to the change of business needs. It can work without or with cloud computing. The advantages of using SOA is that it is easy to maintain, platform independent, and highly scalable.

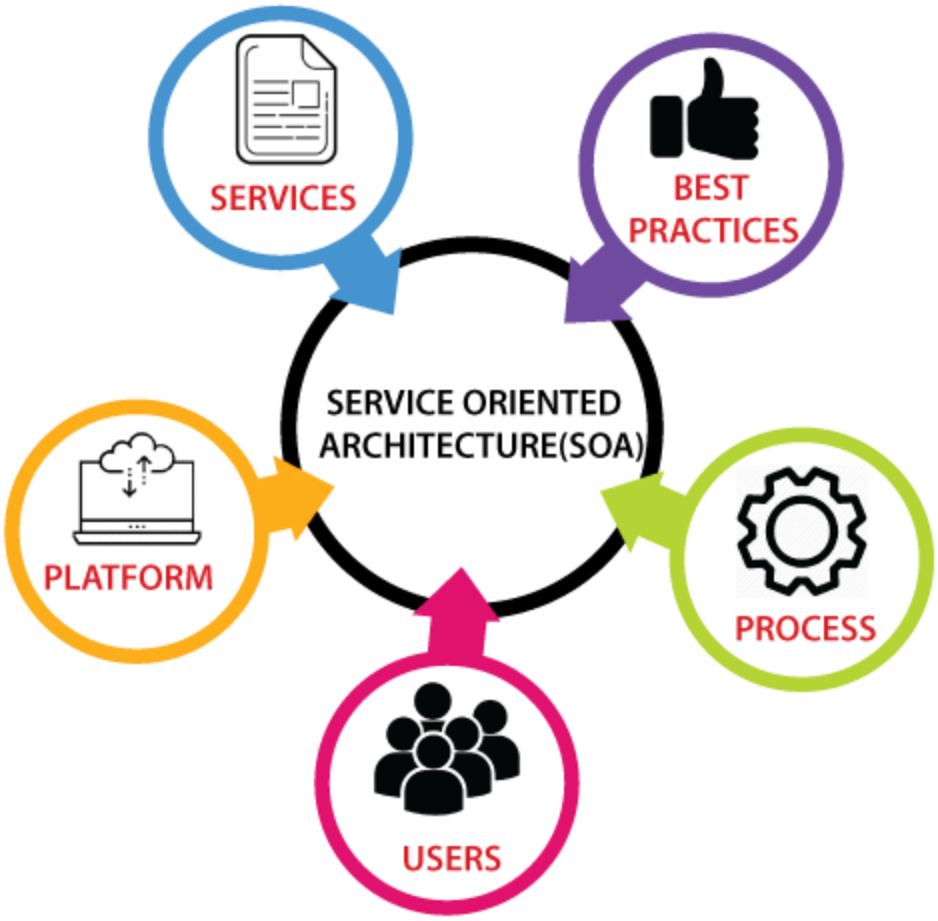
Service Provider and Service consumer are the two major roles within SOA.

Applications of Service-Oriented Architecture

There are the following applications of Service-Oriented Architecture -

- It is used in the healthcare industry.
- It is used to create many mobile applications and games.
- In the air force, SOA infrastructure is used to deploy situational awareness systems.

The service-oriented architecture is shown below:



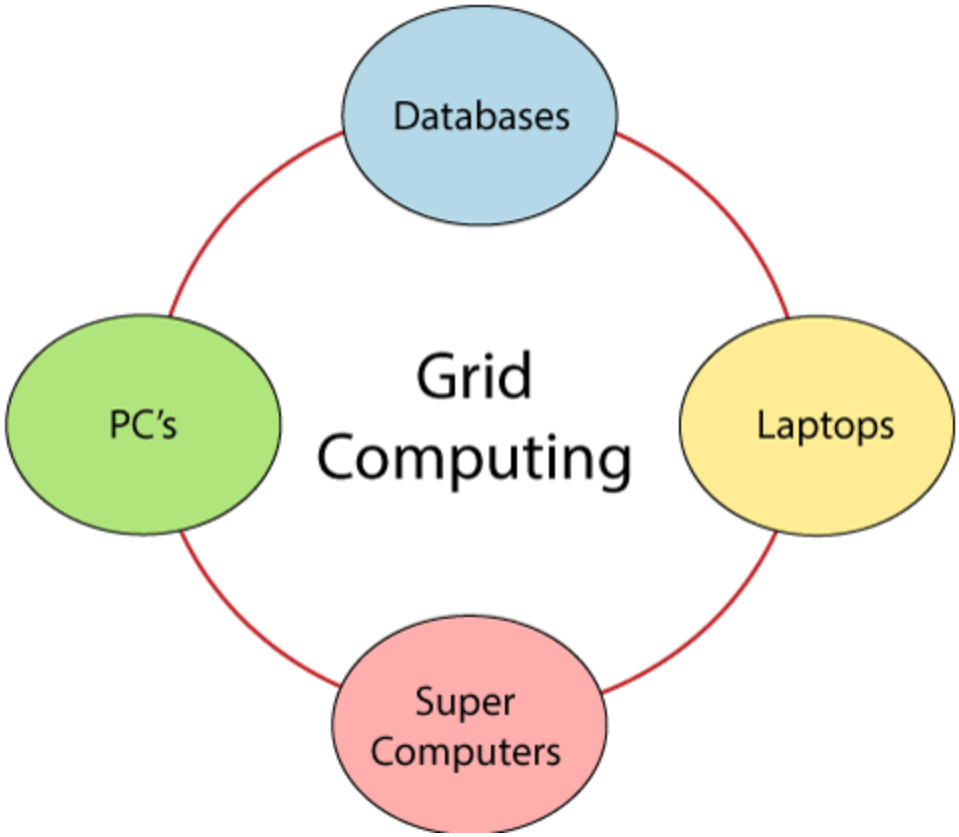
Grid Computing

Grid computing is also known as **distributed computing**. It is a processor architecture that combines various different computing resources from multiple locations to achieve a common goal. In grid computing, the grid is connected by parallel nodes to form a computer cluster. These computer clusters are in different sizes and can run on any operating system.

Grid computing contains the following three types of machines -

1. **Control Node:** It is a group of server which administrates the whole network.
2. **Provider:** It is a computer which contributes its resources in the network resource pool.
3. **User:** It is a computer which uses the resources on the network.

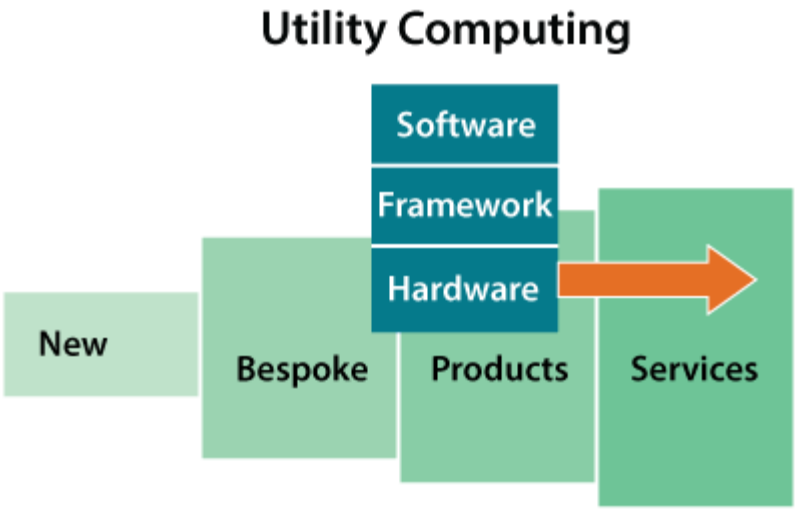
Mainly, grid computing is used in the **ATMs, back-end infrastructures,** and **marketing research**.



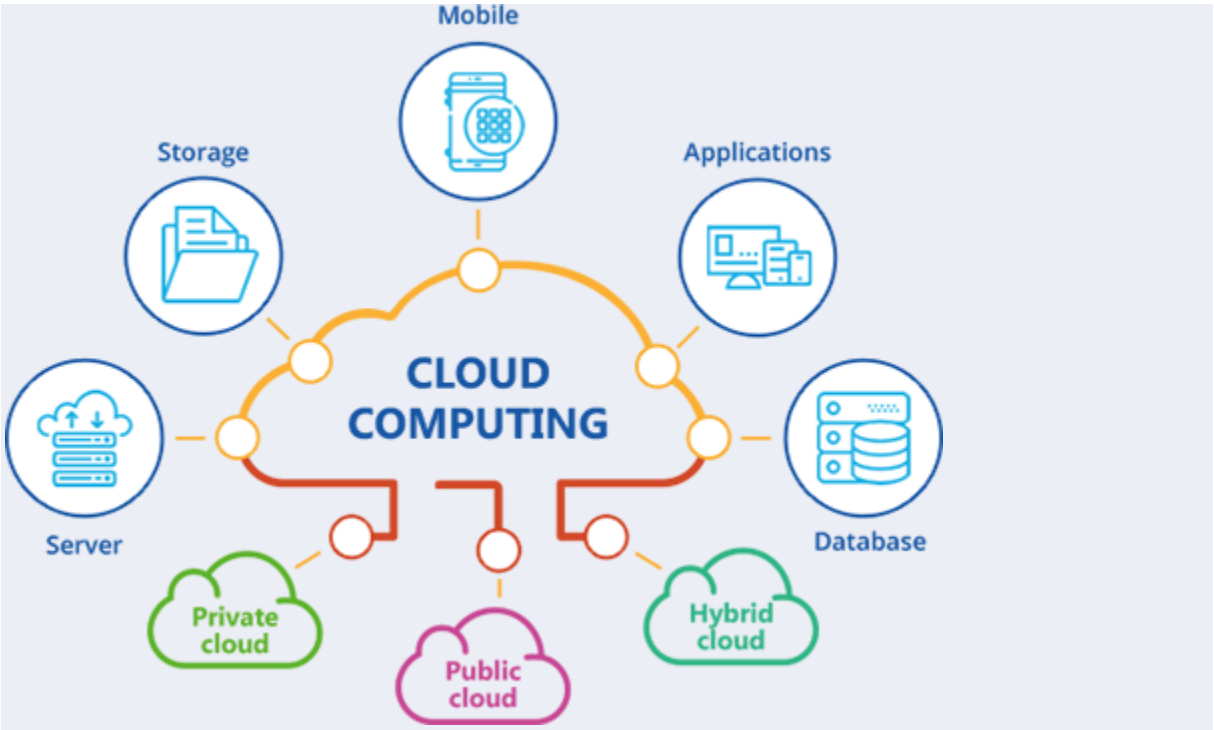
Utility Computing

Utility computing is the most trending IT service model. It provides on-demand computing resources (computation, storage, and programming services via API) and infrastructure based on the **pay per use** method. It minimizes the associated costs and maximizes the efficient use of resources. The advantage of utility computing is that it reduced the IT cost, provides greater flexibility, and easier to manage.

Large organizations such as **Google** and **Amazon** established their own utility services for computing storage and application.



System models for distributed and cloud computing:



Distributed and cloud computing systems are built over a large number of autonomous computer nodes. These node machines are interconnected by SANs, LANs, or WANs in a hierarchical manner.

With today's networking technology, a 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. Massive systems are considered highly scalable, and can reach web-scale connectivity, either physically or logically.

Massive systems are classified into four groups:

1. **Clusters** : A distributed systems cluster is a group of machines that are virtually or geographically separated and that work together to provide the same service or application to clients. It is possible that many of the services you run in your network today are part of a distributed systems *Cluster Distributed Services:*

- Domain Naming System
- Windows Internet Naming Service
- Active Directory

2. P2P Networks : 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. This implies that no master-slave relationship exists among the peers. No central coordination or central database is needed. The system is self-organizing with distributed control.

3. Computing Grids : This is the use of widely distributed computer resources to reach a common goal. A computing grid can be thought of as a distributed system with non-interactive workloads that involve many files. Grid computing is distinguished from conventional high-performance computing systems such as cluster computing in that grid computers have each node set to perform a different task/application. Grid computers also tend to be more heterogeneous and geographically dispersed than cluster computers.

4. Internet clouds : 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. Cloud computing intends to satisfy many user Virtualized resources from data centers to form an Internet cloud, provisioned with hardware, software, storage, network, and services for paid users to run their applications.

NIST Cloud Computing Reference Architecture:

NIST Cloud Computing reference architecture defines five major performers:

- Cloud Provider
- Cloud Carrier
- Cloud Broker
- Cloud Auditor
- Cloud Consumer

Each performer is an object (a person or an organization) that contributes to a transaction or method and/or performs tasks in Cloud computing. There are five major actors defined in the NIST cloud computing reference architecture, which are described below:

1. Cloud Service Providers: A group or object that delivers cloud services to cloud consumers or end-users. It offers various components of cloud computing. Cloud computing consumers purchase a growing variety of cloud services from cloud service providers. There are various categories of cloud-based services mentioned below:

- **IaaS Providers:** In this model, the cloud service providers offer infrastructure components that would exist in an on-premises data center. These components consist of servers, networking, and storage as well as the virtualization layer.
- **SaaS Providers:** In Software as a Service (SaaS), vendors provide a wide sequence of business technologies, such as Human resources management (HRM) software, customer relationship management (CRM) software, all of which the SaaS vendor hosts and provides services through the internet.
- **PaaS Providers:** In Platform as a Service (PaaS), vendors offer cloud infrastructure and services that can access to perform many functions. In PaaS, services and products are mostly utilized in software development. PaaS providers offer more services than IaaS providers. PaaS providers provide operating system and middleware along with application stack, to the underlying infrastructure.

2. Cloud Carrier: The mediator who provides offers connectivity and transport of cloud services within cloud service providers and cloud consumers. It allows access to the services of the cloud through Internet networks, telecommunication, and other access devices. Network and telecom carriers or a transport agent can provide distribution. A consistent level of services is provided when cloud providers set up Service Level Agreements (SLA) with a cloud carrier. In general, Carrier may be required to offer dedicated and encrypted connections.

3. Cloud Broker: An organization or a unit that manages the performance, use, and delivery of cloud services by enhancing specific capability and offers value-added services to cloud consumers. It combines and integrates various services into one or more new services. They provide service arbitrage which allows flexibility and opportunistic choices. There are major three services offered by a cloud broker:

- Service Intermediation.
- Service Aggregation.
- Service Arbitrage.

4. Cloud Auditor: An entity that can conduct independent assessment of cloud services, security, performance, and information system operations of the cloud implementations. The services that are provided by Cloud Service Providers (CSP) can be evaluated by service auditors in terms of privacy impact, security control, and performance, etc. Cloud Auditor can make an

assessment of the security controls in the information system to determine the extent to which the controls are implemented correctly, operating as planned and constructing the desired outcome with respect to meeting the security necessities for the system. There are three major roles of Cloud Auditor which are mentioned below:

- Security Audit.
- Privacy Impact Audit.
- Performance Audit.

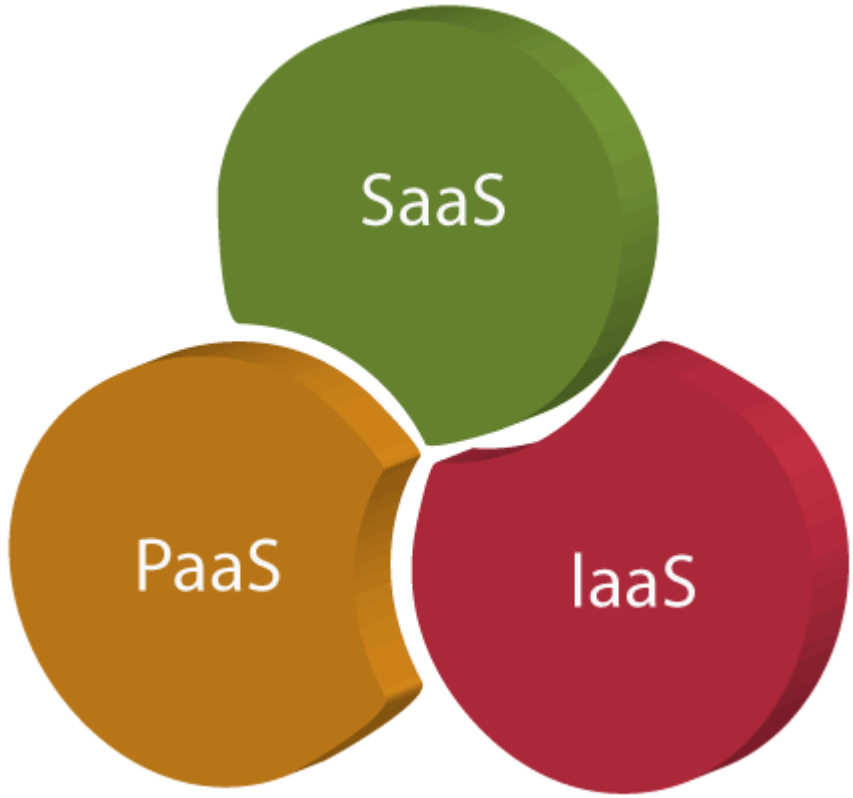
5. Cloud Consumer: A cloud consumer is the end-user who browses or utilizes the services provided by Cloud Service Providers (CSP), sets up service contracts with the cloud provider. The cloud consumer pays peruse of the service provisioned. Measured services utilized by the consumer. In this, a set of organizations having mutual regulatory constraints performs a security and risk assessment for each use case of Cloud migrations and deployments.
Cloud consumers use Service-Level Agreement (SLAs) to specify the technical performance requirements to be fulfilled by a cloud provider. SLAs can cover terms concerning the quality of service, security, and remedies for performance failures. A cloud provider may also list in the SLAs a set of limitations or boundaries, and obligations that cloud consumers must accept. In a mature market environment, a cloud consumer can freely pick a cloud provider with better pricing and more favourable terms. Typically, a cloud provider's public pricing policy and SLAs are non-negotiable, although a cloud consumer who assumes to have substantial usage might be able to negotiate for better contracts.

Cloud models:

There are the following three types of cloud service models -

1. Infrastructure as a Service (IaaS)
2. Platform as a Service (PaaS)

3. Software as a Service (SaaS)



Infrastructure as a Service (IaaS)

IaaS is also known as **Hardware as a Service (HaaS)**. It is a computing infrastructure managed over the internet. The main advantage of using IaaS is that it helps users to avoid the cost and complexity of purchasing and managing the physical servers.

Characteristics of IaaS

There are the following characteristics of IaaS -

- Resources are available as a service
- Services are highly scalable
- Dynamic and flexible
- GUI and API-based access
- Automated administrative tasks

Example: DigitalOcean, Linode, Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE), Rackspace, and Cisco Metacloud.

To know more about the IaaS, **click here**.

Platform as a Service (PaaS)

PaaS cloud computing platform is created for the programmer to develop, test, run, and manage the applications.

Characteristics of PaaS

There are the following characteristics of PaaS -

- Accessible to various users via the same development application.
- Integrates with web services and databases.
- Builds on virtualization technology, so resources can easily be scaled up or down as per the organization's need.
- Support multiple languages and frameworks.
- Provides an ability to "**Auto-scale**".

Example: AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, Magento Commerce Cloud, and OpenShift.

To know more about PaaS, click here.

Software as a Service (SaaS)

SaaS is also known as "**on-demand software**". It is a software in which the applications are hosted by a cloud service provider. Users can access these applications with the help of internet connection and web browser.

Characteristics of SaaS

There are the following characteristics of SaaS -

- Managed from a central location
- Hosted on a remote server
- Accessible over the internet

- Users are not responsible for hardware and software updates. Updates are applied automatically.
- The services are purchased on the pay-as-per-use basis

Example: BigCommerce, Google Apps, Salesforce, Dropbox, ZenDesk, Cisco WebEx, ZenDesk, Slack, and GoToMeeting.

To know more about the SaaS, **click here**.

Difference between IaaS, PaaS, and SaaS

The below table shows the difference between IaaS, PaaS, and SaaS -

IaaS	Paas	SaaS
It provides a virtual data center to store information and create platforms for app development, testing, and deployment.	It provides virtual platforms and tools to create, test, and deploy apps.	It provides w apps to comple
It provides access to resources such as virtual machines, virtual storage, etc.	It provides runtime environments and deployment tools for applications.	It provides soft to the end-user
It is used by network architects.	It is used by developers.	It is used by en
IaaS provides only Infrastructure.	PaaS provides Infrastructure+Platform.	SaaS Infrastructure+ +Software.

Characteristics of Cloud Computing:

There are basically 5 essential characteristics of Cloud Computing.

1. **On-demand self-services:** The Cloud computing services does not require any human administrators, user themselves are able to provision, monitor and manage computing resources as needed.
2. **Broad network access:** The Computing services are generally provided over standard networks and heterogeneous devices.
3. **Rapid elasticity:** The Computing services should have IT resources that are able to scale out and in quickly and on as

needed basis. Whenever the user require services it is provided to him and it is scale out as soon as its requirement gets over.
4. **Resource pooling:** The IT resource (e.g., networks, servers, storage, applications, and services) present are shared across multiple applications and occupant in an uncommitted manner. Multiple clients are provided service from a same physical resource.
5. **Measured service:** The resource utilization is tracked for each application and occupant, it will provide both the user and the resource provider with an account of what has been used. This is done for various reasons like monitoring billing and effective use of resource.
6. **Multi-tenancy:** Cloud computing providers can support multiple tenants (users or organizations) on a single set of shared resources.
7. **Virtualization:** Cloud computing providers use virtualization technology to abstract underlying hardware resources and present them as logical resources to users.

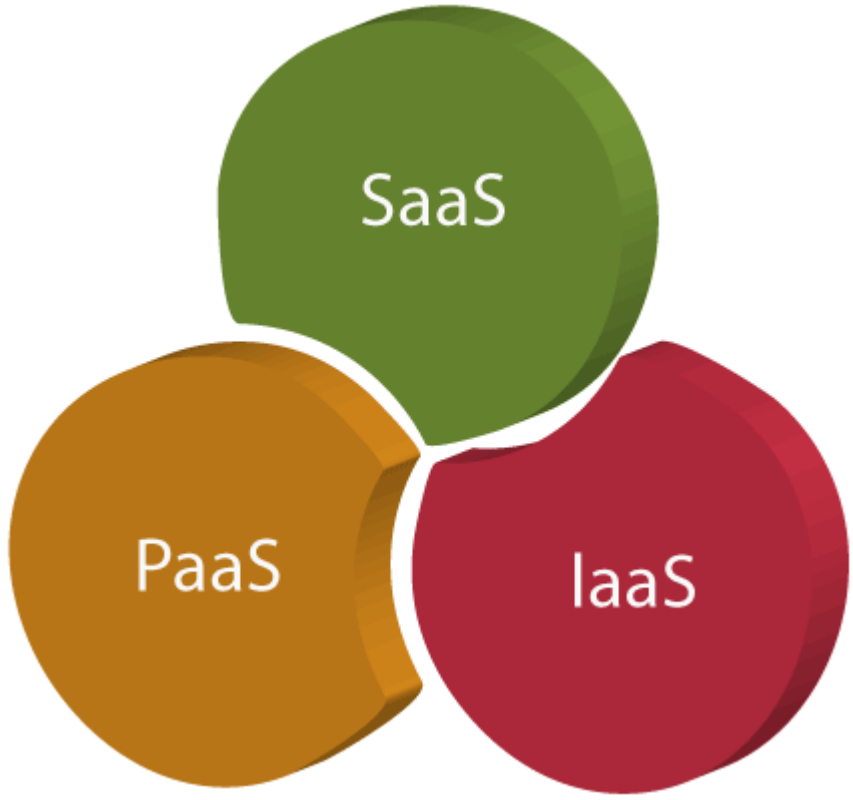
Cloud Services:

The term "cloud services" refers to a wide range of services delivered on demand to companies and customers over the internet. These services are designed to provide easy, affordable access to applications and resources, without the need for internal infrastructure or hardware. From checking email to collaborating on documents, most employees use cloud services throughout the workday, whether they're aware of it or not.
Cloud services are fully managed by cloud computing vendors and service providers. They're made available to customers from the providers' servers, so there's no need for a company to host applications on its own on-premises servers.

Cloud models (IaaS, PaaS, SaaS):

There are the following three types of cloud service models -

1. Infrastructure as a Service (IaaS)
2. Platform as a Service (PaaS)
3. Software as a Service (SaaS)



Infrastructure as a Service (IaaS)

IaaS is also known as **Hardware as a Service (HaaS)**. It is a computing infrastructure managed over the internet. The main advantage of using IaaS is that it helps users to avoid the cost and complexity of purchasing and managing the physical servers.

Characteristics of IaaS

There are the following characteristics of IaaS -

- Resources are available as a service
- Services are highly scalable
- Dynamic and flexible
- GUI and API-based access
- Automated administrative tasks

Example: DigitalOcean, Linode, Amazon Web Services (AWS), Microsoft Azure, Google Compute Engine (GCE), Rackspace, and Cisco Metacloud.

To know more about the IaaS, **click here**.

Platform as a Service (PaaS)

PaaS cloud computing platform is created for the programmer to develop, test, run, and manage the applications.

Characteristics of PaaS

There are the following characteristics of PaaS -

- Accessible to various users via the same development application.
- Integrates with web services and databases.
- Builds on virtualization technology, so resources can easily be scaled up or down as per the organization's need.
- Support multiple languages and frameworks.
- Provides an ability to "**Auto-scale**".

Example: AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com, Google App Engine, Apache Stratos, Magento Commerce Cloud, and OpenShift.

To know more about PaaS, click here.

Software as a Service (SaaS)

SaaS is also known as "**on-demand software**". It is a software in which the applications are hosted by a cloud service provider. Users can access these applications with the help of internet connection and web browser.

Characteristics of SaaS

There are the following characteristics of SaaS -

- Managed from a central location
- Hosted on a remote server
- Accessible over the internet
- Users are not responsible for hardware and software updates. Updates are applied automatically.
- The services are purchased on the pay-as-per-use basis

Example: BigCommerce, Google Apps, Salesforce, Dropbox, ZenDesk, Cisco WebEx, ZenDesk, Slack, and GoToMeeting.

To know more about the SaaS, **click here**.

Difference between IaaS, PaaS, and SaaS

The below table shows the difference between IaaS, PaaS, and SaaS -

IaaS	Paas	SaaS
It provides a virtual data center to store information and create platforms for app development, testing, and deployment.	It provides virtual platforms and tools to create, test, and deploy apps.	It provides w apps to comple
It provides access to resources such as virtual machines, virtual storage, etc.	It provides runtime environments and deployment tools for applications.	It provides soft to the end-user
It is used by network architects.	It is used by developers.	It is used by en
IaaS provides only Infrastructure.	PaaS provides Infrastructure+Platform.	SaaS Infrastructure+ +Software.

Public Vs Private Cloud:

Cloud computing is a way of providing IT infrastructure to customers, it is not just a set of products to be implemented. For any service to be a cloud service, the following five criteria need to be fulfilled as follows:

1. **On-demand self-service:** Decision of starting and stopping service depends on customers without direct interaction with providers.
2. **Broad Network Access:** Service must be available to any device using any network.
3. **Resource Pooling:** Provider creates a pool of resources and dynamically allocates them to customers.
4. **Rapid Elasticity:** The services provided by the provider must be easily expandable and quick.

5. **Measured Services:** Provider must measure the usage of service and charge it accordingly. Tracking usage is also helpful in improving services.

1. Public Cloud:

Computing in which service provider makes all resources public over the internet. It is connected to the public Internet. Service provider serves resources such as virtual machines, applications, storage, etc to the general public over the internet. It may be free of cost or with minimal pay-per-usage. It is available for public display, Google uses the cloud to run some of its applications like google docs, google drive or YouTube, etc.

It is the most common way of implementing cloud computing. The external cloud service provider owns, operates, and delivers it over the public network.

It is best for the companies which need an infrastructure to accommodate a large number of customers and work on projects which have diverse organizations i.e. research institutions and NGOs etc.

2. Private Cloud:

Computing in which service provider does not makes all resources public over the internet. It only supports connectivity over the private network. It has only authentic users and single-occupant architecture. Google back-end data of the applications like Google Drive, Google docs, YouTube, etc are not available to the public, these types of data and applications run on a private cloud.

The infrastructure and services are maintained and deployed over a private network; hardware and software are dedicated only to a private company i.e. members of the special entity.

It is best for the companies which need an infrastructure that has high performance, high security, and privacy due to its best adaptability and flexibility.

Below is a table of differences between Public Cloud and Private Cloud is as follows:

Public Cloud	Private Cloud
Cloud Computing infrastructure is shared with the public by service providers over the internet. It supports multiple customers i.e, enterprises.	Cloud Computing infrastructure is shared with private organizations by service providers over the internet. It supports one enterprise.
Multi-Tenancy i.e, Data of many enterprises are stored in a shared environment but are isolated. Data is shared as per rule, permission, and security.	Single Tenancy i.e, Data of a single enterprise is stored.

Public Cloud	Private Cloud
Cloud service provider provides all the possible services and hardware as the user-base is the world. Different people and organizations may need different services and hardware. Services provided must be versatile.	Specific services and hardware as per the need of the enterprise are available in a private cloud.
It is hosted at the Service Provider site.	It is hosted at the Service Provider site or enterprise.
It is connected to the public internet.	It only supports connectivity over the private network.
Scalability is very high, and reliability is moderate.	Scalability is limited, and reliability is very high.
Cloud service provider manages the cloud and customers use them.	Managed and used by a single enterprise.
It is cheaper than the private cloud.	It is costlier than the public cloud.
Security matters and dependent on the service provider.	It gives a high class of security.
Performance is low to medium.	Performance is high.
It has shared servers.	It has dedicated servers.
Example: Amazon web service (AWS) and Google AppEngine etc.	Example: Microsoft KVM, HP, Red Hat & VMWare etc.

Cloud Solutions:

Cloud solutions, also known as cloud computing or cloud services, deliver IT resources on demand over the Internet. Cloud service providers such Amazon Web Services, Microsoft Azure and Google Cloud Platform can deliver everything from applications to data centers on a pay-for-use basis to their subscribers. With cloud solutions, IT resources can scale up or down quickly to meet business demands. Cloud solutions enable rapid access to flexible and low-cost IT resources without large upfront investments in hardware or time-consuming installation and maintenance. Businesses can provision exactly the type and size of

computing resources they need to power a new initiative or operate their IT departments more efficiently.
Most cloud solutions fall into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). IaaS allows businesses to essentially rent their IT infrastructure from a cloud provider. PaaS supplies an on-demand environment for software development. SaaS delivers applications over the Internet. Businesses of all sizes use cloud solutions to reduce hardware, software and IT maintenance costs. As cloud solutions evolve beyond IaaS, PaaS and SaaS offerings, enterprises are relying on the cloud for software-defined technology. Data center resources – including compute, storage and network resources – can be virtualized and centrally managed as software-defined pools. Cloud providers are now offering pre-built cloud solutions with the agility to deploy abstracted, software-defined resources to workloads as needed.
Cloud solutions provide convenient, on-demand access to shared pools of IT resources, helping businesses improve efficiency, reduce costs and rebalance capital and operating expenses. Many businesses adopt a mixture of public, private and hybrid cloud solutions. OpenStack is a software-defined infrastructure for developing new business workloads, for DevOps initiatives, or for transforming traditional data centers to a private cloud.

Cloud ecosystem:

A cloud ecosystem is a complex system of interdependent components that all work together to enable cloud services. In nature, an ecosystem is composed of living and nonliving things that are connected and work together. In cloud computing, the ecosystem consists of hardware and software as well as cloud customers, cloud engineers, consultants, integrators and partners.

Werner Vogels, CTO at Amazon, first compared the cloud to an ecosystem in a keynote address at the Cloud Connect 2011 conference. At the time, enterprise cloud computing was usually thought of in terms of three broad service areas -- infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS) and software-as-a-service (SaaS). Vogels proposed that the cloud was really more complex and its description also needed to include the array of service providers that companies rely on to operate in the cloud.

How a cloud ecosystem works

The center of a cloud ecosystem is a public cloud provider. It might be an IaaS provider such as Amazon Web Services (AWS) or a SaaS vendor such as Salesforce. Radiating out from the center of the cloud are software

companies that use the provider's anchor platform, as well as consultants and companies that have formed strategic alliances with the anchor provider.

There is no vendor lock-in because these companies overlap, making the ecosystem more complex. For example, AWS is the center of its own ecosystem, but it's also a part of the Salesforce ecosystem. Salesforce runs a number of its services on AWS's infrastructure, and Salesforce customers can gain access, through devices called connectors, to pieces of AWS, such as its Simple Storage Service (S3).

A robust ecosystem provides a cloud provider's customers with an easy way to find and purchase business applications and respond to changing business needs. When the apps are sold through a provider's app store such as AWS Marketplace, Microsoft Azure Marketplace (for cloud software) or Microsoft AppSource (for business applications), the customer essentially has access to a catalog of different vendors' software and services that have already been vetted and reviewed for security, risk and cost.

The benefits of a cloud ecosystem

Companies can use a cloud ecosystem to build new business models. It becomes relatively easy for a medical device manufacturer, for example, to launch a heart-monitoring service on its cloud service provider's cloud infrastructure and then sell the service alongside its main business of manufacturing heart monitors for hospitals.

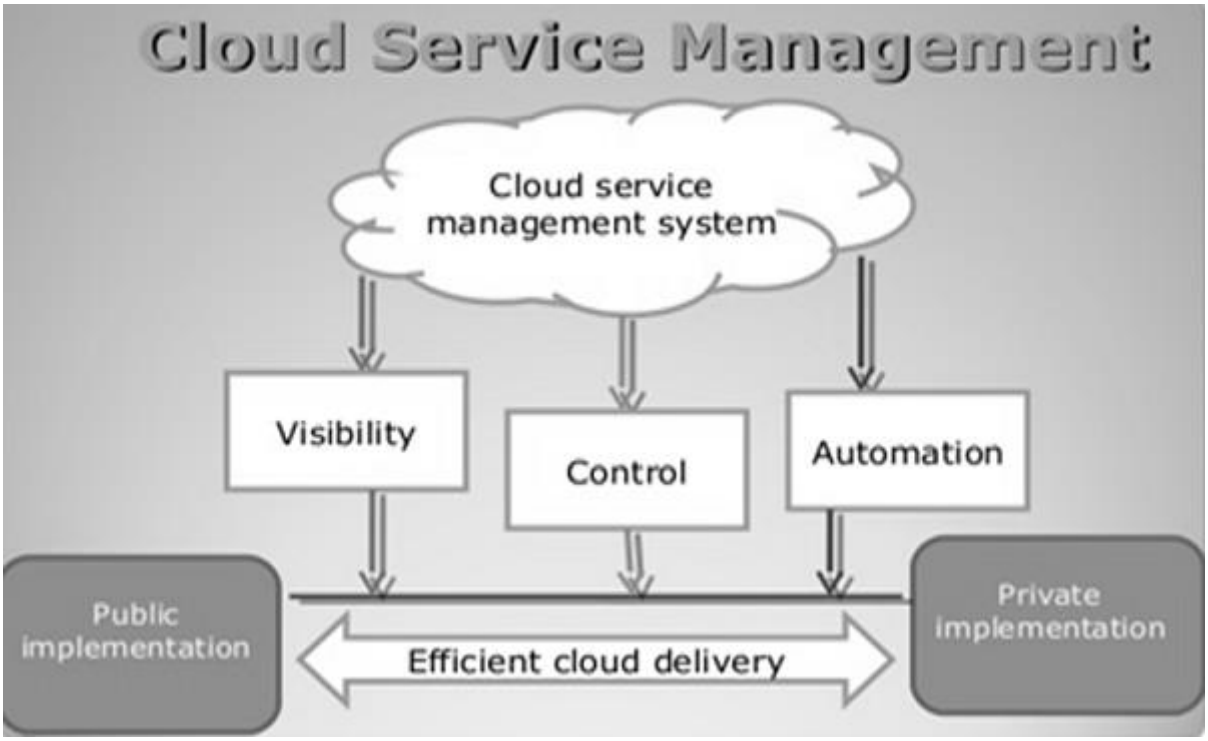
In a cloud ecosystem, it is also easier to aggregate data and analyze how each part of the system affects the other parts. For example, if an ecosystem consists of patient records, smart device logs and healthcare provider records, it becomes possible to analyze patterns across an entire patient population.

Service management:

The management of cloud infrastructure products and services is cloud management. Public clouds are operated by public cloud service providers, which provide the servers, storage, networking and data centre operations of the public cloud environment. With a third-party cloud management tool, users can also choose to manage their public cloud services.

Public cloud service users can typically choose from three categories of specific cloud provisioning:

- **User self-provisioning**: Users, usually via a web form or console interface, buy cloud services directly from the provider. On a per-transaction basis, the client pays.
- **Advanced provisioning**: A pre-determined sum of services scheduled in advance of operation is contracted in advance by customers. A flat fee or a monthly fee is charged by the consumer.
- **Dynamic provisioning**: When the client requires them, the provider allocates resources, and then decommissions them when they are no longer required. On a pay-per-use basis, the client is paid.



The purpose and scope of the management of cloud services are listed below:

- **Purpose**: Establish suitable techniques for managing and running cloud-based services. Insert cloud service management techniques into current frameworks for IT creation and support.

- **Scope**: Oversight of cloud-based service design, development and change. Cloud-based service management and operation.

Characteristics of Cloud service Management

In a design for handling cloud environments, cloud management incorporates applications and technologies. With a range of cloud management platforms and instruments, software developers have responded to the management challenges of cloud computing. These solutions include native tools provided by public cloud providers, as well as third-party tools designed by various cloud providers to provide consistent functionality. With access to various native features within individual cloud platforms, administrators must balance the conflicting requirements of efficient consistency across various cloud platforms. The need for transparent cross-platform management is motivated by increasing public cloud adoption and increased multi-cloud use. For those technical professionals responsible for maintaining IT systems and facilities, the rapid adoption of cloud services presents a new set of management challenges.

In the following categories, cloud-management systems and instruments should be able to have minimum functionality.

- **Service request**: receiving and fulfilling user requests to access and deploy cloud services.
- **Cost management and optimization**: Cloud spending monitors and accurate sizes and aligns resources and efficiency with real demand.
- **Security and compliance**: handling cloud providers' role-based access and implementing security settings.
- **Inventory and classification**: discover and maintain pre-existing cloud infrastructure in the brownfield plus track and handle modifications.

Computing on demand:

On-demand computing is a business computing model in which computing resources are made available to the user on an “as needed” basis. Rather than all at once, on-demand computing allows cloud hosting companies to provide their clients with access to computing resources as they become necessary.

The on-demand computing model was developed to overcome the common challenge that enterprises encountered of not being able to meet unpredictable, fluctuating computing demands efficiently. Businesses today need to be agile and need the ability to

scale resources easily and quickly based on rapidly changing market needs. Because an enterprise's demand for computing resources can vary dramatically from one period to another, maintaining sufficient resources to meet peak requirements can be costly. However, with on-demand computing, companies can cut costs by maintaining minimal computing resources until they run into the need to increase them while only paying for their use.

Industry experts predict on-demand computing to soon be the most widely used computing model for enterprises. In fact, IBM's vice-president of technology and strategy stated, "The technology is at a point where we can start to move into an era of on-demand computing. I give it between two and four years to reach a level of maturity."

Cloud Computing Creates Success in the Manufacturing Industry

According to a recent article in Forbes, many in the manufacturing industry seek cloud-based services to increase efficiency from a supply chain, distribution, and services standpoint. In fact, more than a few manufacturers have already adopted cloud-based applications throughout their companies. According to MintJutras, a research-based consulting firm specializing in analyzing enterprise applications' business impact, SaaS applications make up 22% of all manufacturing and distribution software installed today. Additionally, this number is expected to grow to 45% within ten years.

The main goal of manufacturers is to make themselves easier to work with both externally and internally. Manufacturers are constantly facing pressure to increase accuracy, improve process speed, and effectively utilize their internal intelligence to make every supplier, distributor, and service interaction worthwhile. Cloud-based services are helping to give these companies the chance to do just this.

Experts say that cloud-based services have the potential to streamline key areas of business so that manufacturers have more time to sell their products and invest in new ones. As mentioned in the Forbes article, here are just a few of the ways that cloud computing is revolutionizing the manufacturing industry:

- Capturing and applying company-wide data through the use of analytics, business intelligence, and rules engines.
- Piloting and quickly moving to a full launch of supplier portals and collaboration platforms, complete with quality management dashboards and workflows.
- Accelerating new product development and introduction strategies to attain time-to-market objectives.
- Managing indirect and direct channel sales from a single cloud platform tracks sales results at the individual, group, and divisional levels.
- Automating customer service, support, and common order status inquiries online.
- Increasing reliance on two-tier ERP software strategies to gain greater efficiencies in material planning and supplier management and reduce logistics costs.

Short Answer questions:

1.write the architecture of cloud computing?
2.what are the technologies for network?
3.what are the types of cloud models?
4.Differentiate between public and private cloud computing?
5.explain cloud ecosystem?

Long Answer questions:

1. what are the types of cloud models and explain briefly?
2.Explain about NIST cloud computing?
3.Explain service management?
4.what are the cloud services and explain in detail?
5.Explain about system models for distributed and cloud computing?

UNIT-2

Virtualization:

Virtualization is the "creation of a virtual (rather than actual) version of something, such as a server, a desktop, a storage device, an operating system or network resources".

In other words, Virtualization is a technique, which allows to share a single physical instance of a resource or an application among multiple customers and organizations. It does by assigning a logical name to a physical storage and providing a pointer to that physical resource when demanded.

Creation of a virtual machine over existing operating system and hardware is known as Hardware Virtualization. A Virtual machine provides an environment that is logically separated from the underlying hardware.

The machine on which the virtual machine is going to create is known as **Host Machine** and that virtual machine is referred as a **Guest Machine**

Types of virtualization:

1. Hardware Virtualization.
2. Operating system Virtualization.
3. Server Virtualization.
4. Storage Virtualization.

1) Hardware Virtualization:

When the virtual machine software or virtual machine manager *(VMM) is directly installed on the hardware system* is known as hardware virtualization.

The main job of hypervisor is to control and monitoring the processor, memory and other hardware resources.

After virtualization of hardware system we can install different operating system on it and run different applications on those OS.

Usage:

Hardware virtualization is mainly done for the server platforms, because controlling virtual machines is much easier than controlling a physical server.

2) Operating System Virtualization:

When the virtual machine software or virtual machine manager *(VMM) is installed on the Host operating system* instead of directly on the hardware system is known as operating system virtualization.

Usage:

Operating System Virtualization is mainly used for testing the applications on different platforms of OS.

3) Server Virtualization:

When the virtual machine software or virtual machine manager *(VMM) is directly installed on the Server system* is known as server virtualization.

Usage:

Server virtualization is done because a single physical server can be divided into multiple servers on the demand basis and for balancing the load.

4) Storage Virtualization:

Storage virtualization is the *process of grouping the physical storage from multiple network storage devices so that it looks like a single storage device*.

Storage virtualization is also implemented by using software applications.

Implementation levels of virtualization:

It is not simple to set up virtualization. Your computer runs on an operating system that gets configured on some particular hardware. It is not feasible or easy to run a different operating system using the same hardware.

To do this, you will need a hypervisor. Now, what is the role of the hypervisor? It is a bridge between the hardware and the virtual operating system, which allows smooth functioning.

Talking of the Implementation levels of virtualization in Cloud Computing., there are a total of five levels that are commonly used. Let us now look closely at each of these levels of virtualization implementation in Cloud Computing.

1) Instruction Set Architecture Level (ISA)

ISA virtualization can work through ISA emulation. This is used to run many legacy codes written for a different hardware configuration. These codes run on any virtual machine using the ISA. With this, a binary code that originally needed some additional layers to run is now capable of running on the x86 machines. It can also be tweaked to run on the x64 machine. With ISA, it is possible to make the virtual machine hardware agnostic.

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

processing. This is one of the five implementation levels of virtualization in Cloud Computing..

2) Hardware Abstraction Level (HAL)

True to its name HAL lets the virtualization perform at the level of the hardware. This makes use of a hypervisor which is used for functioning. The virtual machine is formed at this level, which manages the hardware using the virtualization process. It allows the virtualization of each of the hardware components, which could be the input-output device, the memory, the processor, etc.

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

3) Operating System Level

At the level of the operating system, the virtualization model is capable of creating a layer that is abstract between the operating system and the application. This is an isolated container on the operating system and the physical server, which uses the software and hardware. Each of these then functions in the form of a server.

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

4) Library Level

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

5) Application Level

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

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

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

Virtualization Structure:

Virtualization is achieved through the software known as virtual machine monitor or the hypervisor . the software is used in two ways thus forming two different structure of virtualization,namely Bare Metal Virtualization and Hosted Virtualization.

Bare-metal virtualization hypervisors: (TYPE I HYPERVISOR)

• Is deployed as a bare-metal installation (the first thing to be installed on a server as the operating system will be the hypervisor).

• The hypervisor will communicate directly with the underlying physical server hardware, manages all hardware resources and support execution of VMs.

• Hardware support is typically more limited, because the hypervisor usually has limited device drivers built into it.

• Well suited for enterprise data centers, because it usually comes with advanced features for resource management, high availability and security.

• Bare-metal virtualization hypervisors examples: VMware ESX and ESXi, Microsoft Hyper-V, Citrix Systems XenServer.

Hosted virtualization hypervisors: (TYPE II HYPERVISOR)

• The software is not installed onto the bare-metal, but instead is loaded on top of an already live operating system, so it requires you to first install an OS(Host OS).

• The Host OS integrates a hypervisor that is responsible for providing the virtual machines(VMs) with their virtual platform interface and for managing all context switching scheduling, etc.

• The hypervisor will invoke drivers or other component of the Host OS as needed.

• On the Host OS you may run Guest VMs, but you can also run native applications

• This approach provides better hardware compatibility than bare-metal virtualization, because the OS is responsible for the hardware drivers instead of the hypervisor.

• A hosted virtualization hypervisor does not have direct access to hardware and must go through the OS, which increases resource overhead and can degrade virtual machine (VM) performance.

• The latency is minimal and with today's modern software enhancements, the hypervisor can still perform optimally.

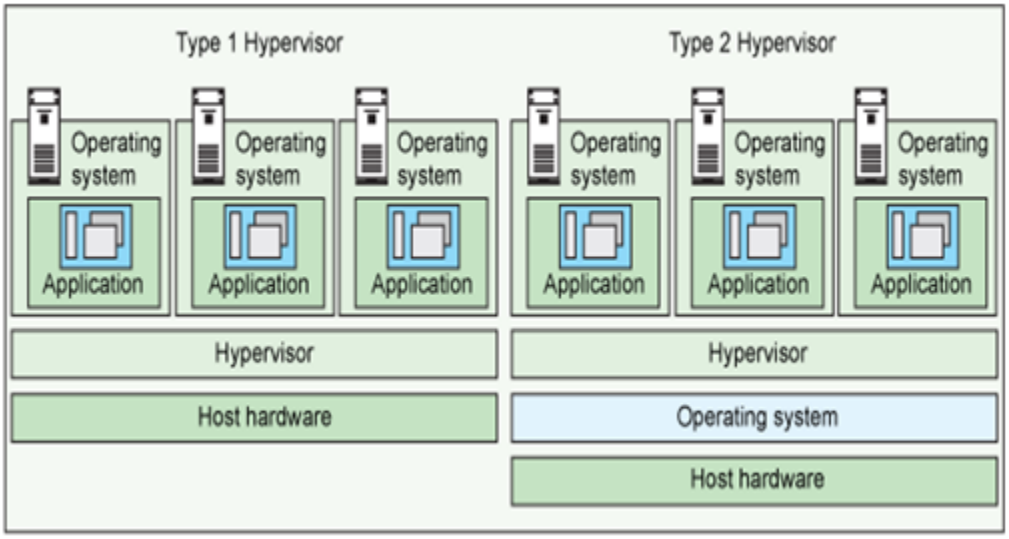
• Common for desktops, because they allow you to run multiple OSes. These virtualization hypervisor types are also popular for developers, to maintain application compatibility on modern OSes.

• Because there are typically many services and applications running on the host OS, the hypervisor often steals resources from the VMs running on it

• The most popular hosted virtualization hypervisors are: VMware Workstation, Server, Player and Fusion; Oracle VM VirtualBox; Microsoft Virtual PC; Parallels Desktop.

• The below figure shows stucture of TYPE I and TYPE II virtualization.

• The below figure shows stucture of TYPE I and TYPE II virtualization.



Virtualization of CPU,Memory I/o devices:

To support virtualization, processors such as the x86 employ a special running mode and instructions, known as hardware-assisted virtualization. In this way, the VMM and guest OS run in different modes and all sensitive instructions of the guest OS and its applications are trapped in the VMM. To save processor states, mode switching is completed by hardware. For the x86 architecture, Intel and AMD have proprietary technologies for hardware-assisted virtualization.

1. Hardware Support for Virtualization

Modern operating systems and processors permit multiple processes to run simultaneously. If there is no protection mechanism in a processor, all instructions from different processes will access the hardware directly and cause a system crash. Therefore, all processors have at least two modes, user mode and supervisor mode, to ensure controlled access of critical hardware. Instructions running in supervisor mode are called privileged instructions. Other instructions are unprivileged instructions. In a virtualized environment, it is more difficult to make OSes and applications run correctly because there are more layers in the machine stack. Example 3.4 discusses Intel's hardware support approach.

At the time of this writing, many hardware virtualization products were available. The VMware Workstation is a VM software suite for x86 and x86-64 computers. This software suite allows users to set up multiple x86 and x86-64 virtual computers and to use one or more of these VMs simultaneously with the host operating system. The VMware Workstation assumes the host-based virtualization. Xen is a hypervisor for use in IA-32, x86-64, Itanium, and PowerPC 970 hosts. Actually, Xen modifies Linux as the lowest and most privileged layer, or a hypervisor.

One or more guest OS can run on top of the hypervisor. KVM (Kernel-based Virtual Machine) is a Linux kernel virtualization infrastructure. KVM can support hardware-assisted virtualization and paravirtualization by using the Intel VT-x or AMD-v and VirtIO framework, respectively. The VirtIO framework includes a paravirtual Ethernet card, a disk I/O controller, a balloon device for adjusting guest memory usage, and a VGA graphics interface using VMware drivers.

Example 3.4 Hardware Support for Virtualization in the Intel x86 Processor

Since software-based virtualization techniques are complicated and incur performance overhead, Intel provides a hardware-assist technique to make virtualization easy and improve performance. Figure 3.10 provides an overview of Intel's full virtualization techniques. For processor virtualization, Intel offers the VT-x or VT-i technique. VT-x adds a privileged mode (VMX Root Mode) and some instructions to processors. This enhancement traps all sensitive instructions in the VMM automatically. For memory virtualization, Intel offers the EPT, which translates the

virtual address to the machine's physical addresses to improve performance. For I/O virtualization, Intel implements VT-d and VT-c to support this.

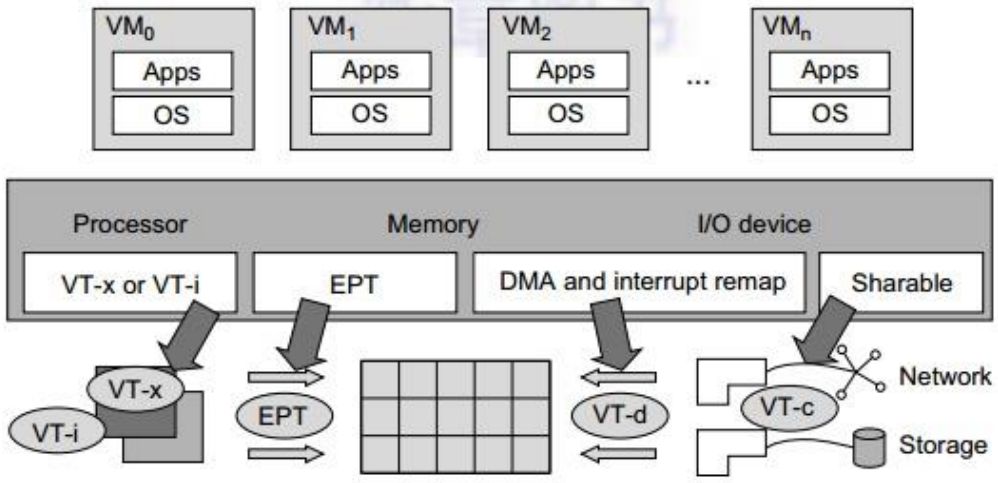


FIGURE 3.10

Intel hardware support for virtualization of processor, memory, and I/O devices.

2. CPU Virtualization

A VM is a duplicate of an existing computer system in which a majority of the VM instructions are executed on the host processor in native mode. Thus, unprivileged instructions of VMs run directly on the host machine for higher efficiency. Other critical instructions should be handled carefully for correctness and stability. The critical instructions are divided into three categories: privileged instructions, control-sensitive instructions, and behavior-sensitive instructions. Privileged instructions execute in a privileged mode and will be trapped if executed outside this mode. Control-sensitive instructions attempt to change the configuration of resources used. Behavior-sensitive instructions have different behaviors depending on the configuration of resources, including the load and store operations over the virtual memory.

A CPU architecture is virtualizable if it supports the ability to run the VM's privileged and unprivileged instructions in the CPU's user mode while the VMM runs in supervisor mode. When the privileged instructions including control- and behavior-sensitive instructions of a VM are exe-cuted, they are trapped in the VMM. In this case, the VMM acts as a unified mediator for hardware access from different VMs to guarantee the correctness and stability of the whole system. However, not all CPU architectures are virtualizable. RISC CPU architectures can be naturally virtualized because all control- and

behavior-sensitive instructions are privileged instructions. On the contrary, x86 CPU architectures are not primarily designed to support virtualization. This is because about 10 sensitive instructions, such as SGDT and SMSW, are not privileged instructions. When these instruc-tions execute in virtualization, they cannot be trapped in the VMM.

On a native UNIX-like system, a system call triggers the 80h interrupt and passes control to the OS kernel. The interrupt handler in the kernel is then invoked to process the system call. On a para-virtualization system such as Xen, a system call in the guest OS first triggers the 80h interrupt nor-mally. Almost at the same time, the 82h interrupt in the hypervisor is triggered. Incidentally, control is passed on to the hypervisor as well. When the hypervisor completes its task for the guest OS system call, it passes control back to the guest OS kernel. Certainly, the guest OS kernel may also invoke the hypercall while it's running. Although paravirtualization of a CPU lets unmodified applications run in the VM, it causes a small performance penalty.

2.1 Hardware-Assisted CPU Virtualization

This technique attempts to simplify virtualization because full or paravirtualization is complicated. Intel and AMD add an additional mode called privilege mode level (some people call it Ring-1) to x86 processors. Therefore, operating systems can still run at Ring 0 and the hypervisor can run at Ring -1. All the privileged and sensitive instructions are trapped in the hypervisor automatically. This technique removes the difficulty of implementing binary translation of full virtualization. It also lets the operating system run in VMs without modification.

Example 3.5 Intel Hardware-Assisted CPU Virtualization

Although x86 processors are not virtualizable primarily, great effort is taken to virtualize them. They are used widely in comparing RISC processors that the bulk of x86-based legacy systems cannot discard easily. Virtuali-zation of x86 processors is detailed in the following sections. Intel's VT-x technology is an example of hardware-assisted virtualization, as shown in Figure 3.11. Intel calls the privilege level of x86 processors the VMX Root Mode. In order to control the start and stop of a VM and allocate a memory page to maintain the

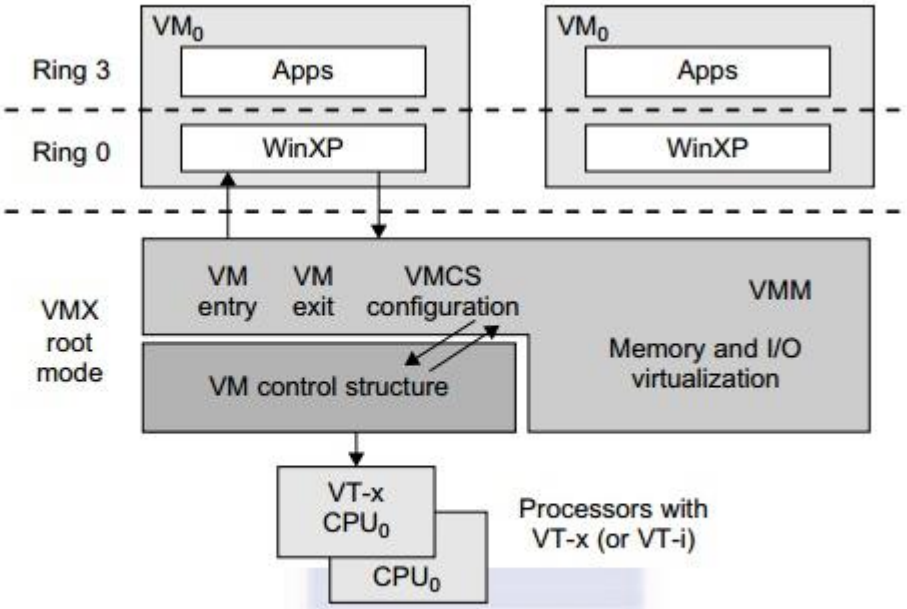


FIGURE 3.11

Intel hardware-assisted CPU virtualization.

CPU state for VMs, a set of additional instructions is added. At the time of this writing, Xen, VMware, and the Microsoft Virtual PC all implement their hypervisors by using the VT-x technology.

Generally, hardware-assisted virtualization should have high efficiency. However, since the transition from the hypervisor to the guest OS incurs high overhead switches between processor modes, it sometimes cannot outperform binary translation. Hence, virtualization systems such as VMware now use a hybrid approach, in which a few tasks are offloaded to the hardware but the rest is still done in software. In addition, para-virtualization and hardware-assisted virtualization can be combined to improve the performance further.

3. Memory Virtualization

Virtual memory virtualization is similar to the virtual memory support provided by modern operat-ing systems. In a traditional execution environment, the operating system maintains mappings of virtual memory to machine memory using page tables, which is a one-stage mapping from virtual memory to machine memory. All modern x86 CPUs include a memory management unit (MMU) and a translation lookaside buffer (TLB) to optimize virtual memory performance. However, in a virtual execution environment, virtual

memory virtualization involves sharing the physical system memory in RAM and dynamically allocating it to the physical memory of the VMs.

That means a two-stage mapping process should be maintained by the guest OS and the VMM, respectively: virtual memory to physical memory and physical memory to machine memory. Furthermore, MMU virtualization should be supported, which is transparent to the guest OS. The guest OS continues to control the mapping of virtual addresses to the physical memory addresses of VMs. But the guest OS cannot directly access the actual machine memory. The VMM is responsible for mapping the guest physical memory to the actual machine memory. Figure 3.12 shows the two-level memory mapping procedure.

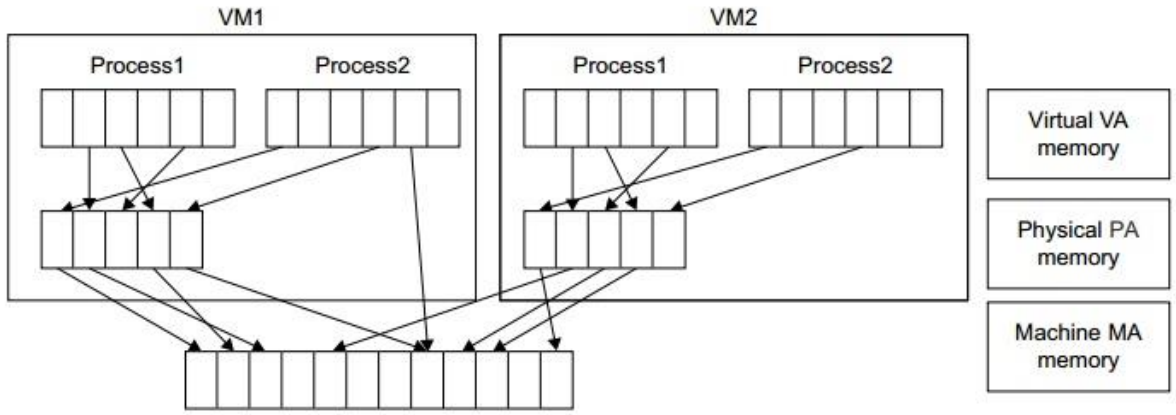


FIGURE 3.12

Two-level memory mapping procedure.

Since each page table of the guest OSes has a separate page table in the VMM corresponding to it, the VMM page table is called the shadow page table. Nested page tables add another layer of indirection to virtual memory. The MMU already handles virtual-to-physical translations as defined by the OS. Then the physical memory addresses are translated to machine addresses using another set of page tables defined by the hypervisor. Since modern operating systems maintain a set of page tables for every process, the shadow page tables will get flooded. Consequently, the perfor-mance overhead and cost of memory will be very high.

VMware uses shadow page tables to perform virtual-memory-to-machine-memory address translation. Processors use TLB hardware to map the virtual memory directly to the machine memory to avoid the two levels of translation on every access. When the guest OS changes the virtual memory to a physical memory mapping, the VMM updates the shadow page tables to enable a direct

lookup. The AMD Barcelona processor has featured hardware-assisted memory virtualization since 2007. It provides hardware assistance to the two-stage address translation in a virtual execution environment by using a technology called nested paging.

Example 3.6 Extended Page Table by Intel for Memory Virtualization

Since the efficiency of the software shadow page table technique was too low, Intel developed a hardware-based EPT technique to improve it, as illustrated in Figure 3.13. In addition, Intel offers a Virtual Processor ID (VPID) to improve use of the TLB. Therefore, the performance of memory virtualization is greatly improved. In Figure 3.13, the page tables of the guest OS and EPT are all four-level.

When a virtual address needs to be translated, the CPU will first look for the L4 page table pointed to by Guest CR3. Since the address in Guest CR3 is a physical address in the guest OS, the CPU needs to convert the Guest CR3 GPA to the host physical address (HPA) using EPT. In this procedure, the CPU will check the EPT TLB to see if the translation is there. If there is no required translation in the EPT TLB, the CPU will look for it in the EPT. If the CPU cannot find the translation in the EPT, an EPT violation exception will be raised.

When the GPA of the L4 page table is obtained, the CPU will calculate the GPA of the L3 page table by using the GVA and the content of the L4 page table. If the entry corresponding to the GVA in the L4

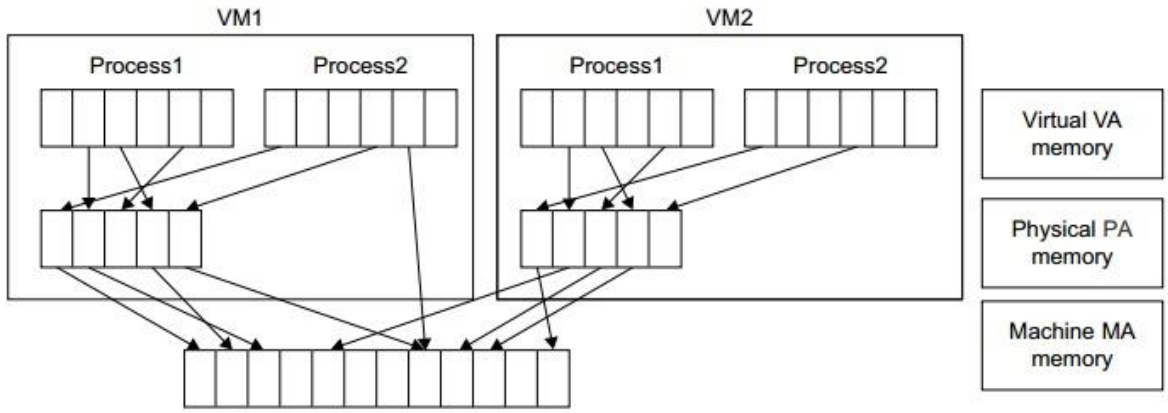


FIGURE 3.12
Two-level memory mapping procedure.

page table is a page fault, the CPU will generate a page fault interrupt and will let the guest OS kernel handle the interrupt. When the PGA of the L3 page table is obtained, the CPU will look for the EPT to get the HPA of the L3 page table, as described earlier. To get the HPA corresponding to a GVA, the

CPU needs to look for the EPT five times, and each time, the memory needs to be accessed four times. There-fore, there are 20 memory accesses in the worst case, which is still very slow. To overcome this short-coming, Intel increased the size of the EPT TLB to decrease the number of memory accesses.

4. I/O Virtualization

I/O virtualization involves managing the routing of I/O requests between virtual devices and the shared physical hardware. At the time of this writing, there are three ways to implement I/O virtualization: full device emulation, para-virtualization, and direct I/O. Full device emulation is the first approach for I/O virtualization. Generally, this approach emulates well-known, real-world devices.

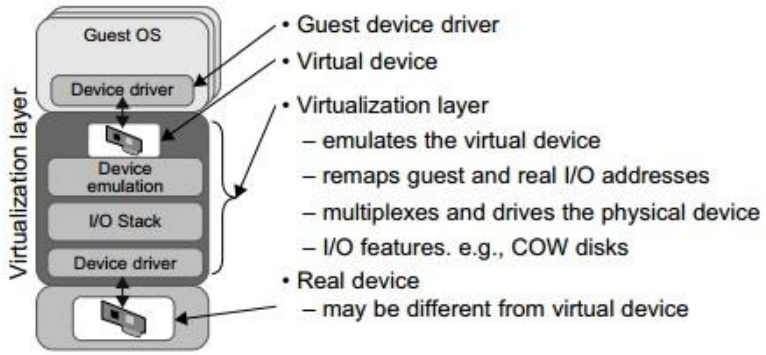


FIGURE 3.14

Device emulation for I/O virtualization implemented inside the middle layer that maps real I/O devices into the virtual devices for the guest device driver to use.

All the functions of a device or bus infrastructure, such as device enumeration, identification, interrupts, and DMA, are replicated in software. This software is located in the VMM and acts as a virtual device. The I/O access requests of the guest OS are trapped in the VMM which interacts with the I/O devices. The full device emulation approach is shown in Figure 3.14.

A single hardware device can be shared by multiple VMs that run concurrently. However, software emulation runs much slower than the hardware it emulates [10,15]. The para-virtualization method of I/O virtualization is typically used in Xen. It is also known as the split driver model consisting of a frontend driver and a backend driver. The frontend driver is running in Domain U and the backend dri-ver is running in Domain 0. They interact with each other via a block of shared memory. The frontend driver manages the I/O requests of the guest OSes and the backend driver is responsible for managing the real I/O

devices and multiplexing the I/O data of different VMs. Although para-I/O-virtualization achieves better device performance than full device emulation, it comes with a higher CPU overhead.

Direct I/O virtualization lets the VM access devices directly. It can achieve close-to-native performance without high CPU costs. However, current direct I/O virtualization implementations focus on networking for mainframes. There are a lot of challenges for commodity hardware devices. For example, when a physical device is reclaimed (required by workload migration) for later reassign-ment, it may have been set to an arbitrary state (e.g., DMA to some arbitrary memory locations) that can function incorrectly or even crash the whole system. Since software-based I/O virtualization requires a very high overhead of device emulation, hardware-assisted I/O virtualization is critical. Intel VT-d supports the remapping of I/O DMA transfers and device-generated interrupts. The architecture of VT-d provides the flexibility to support multiple usage models that may run unmodified, special-purpose, or "virtualization-aware" guest OSes.

Another way to help I/O virtualization is via self-virtualized I/O (SV-IO) [47]. The key idea of SV-IO is to harness the rich resources of a multicore processor. All tasks associated with virtualizing an I/O device are encapsulated in SV-IO. It provides virtual devices and an associated access API to VMs and a management API to the VMM. SV-IO defines one virtual interface (VIF) for every kind of virtua-lized I/O device, such as virtual network interfaces, virtual block devices (disk), virtual camera devices, and others. The guest OS interacts with the VIFs via VIF device drivers. Each VIF consists of two mes-sage queues. One is for outgoing messages to the devices and the other is for incoming messages from the devices. In addition, each VIF has a unique ID for identifying it in SV-IO.

Example 3.7 VMware Workstation for I/O Virtualization

The VMware Workstation runs as an application. It leverages the I/O device support in guest OSes, host OSes, and VMM to implement I/O virtualization. The application portion (VMApp) uses a driver loaded into the host operating system (VMDriver) to establish the privileged VMM, which runs directly on the hardware. A given physical processor is executed in either the host world or the VMM world, with the VMDriver facilitating the transfer of control between the two worlds. The VMware Workstation employs full device emulation to implement I/O virtualization.

Figure 3.15 shows the functional blocks used in sending and receiving packets via the emulated virtual NIC.

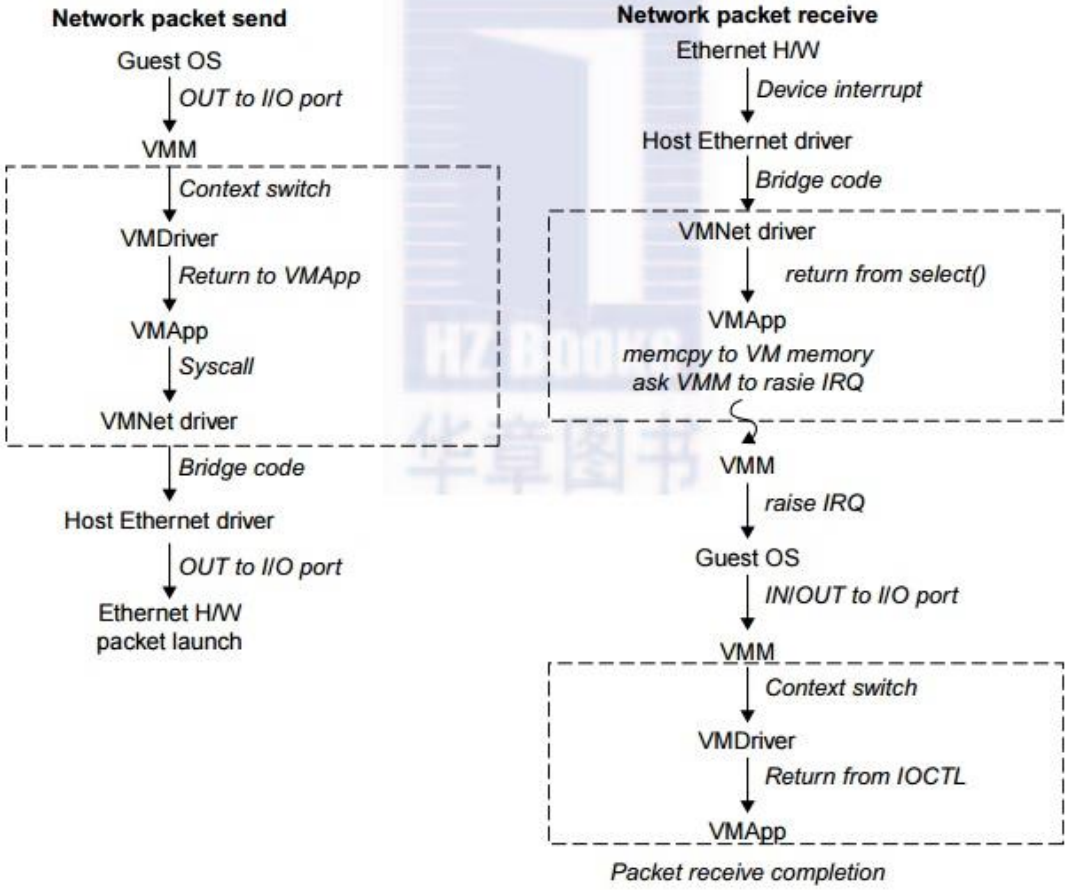


FIGURE 3.15

Functional blocks involved in sending and receiving network packets.

Virtual cluster and Resource Management:

A physical cluster is a collection of servers (physical machines) interconnected by a physical network such as a LAN. In Chapter 2, we studied various clustering techniques on physical machines. Here, we introduce virtual clusters and study its properties as well as explore their potential applications. In this section, we will study three critical design issues of virtual clusters: live migration of VMs, memory and file migrations, and dynamic deployment of virtual clusters.

When a traditional VM is initialized, the administrator needs to manually write configuration information or specify the configuration sources. When more VMs join a network, an inefficient configuration always causes problems

with overloading or underutilization. Amazon's Elastic Compute Cloud (EC2) is a good example of a web service that provides elastic computing power in a cloud. EC2 permits customers to create VMs and to manage user accounts over the time of their use. Most virtualization platforms, including XenServer and VMware ESX Server, support a brid-ging 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.

1. Physical versus Virtual Clusters

Virtual clusters are built with VMs installed at distributed servers from one or more physical clus-ters. The VMs in a virtual cluster are interconnected logically by a virtual network across several physical networks. Figure 3.18 illustrates the concepts of virtual clusters and physical clusters. Each virtual cluster is formed with physical machines or a VM hosted by multiple physical clusters. The virtual cluster boundaries are shown as distinct boundaries.

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

- The virtual cluster nodes can be either physical or virtual machines. Multiple VMs running with different OSes can be deployed on the same physical node.

- A VM runs with a guest OS, which is often different from the host OS, that manages the resources in the physical machine, where the VM is implemented.

- The purpose of using VMs is to consolidate multiple functionalities on the same server. This will greatly enhance server utilization and application flexibility.

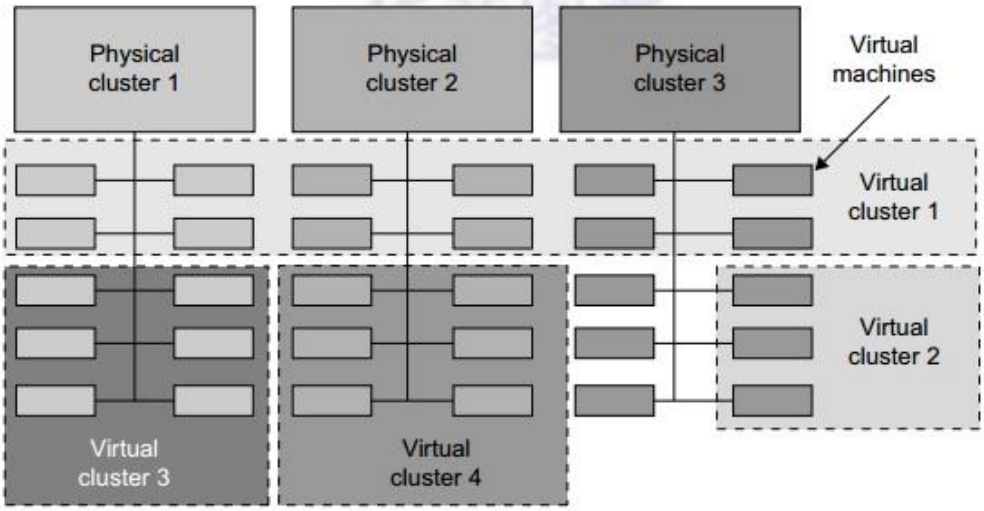


FIGURE 3.18

A cloud platform with four virtual clusters over three physical clusters shaded differently.

- VMs can be colonized (replicated) in multiple servers for the purpose of promoting distributed parallelism, fault tolerance, and disaster recovery.

- The size (number of nodes) of a virtual cluster can grow or shrink dynamically, similar to the way an overlay network varies in size in a peer-to-peer (P2P) network.

- The failure of any physical nodes may disable some VMs installed on the failing nodes. But the failure of VMs will not pull down the host system.

Since system virtualization has been widely used, it is necessary to effectively manage VMs running on a mass of physical computing nodes (also called virtual clusters) and consequently build a high-performance virtualized computing environment. This involves virtual cluster deployment, monitoring and management over large-scale clusters, as well as resource scheduling, load balancing, server consolidation, fault tolerance, and other techniques. The different node colors in Figure 3.18 refer to different virtual clusters. In a virtual cluster system, it is quite important to store the large number of VM images efficiently.

Figure 3.19 shows the concept of a virtual cluster based on application partitioning or customi-zation. The different colors in the figure represent the nodes in different virtual clusters. As a large number of VM images might be present, the most important thing is to determine how to store those images in

the system efficiently. There are common installations for most users or applications, such as operating systems or user-level programming libraries. These software packages can be preinstalled as templates (called template VMs). With these templates, users can build their own software stacks. New OS instances can be copied from the template VM. User-specific components such as programming libraries and applications can be installed to those instances.

Three physical clusters are shown on the left side of Figure 3.18. Four virtual clusters are created on the right, over the physical clusters. The physical machines are also called host systems. In contrast, the VMs are guest systems. The host and guest systems may run with different operating

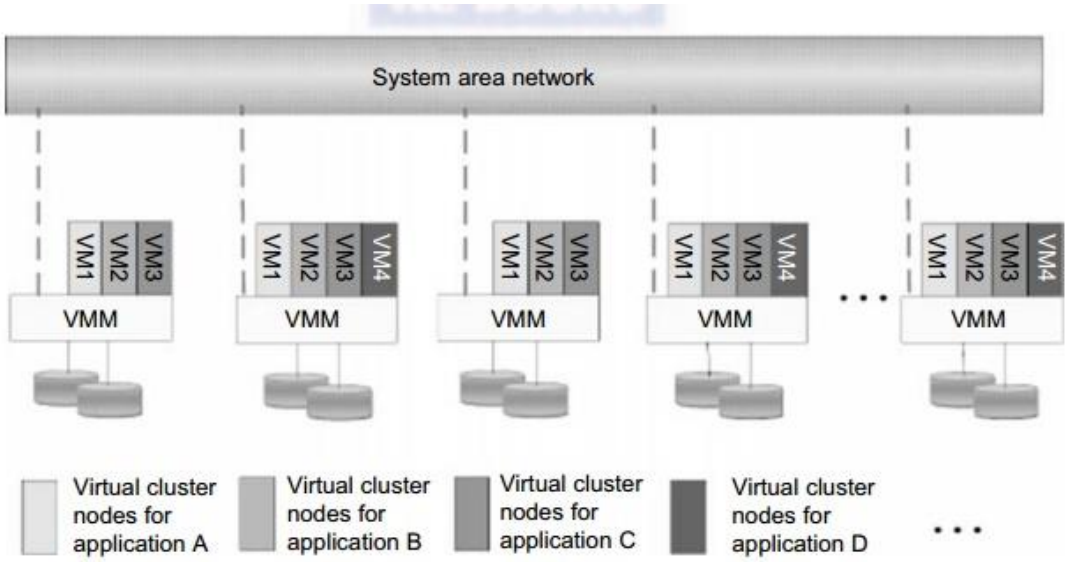


FIGURE 3.19
The concept of a virtual cluster based on application partitioning.

systems. Each VM can be installed on a remote server or replicated on multiple servers belonging to the same or different physical clusters. The boundary of a virtual cluster can change as VM nodes are added, removed, or migrated dynamically over time.

1.1 Fast Deployment and Effective Scheduling

The system should have the capability of fast deployment. Here, deployment means two things: to construct and distribute software stacks (OS, libraries, applications) to a physical node inside clus-ters as fast as possible, and to quickly switch runtime environments from one user's virtual cluster to another user's virtual cluster. If one user finishes using his system, the corresponding

virtual cluster should shut down or suspend quickly to save the resources to run other VMs for other users.

The concept of "green computing" has attracted much attention recently. However, previous approaches have focused on saving the energy cost of components in a single workstation without a global vision. Consequently, they do not necessarily reduce the power consumption of the whole clus-ter. Other cluster-wide energy-efficient techniques can only be applied to homogeneous workstations and specific applications. The live migration of VMs allows workloads of one node to transfer to another node. However, it does not guarantee that VMs can randomly migrate among themselves. In fact, the potential overhead caused by live migrations of VMs cannot be ignored.

Virtualization for data centre Automation:

1. Server Consolidation in Data Centers

In data centers, a large number of heterogeneous workloads can run on servers at various times. These heterogeneous workloads can be roughly divided into two categories: chatty workloads and noninter-active workloads. Chatty workloads may burst at some point and return to a silent state at some other point. A web video service is an example of this, whereby a lot of people use it at night and few peo-ple use it during the day. Noninteractive workloads do not require people's efforts to make progress after they are submitted. High-performance computing is a typical example of this. At various stages, the requirements for resources of these workloads are dramatically different. However, to guarantee that a workload will always be able to cope with all demand levels, the workload is statically allo-cated enough resources so that peak demand is satisfied.

Therefore, it is common that most servers in data centers are underutilized. A large amount of hardware, space, power, and management cost of these servers is wasted. Server consolidation is an approach to improve the low utility ratio of hardware resources by reducing the number of physical servers. Among several server consolidation techniques such as centralized and physical consolida-tion, virtualization-based server consolidation is the most powerful. Data centers need to optimize their resource management. Yet these techniques are performed with the granularity of a full server machine, which makes resource management far from well optimized. Server virtualization enables smaller resource allocation than a physical machine.

In general, the use of VMs increases resource management complexity. This causes a challenge in terms of how to improve resource utilization as well as guarantee QoS in data centers. In detail, server virtualization has the following side effects:

- Consolidation enhances hardware utilization. Many underutilized servers are consolidated into fewer servers to enhance resource utilization. Consolidation also facilitates backup services and disaster recovery.

- This approach enables more agile provisioning and deployment of resources. In a virtual environment, the images of the guest OSes and their applications are readily cloned and reused.

- The total cost of ownership is reduced. In this sense, server virtualization causes deferred purchases of new servers, a smaller data-center footprint, lower maintenance costs, and lower power, cooling, and cabling requirements.

- This approach improves availability and business continuity. The crash of a guest OS has no effect on the host OS or any other guest OS. It becomes easier to transfer a VM from one server to another, because virtual servers are unaware of the underlying hardware.

To automate data-center operations, one must consider resource scheduling, architectural support, power management, automatic or autonomic resource management, performance of analytical mod-els, and so on. In virtualized data centers, an efficient, on-demand, fine-grained scheduler is one of the key factors to improve resource utilization. Scheduling and reallocations can be done in a wide range of levels in a set of data centers. The levels match at least at the VM level, server level, and data-center level. Ideally, scheduling and resource reallocations should be done at all levels. However, due to the complexity of this, current techniques only focus on a single level or, at most, two levels.

Dynamic CPU allocation is based on VM utilization and application-level QoS metrics. One method considers both CPU and memory flowing as well as automatically adjusting resource over-head based on varying workloads in hosted services. Another scheme uses a two-level resource management system to handle the complexity involved. A local controller at the VM level and a

global controller at the server level are designed. They implement autonomic resource allocation via the interaction of the local and global controllers. Multicore and virtualization are two cutting tech-niques that can enhance each other.

However, the use of CMP is far from well optimized. The memory system of CMP is a typical example. One can design a virtual hierarchy on a CMP in data centers. One can consider protocols that minimize the memory access time, inter-VM interferences, facilitating VM reassignment, and supporting inter-VM sharing. One can also consider a VM-aware power budgeting scheme using multiple managers integrated to achieve better power management. The power budgeting policies cannot ignore the heterogeneity problems. Consequently, one must address the trade-off of power saving and data-center performance.

2. Virtual Storage Management

The term “storage virtualization” was widely used before the renaissance of system virtualization. Yet the term has a different meaning in a system virtualization environment. Previously, storage virtualiza-tion was largely used to describe the aggregation and repartitioning of disks at very coarse time scales for use by physical machines. In system virtualization, virtual storage includes the storage managed by VMMs and guest OSes. Generally, the data stored in this environment can be classified into two categories: VM images and application data. The VM images are special to the virtual environment, while application data includes all other data which is the same as the data in traditional OS environments.

The most important aspects of system virtualization are encapsulation and isolation. Traditional operating systems and applications running on them can be encapsulated in VMs. Only one operating system runs in a virtualization while many applications run in the operating system. System virtualization allows multiple VMs to run on a physical machine and the VMs are completely isolated. To achieve encapsulation and isolation, both the system software and the hardware platform, such as CPUs and chipsets, are rapidly updated. However, storage is lagging. The storage systems become the main bottleneck of VM deployment.

In virtualization environments, a virtualization layer is inserted between the hardware and traditional operating systems or a traditional operating system is modified to support virtualization. This procedure complicates storage operations. On the one hand, storage management of the guest OS per-forms as though it is operating in a real hard disk while the guest OSes cannot access the hard disk directly. On the other hand, many guest OSes contest the hard disk when many VMs are running on a single physical machine. Therefore, storage management of the underlying VMM is much more complex than that of guest OSes (traditional OSes).

In addition, the storage primitives used by VMs are not nimble. Hence, operations such as remap-ping volumes across hosts and checkpointing disks are frequently clumsy and esoteric, and sometimes simply unavailable. In data centers, there are often thousands of VMs, which cause the VM images to become flooded. Many researchers tried to solve these problems in virtual storage management. The main purposes of their research are to make management easy while enhancing performance and reducing the amount of storage occupied by the VM images. Parallax is a distributed storage system customized for virtualization environments. Content Addressable Storage (CAS) is a solution to reduce the total size of VM images, and therefore supports a large set of VM-based systems in data centers.

Since traditional storage management techniques do not consider the features of storage in virtualization environments, Parallax designs a novel architecture in which storage features that have traditionally been implemented directly on high-end storage arrays and switchers are relocated into a federation of storage VMs. These storage VMs share the same physical hosts as the VMs that they serve. Figure 3.26 provides an overview of the Parallax system architecture. It supports all popular system virtualization techniques, such as paravirtualization and full virtualization. For each physical machine, Parallax customizes a special storage appliance VM. The storage appliance VM acts as a block virtualization layer between individual VMs and the physical storage device. It provides a virtual disk for each VM on the same physical machine.

Example 3.11 Parallax Providing Virtual Disks to Client VMs from a Large Common Shared Physical Disk

The architecture of Parallax is scalable and especially suitable for use in cluster-based environments. Figure 3.26 shows a high-level view of the structure of a Parallax-based cluster. A cluster-wide

administrative domain manages all storage appliance VMs, which makes storage management easy. The storage appliance

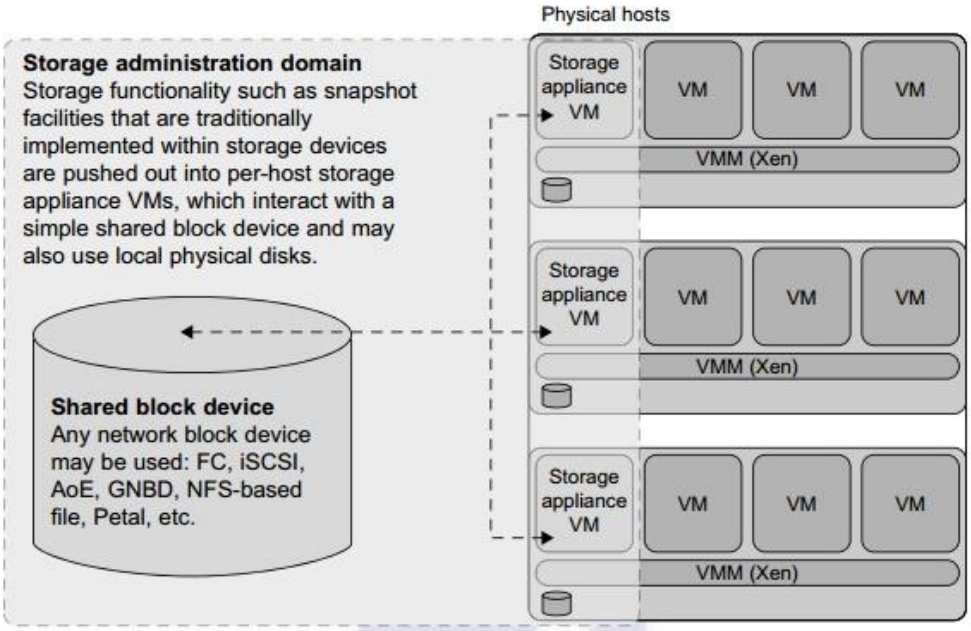


FIGURE 3.26

Parallax is a set of per-host storage appliances that share access to a common block device and presents virtual disks to client VMs.

Virtual box virtualization software:

Oracle VM VirtualBox (formerly **Sun VirtualBox**, **Sun xVM VirtualBox** and **InnoTek VirtualBox**) is a type-2 hypervisor for x86 virtualization developed by Oracle Corporation. VirtualBox was originally created by InnoTek Systemberatung GmbH, which was acquired by Sun Microsystems in 2008, which was in turn acquired by Oracle in 2010.

VirtualBox may be installed on Microsoft Windows, macOS, Linux, Solaris and OpenSolaris. There are also ports to FreeBSD[5] and Genode.[6] It supports the creation and management of guest virtual machines running Windows, Linux, BSD, OS/2, Solaris, Haiku, and OSx86,[7] as well as limited virtualization of macOS guests on Apple hardware.[8][9] For some guest operating systems, a "Guest Additions" package of device drivers and system applications is available,[10][11] which typically improves performance, especially that of graphics, and allows changing the resolution of the guest OS automatically when the window of the virtual machine on the host OS is resized.

Released under the terms of the GNU General Public License and, optionally, the CDDL for most files of the source distribution, VirtualBox is free and open-source software, though the Extension Pack is proprietary software. The License to VirtualBox was relicensed to GPLv3 with linking exceptions to the CDDL and other GPL-incompatible licenses.

Short Answer Questions:

1.What are the basics of virtualization?

2.Define virtualization?

3.What is the structure of virtualization?

4.Define I/O devices in cloud computing?

Long Answer Questions:

1.Explain different types of virtualization?

2.Explain implementation of virtualization?

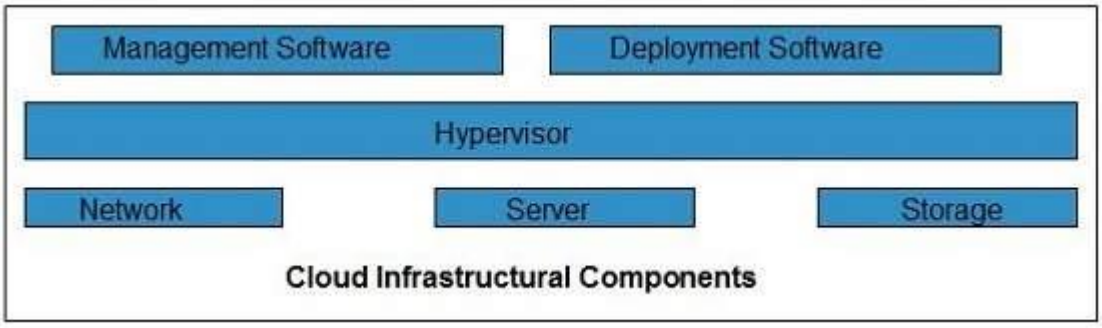
3.Explain virtual box virtualization software?

4.Define virtual cluster and research management?

UNIT-3

Cloud Infrastructure:

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



Hypervisor

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

Management Software

It helps to maintain and configure the infrastructure.

Deployment Software

It helps to deploy and integrate the application on the cloud.

Network

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

Server

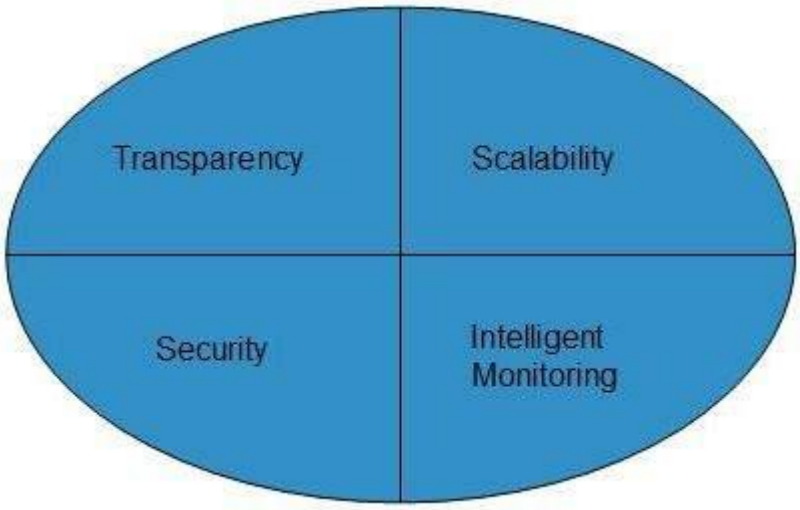
The **server** helps to compute the resource sharing and offers other services such as resource allocation and de-allocation, monitoring the resources, providing security etc.

Storage

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

Infrastructural Constraints

Fundamental constraints that cloud infrastructure should implement are shown in the following diagram:



Transparency

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

Scalability

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

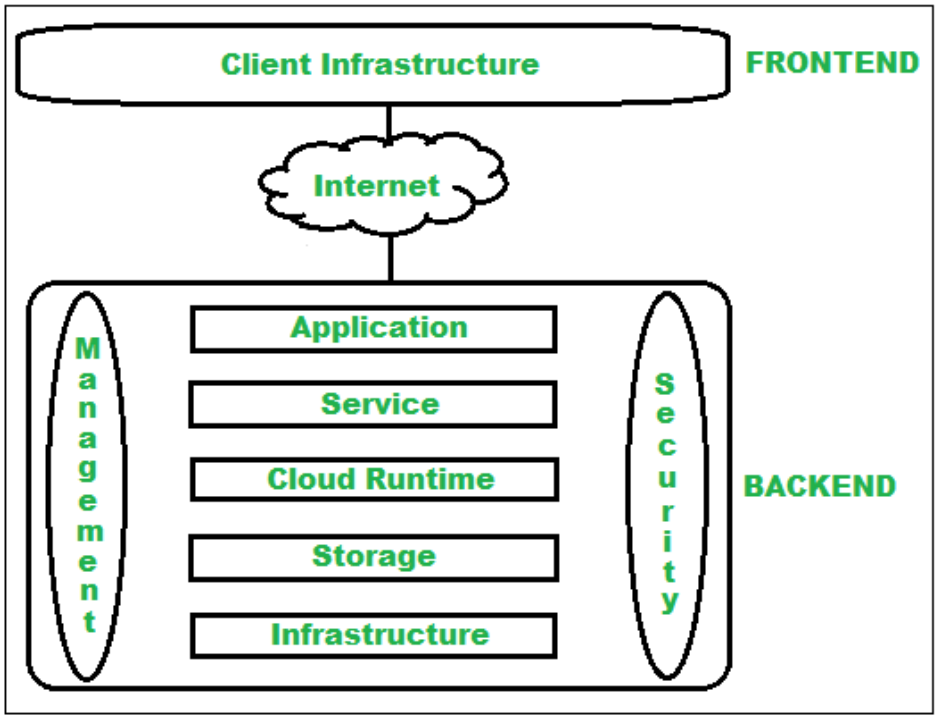
Architectural design of compute and storage clouds:

Cloud Computing , which is one of the demanding technology of the current time and which is giving a new shape to every organization by providing on demand virtualized services/resources. Starting from small to medium and medium to large, every organization use cloud computing services for storing information and accessing it from anywhere and any time only with the help of internet. In this article, we will know more about the internal architecture of cloud computing.
Transparency, scalability, security and intelligent monitoring are some of the most important constraints which every cloud infrastructure should experience. Current research on other important constraints is helping cloud computing system to come up with new features and strategies with a great capability of providing more advanced cloud solutions.

Cloud Computing Architecture :
The cloud architecture is divided into 2 parts i.e.

1. Frontend
2. Backend

The below figure represents an internal architectural view of cloud computing.



Architecture of Cloud Computing

Architecture of cloud computing is the combination of both SOA (Service Oriented Architecture) and EDA (Event Driven Architecture). Client infrastructure, application, service, runtime cloud, storage, infrastructure, management and security all these are the components of cloud computing architecture.

1. Frontend :

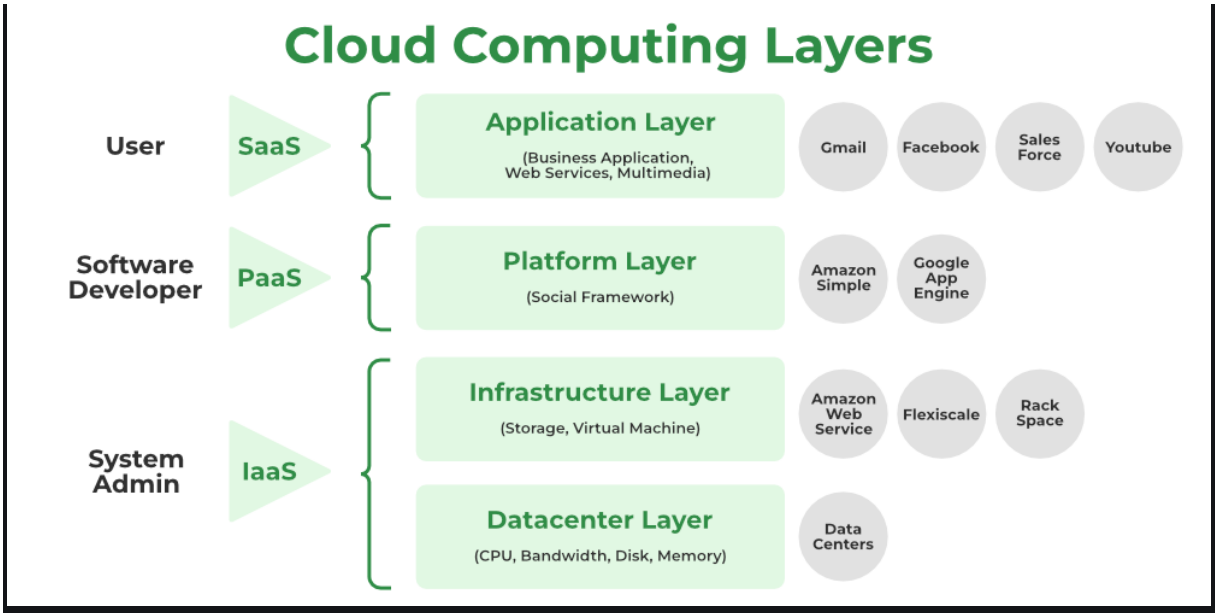
Frontend of the cloud architecture refers to the client side of cloud computing system. Means it contains all the user interfaces and applications which are used by the client to access the cloud computing services/resources. For example, use of a web browser to access the cloud platform.

- **Client Infrastructure –** Client Infrastructure is a part of the frontend component. It contains the applications and user interfaces which are required to access the cloud platform.
- In other words, it provides a GUI(Graphical User Interface) to interact with the cloud.

2. Backend :

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

Layered cloud Architecture development:



Application Layer

1. The application layer, which is at the top of the stack, is where the actual cloud apps are located. Cloud applications, as opposed to traditional applications, can take advantage of the **automatic-scaling** functionality to gain greater performance, availability, and lower operational costs.
2. This layer consists of different Cloud Services which are used by cloud users. Users can access these applications according to their needs. Applications are divided into **Execution layers** and **Application layers**.
3. In order for an application to transfer data, the application layer determines whether communication partners are available. Whether enough cloud resources are accessible for the required communication is decided at the application layer. Applications must cooperate in order to communicate, and an application layer is in charge of this.
4. The application layer, in particular, is responsible for processing IP traffic handling protocols like Telnet and FTP. Other examples of application layer systems include web browsers, SNMP protocols, HTTP protocols, or HTTPS, which is HTTP's successor protocol.

Platform Layer

1. The operating system and application software make up this layer.
2. Users should be able to rely on the platform to provide them with **Scalability, Dependability, and Security Protection** which gives users a space to create their apps, test operational processes, and keep track of execution outcomes and performance. SaaS application implementation's application layer foundation.
3. The objective of this layer is to deploy applications directly on virtual machines.
4. Operating systems and application frameworks make up the platform layer, which is built on top of the infrastructure layer. The platform layer's goal is to lessen the difficulty of deploying programmers directly into VM containers.
5. By way of illustration, Google App Engine functions at the platform layer to provide API support for implementing storage, databases, and business logic of ordinary web apps.

Infrastructure Layer

1. It is a layer of virtualization where physical resources are divided into a collection of virtual resources using virtualization technologies like Xen, KVM, and VMware.
2. **This layer serves as the Central Hub of the Cloud Environment**, where resources are constantly added utilizing a variety of virtualization techniques.

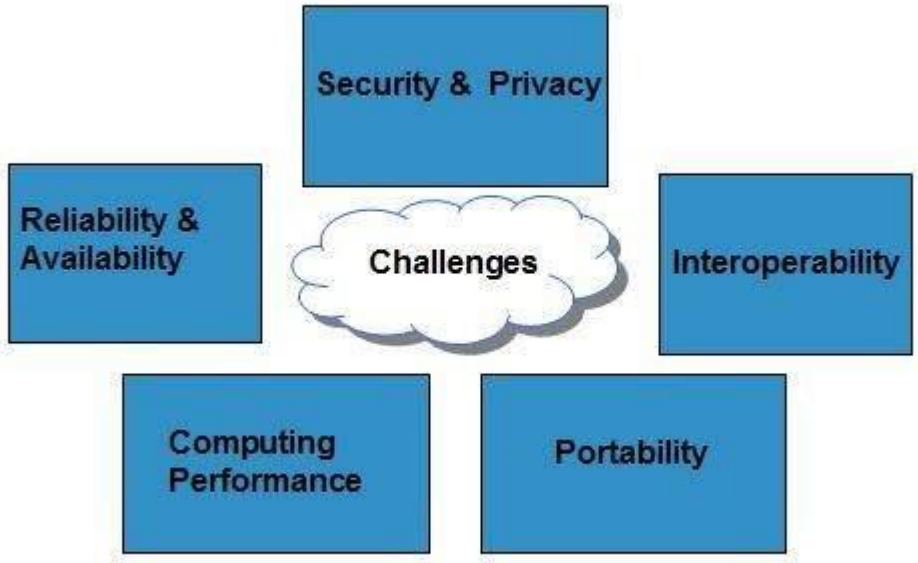
3. A base upon which to create the platform layer. constructed using the virtualized network, storage, and computing resources. Give users the flexibility they want.
4. Automated resource provisioning is made possible by virtualization**,** which also improves infrastructure management.
5. The infrastructure layer sometimes referred to as the virtualization layer, partitions the physical resources using virtualization technologies like **Xen, KVM, Hyper-V, and VMware** to create a pool of compute and storage resources.
6. The infrastructure layer is crucial to cloud computing since virtualization technologies are the only ones that can provide many vital capabilities, like dynamic resource assignment.

Datacenter Layer

- In a cloud environment, this layer is responsible for **Managing Physical Resources** such as servers, switches, routers, power supplies, and cooling systems.
- Providing end users with services requires all resources to be available and managed in data centers.
- Physical servers connect through high-speed devices such as routers and switches to the data center.
- In software application designs, the division of business logic from the persistent data it manipulates is well-established. This is due to the fact that the same data cannot be incorporated into a single application because it can be used in numerous ways to support numerous use cases. The requirement for this data to become a service has arisen with the introduction of microservices.
- A single database used by many microservices creates a very close coupling. As a result, it is hard to deploy new or emerging services separately if such services need database modifications that may have an impact on other services. A data layer containing many databases, each serving a single microservice or perhaps a few closely related microservices, is needed to break complex service interdependencies.

Design challenges:

Cloud computing, an emergent technology, has placed many challenges in different aspects of data and information handling. Some of these are shown in the following diagram:



Security and Privacy

Security and Privacy of information is the biggest challenge to cloud computing. Security and privacy issues can be overcome by employing encryption, security hardware and security applications.

Portability

This is another challenge to cloud computing that applications should easily be migrated from one cloud provider to another. There must not be vendor lock-in. However, it is not yet made possible because each of the cloud provider uses different standard languages for their platforms.

Interoperability

It means the application on one platform should be able to incorporate services from the other platforms. It is made possible via web services, but developing such web services is very complex.

Computing Performance

Data intensive applications on cloud requires high network bandwidth, which results in high cost. Low bandwidth does not meet the desired computing performance of cloud application.

Reliability and Availability

It is necessary for cloud systems to be reliable and robust because most of the businesses are now becoming dependent on services provided by third-party.

Intercloud Resource Management:

A theoretical model for cloud computing services is referred to as the “inter-cloud” or “cloud of clouds.” combining numerous various separate clouds into a single fluid mass for on-demand operations Simply put, the inter-cloud would ensure that a cloud could utilize resources outside of its range using current agreements with other cloud service providers. There are limits to the physical resources and the geographic reach of any one cloud.

Need of Inter-Cloud

Due to their Physical Resource limits, Clouds have certain Drawbacks:

- When a cloud’s computational and storage capacity is completely depleted, it is unable to serve its customers.
- The Inter-Cloud addresses these circumstances when one cloud would access the computing, storage, or any other resource of the infrastructures of other clouds8.

Benefits of the Inter-Cloud Environment include:

- Avoiding vendor lock-in to the cloud client
- Having access to a variety of geographical locations, as well as enhanced application resiliency.
- Better service level agreements (SLAs) to the cloud client
- Expand-on-demand is an advantage for the cloud provider.

Inter-Cloud Resource Management

A cloud’s infrastructure’s processing and storage capacity could be exhausted. combining numerous various separate clouds into a single fluid mass for on-demand operations. Simply put, the intercloud would ensure that a cloud could utilize resources outside of its range combining numerous various separate clouds into a single fluid mass for on-demand operations. Such requests for service allocations received by its clients would still be met by it.

Resource provisioning And Platform Deployment:

Resource provisioning means the

• Selection

• deployment

• run-time management of S/W & H/W resources for ensuring guaranteed performance for applications

Resource provisioning phases

• Reservation phase – reserve resources

• Expending phase - utilize resources

• On-demand phase–provision of more resources

Based on the Application, it may be

Static Provisioning –unchanging demands

Dynamic Provisioning

- demands may change or vary

Self Provisioning –

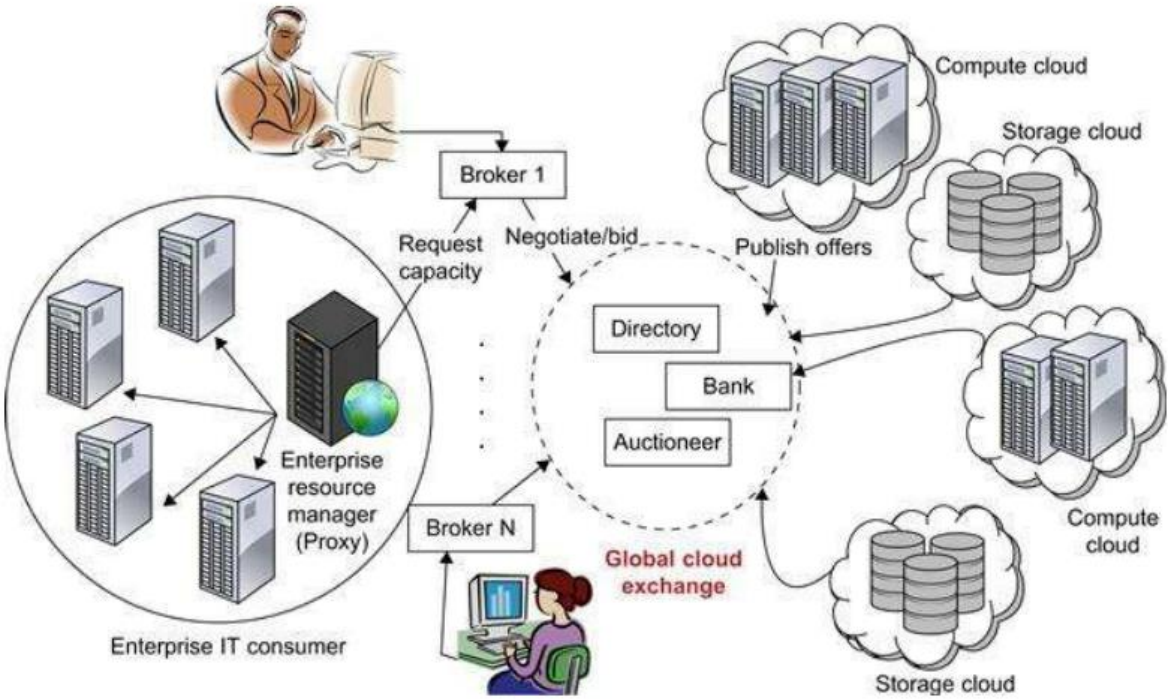
customer purchase computing resource from provider

Platform Deployment:

- Platform deployment options include manual installation, unattended or scripted installs, server cloning, and a newer approach called server provisioning.
- ☑
- Some organizations may want to use a mixed or hybrid approach that uses the best pieces of these various solutions. Selecting the best platform deployment option for your environment involves a detailed review of the requirements for each with special emphasis on the additional software and engineer expertise required. Most environments may lean heavily on one of these solutions, but may ultimately be categorized as a hybrid due to the nuances required integrating into the existing networking environment.

Global Exchange of Cloud Resources:

Cloud Exchange (CEx) serves as a market maker, bringing service providers and users together. The University of Melbourne proposed it under Intercloud architecture (Cloudbus). It supports brokering and exchanging cloud resources for scaling applications across multiple clouds. It aggregates the infrastructure demands from application brokers and evaluates them against the available supply. It supports the trading of cloud services based on competitive economic models such as commodity markets and auctions.



Global exchange of cloud resources
Source: https://snscourseware.org/snsctnew/files/1583815568.pdf

Now we will talk about various entities of the Global exchange of cloud resources.

Entities of the Global exchange of cloud resources

Now we will talk about the various entities of the global exchange of cloud resources.

Market directory

A market directory is an extensive database of resources, providers, and participants using the resources. Participants can use the market directory to find providers or customers with suitable offers.

Auctioneers

Auctioneers clear bids and ask from market participants regularly. Auctioneers sit between providers and customers and grant the resources available in the Global exchange of cloud resources to the highest bidding customer.

Brokers

Brokers mediate between consumers and providers by buying capacity from the provider and sub-leasing these to the consumers. They must select consumers whose apps will provide the most utility. Brokers may also communicate with resource providers and other brokers to acquire or trade resource shares. To make decisions, these brokers are equipped with a negotiating module informed by the present conditions of the resources and the current demand.

Service-level agreements(SLAs)

The service level agreement (SLA) highlights the details of the service to be provided in terms of metrics that have been agreed upon by all parties, as well as penalties for meeting and failing to meet the expectations.

The consumer participates in the utility market via a resource management proxy that chooses a set of brokers based on their offering.SLAs are formed between the consumer and the brokers, which bind the latter to offer the guaranteed resources.

After that, the customer either runs their environment on the leased resources or uses the provider's interfaces to scale their applications.

Providers

A provider has a **price-setting mechanism** that determines the current price for their source based on market conditions, user demand, and the current degree of utilization of the resource.
Based on an initial estimate of utility, an admission-control mechanism at a provider's end selects the auctions to participate in or to negotiate with the brokers.

Four Levels of Federation:

The implementation and management of several internal and external cloud computing services to meet business demands is known as cloud federation, sometimes known as federated cloud. A global cloud system combines community, private, and public clouds into scalable computing platforms. By utilizing a common standard to link the cloud environments of several cloud providers, a federated cloud is built.

Levels of Cloud Federation

Cloud Federation stack

Each level of the cloud federation poses unique problems and functions at a different level of the IT stack. Then, several strategies and technologies are needed. The answers to the problems encountered at each of these levels when combined form a reference model for a cloud federation.

Conceptual Level

The difficulties in presenting a cloud federation as an advantageous option for using services rented from a single cloud provider are addressed at the conceptual level. At this level, it's crucial to define the new opportunities that a federated environment brings in comparison to a single-provider solution and to explicitly describe the benefits of joining a federation for service providers or service users.

At this level, the following factors need attention:

- The reasons that cloud providers would want to join a federation.
- Motivations for service users to use a federation.
- Benefits for service providers who rent their services to other service providers. Once a provider joins the federation, they have obligations.
- Agreements on trust between suppliers.
- Consumers versus transparency.

Future of Federation:

Federation creates a hybrid cloud environment with an increased focus on maintaining the integrity of corporate policies and data integrity. Think of federation as a pool of clouds connected through a channel of gateways; gateways which can be used to optimize a cloud for a service or set of specific services. Such gateways can be used to segment service audiences or to limit access to specific data sets. In essence, federation has the ability for enterprises to service their audiences with economy of scale without exposing critical applications or vital data through weak policies or vulnerabilities.

Many would raise the question: if Federation creates multiples of clouds, doesn't that mean cloud benefits are diminished? I believe the answer is no, due to the fact that a fundamental change has transformed enterprises through the original adoption of cloud computing, namely the creation of a flexible environment able to adapt rapidly to changing needs based on policy and automation.

We can't time travel like the movies or even predict the future with 100 percent accuracy but hybrid cloud is well on its way, and federation is what the preferred cloud model will be. Through a federated model, hybrid cloud adoption will dramatically accelerate. A future comprised of many clouds simply makes sense – and Federation casts a true vision of global computing.

Short Answer Questions:

1.What is cloud infrastructure?

2.Define Design Challenges?

3.What is future of federation?

4.Write any two applications of federated services?

5.Write four levels of ferderation?

Long Answer Questions:

1.Define Architectural design of compute and storage clouds?

2.Describe Layered cloud Architecture development?

3.Describe global exchange of Cloud resources?

4.Define federated services and its applications?

UNIT-4

Programming Model:

This article is about the definition of the term 'programming model'. For classification of programming languages, see Programming paradigm.

A **programming model** is an execution model coupled to an API or a particular pattern of code. In this style, there are actually two execution models in play: the execution model of the base programming language and the execution model of the programming model. An example is Spark where Java is the base language, and Spark is the programming model. Execution may be based on what appear to be library calls. Other examples include the POSIX Threads library and Hadoop's MapReduce.[1] In both cases, the execution model of the programming model is different from that of the base language in which the code is written. For example, the C programming language has no behavior in its execution model for input/output or thread behavior. But such behavior can be invoked from C syntax, by making what appears to be a call to a normal C library.

What distinguishes a programming model from a normal library is that the behavior of the call cannot be understood in terms of the language the program is written in. For example, the behavior of calls to the POSIX thread library cannot be understood in terms of the C language. The reason is that the call invokes an execution model that is different from the execution model of the language. This invocation of an outside execution model is the defining characteristic of a programming *model*, in contrast to a programming *language*.

In parallel computing, the execution model often must expose features of the hardware in order to achieve high performance. The large amount of variation in parallel hardware causes a concurrent need for a similarly large number of parallel execution models. It is impractical to make a new language for each execution model, hence it is a common practice to invoke the behaviors of the parallel execution model via an API. So, most of the programming effort is done via parallel programming models rather than parallel languages. Unfortunately, the terminology around such programming models tends to focus on the details of the hardware that inspired the execution model, and in that insular world the mistaken belief is formed that a programming model is only for the case when an execution model is closely matched to hardware features.

Parallel and Distributed Programming Paradigms:

PARALLEL AND DISTRIBUTED PROGRAMMING PARADIGMS 1

• Parallel and distributed program as a parallel program running on a set of computing engines or a distributed computing system. • A distributed computing system is a set of computational engines connected by a network to achieve a common goal of running a job or an application. • A computer cluster or network of workstations is an example of a distributed computing system. • Parallel computing is the simultaneous use of more than one computational engine to run a job or an application. • Running a parallel program on a distributed system decreases application response time and increases throughput and resource utilization. 2

PARALLEL COMPUTING AND PROGRAMMING PARADIGMS • Partitioning • Computation partitioning - Job into smaller tasks • Data partitioning - Input into smaller pieces • Mapping - assigns the either smaller parts of a program or the smaller pieces of data to underlying resources. • Synchronization - synchronization and coordination among workers is necessary to avoid race conditions and data dependency • Communication - communication is always

triggered when the intermediate data is sent to workers. • Scheduling - A scheduler selects a sequence of tasks or data pieces to be assigned to the workers. 3

Motivation for Programming Paradigms • Handling the whole data flow of parallel and distributed programming is very time-consuming • Aim to provide an abstraction layer to hide implementation details of the data flow which users formerly ought to write codes for. • Parallel and distributed programming models are used to (1) to improve productivity of programmers (2) to decrease programs time to market (3) to leverage underlying resources more efficiently (4) to increase system throughput (5) to support higher levels of abstraction • MapReduce, Hadoop, and Dryad are the most recently proposed parallel and distributed programming models. 4

MAPREDUCE • MapReduce, is a software framework which supports parallel and distributed computing on large data sets. • This software framework abstracts the data flow of running a parallel program on a distributed computing system • It provides two interfaces in the form of two functions: Map and Reduce. • Users can override these two functions to interact with and manipulate the data flow of running their programs. 5

MapReduce Framework 7

Structure of the user program Map Function (….) { … } Reduce Function (….) { … } Main Function (….) { Initialize Spec object … MapReduce (Spec, & Results) } 8

MAPREDUCE LOGICAL DATA FLOW The input data to the Map function is in the form of a (key, value) pair. The user-defined Map function processes each input (key, value) pair It produces a number of (zero, one, or more) intermediate (key, value) pairs. The Reduce function receives the intermediate (key, value) pairs in the form of a group of intermediate values associated with one intermediate key, (key, [set of values]). • MapReduce framework forms these groups by first sorting the intermediate (key, value) pairs and then grouping values with the same key. • The Reduce function processes each (key, [set of values]) group and produces a set of (key, value) pairs as output. 9

Logical Data Flow of MapReduce 10

Data Flow of the Word-Count Problem 11

STRATEGY TO SOLVE MAPREDUCE PROBLEMS After grouping all the intermediate data, the values of all occurrences of the same key are sorted and grouped together. After grouping, each key becomes unique in all intermediate data. Finding unique keys is the starting point to solve a typical MapReduce problem. Then the intermediate (key, value) pairs as the output of the Map function will be automatically found. 12

Example for solving MapReduce Problems Problem 1: Counting the number of occurrences of each word in a collection of documents Solution: unique “key”: each word, intermediate “value”: number of occurrences Problem 2: Counting the number of occurrences of words having the same size, or the same number of letters, in a collection of documents Solution: unique “key”: each word, intermediate “value”: size of the word Problem 3: Counting the number of occurrences of anagrams in a collection of documents. Anagrams are words with the same set of letters but in a different order (e.g., the words “listen” and “silent”). Solution: unique “key”: alphabetically sorted sequence of letters for each word (e.g., “eilnst”), intermediate “value”: number of occurrences 13STEPS FOR MAPREDUCE ACTUAL DATA AND CONTROL FLOW 1.Data partitioning 2.Computation partitioning 3.Determining the master and workers 4.Reading the input 5.Map function 6.Combiner function 7.Partitioning function 8.Synchronization 9.Communication 10.Sorting and Grouping 11.Reduce function 14

Use of MapReduce partitioning function to link the Map and Reduce workers 15

Data flow implementation of many functions in Map and in Reduce workers 16

Control flow implementation of MapReduce functionalities in Map and Reduce workers 17

Mapping Applications ClassDescriptionMachine Architecture SynchronousSimilar to instruction level operation as SIMD architecture SIMD Loosely Synchronous Independent compute operations for each CPU MIMD or MPP AsynchronousSupports Combinatorial computingShared memory Pleasingly parallelEach component is independentGrid computing to cloud computing MetaproblemsCoarse grained combinations of categories Grids of clusters MapReduce++ (Twister) Pleasing parallel map Map followed by reductions Iterative map Master-worker or mapreduce Mapreduce Twister 18

PROGRAMMING THE GOOGLE APP ENGINE • A client environment that includes an Eclipse plug-in for Java allows to debug GAE on the local machine. • GWT Google Web Toolkit is available for Java web application developers. • Developers can use this, or any other language using a JVMbased interpreter or compiler, such as JavaScript or Ruby. • Python is often used with frameworks • Data store is a NOSQL data management system • Java offers Java Data Object (JDO) and Java Persistence API (JPA) interfaces implemented by the open source Data Nucleus Access platform • Google SDC Secure Data Connection can tunnel through the Internet and link your intranet to an external GAE application. • URL Fetch operation provides the ability for applications to fetch resources and communicate with other hosts over the Internet using HTTP and HTTPS requests. • A GAE application is configured to consume resources up to certain limits or quotas. 23

Programming Environment of GoogleApp Engine 24

PROGRAMMING ON AMAZON AWS • Amazon offers a Relational Database Service (RDS) with a messaging interface • Amazon has NOSQL support in SimpleDB • It offers the Simple Queue Service (SQS) and Simple Notification Service (SNS), which are the cloud implementations of services • Auto-scaling and elastic load balancing services are provided in Amazon • Auto-scaling enables to automatically scale Amazon EC2 capacity up or down according to the given conditions • Elastic load balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances • CloudWatch is a web service that provides monitoring for AWS cloud resources used for both auto scaling and elastic load balancing 25

PROGRAMMING ON AMAZON EC2 • Customers can rent VMs instead of physical machines to run their own applications. • The elastic feature of such a service is that a customer can create, launch, and terminate server instances as needed, paying by the hour for active servers. • Instances are often called Amazon Machine Images (AMIs) which are preconfigured with operating systems Types of AMI • Private AMI - Images created are private by default. • Public AMI - Images created by users and released to the AWS community • Paid AMI - Images can be created by providing specific functions that can be launched by anyone willing to pay per usage 26

Amazon EC2 Execution Environment 27

Types of IaaS instances in AWS Platform • Standard instances are well suited for most applications. • Micro instances provide a small number of consistent CPU resources • High-memory instances offer large memory sizes for high-throughput applications • High-CPU instances have proportionally more CPU resources than memory and are well suited for compute- intensive applications. • Cluster compute instances provide proportionally high CPU resources with increased network performance 28

AMAZON SIMPLE STORAGE SERVICE • Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web. • S3 provides the object-oriented storage service for users. • Users can access their objects through Simple Object Access Protocol (SOAP) • The fundamental operation unit of S3

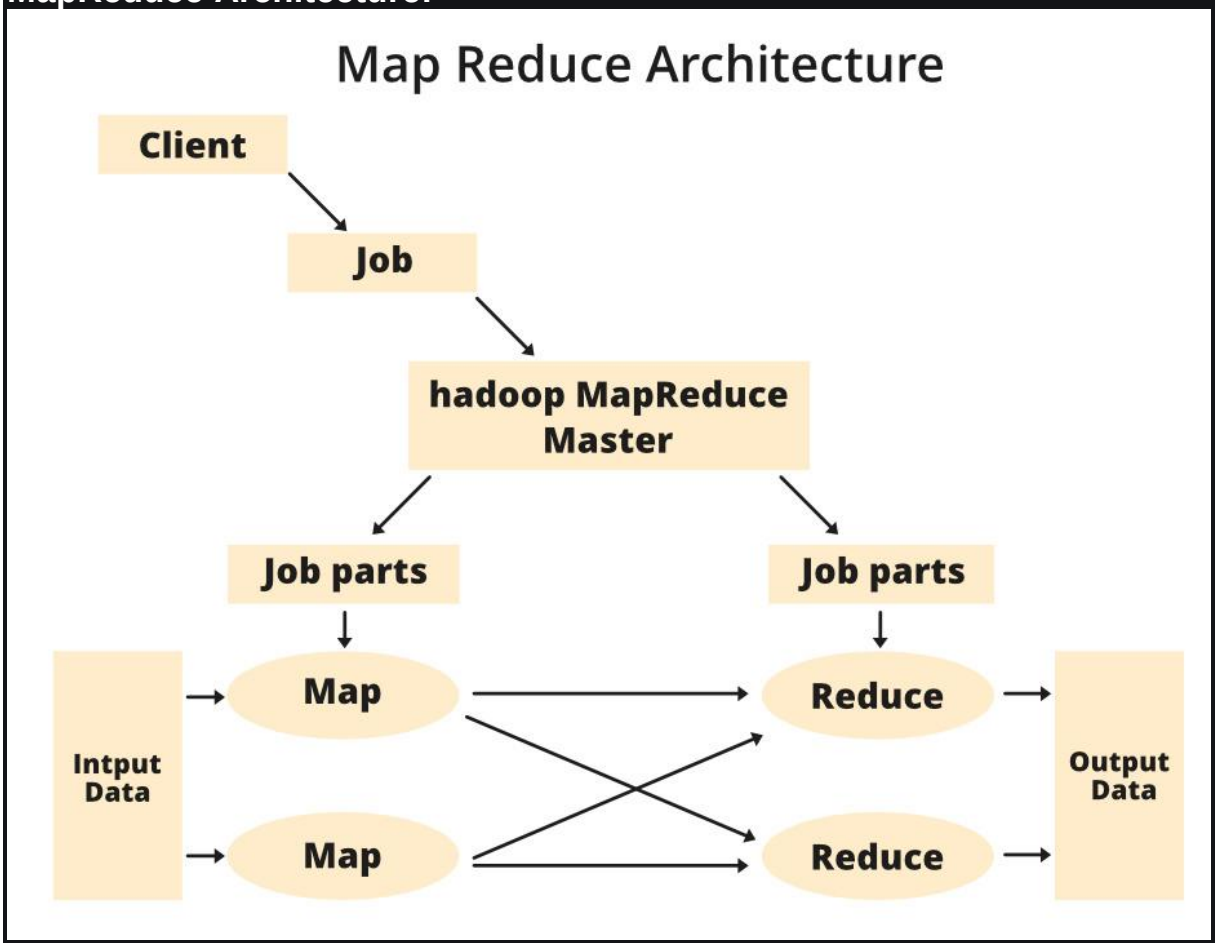
is called an object. • Each object is stored in a bucket and retrieved via a unique order • The storage provided by S3 can be viewed as a very coarse-grained key- value pair. • Through the key-value programming interface, users can write, read, and delete objects • Web service interfaces used here are SOAP and REST 29

Features of S3 • Redundant through geographic dispersion. • High availability of objects with cheaper reduced redundancy storage (RRS). • Authentication mechanisms to ensure that data is kept secure from unauthorized access. • Per-object URLs and access control lists (ACL). • Default download protocol of HTTP. • There is no data transfer charge for data transferred between Amazon EC2 and Amazon S3 .

Mapreduce, Twister and Iterative Mapreduce:

MapReduce and HDFS are the two major components of Hadoop which makes it so powerful and efficient to use. MapReduce is a programming model used for efficient processing in parallel over large data-sets in a distributed manner. The data is first split and then combined to produce the final result. The libraries for MapReduce is written in so many programming languages with various different-different optimizations. The purpose of MapReduce in Hadoop is to Map each of the jobs and then it will reduce it to equivalent tasks for providing less overhead over the cluster network and to reduce the processing power. The MapReduce task is mainly divided into two phases Map Phase and Reduce Phase.

MapReduce Architecture:



Components of MapReduce Architecture:

1. **Client:** The MapReduce client is the one who brings the Job to the MapReduce for processing. There can be multiple clients available that continuously send jobs for processing to the Hadoop MapReduce Manager.
2. **Job:** The MapReduce Job is the actual work that the client wanted to do which is comprised of so many smaller tasks that the client wants to process or execute.
3. **Hadoop MapReduce Master:** It divides the particular job into subsequent job-parts.
4. **Job-Parts:** The task or sub-jobs that are obtained after dividing the main job. The result of all the job-parts combined to produce the final output.
5. **Input Data:** The data set that is fed to the MapReduce for processing.
6. **Output Data:** The final result is obtained after the processing.

In **MapReduce**, we have a client. The client will submit the job of a particular size to the Hadoop MapReduce Master. Now, the MapReduce master will divide this job into further equivalent job-parts. These job-parts are then made available for the Map and Reduce Task. This Map and Reduce task will contain the program as per the requirement of the use-case that the particular company is solving. The developer writes their logic to fulfill the requirement that the industry requires. The input data which we are using is then fed to the Map Task and the Map will generate intermediate key-value pair as its output. The output of Map i.e. these key-value pairs are then fed to the Reducer and the final output is stored on the HDFS. There can be n number of Map and Reduce tasks made available for processing the data as per the requirement. The algorithm for Map and Reduce is made with a very optimized way such that the time complexity or space complexity is minimum.

Let's discuss the MapReduce phases to get a better understanding of its architecture:

The MapReduce task is mainly divided into **2 phases** i.e. Map phase and Reduce phase.

1. **Map:** As the name suggests its main use is to map the input data in key-value pairs. The input to the map may be a key-value pair where the key can be the id of some kind of address and value is the actual value that it keeps. The *Map()* function will be executed in its memory repository on each of these input key-value pairs and generates the intermediate key-value pair which works as input for the Reducer or *Reduce()* function.

2. **Reduce:** The intermediate key-value pairs that work as input for Reducer are shuffled and sort and send to the *Reduce()* function.

Reducer aggregate or group the data based on its key-value pair as per the reducer algorithm written by the developer.

How Job tracker and the task tracker deal with MapReduce:

1. **Job Tracker:** The work of Job tracker is to manage all the resources and all the jobs across the cluster and also to schedule each map on the Task Tracker running on the same data node since there can be hundreds of data nodes available in the cluster.

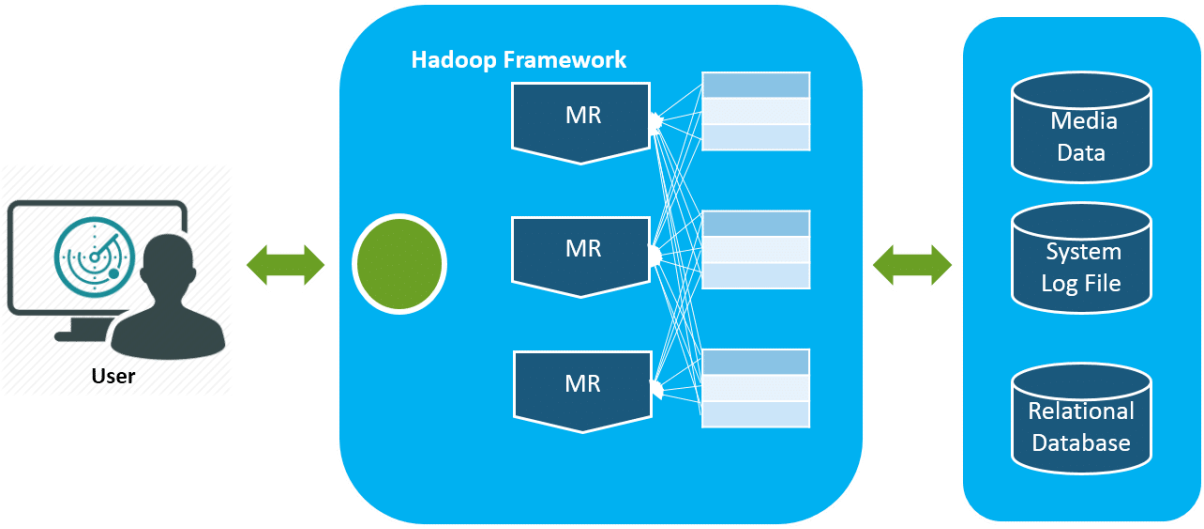
2. **Task Tracker:** The Task Tracker can be considered as the actual slaves that are working on the instruction given by the Job Tracker. This Task Tracker is deployed on each of the nodes available in the cluster that executes the Map and Reduce task as instructed by Job Tracker.

There is also one important component of MapReduce Architecture known as **Job History Server**. The Job History Server is a daemon process that saves and stores historical information about the task or application, like the logs which are generated during or after the job execution are stored on Job History Server.

Hadoop Library from Apache:

Apache Hadoop was born to enhance the usage and solve major issues of big data. The web media was generating loads of information on a daily basis, and it was becoming very difficult to manage the data of around one billion pages of content. In order of revolutionary, Google invented a new methodology of processing data popularly known as MapReduce. Later after a year Google published a white paper of Map Reducing framework where Doug Cutting and Mike Cafarella, inspired by the white paper and thus created Hadoop to apply these concepts to an open-source software framework that supported the Nutch search engine project. Considering the original case study, Hadoop was designed with much simpler storage infrastructure facilities.

Apache Hadoop is the most important framework for working with Big Data. Hadoop biggest strength is scalability. It upgrades from working on a single node to thousands of nodes without any issue in a seamless manner.

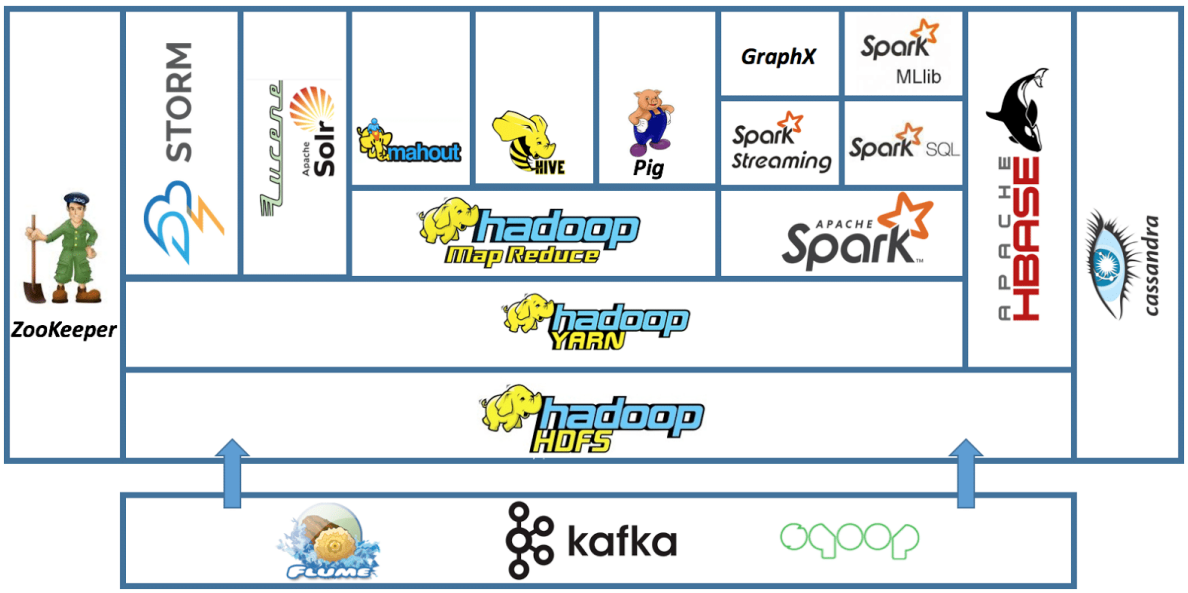


The different domains of Big Data means we are able to manage the data's are from videos, text medium, transactional data, sensor information, statistical data, social media conversations, search engine queries, ecommerce data, financial information, weather data, news updates, forum discussions, executive reports, and so on

Google's **Doug Cutting** and his team members developed an Open Source Project namely known as HADOOP which allows you to handle the very large amount of data. Hadoop runs the applications on the basis of MapReduce where the data is processed in parallel and accomplish the entire statistical analysis on large amount of data.

It is a framework which is based on java programming. It is intended to work upon from a single server to thousands of machines each offering local computation and storage. It supports the large collection of data set in a distributed computing environment.

The Apache Hadoop software library based framework that gives permissions to distribute huge amount of data sets processing across clusters of computers using easy programming models.

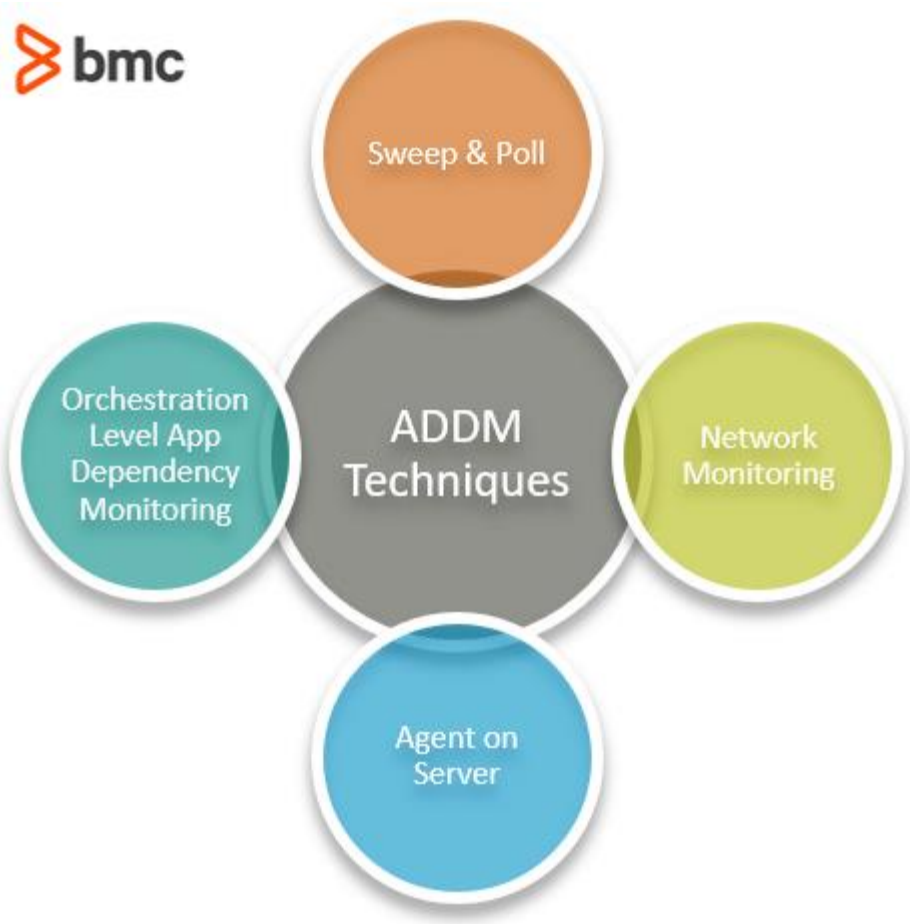


Mapping Applications:

Application mapping is the process of discovering and identifying the interactions and interdependencies between application components and their underlying hardware infrastructure.
To ensure that apps perform optimally, it's important to discover and map the underlying dependencies. The technology that enables this capability is common called "application mapping", but Application Discovery and Dependency Mapping (ADDM) is another word for it. Application mapping solutions are:
A management solution that discovers the relationships of app components and the underlying components and maps them to deliver a comprehensive insight into the resources running in the IT infrastructure and their dependencies.

How application mapping works

Application mapping can be implemented in several ways, ranging from brainstorming and manual element polling to automated discovery of the entire IT ecosystem. The techniques can be agent-based or agentless monitoring, as described below:



Sweep and poll

The agentless monitoring technique is the traditional way of discovering IT assets: by pinging IP addresses and identifying the responding devices. The technique identifies app components, devices, and server systems based on information such as discovery ping rate and device group information, established based on known device data.
The sweep and poll method is lightweight and allows users to sweep the entire network from a single connected node location.

The process may be slow for large-scale data centers, which is a drawback considering that dependencies can change during the process and leave critical assets undiscovered. Furthermore, discovering app components, particularly in dynamic virtualized and cloud-based IT environments with limited visibility and control, may be particularly challenging.

Network monitoring

Network monitoring looks at real-time packet information and captures accurate data on application dependencies. The Netflow protocol contains IP traffic information such as:

- Volume
- Path
- Source and destination nodes
- Other IP flow attributes

Using protocols such as Netflow for traffic monitoring has its disadvantages.

Netflow implementation can impact the performance of devices given its large bandwidth requirements. To address this issue, the data is sampled at intervals, which reduces the demand for bandwidth consumption but also collects fewer packet information since the unsampled data is not monitored.

Additionally, Netflow containing IP address and TCP port data cannot differentiate application-level dependencies. As an alternative, data packets can be captured but still provide only limited information collected at the time of probing.

Software Environment for Service Deployment:

Software deployment includes all of the steps, processes, and activities that are required to make a software system or update available to its intended users. Today, most IT organizations and software developers deploy software updates, patches and new applications with a combination of manual and automated processes. Some of the most common activities of software deployment include software release, installation, testing, deployment, and performance monitoring.

Key takeaways

- Software deployment refers to the process of making the application work on a target device, whether it be a test server, production environment or a user's computer or mobile device.
- Software and application deployment are terms that can be used interchangeably.
- Today, most IT organizations and software developers deploy software updates, patches and new applications with a combination of manual and automated processes.
- Software developers have created workflows that enable faster and more frequent deployment of software updates to the production environment where they can be accessed by users.

Why is software deployment important

Software deployment is one of the most important aspects of the software development process. Deployment is the mechanism through which applications, modules, updates, and patches are delivered from developers to users. The methods used by developers to build, test and deploy new code will impact how fast a product can respond to changes in customer preferences or requirements and the quality of each change.

Software development teams that streamline the process of building, testing and deploying new code can respond more quickly to customer demand with new updates and deliver new features more frequently to drive customer satisfaction, satisfy user needs and take advantage of economic opportunities.

Software development teams have innovated heavily over the past two decades, creating new paradigms and working methods for software delivery that are designed to meet the changing demands of consumers in an increasingly connected world. In particular, software developers have created workflows that enable faster and more frequent deployment of software updates to the production environment where they can be accessed by users.

Software deployment vs. software release - what's the difference

For the uninitiated, **software deployment** and **software release** may sound like very much the same thing. In fact, these terms describe two separate aspects of the overall software deployment process that should be understood separately.

The software release process

The **software release** cycle refers to the stages of development for a piece of computer software, whether it is released as a piece of physical media, online, or as a web-based application (SaaS). When a software development team prepares a new software release, it typically includes a specific version of the code and associated resources that have been assigned a version number. When the code is updated or modified with bug fixes, a new version of the code may be packaged with supporting resources and assigned a new release number. Versioning new software releases in this way helps to differentiate between different versions and identify the most up-to-date software release.

The software deployment process

Software deployment refers to the process of running an application on a server or device. A software update or application may be deployed to a test server, a testing machine, or into the live environment, and it may be deployed several times during the development process to verify its proper functioning and check for errors. Software deployment is the process of running an application on a server or device. Software and application deployment are terms that can be used interchangeably. Another example of software deployment could be when a user downloads a mobile application from the Integration Store and installs it onto their mobile device.

To summarize, a software release is a specific version of a code and its dependencies that are made available for deployment. Software deployment refers to the process of making the application work on a target device, whether it be a test server, production environment or a user's computer or mobile device.

Different types of application or software deployment strategies are:

- Basic
- Multi-service
- Rolling
- Blue Green Deployment
- Canary
- A/B Testing

While many development teams still choose to host applications using on-premises IT infrastructure, cloud service providers like Amazon Web Services (AWS), Google Cloud Platform and Microsoft Azure now offer IT Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS) products that help developers deploy applications into live environments without the additional financial and administrative burden of managing their own storage and virtualization servers.

Software deployment methodologies

DevOps is a methodology and a set of best practices for software development whose primary goals are to shorten delivery times for new software updates while maintaining high quality. In the DevOps framework, there are seven .

Cloud Environments:

Cloud computing is the on-demand delivery of computing services, like storage, software, analytics, and databases over the Internet to offer flexible resources and economies of scale. It is a paradigm shift from the traditional way businesses think about computing and IT resources.

Instead of owning computing infrastructure or datacenters, organizations can rent access to cloud solutions and pay for only the solutions they use. Thus, organizations can avoid heavy upfront investments and the complexity of maintaining the IT infrastructure. Cloud computing service providers can benefit by delivering their IT services to multiple tenants over the Internet.

Prominent public cloud service providers include Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure.

Types of Cloud Computing Environments

Public cloud: The third-party cloud service providers own the public cloud and the supporting infrastructure like hardware and software. They deliver computing resources to organizations on the basis of a subscription fee. Users can get access to these services and manage their profiles using web browsers. This cloud deployment

model is completely virtualized, which allows users to share resources while maintaining the security and privacy of each user's data.

Private cloud: In this internal or corporate cloud environment, computing services are offered via a private network and used exclusively by a single customer or business organization. Companies can maintain their private cloud on their on-premise datacenters or subscribe to third-party service providers to host their private cloud. Private cloud not only offers several benefits of cloud computing, like flexibility and scalability but also enhances security and resource optimization of on-premise infrastructure.

Hybrid cloud: It is a combination of both public and private cloud, or like an on-premise data center and a public cloud, where data and applications can be shared securely between them. Most organizations do not depend entirely on the public cloud. A hybrid cloud environment offers greater flexibility, and more deployment options, and helps in maintaining the privacy and regulatory compliance

Cloud Storage:

Cloud storage is a cloud computing model that enables storing data and files on the internet through a cloud computing provider that you access either through the public internet or a dedicated private network connection. The provider securely stores, manages, and maintains the storage servers, infrastructure, and network to ensure you have access to the data when you need it at virtually unlimited scale, and with elastic capacity. Cloud storage removes the need to buy and manage your own data storage infrastructure, giving you agility, scalability, and durability, with any time, anywhere data access.
Cloud storage delivers cost-effective, scalable storage. You no longer need to worry about running out of capacity, maintaining storage area networks (SANs), replacing failed devices, adding infrastructure to scale up with demand, or operating underutilized hardware when demand decreases. Cloud storage is elastic, meaning you scale up and down with demand and pay only for what you use. It is a way for organizations to save data securely online so that it can be accessed anytime from any location by those with permission.

Whether you are a small business or a large enterprise, cloud storage can deliver the agility, cost savings, security, and simplicity to focus on your core business growth. For small businesses, you no longer have to worry about devoting valuable resources to manage storage yourself, and cloud storage gives you the ability to scale as the business grows.

For large enterprises with billions of files and petabytes of data, you can rely on the scalability, durability, and cost savings of cloud storage to create centralized data lakes to make your data accessible to all who need it.

Cost effectiveness

With cloud storage, there is no hardware to purchase, no storage to provision, and no extra capital being used for business spikes. You can add or remove storage capacity on demand, quickly change performance and retention characteristics, and only pay for storage that you actually use. As data becomes infrequently and rarely accessed, you can even automatically move it to lower-cost storage, thus creating even more cost savings. By moving storage

workloads from on premises to the cloud, you can reduce total cost of ownership by removing overprovisioning and the cost of maintaining storage infrastructure.

Increased agility

With cloud storage, resources are only a click away. You reduce the time to make those resources available to your organization from weeks to just minutes. This results in a dramatic increase in agility for your organization. Your staff is largely freed from the tasks of procurement, installation, administration, and maintenance. And because cloud storage integrates with a wide range of analytics tools, your staff can now extract more insights from your data to fuel innovation.

Faster deployment

When development teams are ready to begin, infrastructure should never slow them down. Cloud storage services allow IT to quickly deliver the exact amount of storage needed, whenever and wherever it's needed. Your developers can focus on solving complex application problems instead of having to manage storage systems.

Efficient data management

By using cloud storage lifecycle management policies, you can perform powerful information management tasks including automated tiering or locking down data in support of compliance requirements. You can also use cloud storage to create multi-region or global storage for your distributed teams by using tools such as replication. You can organize and manage your data in ways that support specific use cases, create cost efficiencies, enforce security, and meet compliance requirements.

Virtually unlimited scalability

Cloud storage delivers virtually unlimited storage capacity, allowing you to scale up as much and as quickly as you need. This removes the constraints of on-premises storage capacity. You can efficiently scale cloud storage up and down as required for analytics, data lakes, backups, or cloud native applications. Users can access storage from anywhere, at any time, without worrying about complex storage allocation processes, or waiting for new hardware.

Business continuity

Cloud storage providers store your data in highly secure data centers, protecting your data and ensuring business continuity. Cloud storage services are designed to handle concurrent device failure by quickly detecting and repairing any lost redundancy. You can further protect your data by using versioning and replication tools to more easily recover from both unintended user actions or application failures.

With cloud storage services, you can:

- Cost-effectively protect data in the cloud without sacrificing performance.
- Scale up your backup resources in minutes as data requirements change.

- Protect backups with a data center and network architecture built for security-sensitive organizations.

Cloud storage is delivered by a cloud services provider that owns and operates data storage capacity by maintaining large datacenters in multiple locations around the world. Cloud storage providers manage capacity, security, and durability to make data accessible to your applications over the internet in a pay-as-you-go model. Typically, you connect to the storage cloud either through the internet or through a dedicated private connection, using a web portal, website, or a mobile app. When customers purchase cloud storage from a service provider, they turn over most aspects of the data storage to the vendor, including capacity, security, data availability, storage servers and computing resources, and network data delivery. Your applications access cloud storage through traditional storage protocols or directly using an application programming interface (API). The cloud storage provider might also offer services designed to help collect, manage, secure, and analyze data at a massive sc
There are three main cloud storage types: object storage, file storage, and block storage. Each offers its own advantages and has its own use cases.

Object storage

Organizations have to store a massive and growing amount of unstructured data, such as photos, videos, machine learning (ML), sensor data, audio files, and other types of web content, and finding scalable, efficient, and affordable ways to store them can be a challenge. Object storage is a data storage architecture for large stores of unstructured data. Objects store data in the format it arrives in and makes it possible to customize metadata in ways that make the data easier to access and analyze. Instead of being organized in files or folder hierarchies, objects are kept in secure buckets that deliver virtually unlimited scalability. It is also less costly to store large data volumes.

Applications developed in the cloud often take advantage of the vast scalability and metadata characteristics of object storage. Object storage solutions are ideal for building modern applications from scratch that require scale and flexibility, and can also be used to import existing data stores for analytics, backup, or archive.

File storage

File-based storage or file storage is widely used among applications and stores data in a hierarchical folder and file format. This type of storage is often known as a network-attached storage (NAS) server with common file level protocols of Server Message Block (SMB) used in Windows instances and Network File System (NFS) found in Linux.

Block storage

Enterprise applications like databases or enterprise resource planning (ERP) systems often require dedicated, low-latency storage for each host. This is analogous to direct-attached storage (DAS) or a storage area network (SAN). In this case, you can use a cloud storage service that stores data in the form of blocks. Each block has its own unique identifier for quick storage and retrieval.
Ensuring your company’s critical data is safe, secure, and available when needed is essential. There are several fundamental requirements when considering storing data in the cloud.

Durability and availability

Cloud storage simplifies and enhances traditional data center practices around data durability and availability. With cloud storage, data is redundantly stored on multiple devices across one or more data centers.

Security

With cloud storage, you control where your data is stored, who can access it, and what resources your organization is consuming at any given moment. Ideally, all data is encrypted, both at rest and in transit. Permissions and access controls should work just as well in the cloud as they do for on-premises storage.

Cloud storage has several use cases in application management, data management, and business continuity. Let's consider some examples below.

Analytics and data lakes

Traditional on-premises storage solutions can be inconsistent in their cost, performance, and scalability — especially over time. Analytics demand large-scale, affordable, highly available, and secure storage pools that are commonly referred to as data lakes.

Data lakes built on object storage keep information in its native form and include rich metadata that allows selective extraction and use for analysis. Cloud-based data lakes can sit at the center of multiple kinds of data warehousing and processing, as well as big data and analytical engines, to help you accomplish your next project in less time and with more targeted relevance.

Backup and disaster recovery

Backup and disaster recovery are critical for data protection and accessibility, but keeping up with increasing capacity requirements can be a constant challenge. Cloud storage brings low cost, high durability, and extreme scale to data backup and recovery solutions. Embedded data management policies can automatically migrate data to lower-cost storage based on frequency or timing settings, and archival vaults can be created to help comply with legal or regulatory requirements. These benefits allow for tremendous scale possibilities within industries such as financial services, healthcare and life sciences, and media and entertainment that produce high volumes of unstructured data with long-term retention needs.

Advantages of Cloud Storage:

Below are the advantages of cloud storage:

1. Cost Saving

By using cloud storage, there is no need to buy as many hard drives, enclosures to house them in, RAID cards to enable data redundancy, electricity to power them, or hardware warranty services to safeguard them. However, it also cuts management expenses by decreasing the need for in-depth capacity planning, streamlining monitoring, and minimizing on-premise hardware and software management. Administrators can instead concentrate on other, more crucial activities.

2. Data Redundancy and Replication

The majority of cloud storage providers maintain numerous copies of data, even inside of a single “Data Center”, and they provide excellent object durability to lower the risk of data loss. Geographic replication options, however, can spread out several copies of data across areas if you’re searching for even greater security. Others provide replication services that swiftly transfer data between data centers, while some offer geo-replication as a storage class option. Your backups are adequately shielded.

3. Data Tiering for Cost Savings

Various storage classes and data tiers are offered by numerous cloud storage providers. Choose based on how regularly and rapidly one restores backups, as well as how long one wants to retain the backups. Consider using the vendor’s hot storage for backups that require quick and/or frequent restores because it offers the fastest and most economical retrieval. Consider shifting data to archive storage for long-term archiving, Although data retrieval may be more time-consuming and expensive, storage expenses are far lower, especially if one intends to preserve backups for many years. The ability to automatically shift data between tiers is a feature that some vendors offer. This minimizes administration and makes it easier to obtain cost savings.

4. Regulatory Compliance

For regulatory compliance, keeping backups in the same area as the data’s origin may be the best option. Worldwide alternatives for data centers are provided by many cloud suppliers. Look for a cloud storage provider that can accommodate if one needs to store EU client data in an EU data canter. Moving data to cloud storage within the same region is also advantageous for performance. Even if you are not subject to regulation, the enhanced performance might be valuable to you.

5. Ransomware/Malware Protection

Ransomware is plain nasty. Unfortunately, it frequently makes the news. The malware will search the network for shares that contain files and documents to encrypt in addition to the locally infected computer, which is one of the more frightening characteristics of ransomware. You might be relieved to learn that your cloud storage can assist prevent ransomware by providing some backup security advantages because it’s more difficult to access without proper authentication if you’re hit by ransomware or another type of malware that is encrypting or destroying files.

6. Usability or Accessibility

The key advantages of cloud storage are accessibility and usability. You can rapidly upload your file to your online drive even if you lack technical ability because they both have simple user interfaces. Most cloud data storage

providers include drag-and-drop functionality and an intuitive user interface. For instance, if you saved a file to a disc on a mobile device, you can access that file on a computer or any other device with internet access. It doesn't matter where you are right now. Your files, which are kept online in one of the data centers, can be accessed if you have a strong internet connection.

7. Flexibility

In general, using the cloud gives businesses more flexibility than hosting on a local server. Additionally, a cloud-based solution could be able to quickly meet your need for more bandwidth without necessitating a difficult (and expensive) update to your IT infrastructure. This improved independence and flexibility may considerably raise the overall effectiveness of your firm. You won't be able to focus on achieving your company's goals and satisfying consumers if your present IT solutions need you to devote too much of your time to computer and data-storage concerns. However, if you rely on a third party to manage the IT hosting and infrastructure, you'll free up more time for the areas of your organization that directly impact your bottom line.

8. Automation

A cloud storage service may be used by multiple users, and as everything is handled and automated by the cloud provider vendor, one user's current task would not influence that of another. When you want to store a file in the cloud, cloud storage services function like a hard drive on your computer and won't interfere with any ongoing tasks.

9. Scalable

You can upgrade the service plan if the storage included in the current plan is insufficient. Additionally, the additional space will be provided to your data storage environment with some new capabilities, so you won't need to migrate any data from one place to another. Scalable and adaptable cloud storage is offered.

10. Reliability

Many people create a cloud backup of their hard disc in case their hard drive fails. The comfort that comes from knowing that data won't suddenly vanish one day may be well worth the small price.

10 Disadvantages of Cloud Storage

Below are the disadvantages of cloud storage:

1. Vulnerability

The majority of PCs and servers that save data require an internet connection. Cloud solutions are internet-based by nature, which means that they are linked to other computers and servers. Thus, making them vulnerable to attacks by malicious users on the network.

2. Internet Dependency

The internet starts to be reliant on your storage. Due to the fact that the internet will govern our world in 2022, this disadvantage will diminish. One can always save files while offline and access them later. However, an internet connection will be required for the update and sync.

3. Issues in Security and Privacy

Uncertainties about privacy and security on the cloud are the next significant point to be made in relation to the drawbacks of cloud storage. Confidential data must be given over to a third-party organization in order to be stored in the cloud. One must therefore have complete faith in the cloud vendor.

4. Limitations on Control

After the user moves data to the cloud, the vendor is now in charge of it. This implies that users must rely on the vendors to maintain their services in a safe, stable, up-and-running, and fully functional manner. This limits the influence on data safety Nearly all reliability is left up to the storage vendor, along with accessibility.

5. Cost

Although cloud storage options are normally a cost-effective choice, they could not be available if utilized for short-term or very small-scale projects, depending on the cloud vendor. Users can be on the hook for 18 more months than needed, which is not financially feasible if the demand for data storage is for 6 months but the vendor's minimum offer is 2 years.

6. Migration

The ability to switch to another cloud service has grown much less well than other aspects of cloud storage. As a result, many clients continue to find this scenario to be quite difficult. Although several of these options are currently unavailable, one may choose to migrate data to another solution if one discovers that the cloud provider does not adequately address all of the users' needs. If there are such options, they might be pricey. As a result, additional expenses are now necessary, most often in the form of signing a deal with a second cloud-storage vendor while being bound by the terms of the first.

7. Regulatory Compliances

Only cloud storage options with the highest levels of security protection may be trusted by financial organizations. When penalties, fines, and lawsuits are imposed because authorities think it necessary for specific criteria for data protection, the company will be the target, not the cloud provider.

8. Minimal Support
Even when using a simple, rapid, and secure cloud solution, problems can occur because every user and business has different wants and specifics. The absence of support for cloud storage is one of the main difficulties there.

9. Features
Each cloud service provider is unique from the others. Users may occasionally only be able to use the basic kinds of cloud storage that cloud providers provide. As a result, one cannot modify certain features or take advantage of all of their benefits.

10. Data Management

Given that cloud storage systems have their own organizational systems, managing data may be a pain. The system used by a cloud vendor might not work with the way one manages their storage currently.

Short Answer Questions:

1.define mapreduce and iterative reduce?

2.explain the advantages of cloud storage?

3.define cloud service?

4.Explain openstack cloud Environment?

Long Answer Questions:

1.Explain about software environments for service deployment?

2.Explain parallel and distributed programming paradigms?

3.Define Hadoop Library from Apache?

4.define Cloud Environments in Detail?

Unit-5

Security in the cloud:

Cloud security is a responsibility that is shared between the cloud provider and the customer. There are basically three categories of responsibilities in the Shared Responsibility Model: responsibilities that are *always* the provider's, responsibilities that are *always* the customer's, and responsibilities that *vary depending on the service model*: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), or Software as a Service (SaaS), such as cloud email.

Because the public cloud does not have clear perimeters, it presents a fundamentally different security reality. This becomes even more challenging when adopting modern cloud approaches such as automated Continuous Integration and Continuous Deployment (CI/CD) methods, distributed serverless architectures, and ephemeral assets like Functions as a Service and containers.

Some of the advanced cloud-native security challenges and the multiple layers of risk faced by today's cloud-oriented organizations include:

1.Increased Attack Surface

> The public cloud environment has become a large and highly attractive attack surface for hackers who exploit poorly secured cloud ingress ports in order to access and disrupt workloads and data in the cloud. Malware, Zero-Day, Account Takeover and many other malicious threats have become a day-to-day reality.

2.Lack of Visibility and Tracking

> In the IaaS model, the cloud providers have full control over the infrastructure layer and do not expose it to their customers. The lack of visibility and control is further extended in the PaaS and SaaS cloud models. Cloud customers often cannot effectively identify and quantify their cloud assets or visualize their cloud environmets.

3.Ever-Changing Workloads

Cloud assets are provisioned and decommissioned dynamically—at scale and at velocity. Traditional security tools are simply incapable of enforcing protection policies in such a flexible and dynamic environment with its ever-changing and ephemeral workloads.

4.DevOps, DevSecOps and Automation

Organizations that have embraced the highly automated DevOps CI/CD culture must ensure that appropriate security controls are identified and embedded in code and templates early in the development cycle. Security-related changes implemented *after* a workload has been deployed in production can undermine the organization's security posture as well as lengthen time to market.

5.Granular Privilege and Key Management

Often cloud user roles are configured very loosely, granting extensive privileges beyond what is intended or required. One common example is giving database delete or write permissions to untrained users or users who have no business need to delete or add database assets. At the application level, improperly configured keys and privileges expose sessions to security risks.

6.Complex Environments

Managing security in a consistent way in the hybrid and multicloud environments favored by enterprises these days requires methods and tools that work seamlessly across public cloud providers, private cloud providers, and on-premise deployments—including branch office edge protection for geographically distributed organizations.

7.Cloud Compliance and Governance

All the leading cloud providers have aligned themselves with most of the well-known accreditation programs such as PCI 3.2, NIST 800-

53, HIPAA and GDPR. However, customers are responsible for ensuring that their workload and data processes are compliant. Given the poor visibility as well as the dynamics of the cloud environment, the compliance audit process becomes close to mission impossible unless tools are used to achieve continuous compliance checks and issue real-time alerts about misconfigurations.

Cloud Security Challenges and Risks:

All companies face security risks, threats, and challenges every day. Many think these terms all mean the same thing, but they're more nuanced. Understanding the subtle differences between them will help you better protect your cloud assets.

What is the difference between risks, threats, and challenges?

- A **risk** is a potential for loss of data or a weak spot.
- A **threat** is a type of attack or adversary.
- A **challenge** is an organization's hurdles in implementing practical cloud security.

Let's consider an example: An API endpoint hosted in the cloud and exposed to the public Internet is a **risk**, the attacker who tries to access sensitive data using that API is the **threat** (along with any specific techniques they could try), and your organization's **challenge** is effectively protecting public APIs while keeping them available for legitimate users or customers who need them.

A complete cloud security strategy addresses all three aspects, so no cracks exist within the foundation. You can think of each as a different lens or angle with which to view cloud security. A solid strategy must mitigate risk (security controls), defend against threats (secure coding and deployment), and overcome challenges (implement cultural and technical solutions) for your business to use the cloud to grow securely.

4 Cloud Security Risks

You cannot completely eliminate risk; you can only manage it. Knowing common risks ahead of time will prepare you to deal with them within your environment. **What are four cloud security risks?**

1. Unmanaged Attack Surface
2. Human Error
3. Misconfiguration
4. Data Breach

1. Unmanaged Attack Surface

An attack surface is your environment's total exposure. The adoption of microservices can lead to an explosion of publicly available workload. Every

workload adds to the attack surface. Without close management, you could expose your infrastructure in ways you don't know until an attack occurs.

No one wants that late-night call.

Attack surface can also include subtle information leaks that lead to an attack. For example, CrowdStrike's team of threat hunters found an attacker using sampled DNS request data gathered over public WiFi to work out the names of S3 buckets. CrowStrike stopped the attack before the attackers did any damage, but it's a great illustration of risk's ubiquitous nature. Even strong controls on the S3 buckets weren't enough to completely hide their existence. As long as you use the public Internet or cloud, you're automatically exposing an attack surface to the world.

Your business may need it to operate, but keep an eye on it.

2. Human Error

According to Gartner, through 2025, 99% of all cloud security failures will be due to some level of human error. Human error is a constant risk when building business applications. However, hosting resources on the public cloud magnifies the risk.

The cloud's ease of use means that users could be using APIs you're not aware of without proper controls and opening up holes in your perimeter. Manage human error by building strong controls to help people make the right decisions.

One final rule — don't blame people for errors. Blame the process. Build processes and guardrails to help people do the right thing. Pointing fingers doesn't help your business become more secure.

3. Misconfiguration

Cloud settings keep growing as providers add more services over time. Many companies are using more than one provider.

Providers have different default configurations, with each service having its distinct implementations and nuances. Until organizations become proficient at securing their various cloud services, adversaries will continue to exploit misconfigurations.

4. Data Breaches

A data breach occurs when sensitive information leaves your possession without your knowledge or permission. Data is worth more to attackers than anything else, making it the goal of most attacks. Cloud misconfiguration and lack of runtime protection can leave it wide open for thieves to steal.

The impact of data breaches depends on the type of data stolen. Thieves sell personally identifiable information (PII) and personal health information (PHI) on

the dark web to those who want to steal identities or use the information in phishing emails.

Other sensitive information, such as internal documents or emails, could be used to damage a company's reputation or sabotage its stock price. No matter the reason for stealing the data, breaches continue to be an imposing threat to companies using the cloud.

How To Manage Cloud Security Risks

Follow these tips to manage risk in the cloud:

- Perform regular risk assessments to find new risks.
- Prioritize and implement security controls to mitigate the risks you've identified (CrowdStrike can help).
- Document and revisit any risks you choose to accept.

Challenges are the gap between theory and practice. It's great to know you need a cloud security strategy. But where do you start? How do you tackle cultural change? What are the daily practical steps to make it happen?

1. Lack of Cloud Security and Skills
2. Identity and Access Management
3. Shadow IT
4. Cloud Compliance

1. Lack Of Cloud Security Strategy and Skills

Traditional data center security models are not suitable for the cloud. Administrators must learn new strategies and skills specific to cloud computing.

Cloud may give organizations agility, but it can also open up vulnerabilities for organizations that lack the internal knowledge and skills to understand security challenges in the cloud effectively. Poor planning can manifest itself in misunderstanding the implications of the shared responsibility model, which lays out the security duties of the cloud provider and the user. This misunderstanding could lead to the exploitation of unintentional security holes.

2. Identity and Access Management

Identity and Access Management (IAM) is essential. While this may seem obvious, the challenge lies in the details.

It's a daunting task to create the necessary roles and permissions for an enterprise of thousands of employees. There are three steps to a holistic IAM strategy: role design, privileged access management, and implementation.

Begin with a solid role design based on the needs of those using the cloud. Design the roles outside of any specific IAM system. These roles describe the work your employees do, which won't change between cloud providers.

Next, a strategy for privileged access management (PAM) outlines which roles require more protection due to their privileges. Tightly control who has access to privileged credentials and rotate them regularly.

Finally, it's time to implement the designed roles within the cloud provider's IAM service. This step will be much easier after developing these ahead of time.

3. Shadow IT

Shadow IT challenges security because it circumvents the standard IT approval and management process.

Shadow IT is the result of employees adopting cloud services to do their jobs. The ease with which cloud resources can be spun up and down makes controlling its growth difficult. For example, developers can quickly spawn workloads using their accounts. Unfortunately, assets created in this way may not be adequately secured and accessible via default passwords and misconfigurations.

The adoption of DevOps complicates matters. Cloud and DevOps teams like to run fast and without friction. However, obtaining the visibility and management levels that the security teams require is difficult without hampering DevOps activities. DevOps needs a frictionless way to deploy secure applications and directly integrate with their continuous integration/continuous delivery (CI/CD) pipeline. There needs to be a unified approach for security teams to get the information they need without slowing down DevOps. IT and security need to find solutions that will work for the cloud — at DevOps' velocity.

4. Cloud Compliance

Organizations have to adhere to regulations that protect sensitive data like PCI DSS and HIPAA. Sensitive data includes credit card information, healthcare patient records, etc. To ensure compliance standards are met, many organizations limit access and what users can do when granted access. If access control measures are not set in place, it becomes a challenge to monitor access to the network.

How to Overcome Cloud Security Challenges

Each challenge is different and therefore requires unique solutions. Take the time to plan before making use of any cloud services. A sound strategy takes into consideration any common cloud challenges like the ones we've discussed here. Then you'll have a plan of action for each anticipated challenge.

Security Governance:

Enterprises are increasingly pursuing the business advantages of migrating technology platforms and services into the cloud environment leveraging one or

more of the three main cloud service areas – Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). These advantages include but are not limited to rapid information system deployment, significantly reduced operating costs, massive economies of scale, processing speed, and agility. However, subscription to these services often imply security and compliance challenges for enterprises who are often unprepared to resolve them.

Data breaches, system vulnerabilities, insufficient identity, and credential and access management are some of the typical security challenges in the cloud environment that subscriber enterprises must address. In some situations, an enterprise may lack adequate operationalization and enforcement of policies, procedures, a formal operating model, or even a properly constituted organizational function to effectively manage security in the cloud. In other situations, the enterprise may also not sufficiently exercise its responsibility to protect data in the cloud or may lack the means for senior management visibility into cloud security performance and risks. These issues may prevail even when an enterprise stands to gain significant business benefits from transforming its service delivery model via the use of cloud computing platforms.

The underlying business problem leading to these challenges is the lack of effective governance of cloud security. In this blog, I explore cloud security governance, common challenges, and review key targets that can help enterprises optimize the business benefits of cloud security programs.

What Is Cloud Security Governance?

Cloud security governance refers to the management model that facilitates effective and efficient security management and operations in the cloud environment so that an enterprise's business targets are achieved. This model incorporates a hierarchy of executive mandates, performance expectations, operational practices, structures, and metrics that, when implemented, result in the optimization of business value for an enterprise. Cloud security governance helps answer leadership questions such as:

- Are our security investments yielding the desired returns?
- Do we know our security risks and their business impact?
- Are we progressively reducing security risks to acceptable levels?
- Have we established a security-conscious culture within the enterprise?

Strategic alignment, value delivery, risk mitigation, effective use of resources, and performance measurement are key objectives of any IT-related governance model, security included. To successfully pursue and achieve these objectives, it is important to understand the operational culture and business and customer profiles of an enterprise, so that an effective security governance model can be customized for the enterprise.

Cloud Security Governance Challenges

Whether developing a governance model from the start or having to retrofit one on existing investments in cloud, these are some of the common challenges:

Lack of senior management participation and buy-in

The lack of a senior management influenced and endorsed security policy is one of the common challenges facing cloud customers. An enterprise security policy is intended to set the executive tone, principles and expectations for security management and operations in the cloud. However, many enterprises tend to author security policies that are often laden with tactical content, and lack executive input or influence. The result of this situation is the ineffective definition and communication of executive tone and expectations for security in the cloud. To resolve this challenge, it is essential to engage enterprise executives in the discussion and definition of tone and expectations for security that will feed a formal enterprise security policy. It is also essential for the executives to take full accountability for the policy, communicating inherent provisions to the enterprise, and subsequently enforcing compliance

Lack of embedded management operational controls

Another common cloud security governance challenge is lack of embedded management controls into cloud security operational processes and procedures. Controls are often interpreted as an auditor's checklist or repackaged as procedures, and as a result, are not effectively embedded into security operational processes and procedures as they should be, for purposes of optimizing value and reducing day-to-day operational risks. This lack of embedded controls may result in operational risks that may not be apparent to the enterprise. For example, the security configuration of a device may be modified (change event) by a staffer without proper analysis of the business impact (control) of the modification. The net result could be the introduction of exploitable security weaknesses that may not have been apparent with this modification. The enterprise would now have to live with an inherent operational risk that could have been avoided if the control had been embedded in the change execution process.

Lack of operating model, roles, and responsibilities

Many enterprises moving into the cloud environment tend to lack a formal operating model for security, or do not have strategic and tactical roles and responsibilities properly defined and operationalized. This situation stifles the effectiveness of a security management and operational function/organization to support security in the cloud. Simply, establishing a hierarchy that includes designating an accountable official at the top, supported by a stakeholder committee, management team, operational staff, and third-party provider support (in that order) can help an enterprise to better manage and control security in the cloud, and protect associated investments in accordance with enterprise business goals. This hierarchy can be employed in an in-sourced, out-sourced, or co-sourced model depending on the culture, norms, and risk tolerance of the enterprise.

Lack of metrics for measuring performance and risk

Another major challenge for cloud customers is the lack of defined metrics to measure security performance and risks – a problem that also stifles executive

visibility into the real security risks in the cloud. This challenge is directly attributable to the combination of other challenges discussed above. For example, a metric that quantitatively measures the number of exploitable security vulnerabilities on host devices in the cloud over time can be leveraged as an indicator of risk in the host device environment. Similarly, a metric that measures the number of user-reported security incidents over a given period can be leveraged as a performance indicator of staff awareness and training efforts. Metrics enable executive visibility into the extent to which security tone and expectations (per established policy) are being met within the enterprise and support prompt decision-making in reducing risks or rewarding performance as appropriate.

The challenges described above clearly highlight the need for cloud customers to establish a framework to effectively manage and support security in cloud management, so that the pursuit of business targets are not potentially compromised. Unless tone and expectations for cloud security are established (via an enterprise policy) to drive operational processes and procedures with embedded management controls, it is very difficult to determine or evaluate business value, performance, resource effectiveness, and risks regarding security operations in the cloud. Cloud security governance facilitates the institution of a model that helps enterprises explicitly address the challenges described above.

Key Objectives for Cloud Security Governance

Building a cloud security governance model for an enterprise requires strategic-level security management competencies in combination with the use of appropriate security standards and frameworks (e.g., NIST, ISO, CSA) and the adoption of a governance framework (e.g., COBIT). The first step is to visualize the overall governance structure, inherent components, and to direct its effective design and implementation. The use of appropriate security standards and frameworks allow for a minimum standard of security controls to be implemented in the cloud, while also meeting customer and regulatory compliance obligations where applicable. A governance framework provides referential guidance and best practices for establishing the governance model for security in the cloud. The following represents key objectives to pursue in establishing a governance model for security in the cloud. These objectives assume that appropriate security standards and a governance framework have been chosen based on the enterprise's business targets, customer profile, and obligations for protecting data and other information assets in the cloud environment.

1. **Strategic Alignment**
   Enterprises should mandate that security investments, services, and projects in the cloud are executed to achieve established business goals (e.g., market competitiveness, financial, or operational performance).

2. **Value Delivery**
   Enterprises should define, operationalize, and maintain an appropriate

security function/organization with appropriate strategic and tactical representation, and charged with the responsibility to maximize the business value (Key Goal Indicators, ROI) from the pursuit of security initiatives in the cloud.

3. **Risk Mitigation**
   Security initiatives in the cloud should be subject to measurements that gauge effectiveness in mitigating risk to the enterprise (Key Risk Indicators). These initiatives should also yield results that progressively demonstrate a reduction in these risks over time.

4. **Effective Use of Resources**
   It is important for enterprises to establish a practical operating model for managing and performing security operations in the cloud, including the proper definition and operationalization of due processes, the institution of appropriate roles and responsibilities, and use of relevant tools for overall efficiency and effectiveness.

5. **Sustained Performance**
   Security initiatives in the cloud should be measurable in terms of performance, value and risk to the enterprise (Key Performance Indicators, Key Risk Indicators), and yield results that demonstrate attainment of desired targets (Key Goal Indicators) over time.

Risk Management:

Risk management is the process of identifying, assessing and controlling financial, legal, strategic and security risks to an organization's capital and earnings. These threats, or risks, could stem from a wide variety of sources, including financial uncertainty, legal liabilities, strategic management errors, accidents and natural disasters.

If an unforeseen event catches your organization unaware, the impact could be minor, such as a small impact on your overhead costs. In a worst-case scenario, though, it could be catastrophic and have serious ramifications, such as a significant financial burden or even the closure of your business.

To reduce risk, an organization needs to apply resources to minimize, monitor and control the impact of negative events while maximizing positive events. A consistent, systemic and integrated approach to risk management can help determine how best to identify, manage and mitigate significant risks.

The risk management process

At the broadest level, risk management is a system of people, processes and technology that enables an organization to establish objectives in line with values and risks.

A successful risk assessment program must meet legal, contractual, internal, social and ethical goals, as well as monitor new technology-related regulations. By focusing attention on risk and committing the necessary resources to control and mitigate risk, a business will protect itself from uncertainty, reduce costs and increase the likelihood of business continuity and success.
Three important steps of the risk management process are risk identification, risk analysis and assessment, and risk mitigation and monitoring.
Identifying risks

Risk identification is the process of identifying and assessing threats to an organization, its operations and its workforce. For example, risk identification may include assessing IT security threats such as malware and ransomware, accidents, natural disasters and other potentially harmful events that could disrupt business operations.
Risk analysis and assessment

Risk analysis involves establishing the probability that a risk event might occur and the potential outcome of each event. Risk evaluation compares the magnitude of each risk and ranks them according to prominence and consequence.
Risk mitigation and monitoring

Risk mitigation refers to the process of planning and developing methods and options to reduce threats to project objectives. A project team might implement risk mitigation strategies to identify, monitor and evaluate risks and consequences inherent to completing a specific project, such as new product creation. Risk mitigation also includes the actions put into place to deal with issues and effects of those issues regarding a project.

Risk management is a nonstop process that adapts and changes over time. Repeating and continually monitoring the processes can help assure maximum coverage of known and unknown risks.

Risk response strategies

There are five commonly accepted strategies for addressing risk. The process begins with an initial consideration of risk avoidance then proceeds to three additional avenues of addressing risk (transfer, spreading and reduction). Ideally, these three avenues are employed in concert with one another as part of a comprehensive strategy. Some residual risk may remain.
Risk avoidance
Avoidance is a method for mitigating risk by not participating in activities that may negatively affect the organization. Not making an investment or starting a product line are examples of such activities as they avoid the risk of loss.
Risk reduction
This method of risk management attempts to minimize the loss, rather than completely eliminate it. While accepting the risk, it stays focused on keeping the loss contained and

preventing it from spreading. An example of this in health insurance is preventative care.

Risk sharing

When risks are shared, the possibility of loss is transferred from the individual to the group. A corporation is a good example of risk sharing — a number of investors pool their capital and each only bears a portion of the risk that the enterprise may fail.

Transferring risk

Contractually transferring a risk to a third-party, such as, insurance to cover possible property damage or injury shifts the risks associated with the property from the owner to the insurance company.

Risk acceptance and retention

After all risk sharing, risk transfer and risk reduction measures have been implemented, some risk will remain since it is virtually impossible to eliminate all risk (except through risk avoidance). This is called residual risk.

Security Monitoring:

Security monitoring, sometimes referred to as "security information monitoring (SIM)" or "security event monitoring (SEM)," involves collecting and analysing information to detect suspicious behavior or unauthorised system changes on your network, defining which types of behavior should trigger alerts, and taking action on alerts as needed.

From hackers and malware, to disgruntled or careless employees, to outdated or otherwise vulnerable devices and operating systems, to mobile and public cloud computing, to third-party service providers, most companies are routinely exposed to security threats of varying severity in the normal course of conducting business. Given the ubiquitous, unavoidable nature of security risks, quick response time is essential to maintaining system security, and automated, continuous security monitoring is key to quick threat detection and response.

Security Architecture Design:

This secure architecture design is the result of an evolutionary process of technology advancement and increasing cyber vulnerability presented in the Recommended Practice document, *Control Systems Defense in Depth Strategies*.

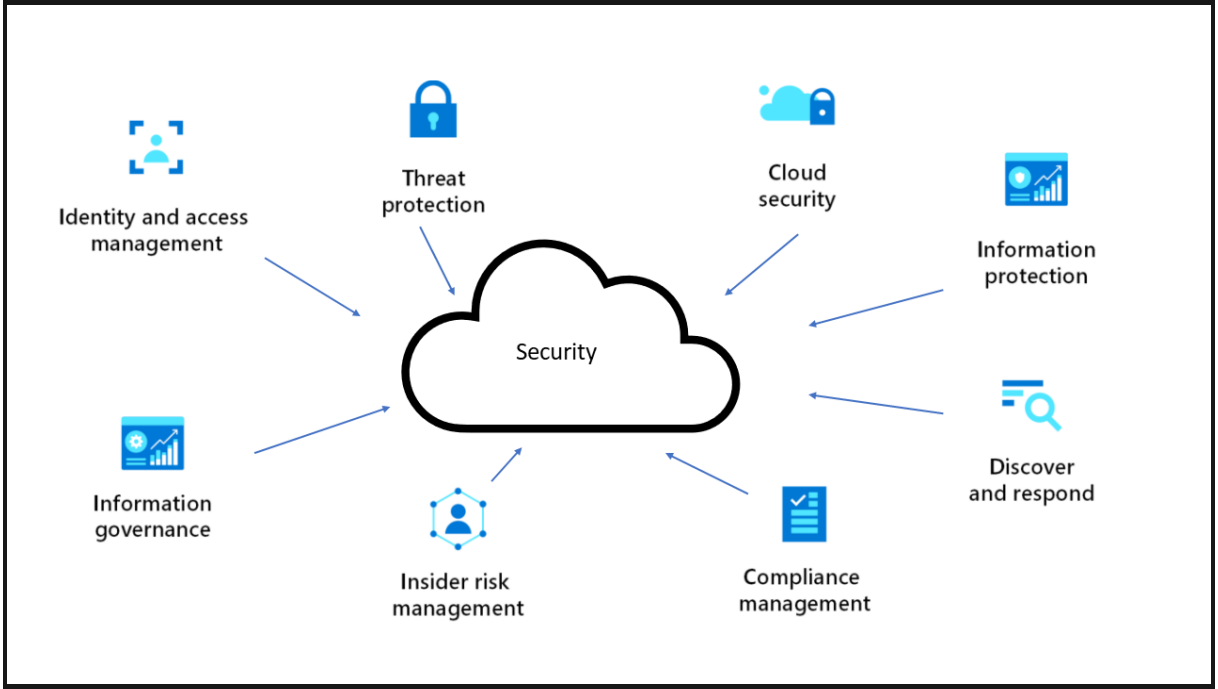
Hover over the various areas of the graphic and click inside the Box for additional information associated with the system elements. Information security has always been a complex subject, and it evolves quickly with the creative ideas and implementations of attackers and security researchers.

Security is one of the most important aspects of any architecture. Good security provides confidentiality, integrity, and availability assurances against deliberate attacks and abuse of your valuable data and systems. Losing these assurances can harm your business operations and revenue, and your organization's reputation.

Note

Learn how cloud security is an ongoing journey of incremental progress and maturity, in **Security in the Microsoft Cloud Adoption Framework for Azure**. Learn how to build security into your solution, in the Azure Well-Architected Framework **Overview of the security pillar**.

Here are some broad categories to consider when you design a security system:



Azure provides a wide range of security tools and capabilities. These are just some of the key security services available in Azure:

- Microsoft Defender for Cloud. A unified infrastructure security management system that strengthens the security posture of your datacenters. It also provides advanced threat protection across your hybrid workloads in the cloud and on-premises.
- Azure Active Directory (Azure AD). The Microsoft cloud-based identity and access management service.

- Azure Front Door. A global, scalable entry-point that uses the Microsoft global edge network to create fast, highly secure, and widely scalable web applications.
- Azure Firewall. A cloud-native, intelligent network firewall security service that provides threat protection for your cloud workloads that run in Azure.
- Azure Key Vault. A high-security secret store for tokens, passwords, certificates, API keys, and other secrets. You can also use Key Vault to create and control the encryption keys used to encrypt your data.
- Azure Private Link. A service that enables you to access Azure PaaS services, Azure-hosted services that you own, or partner services over a private endpoint in your virtual network.
- Azure Application Gateway. An advanced web traffic load balancer that enables you to manage traffic to your web applications.
- Azure Policy. A service that helps you enforce organizational standards and assess compliance.

Data Security:

Data security is the practice of protecting digital information from unauthorized access, corruption, or theft throughout its entire lifecycle. It's a concept that encompasses every aspect of information security from the physical security of hardware and storage devices to administrative and access controls, as well as the logical security of software applications. It also includes organizational policies and procedures.

When properly implemented, robust data security strategies will protect an organization's information assets against cybercriminal activities, but they also guard against insider threats and human error, which remains among the leading causes of data breaches today. Data security involves deploying tools and technologies that enhance the organization's visibility into where its critical data resides and how it is used. Ideally, these tools should be able to apply protections like encryption, data masking, and redaction of sensitive files, and should automate reporting to streamline audits and adhering to regulatory requirements.

Business challenges

Digital transformation is profoundly altering every aspect of how today's businesses operate and compete. The sheer volume of data that enterprises create, manipulate, and store is growing, and drives a greater need for data governance. In addition, computing environments are more complex than they once were, routinely spanning the public cloud, the enterprise data center, and numerous edge devices ranging from Internet of Things (IoT) sensors to robots and remote servers. This complexity creates an expanded attack surface that's more challenging to monitor and secure.

At the same time, consumer awareness of the importance of data privacy is on the rise. Fueled by increasing public demand for data protection initiatives, multiple new privacy regulations have recently been enacted, including Europe's General Data Protection Regulation (GDPR) and the California Consumer Protection Act (CCPA). These rules join

longstanding data security provisions like the Health Insurance Portability and Accountability Act (HIPAA), protecting electronic health records, and the Sarbanes-Oxley Act (SOX), protecting shareholders in public companies from accounting errors and financial fraud. With maximum fines in the millions of dollars, every enterprise has a strong financial incentive to ensure it maintains compliance.

The business value of data has never been greater than it is today. The loss of trade secrets or intellectual property (IP) can impact future innovations and profitability. So, trustworthiness is increasingly important to consumers, with a full 75% reporting that they will not purchase from companies they don't trust to protect their data.

Application Security:

Application security describes security measures at the application level that aim to prevent data or code within the app from being stolen or hijacked. It encompasses the security considerations that happen during application development and design, but it also involves systems and approaches to protect apps after they get deployed.

Application security may include hardware, software, and procedures that identify or minimize security vulnerabilities. A router that prevents anyone from viewing a computer's IP address from the Internet is a form of hardware application security. But security measures at the application level are also typically built into the software, such as an application firewall that strictly defines what activities are allowed and prohibited. Procedures can entail things like an application security routine that includes protocols such as regular testing.

Application security definition

Application security is the process of developing, adding, and testing security features within applications to prevent security vulnerabilities against threats such as unauthorized access and modification.

application security is important

Application security is important because today's applications are often available over various networks and connected to the cloud, increasing vulnerabilities to security threats and breaches. There is increasing pressure and incentive to not only ensure security at the network level but also within applications themselves. One reason for this is because hackers are going after apps with their attacks more today than in the past. Application security testing can reveal weaknesses at the application level, helping to prevent these attacks.

Types of application security

Different types of application security features include authentication, authorization, encryption, logging, and application security testing. Developers can also code applications to reduce security vulnerabilities.

- **Authentication**: When software developers build procedures into an application to ensure that only authorized users gain access to it. Authentication procedures ensure that a user is who they say they are. This can be accomplished by requiring the user to provide a user name and password when logging in to an application. Multi-factor authentication requires more than one form of authentication—the factors might include something you know (a password), something you have (a mobile device), and something you are (a thumb print or facial recognition).
- **Authorization**: After a user has been authenticated, the user may be authorized to access and use the application. The system can validate that a user has permission to access the application by comparing the user's identity with a list of authorized users. Authentication must happen before authorization so that the application matches only validated user credentials to the authorized user list.
- **Encryption**: After a user has been authenticated and is using the application, other security measures can protect sensitive data from being seen or even used by a cybercriminal. In cloud-based applications, where traffic containing sensitive data travels between the end user and the cloud, that traffic can be encrypted to keep the data safe.
- **Logging**: If there is a security breach in an application, logging can help identify who got access to the data and how. Application log files provide a time-stamped record of which aspects of the application were accessed and by whom.
- **Application security testing**: A necessary process to ensure that all of these security controls work properly.

Application security in the cloud

Application security in the cloud poses some extra challenges. Because cloud environments provide shared resources, special care must be taken to ensure that users only have access to the data they are authorized to view in their cloud-based applications. Sensitive data is also more vulnerable in cloud-based applications because that data is transmitted across the Internet from the user to the application and back.

Mobile application security

Mobile devices also transmit and receive information across the Internet, as opposed to a private network, making them vulnerable to attack. Enterprises can use virtual private networks (VPNs) to add a layer of mobile application security for employees who log in to applications remotely. IT departments may also decide to vet mobile apps and make sure they conform to company security policies before allowing employees to use them on mobile devices that connect to the corporate network.

Virtual machine Security:

The term **"Virtualized Security,"** sometimes known as "security virtualization," describes security solutions that are software-based and created to operate in a virtualized IT environment. This is distinct from conventional hardware-based network security, which is static and is supported by equipment like conventional switches, routers, and firewalls.
Virtualized security is flexible and adaptive, in contrast to hardware-based security. It can be deployed anywhere on the network and is frequently cloud-based so it is not bound to a specific device.

In Cloud Computing**,** where operators construct workloads and applications on-demand, virtualized security enables security services and functions to move around with those on-demand-created workloads. This is crucial for virtual machine security. It's crucial to protect virtualized security in cloud computing technologies such as isolating multitenant setups in public cloud settings. Because data and workloads move around a complex ecosystem including several providers, virtualized security's flexibility is useful for securing hybrid and multi-cloud settings.

Types of Hypervisors

Type-1 Hypervisors

Its functions are on unmanaged systems. Type 1 hypervisors include **Lynx Secure, RTS Hypervisor, Oracle VM, Sun xVM Server, and Virtual Logic VLX**. Since they are placed on bare systems, type 1 hypervisor do not have any host operating systems.

Type-2 Hypervisor

It is a software interface that simulates the hardware that a system typically communicates with. Examples of Type 2 hypervisors include **containers, KVM, Microsoft Hyper V, VMWare Fusion, Virtual Server 2005 R2, Windows Virtual PC, and VMware workstation 6.0**.

Type I Virtualization

In this design, the **Virtual Machine Monitor (VMM)** sits directly above the hardware and eavesdrops on all interactions between the VMs and the hardware. On top of the VMM is a management VM that handles other guest VM management and handles the majority of a hardware connections. The Xen system is a common illustration of this kind of virtualization design.

Type II virtualization

In these architectures, like VMware Player, allow for the operation of the VMM as an application within the host operating system (OS). I/O drivers and guest VM management are the responsibilities of the host OS.

Identity management and Access Control:

n a recent study by Verizon, 63% of the confirmed data breaches are due to either weak, stolen, or default passwords used. There is a saying in the cybersecurity world that goes like this “No matter how good your chain is it’s only as strong as your weakest link.” and exactly hackers use the weakest links in the organization to infiltrate. They usually use phishing attacks to infiltrate an organization and if they get at least one person to fall for it, it’s a serious turn of events from thereon. They use the stolen credentials to plant back doors, install malware or exfiltrate confidential data, all of which will cause serious losses for an organization. And so Identity and Access Management (IAM) is a combination of policies and technologies that allows organizations to identify users and provide the right form of access as and when required. There has been a burst in the market with new applications, and the requirement for an organization to use these applications has increased drastically. The services and resources you want to access can be specified in IAM. IAM doesn’t provide any replica or backup. IAM can be used for many purposes such as, if one want’s to control access of individual and group access for your AWS resources. With IAM policies, managing permissions to your workforce and systems to ensure least-privilege permissions becomes easier. The AWS IAM is a global service.

Components of IAM

- Users
- Roles
- Groups
- Policies

With these new applications being created over the cloud, mobile and on-premise can hold sensitive and regulated information. It’s no longer acceptable and feasible to just create an Identity server and provide access

based on the requests. In current times an organization should be able to track the flow of information and provide least privileged access as and when required, obviously with a large workforce and new applications being added every day it becomes quite difficult to do the same. So organizations specifically concentrate on managing identity and its access with the help of a few IAM tools. It's quite obvious that it is very difficult for a single tool to manage everything but there are multiple IAM tools in the market that help the organizations with any of the few services given below.

Section-A

Answer any five from the following 5*2=10m

1.Define cloud architecture and Model?

2.Define virtualization?

3.Differentiate between private and public cloud?

4.Explain cloud ecosystem?

5.Define virtualization of cpu?

6.Define VMWare?

7.Define virtual clusters?

8.Explain cloud Architecture?

Section-B

Answer any one question from each unit. 2*10=20m

UNIT-1

9a). Explain about NIST cloud computing reference architecture?

b).Explain about cloud Models(IaaS, PaaS, SaaS)?

10a). What are the different technologies for network based system?

b). Explain about cloud ecosystem, service management and computing on demand?

UNIT-2

11a). What are the different types of virtualization and explain in detail?

b). Define virtualization structure, tools and mechanisms?

12a).Explain about virtualization of CPU, memory, I/O devices?

b).Explain about VMWare, virtualbox virtualization software?

Section-A

Answer any five from the following 5*2=10m

1.Define cloud Architecture?

2.Explain about four levels of federation?

3.Define MapReduce , Twister and iterative MapReduce?

4.Explain advantages of cloud storage?

5.what are the cloud environments and explain?

6.what are the cloud security challenges and risks?

7.Define Data security?

8.Explain about future of Federation?

Section-B

Answer any one question from each unit. 2*10=20m

UNIT-1

9a).Explain about Mapping Applications in cloud computing?

b).define Risk Management?and explain abriefly about Security Monitoring?

10a).define feature of federation and describe about cloud storage?

b)what is Cloud Environment? And explain about Eucalyptus I Environments?

11a).Explain about Parallel and Distributed Programming Paradigms?

b)Evaluate Virtual Machine Security?

12a).define about Security Governance and also describe about Design Challenges?

MASTER OF COMPUTER APPLICATIONS DEGREE EXAMINATION

FOURTH SEMESTER

Paper MCA 401A: Cloud Computing

(Under C.B.S.C Revised Regulations w.e.f.2021-2023)

(Common paper to University and all Affliated Colleges)

Time:3 hours Max.Marks:70

PART-A

(Compulsory)

Answer any five of the following questions each question carries 2 marks(5*2=10)

1.a)define cloud computing?

b)write about applicatiions of cloud computing?

c)define Risk Management?

d)write about cloud services?

e)what is cloud Environment?

f)define Cloud Architecture?

g) Define Data security?

h) Define VMWare?

i) Define virtualization?

j) Define cloud architecture and Model?

PART-B

Answer any ONE full question from each unit

Each question carries 12 Marks (5*12=60)

Unit-1

2. what is Cloud Environment? And explain about Eucalyptus I Environments?

(or)

3. Explain about Mapping Applications in cloud computing?

Unit-2

4. Explain about VMWare, virtualbox virtualization software?

(or)

5. Explain about cloud ecosystem, service management and computing on demand?

Unit-3

6. describe Architectural Design of Compute and Storage Clouds?

(or)

7. Evaluate Inter Cloud Resource Management?

Unit-4

8. describe Parallel and Distributed Programming Paradigms?

(OR)

9.define Cloud Storage and give its Advantages?

Unit-5

10. Describe about Identity Management and Access Control.?

(or)

11.define Cloud Security and in detail.?