UNIT-1 CLOUD COMPUTING

1.1 Learning objectives

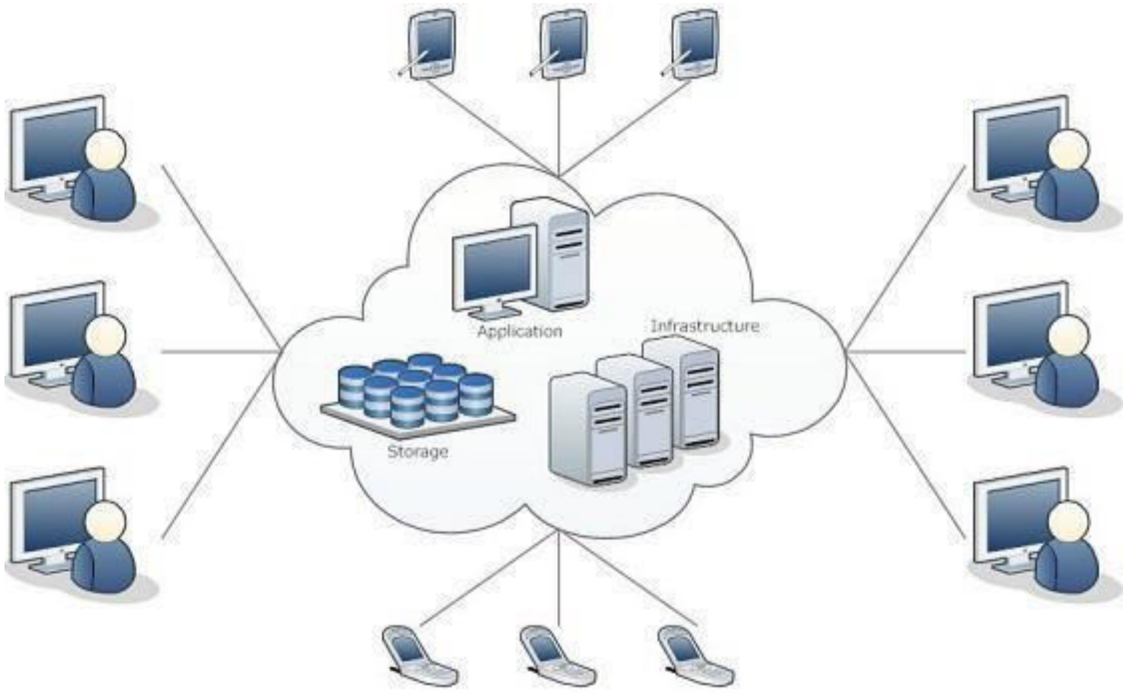
- Describe cloud computing.
- Know about advantages and applications of cloud computing.
- Elaborate about history of cloud computing.
- Know about challenges of cloud computing.

1.2 What is Cloud?

The term Cloud refers to a Network or Internet. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e., on public networks or on private networks, i.e., WAN, LAN or VPN. Applications such as e-mail, web conferencing, customer relationship management (CRM), all run in cloud.

1.3 What is Cloud Computing?

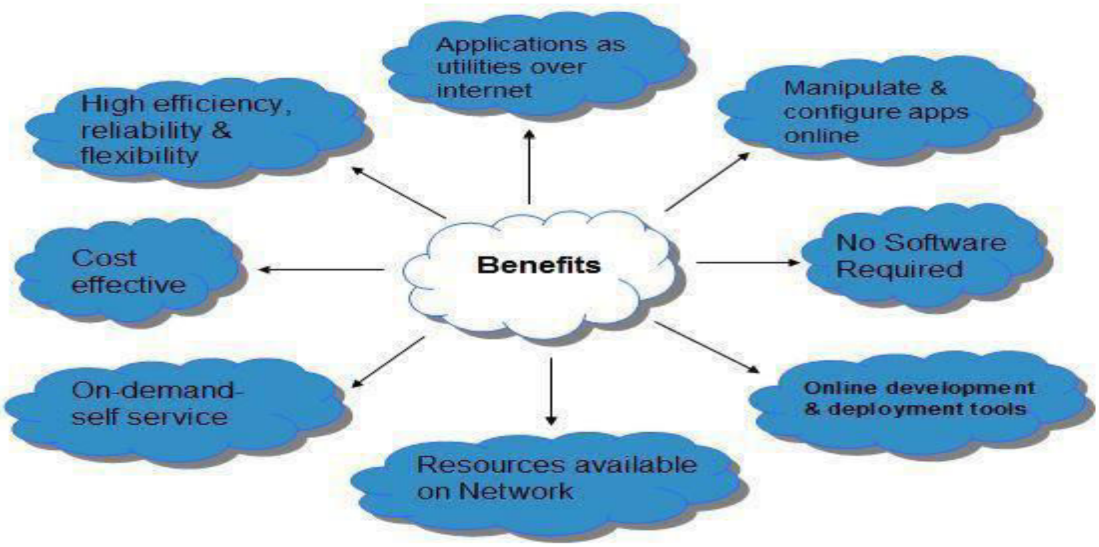
Cloud Computing refers to manipulating, configuring, and accessing the applications online. It offers online data storage, infrastructure and application.



1.4 Advantages of Cloud Computing

Cloud Computing has numerous advantages. Some of them are listed below:

- One can access applications as utilities, over the Internet.
- Manipulate and configure the application online at any time.
- It does not require installing a specific piece of software to access or manipulating cloud application.
- Cloud Computing offers online development and deployment tools, programming runtime environment through Platform as a Service model


Applications as utilities over internet
High efficiency, reliability & flexibility
Manipulate & configure apps online
Cost effective
Benefits
No Software Required
On-demand-self service
Online development & deployment tools
Resources available on Network

Cloud resources are available over the network in a manner that provides platform independent access to any type of clients.

- Cloud Computing offers on-demand self-service. The resources can be used without interaction with cloud service provider.

- Cloud Computing is highly cost effective because it operates at higher efficiencies with greater utilization. It just requires an Internet connection.

- Cloud Computing offers load balancing that makes it more reliable.

1.5 Applications

With its advent, cloud computing has occupied a very significant position in the IT industry. Distinct practical applications are making use of the services provided by it. Arenas like medical research to agriculture; educational institutions to industries are availing its services.

a. Educational institutions: Cloud computing has truly revolutionized the erudition sector. The conventional face to face classroom techniques are increasingly being replaced with other cloud mediated exercises like smart classes using pictorial and auditory illustrations. The cloud based model gives an edge to the redundant study routine people has in their daily lives. It has also enabled remote access of erudition material and has done a great deal in assisting the progress of rustic India whilst making learning easy for them.

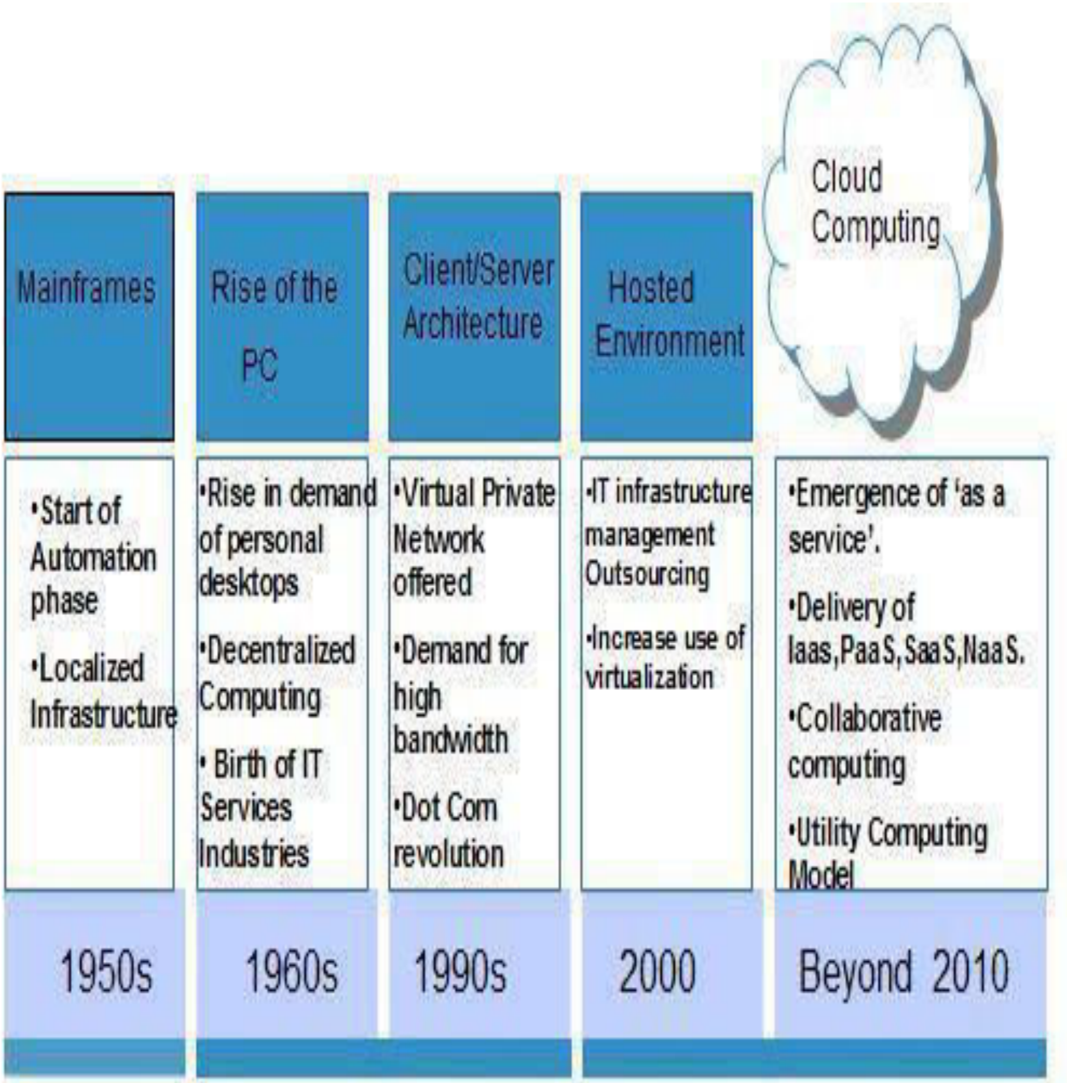
b. Industries: Cloud computing has empowered the industries to prevent varied technical and business problems that can occur while executing their own data centres and save money by incorporating a pay-per-use facility. Additional costs for running their own data centers are reduced thus saving overheads and simultaneously availing cloud services. It also allows them to increase their resources. It has management of data and the records very easy for the companies like never before. They now have access to a plethora of software and hardware services without having the need to buy them all thus improving the quality of services.

c. Medical fields: In hospitals a cloud assists in procuring patient’s information by the medical professionals which enables them to access the data remotely instead of having to go through a hospital’s computers. This aids in updating professionals about their patient’s condition even if they are not present in the hospitals. Cloud computing is still emerging in this field. There is a lot more to come.

d. Banking Industry: All the banking companies across the world have become automated and are now increasingly availing cloud services. Though adoption of cloud in this sector is relatively low on account of the security issues that prevail. With new measures being taken this industry is now increasingly employing cloud services so as to reduce their cost of ownership. Core banking, communication services, on demand BI is some ways in which banks make use of cloud computing

1.6 History

The concept of **Cloud Computing** came into existence in 1950 with implementation of mainframe computers, accessible via **thin/static clients**. Since then, cloud computing has been evolved from static clients to dynamic ones from software to services. The following diagram explains the evolution of cloud computing:

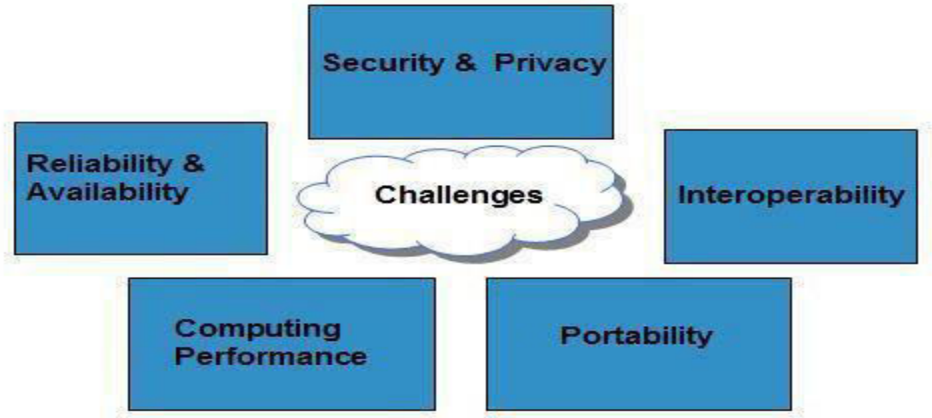


Mainframes	Rise of the PC	Client/Server Architecture	Hosted Environment	Cloud Computing
•Start of Automation phase •Localized Infrastructure	•Rise in demand of personal desktops •Decentralized Computing • Birth of IT Services Industries	•Virtual Private Network offered •Demand for high bandwidth •Dot Com revolution	•IT infrastructure management Outsourcing •Increase use of virtualization	•Emergence of 'as a service'. •Delivery of Iaas,PaaS,SaaS,NaaS. •Collaborative computing •Utility Computing Model
1950s	1960s	1990s	2000	Beyond 2010

1.7 Cloud Computing Challenges

Cloud Computing, an emergence technology, has placed many challenges in different aspects.

Some of these are shown in the following diagram:



1.7.1 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.

1.7.2 Portability

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

1.7.3 Interoperability

Application on one platform should be able to incorporate services from other platform. It is made possible via web services. But writing such web services is very complex.

1.7.4 Computing Performance

To deliver data intensive applications on cloud requires high network bandwidth, which results in high cost. If done at low bandwidth, then it does not meet the required computing performance of cloud application.

1.7.5 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.

VERY SHORT QUESTIONS

1. What is cloud?
2. What is cloud computing?
3. What is on-demand self service?
4. What is Mail chimp?
5. What type of service is provided by Mozy application?
6. What is interoperability?

SHORT QUESTIONS

1. Explain four advantages of cloud computing?
2. Explain any four characteristics of cloud computing?
3. Explain cloud computing with the help of diagram?
4. What is the history of cloud computing?

LONG QUESTIONS

1. What are the challenges of cloud computing?
2. What are the applications of cloud computing?
3. Explain advantages of cloud computing?

UNIT 2 CLOUD COMPUTING SERVICE MODELS AND DEPLOYMENT MODELS

2.1 Learning objectives

- Describe different service models.
- Elaborate benefits of service models and issues related to them.
- Describe different deployment models.
- Know advantages and disadvantages of deployment models.

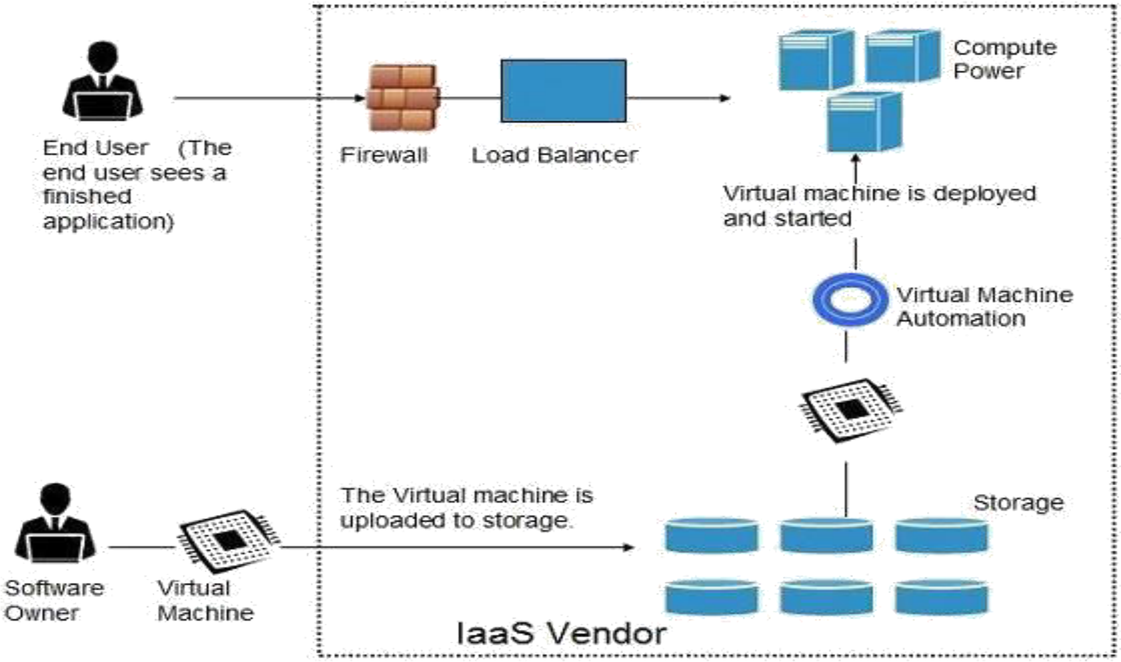
2.2. SERVICE MODELS

2.2.1. Infrastructure as a service

Iaas provider's access to fundamental resources such as physical machines, virtual machines, virtual storage, etc., Apart from these resources, the IaaS also offers:

- Virtual machine disk storage
- Virtual local area network (VLANs)
- Load balancers
- IP addresses
- Software bundles

All of the above resources are made available to end user via server virtualization. Moreover, these resources are accessed by the customers as if they own them.

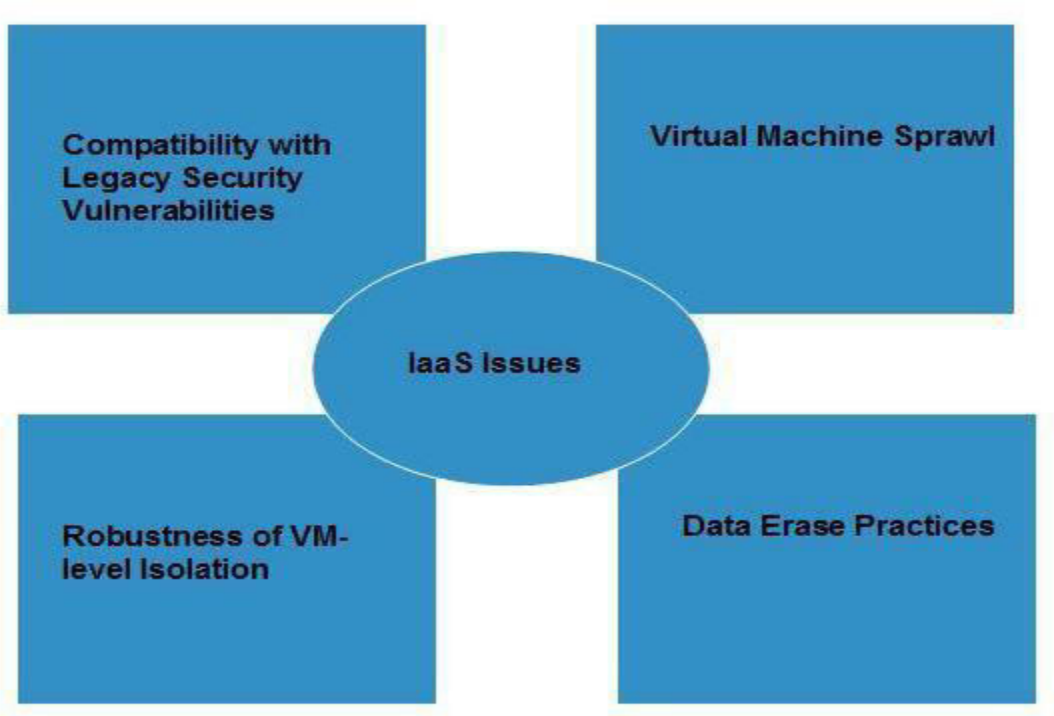

End User (The end user sees a finished application)
Firewall
Load Balancer
Compute Power
Virtual machine is deployed and started
Virtual Machine Automation
The Virtual machine is uploaded to storage.
Storage
Software Owner
Virtual Machine
IaaS Vendor

Benefits:

IaaS allows the cloud provider to freely locate the infrastructure over the Internet in a cost-effective manner. Some of the key benefits of IaaS are listed below:

- Full Control of the computing resources through Administrative Access to VMs.
- Flexible and Efficient renting of Computer Hardware.
- Portability, Interoperability with Legacy Applications.

Issues:
IaaS shares issues with PaaS and SaaS, such as Network dependence and browser based risks. It also have some specific issues associated with it. These issues are mentioned in the following diagram:


Compatibility with Legacy Security Vulnerabilities
Virtual Machine Sprawl
IaaS Issues
Robustness of VM-level Isolation
Data Erase Practices

- Compatibility With Legacy Security Vulnerabilities

Because IaaS offers the consumer to run legacy software in provider's infrastructure, therefore it exposes consumers to all of the security vulnerabilities of such legacy software.

- Virtual Machine Sprawl

The VM can become out of date with respect to security updates because IaaS allows the consumer to operate the virtual machines in running, suspended and off state. However, the provider can automatically update such VMs, but this mechanism is hard and complex.

- Robustness Of Vm-Level Isolation

IaaS offers an isolated environment to individual consumers through hypervisor. Hypervisor is a software layer that includes hardware support for virtualization to split a physical computer into multiple virtual machines.

- Data Erase Practices
  The consumer uses virtual machines that in turn use the common disk resources provided by the cloud provider. When the consumer releases the resource, the cloud provider must ensure that next consumer to rent the resource does not observe data residue from previous consumer.

Characteristics:

Here are the characteristics of IaaS service model:

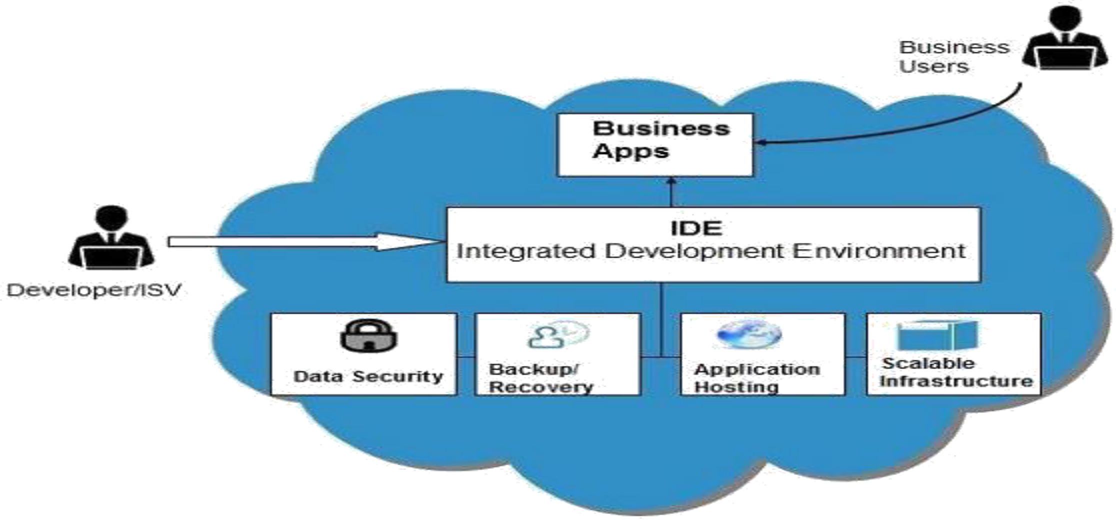
- Virtual machines with pre-installed software.
- Virtual machines with pre-installed Operating Systems such as Windows, Linux, and Solaris.
- On-demand availability of resources.
- Allows to store copies of particular data in different locations.
- The computing resources can be easily scaled up and down.

2.2.2 Platform as a Service(PAAS)-

It also offers development & deployment tools, required to develop applications. PaaS has a feature of point-and-click tools that enables non-developers to create web applications.Google's App Engine, Force.com are examples of PaaS offering vendors. Developer may log on to thesewebsites and use the built-in API to create web-based applications.

But the disadvantage of using PaaS is that the developer lock-in with a particular vendor. For example, an application written in Python against Google's API using Google's App Engine is likely to work only in that environment. Therefore, the vendor lock-in is the biggest problem in PaaS.

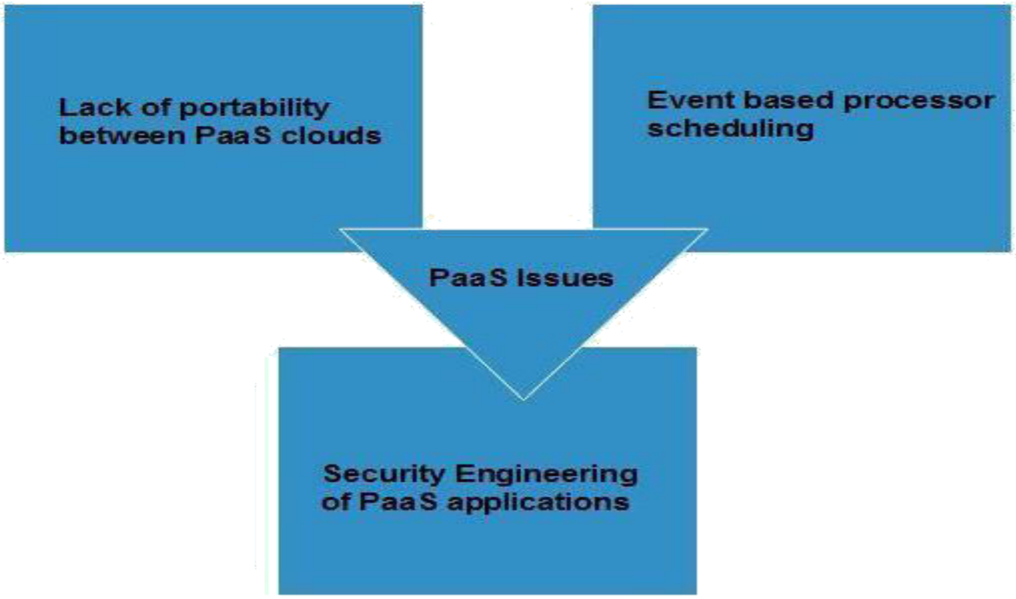
The following diagram shows how PaaS offers an API and development tools to the developers and how it helps the end user to access business applications.



Benefits:

- Lower Administrative Overhead
  Consumer need not to bother much about the administration because it's the responsibility of cloud provider.
- Lower Total Cost Of Ownership
  Consumer need not purchase expensive hardware, servers, power and data storage.
- Scalable Solutions
  It is very easy to scale up or down automatically based on application resource demands.
- More Current System Softwar
  It is the responsibility of the cloud provider to maintain software versions and patch installations.

Issues:


Lack of portability between PaaS clouds
Event based processor scheduling
PaaS Issues
Security Engineering of PaaS applications

- Lack Of Portability Between Paas Clouds
  Although standard languages are used yet the implementations of platforms services may vary. For example, file, queue, or hash table interfaces of one platform may differ from another, making it difficult to transfer workloads from one platform to another.
- Event Based Processor Scheduling

The PaaS applications are event oriented which poses resource constraints on applications, i.e., they have to answer a request in a given interval of time.

- Security Engineering Of Paas Applications
  Since the PaaS applications are dependent on network, PaaS applications must explicitly use cryptography and manage security exposures.

Characteristics:

Here are the characteristics of PaaS service model:

- PaaS offers browser based development environment. It allows the developer to create database and edit the application code either via Application Programming Interface or point-and-click tools.
- PaaS provides built-in security, scalability, and web service interfaces.
- PaaS provides built-in tools for defining workflow and approval processes and defining business rules.
- It is easy to integrate with other applications on the same platform.
- PaaS also provides web services interfaces that allow us to connect the applications outside the platform.

2.2.3 Software as a Service(SaaS)

This model allows providing software application as a service to the end users. It refers to a software that is deployed on a hosted service and is accessible via Internet. There are several

SaaS applications, some of them are listed below:

- Billing and Invoicing System
- Customer Relationship Management (CRM) applications
- Help Desk Applications
- Human Resource (HR) Solutions

Some of the SaaS applications are not customizable such as an Office Suite. But SaaS provides us Application Programming Interface (API), which allows the developer to develop a customized application.

Characteristics:

Here are the characteristics of SaaS service model:

- SaaS makes the software available over the Internet.
- The Software are maintained by the vendor rather than where they are running.
- The license to the software may be subscription based or usage based. And it is billed on recurring basis.
- SaaS applications are cost effective since they do not require any maintenance at end user side.
- They are available on demand.

- They can be scaled up or down on demand.
- They are automatically upgraded and updated.
- SaaS offers share data model. Therefore, multiple users can share single instance of infrastructure. It is not required to hard code the functionality for individual users.
- All users are running same version of the software.

Benefits:

Using SaaS has proved to be beneficial in terms of scalability, efficiency, performance and much more. Some of the benefits are listed below:

- Modest Software Tools
- Efficient use of Software Licenses
- Centralized Management & Data
- Platform responsibilities managed by provider
- Multitenant solutions

Issues:

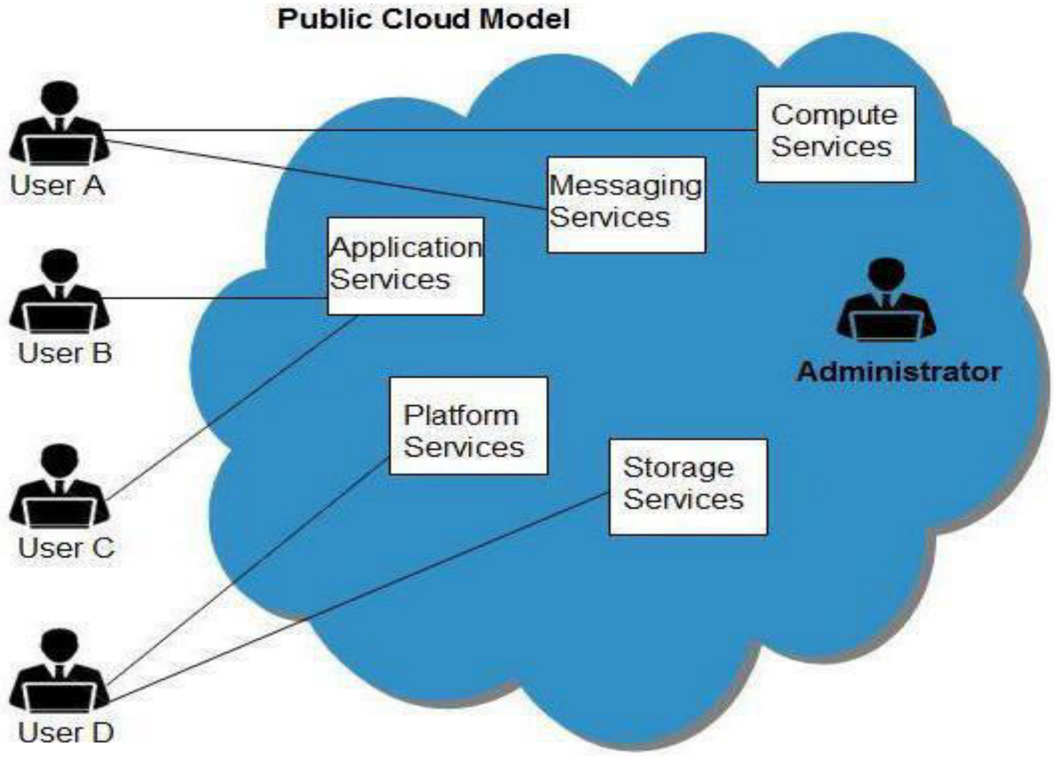
There are several issues associated with SaaS, some of them are listed below:

- Browser Based Risks
  If the consumer visits malicious website and browser becomes infected, and the subsequent access to SaaS application might compromise the consumer's data. To avoid such risks, the consumer can use multiple browsers and dedicate a specific browser to access SaaS applications or can use virtual desktop while accessing the SaaS applications.
- Network Dependence
  The SaaS application can be delivered only when network is continuously available. Also network should be reliable but the network reliability cannot be guaranteed either by cloud provider or the consumer.
- Lack Of Portability Between Saas Clouds
  Transferring workloads from one SaaS cloud to another is not so easy because work flow, business logics, user interfaces, support scripts can be provider specific

2.3 Deployment Models

2.3.1 Public clouds

The Public Cloud allows systems and services to be easily accessible to general public, e.g., Google, Amazon, Microsoft offers cloud services via Internet.



Benefits:

- Cost Effective
  Since public cloud share same resources with large number of consumer, it has low cost.
- Reliability:
  Since public cloud employs large number of resources from different locations, if any of the resource fail, public cloud can employ another one.
- Flexibility
  It is also very easy to integrate public cloud with private cloud, hence gives consumers a flexible approach.
- Location Independence
  Since, public cloud services are delivered through Internet, therefore ensures location independence.
- Utility Style Costing
  Public cloud is also based on pay-per-use model and resources are accessible whenever consumer needs it.
- High Scalability

Cloud resources are made available on demand from a pool of resources, i.e., they can be scaled up or down according the requirement.

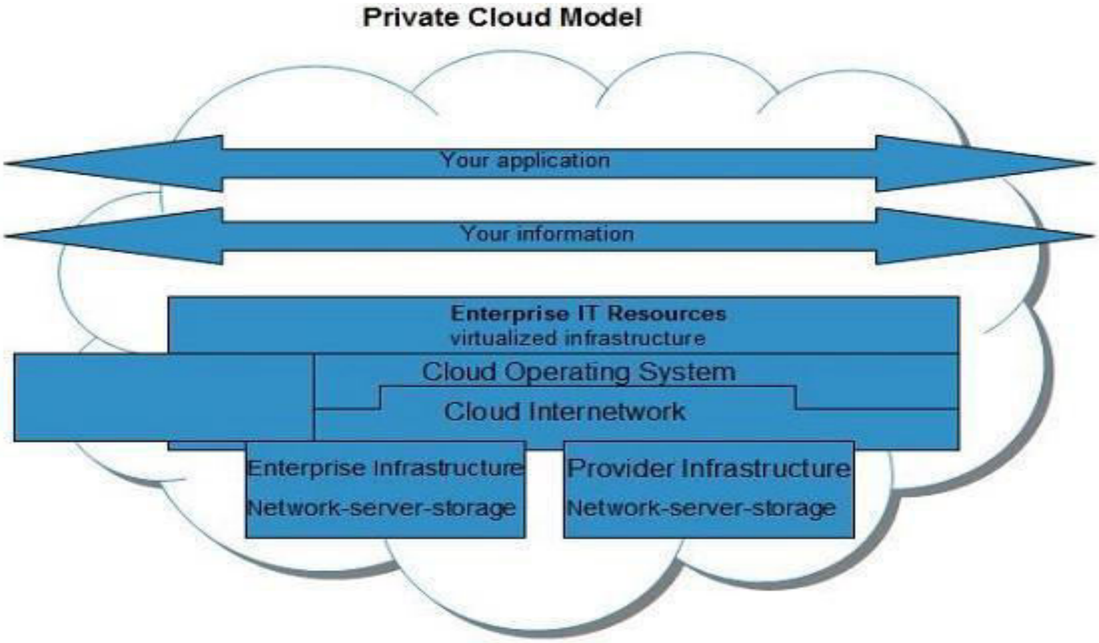
Disadvantages:

Here are the disadvantages of public cloud model:

- Low Security
  In public cloud model, data is hosted off-site and resources are shared publicly, therefore does not ensure higher level of security.
- Less Customizable
  It is comparatively less customizable than private cloud.

2.3.2 Private cloud

The Private Cloud allows systems and services to be accessible within an organization. The Private Cloud is operated only within a single organization. However, it may be managed internally or by third-party.



Benefits:

- Higher Security And Privacy
  Private cloud operations are not available to general public and resources are shared from distinct pool of resources, therefore, ensures high security and privacy.
- More Control
  Private clouds have more control on its resources and hardware than public cloud because it is accessed only within an organization.
- Cost And Energy Efficiency
  Private cloud resources are not as cost effective as public clouds but they offer more efficiency than public cloud.

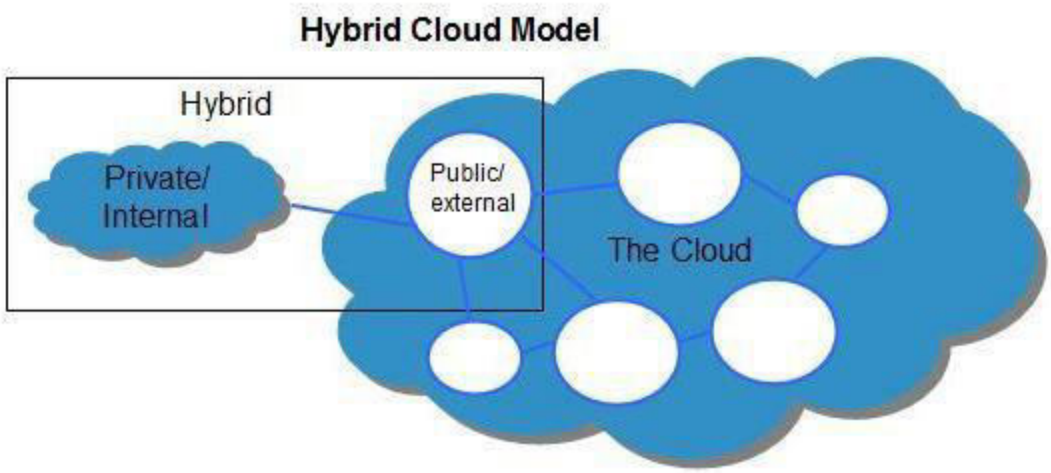
Disadvantages:

Here are the disadvantages of using private cloud model:

- Restricted Area
  Private cloud is only accessible locally and is very difficult to deploy globally.
- Inflexible Pricing
  In order to fulfill demand, purchasing new hardware is very costly.
- Limited Scalability
  Private cloud can be scaled only within capacity of internal hosted resources.

2.3.3 Hybrid cloud

The Hybrid Cloud is a mixture of public and private cloud. Non-critical activities are performed using public cloud while the critical activities are performed using private cloud.



Benefits:

- Scalability
  It offers both features of public cloud scalability and private cloud scalability.
- Flexibility
  It offers both secure resources and scalable public resources.

- Cost Efficiencies

Public cloud are more cost effective than private, therefore hybrid cloud can have this saving.

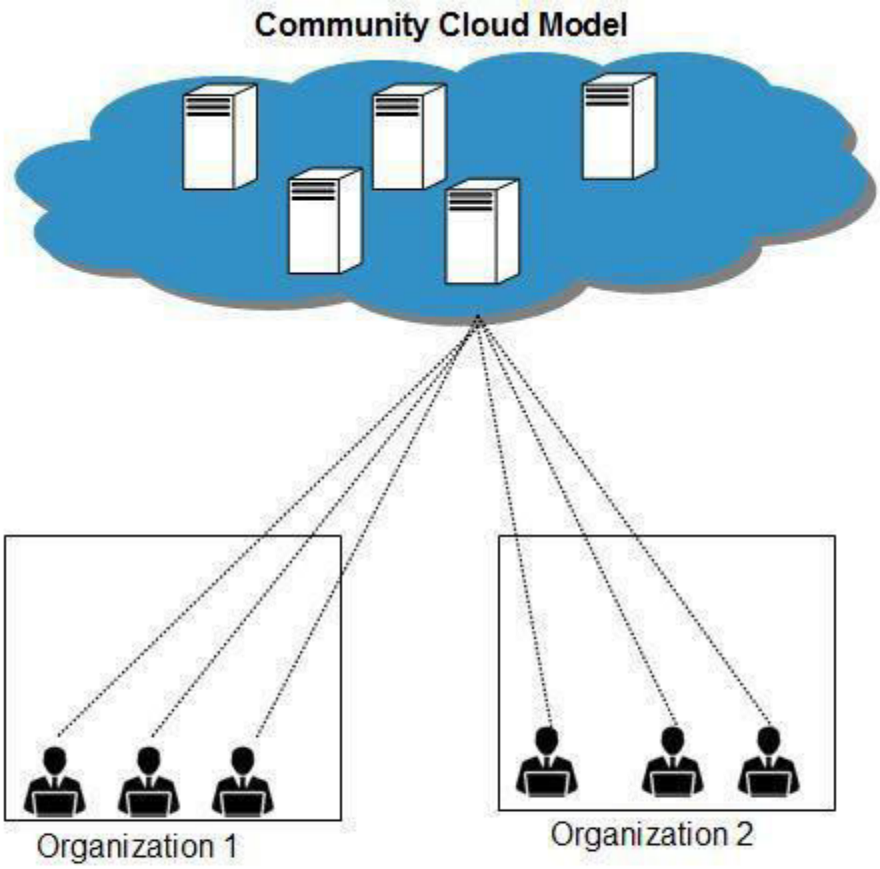
- Security
  Private cloud in hybrid cloud ensures higher degree of security.

Disadvantages:

- Networking Issues
  Networking becomes complex due to presence of private and public cloud.
- Security Compliance
  It is necessary to ensure that cloud services are compliant with organization's security policies.

2.3.4 Community cloud

The community Cloud allows system and services to be accessible by group of organizations. It shares the infrastructure between several organizations from a specific community. It may be managed internally or by the third-party.



Benefits:

There are many benefits of deploying cloud as community cloud model. The following diagram shows some of those benefits:

- Cost Effective
  Community cloud offers same advantage as that of private cloud at low cost.
- Sharing Between Organizations
  Community cloud provides an infrastructure to share cloud resources and capabilities among several organizations.
- Security
  Community cloud is comparatively more secure than the public cloud.

Issues:

- Since all data is housed at one location, one must be careful in storing data in community cloud because it might be accessible by others.
- It is also challenging to allocate responsibilities of governance, security and cost.

Very short questions

1. What are different service models? Name them.
2. Name fundamental resources of Iaas?
3. Give two benefits of Iaas?
4. Paas is also known as.
5. What is private cloud?
6. What is hybrid cloud?
7. What is community cloud?

Short questions

1. Name different cloud service models with suitable example?
2. Explain characteristics of Iaas model?
3. Define private cloud along with diagram.
4. Explain benefits of private cloud?
5. What are the issues found with Paas?

Long questions

1. Explain Iaas along with four benefits, issues and characteristics.
2. What is the difference between public cloud, private cloud, hybrid cloud, community cloud?
3. Explain cloud computing deployment models in detail.

UNIT 3 SERVICE LEVEL AGREEMENT MANAGEMENT

3.1 Learning objectives

- To know about SLA
- Elaborate types of SLA.
- To discuss SLA management process.
- To discuss SLA life cycle

3.2 What is Service Level Agreement

If you've enrolled in an ITIL Training you won't have to ask what a service level agreement (SLA) is. However, if you haven't done an ITIL online course yet, we'll help you out. **A service level agreement (SLA) is an agreement between an IT Service provider and a customer.**

For instance, you are a customer of a bank and the bank provides services to you. A service level agreement between you and the bank describes the services provided and the service levels at which they will be provided. For example, the bank will allow you to withdraw money from an ATM and the transaction will last no longer than 10 seconds. That is an example of a service level agreement and it is part of service level management.

3.3 Types of SLA

There are three types of service level agreements that can be documented.

- Service level SLA

The first type of service level agreement structure is the **service-based SLA**. A service based SLA covers one service for all customers. Let's consider that the IT service provider provides customer query service for many customers. In a service based service level agreement, the service level of the customer query service will be same for all customers that will be using this service. For instance, if the finance department and the human resources department are two customers which will be using this service, the same SLA will be valid between the IT service provider and these two departments since it is a service based SLA.

- Customer based SLA

The second type of service level agreement structure is the customer based SLA. A customer based SLA is an agreement with one customer, covering all the services used by this customer. Let's consider the relationship between you and your telecom operator. You use the voice services, SMS services, data services, and several other services of the telecom operator. For all these services, you have only one contract between you and the telecom operator. Similarly, if the IT service provider provides several services for the business and the customers, and if all the service levels are documented in one service level agreement for the provided services, it will be a customer based SLA.

- Multi-level SLA

The third and the last type of service level agreement is the multi-level SLA. In multi-level SLA, aspects of SLA are defined according to the organization of the customer using some kind of inheritance with overall definitions with relevance for all subordinate levels. This SLA focuses on the organization of the customer. All services and their interrelationships with subordinate services are used when defining the multi-level service level agreement structure.

Maintaining service level agreements are part of service level management. Every time a service change, or the service level target of a service change, the service level agreement needs to be reviewed and revised. The new service level agreement needs to reflect the changes made to the service or the service level targets. Therefore, the management of service level agreements is an important part of ITIL continual service improvement.

3.4 Service Level Management Process

There are nine key activities that form part of the service level management process.

- Designing of SLA Frameworks

The design of service level agreement (SLA) frameworks is an activity of the service level management process. Depending on the size of the business and complexity of the IT organization, there can be several services that need to be supported and provided. Therefore, determining the appropriate SLA structure is one of the main activities of the service level management process.

- Developing of service level requirement

Determining, documentation and agreement of requirements for new or changed services and the development of service level requirements (SLRs) are managed in the ITIL service level management process. The main objective of the service level management process is to:

- Determine the expectations of the business and customer
- Evaluate the resources and capabilities of the IT service provider,
- Agree on what service levels will be provided by the IT service provider to the customer and business.

During these steps, documenting service level requirements and the expectations of the business and, the agreement on the service levels for new or changed services are activities of the service level management process.

- Monitoring service performance

Based on the agreed service levels between the IT service provider and the business or customer of the services, periodic and progressive measurements are done in order to check whether agreed service levels are achieved by the IT service provider. For instance, if 99% availability is a service level requirement of the business for a service, this availability service level requirement should be measured and checked properly. Therefore, monitoring service performance achievements against targets set in an SLA is a process activity of service level management process.

- Improving customer satisfaction

Collating, measuring and improving customer satisfaction is also one of the activities of the service level management process. The service level requirements of the business and agreed upon service level targets for services are all necessary for ensuring and improving customer satisfaction. Customer satisfaction ultimately leads to increased revenue and value to the company. The business and IT departments should always strive to find new ways to improve customer satisfaction, lest the customer finds better services elsewhere, resulting in a loss of revenue.

- Reviewing contracts and agreements

In order to meet the service level targets for the services of an IT service provider, all suppliers and partners serving to the IT service provider must meet their agreed service levels as well. Otherwise, even if other services and steps in the chain can meet their targets, if one service which is acquired or provided externally is not meeting its agreed service targets, the overall service levels will fail to meet its targets. Therefore, the review and revision of all underpinning contracts or agreements with the suppliers and partners who are providing external services to the IT service provider qualify as a service level management process activity.

- Producing service reports

Service reports are produced as a process activity of service level management process. In order to check whether the agreed service levels can be met by the IT service provider, provided services are monitored. Periodic reports about the performance of the agreed services are provided and this is an activity of the service level management process.

- Conducting service reviews

Conducting service reviews and instigating improvements with an overall service improvement plan is done in the within the scope of the service level management process as well. After the service level requirements and the service level targets are agreed upon between the business and IT service provider, the performance of the services are monitored and reported periodically. And based on the performance output, weakness and strengths of the provided services are assessed. Based on the analysis, a service improvement plan is created and revised regularly to achieve better service performance.

- Reviewing of agreements

Based on the performance outputs, reviewing and revision of SLAs, service scope and, underpinning agreements can take place as part of the service level management process. For instance, let's say that the requested service level for a service is serving up to one thousand users concurrently and with 98% availability in the beginning. Due to a high demand from the customers and business, new service levels can increase to serve up to two thousand users concurrently and to ensure 99% availability for this service. In this case, SLAs, service scope, and underpinning contracts with suppliers and partners must be reviewed and revised.

- Developing contracts and relationships

The last activity of the service level management process is developing contacts and relationships. During the gathering of service level requirements from the customer and business, agreeing on the service levels and targets between the IT service provider and the business through SLAs, and agreeing on the service levels for the external services provided by suppliers and partners, contacts and relationships are developed regularly as a natural outcome of the Service Level management process.

3.5 SLA Life Cycle

- Cloud service lifecycle: Acquisition

A prospective cloud customer can use service offerings published by the cloud service provider to check whether it meets her/his requirements, for example, security, personal data protection, performance etc., and see how one offering compares with another in the market. Why is it important? This phase is crucial for establishing an SLA between the cloud customer and the cloud service provider.

Assessment

Any relationship starts with pre-assessing what one would like, why, when and with whom (for instance one or more CSPs), so does the first Cloud SLA lifecycle phase, Assessment. This includes for instance doing market intelligence, checking specific needs, offerings, CSPs, performance of CSPs and setting up a business case...

Preparation

This second Cloud SLA lifecycle phase, includes for instance, the first contact and conversation with possible CSPs, further assessment, pre-evaluation and fine-tuning goals and assumptions...

Negotiation & Contracting

This phase can include preparing for negotiation and the actual negotiation and deal making with one or more CSPs, including sharing concerns, discuss in-scope and out-of-scope (cloud) services, debating about trade-offs and finding common grounds, reaching agreement, double-checking needs, goals and assumptions, and of course documenting the contractual arrangements, and signing thereof...

- Cloud service lifecycle: Operation

This phase determines whether a cloud service meets the committed service level objective (SLO) during the provisioning of the cloud service. This might imply that cloud service providers taking corrective actions to avoid SLA violations. Why is it important? SLAs can be used to monitor the cloud service provider in order to assess the correct fulfilment of the cloud service, or detect potential violations in which case remediation may take place.

Execution & Operation

This phase includes the actual start of setting up the cloud services, populating the respective cloud service with relevant data, on boarding and training users, setting up communication channels and further operational activities while using the respective cloud services...

Updates &

This phase includes updated or otherwise amended needs, goals and

Amendments

assumptions by the Cloud Service Customer during the term of the ongoing cloud services arrangements, as well as improved or added cloud services by the CSP there under. It also includes optimisation of the respective cloud services by CSP as per (contractual or other) non-compliance, breaches and other incidents during that term...

Escalation

This phase deals with contractual or other) non-compliance, breaches and other incidents during the term of the ongoing cloud services arrangements that have resulted in a dispute that needs escalation, (perhaps even litigation as a last resort), negotiation and resolution, either by parties themselves or by arbitration, court or otherwise...

- Cloud service lifecycle: Termination

Why is it important? You should already think about termination in phase 1, as an SLA can be used to arrange the conditions under which the Cloud customer's data (including but not limited to for instance Personal Identifiable Information or PII) will be exported and returned to the cloud customer, and not retained by the cloud service provider (to the extent mandatorily possible).

Termination & Consequences of Termination

This phase deals with the end of the relationship between CSP and CSC, including the end of the legal relationship even though the latter will generally continue for several years after any termination as per mandatory laws and legislation. This last phase for instance includes the assessment of alternatives, settlement and termination arrangements, cloud services transition projects and services, data export, customer and (end)use care and diligence, and adequate data deletion...

VERY SHORT QUESTIONS

1. What does SLA stands for?
2. What is SLA?
3. Name the phases in SLA life cycle.
4. Name the activities in SLA management process.
5. Name the types of SLA.

SHORT QUESTIONS

1. Explain four activities in SLA management process.

2. Elaborate briefly SLA life cycle?

LONG QUESTIONS

1. Explain SLA management process.
2. Explain SLA life cycle.
3. What are the types of SLA?

UNIT 4 VIRTUALIZATION

4.1 Learning objectives

- Know about virtualization.
- Describe hypervisor.
- Elaborate types of hypervisor

4.2 Virtualization

Virtualization is a technique, which allows sharing single physical instance of an application or resource among multiple organizations or tenants (customers). It does so by assigning a logical name to a physical resource and providing a pointer to that physical resource when demanded.

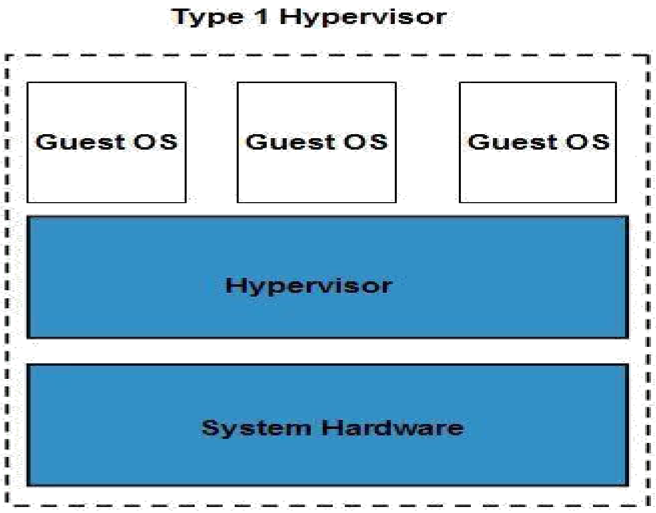
4.2.1 Virtualization Concept

Creating a virtual machine over existing operating system and hardware is referred as Hardware Virtualization. Virtual Machines provide an environment that is logically separated from the underlying hardware. The machine on which the virtual machine is created is known as host machine and virtual machine is referred as a guest machine. This virtual machine is managed by a software or firmware, which is known as hypervisor.

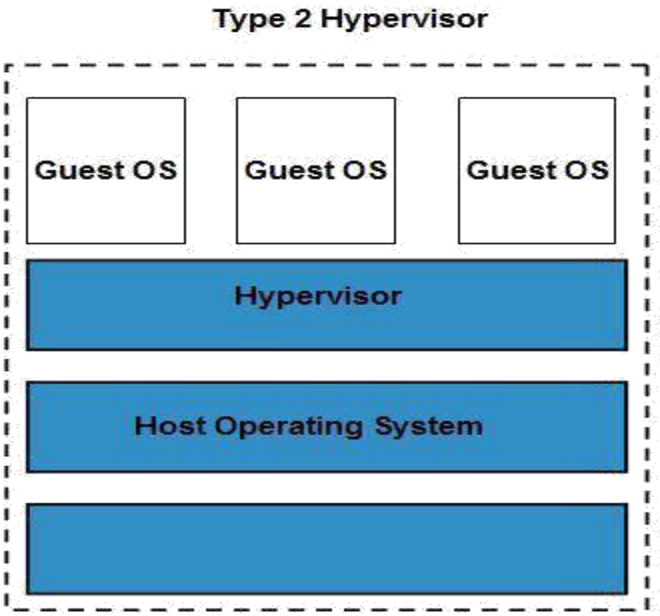
4.2.2 Hypervisor

Hypervisor is a firmware or low-level program that acts as a Virtual Machine Manager. There are two types of hypervisor:

Type 1 hypervisor runs on bare system. LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, Virtual Logic VLX are examples of Type 1 hypervisor. The following diagram shows the Type 1 hypervisor.



The type1 hypervisor does not have any host operating system because they are installed on a bare system. Type 2 hypervisor is a software interface that emulates the devices with which a system normally interacts. Containers, KVM, Microsoft Hyper V, VMWare Fusion, Virtual Server 2005 R2, Windows VirtualPC and VMWare workstation 6.0 are examples of Type 2 hypervisor. The following diagram shows the Type 2hypervisor



3

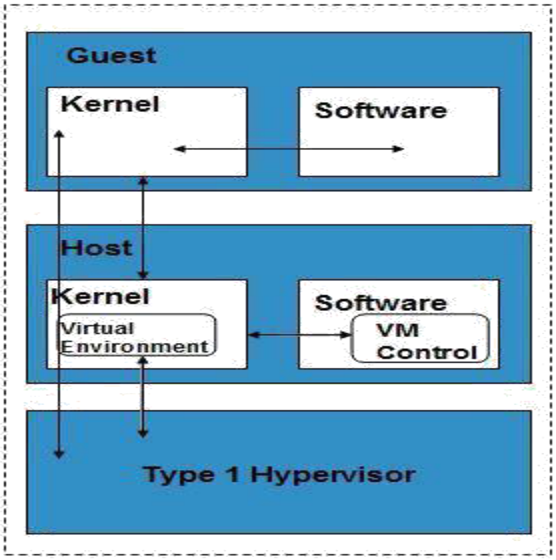
4.3 Types of Hardware Virtualization

Here are the three types of hardware virtualization:

1. Full Virtualization
2. Emulation Virtualization
3. Para virtualization

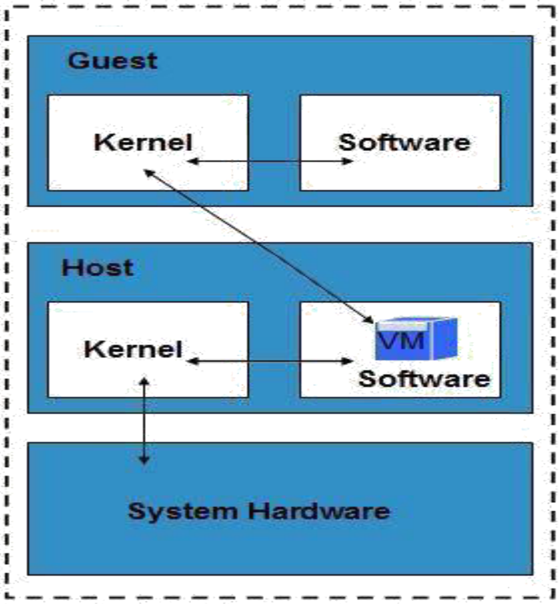
4.3.1 FULL VIRTUALIZATION

In Full Virtualization, the underlying hardware is completely simulated. Guest software does not require any modification to run.



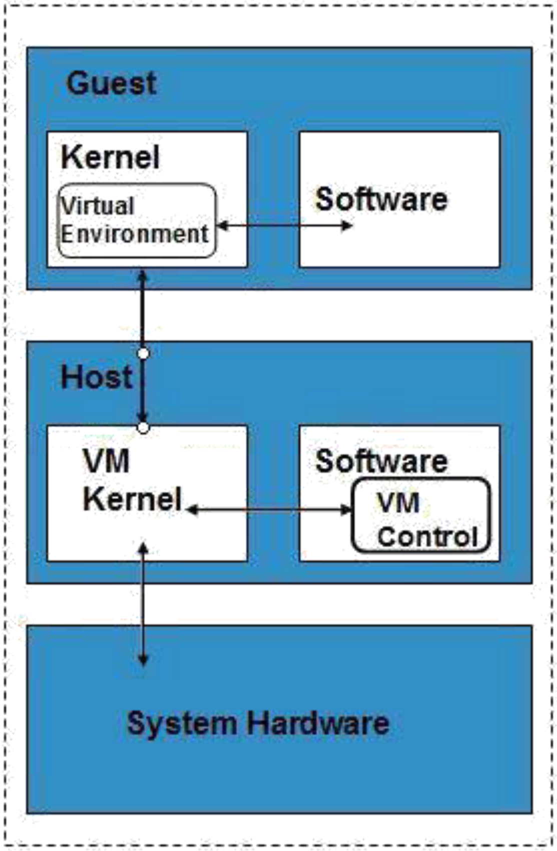
4.3.2 EMULATION VIRTUALIZATION

In Emulation, the virtual machine simulates the hardware and hence become independent of the it. In this, the guest operating system does not require modification.



4.3.3 PARAVIRTUALIZATION

In Para virtualization, the hardware is not simulated. The guest software run their own isolated domains



VMware vSphere is highly developed infrastructure that offers a management infrastructure framework for virtualization. It virtualizes the system, storage and networking hardware.

VERY SHORT QUESTIONS

1. What is virtualization?
2. Name type of virtualization?
3. Define full virtualization?
4. What is hypervisor?
5. Name the types of hypervisor?

SHORT QUESTIONS

1. Define full virtualization along with diagram?
2. Define emulation virtualization along with diagram?

3. What is type 1 hypervisor?
4. What is para virtualization?

LONG QUESTIONS

1. What are types of hypervisor?
2. What is virtualization and its types?

UNIT 5 CLOUD SEURITY

5.1 Learning Objectives

To know about cloud security.
To elaborate data security and infrastructure security.
To know legal issues.

5.2 Cloud Security

Migrating to a cloud computing platform means your responsibility for data security goes up considerably. Data with various levels of sensitivity is moving out of the confines of your firewall. You no longer have control – your data could reside anywhere in the world, depending on which cloud company you use.

Moving to the public cloud or using a hybrid cloud means the potential for cloud security issues is everywhere along the chain. It can happen as the data is prepped for migration, during migration, or potentially within the cloud after the data arrives. And you need to be prepared to address this every step of the way.

Data security has been incumbent on the cloud service providers, and they have risen to the occasion. No matter which platform you select in the debate between AWS vs. Azure vs. Google, all sport various compliances to standards like HIPAA, ISO, PCI DSS, and SOC.

However, just because the providers offer compliance doesn't give customers the right to abdicate their responsibilities. They have some measure of responsibility as well, which creates a significant cloud computing challenge. So here are eight critical concepts for data security in the cloud.

5.3 Data Security

- Privacy Protection

Your data should be protected from unauthorized access regardless of your cloud decisions, which includes data encryption and controlling who sees and can access what. There may also situations where you want to make data available to certain personnel under certain circumstances. For example, developers need live data for testing apps but they don't necessarily need to see the data, so you would use a redaction solution. Oracle, for example, has a Data Redact tool for its databases.

The first step is something you should have done already: identify the sensitive data types and define them. Discover where the sensitive data resides, classify and define the data types, and create policies based on where the data is and which data types can go into the cloud and which cannot. Too many early adopters of the cloud rushed to move all their

- Preserve Data Integrity

Data integrity can be defined as protecting data from unauthorized modification or deletion. This is easy in a single database, because there is only one way in or out of the database, which you can control. But in the cloud, especially a multicloud environment, it gets tricky.

Because of the large number of data sources and means to access, authorization becomes crucial in assuring that only authorized entities can interact with data. This means stricter means of access, like two-factor authorization, and logging to see who accessed what. Another potential means of security is a trusted platform module (TPM) for remote data checks.

- Data Availability

Downtime is a fact of life and all you can do is minimize the impact. That’s of considerable importance with cloud storage providers because your data is on someone else’s servers. This is where the service-level agreement (SLA) is vital and paying a close eye to details really matters.

For example, Microsoft offers 99.9% availability for major Azure storage options but AWS offers 99.99% availability for stored objects. This difference is not trivial. Also, make sure your SLA allows you to specify where the data is stored. Some providers, like AWS, allow you to dictate in what region data is stored. This can be important for issues of compliance and response time/latency.

Every cloud storage service has a particular strength: Amazon’s Glacier is ideal for mass cold storage of rarely accessed data, Microsoft's Azure blob storage is ideal for most unstructured data, while Google Cloud's SQL is tuned for MySQL databases.

- Data Privacy

A huge raft of privacy laws, national and international, have forced more than a few companies to say no to the cloud because they can’t make heads or tails of the law or it’s too burdensome. And it’s not hard to see why.

Many providers may store data on servers not physically located in a region as the data owner and the laws may be different. This is a problem for firms under strict data residency laws. Not to mention that the cloud service provider will likely absolve themselves of any responsibility in the SLA. That leaves the customers with full liability in the event of a breach.

As said above, there are national and international data residency laws that are your responsibility to know. In the U.S. that includes the Health Information Portability and Accountability Act (HIPAA), The Payment Card Industry Data Security Standards (PCI DSS), the International Traffic in Arms Regulations (ITAR) and the Health Information Technology for Economic and Clinical Health Act (HITECH).

In Europe you have the very burdensome General Data Protection Regulation (GDPR) with its wide ranging rules and stiff penalties, plus many European Union (EU) countries now that dictate that sensitive or private information may not leave the physical boundaries of the country or region from which they originate. There are also the United Kingdom Data Protection Law, the Swiss Federal Act on Data Protection, the Russian Data Privacy Law and the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA).

All of these protect the interest of the data owner, so it is in your best interest to know them and know how well your provider complies with them.

- Encryption

Encryption is the means for which data privacy is protected and insured, and encryption technologies are fairly mature. Encryption is done via key-based algorithms and the keys are stored by the cloud provider, although some business-related apps, like Salesforce and Dynamix, use tokenization instead of keys. This involves substituting specific token fields for anonymous data tokens.

Virtually every cloud storage provider encrypts the data while it is in transfer. Most do it through browser interfaces, although there are some cloud storage providers like Mega and SpiderOak that use a dedicated client to perform the encryption. This should all be spelled out in the SLA.

Many cloud services offer key management solutions that allow you to control access because the encryption keys are in your hands. This may prove to be a better or at least more reassuring risk because you are in control over who has the keys. Again, this should be spelled out in the SLA.

- Threats

If you are online you are under threat of attack, that is a fact of life. The old style of attacks, like DDoS attacks, SQL injection, and cross-site scripting, have faded into new fears. Cloud service providers have a variety of security tools and policies in place but problems still happen, usually originating in human error.

- **Data breaches:** This can happen any number of ways, from the usual means – a hacked account or a lost password/laptop – to means unique to the cloud. For example, it is possible for a user on one virtual machine to listen for the signal that an encryption key has arrived on another VM on the same host. It's called the "side channel timing exposure," and it means the victim’s security credentials in the hands of someone else.

- **Data loss:** While the chance of data loss is minimal short of someone logging in and erasing everything, it is possible. You can mitigate this by insuring your applications and data are distributed across several zones and you backup your data using off-site storage.

- **Hijacked accounts:** All it takes is one lost notebook for someone to get into your cloud provider. Secure, tough passwords and two-factor authentication can prevent this. It also helps to have policies that look for and alert to unusual activity, like copying mass amounts of data or deleting it.

  - **Cryptojacking:** Cryptojacking is the act of surreptitiously taking over a computer to farm cryptocurrency, which is a very compute-intensive process. Cryptojacking spiked in 2017 and 2018 and the cloud was a popular target because there is more compute resources available. Monitoring for unusual compute activity is the key way to stop this.

- Data Security and Staff
  - Most employee-related incidents are not malicious. According to the Ponemon Institute's *2016 Cost of Insider Threats Study*, 598 of the 874 insider related incidents in 2016 were caused by careless employees or contractors.
  - However, it also found 85 incidents due to imposters stealing credentials and 191 were by malicious employees and criminals. Bottom line: your greatest threat is inside your walls. Do you know your employees well enough?
- Contractual Data Security

The SLA should include a description of the services to be provided and their expected levels of service and reliability, along with a definition of the metrics by which the services are measured, the obligations and responsibilities of each party, remedies or penalties for failure to meet those metrics, and rules for how to add or remove metrics.

Don't just sign your SLA. Read it, have a lot of people read it, including in-house attorneys. Cloud service providers are not your friend and are not going to fall on their sword for liability. There are multiple checkmarks for a SLA.

- Specifics of services provided, such as uptime and response to failure.

- Definitions of measurement standards and methods, reporting processes, and a resolution process.
- An indemnification clause protecting the customer from third-party litigation resulting from a service level breach.

This last point is crucial because it means the service provider agrees to indemnify the customer company for any breaches, so the service provider is on the hook for any third-party litigation costs resulting from a breach. This gives the provider a major incentive to hold up their end of the security bargain.

5.4 Infrastructure Security

IaaS application providers treat the applications within the customer virtual instance as a black box and therefore are completely in different to the opera-tions and management of a applications of the customer [13]. The entire pack(customer application and run time application) is run on the customers' server on provider infrastructure and is managed by customers themselves. For this reason it is important to note that the customer must take full responsibility for securing their cloud deployed applications [7], [8], [12].

- Cloud deployed applications must be designed for the internet threat model.

- They must be designed with standard security countermeasures to guard against the common web vulnerabilities.

- Customers are responsible for keeping their applications up to date - and must therefore ensure they have a patch strategy to ensure their applications are screened from malware and hackers scanning for vulnerabilities to gain unauthorized access to their data within the cloud.

- Customers should not be tempted to use custom implementations of Au-thentication, Authorization and Accounting as these can become weak if not properly implemented.

The foundational infrastructure for a cloud must be inherently secure whether it is a private or public cloud or whether the service is SAAS, PAAS or IAAS. It will require [7], [9]:

- Inherent component-level security: The cloud needs to be architected to be secure, built with inherently secure components, deployed and provisioned securely with strong interfaces to other components and supported securely, with vulnerability-assessment and change-management processes that pro-duce management information and service-level assurances that build trust.

- Stronger interface security: The points in the system where interaction takes place (user-to-network, server-to application) require stronger security poli-cies and controls that ensure consistency and accountability.

- Resource lifecycle management: The economics of cloud computing are based on multi-tenancy and the sharing of resources. As the needs of the customers and requirements will change, a service provider must provision and decom-mission correspondingly those resources - bandwidth, servers, storage and security. This lifecycle process must be managed in order to build trust.

The infrastructure security can be viewed, assessed and implemented according its building levels - the network, host and application levels

5.4.1Infrastructure Security – The Network Level

When looking at the network level of infrastructure security, it is important to distinguish between public clouds and private clouds. important to distinguish between public clouds and private clouds. With private clouds, there are no new attacks, vulnerabilities, or changes in risk specific to this topology that infor-mation security personnel need to consider. If public cloud services are chosen, changing security requirements will require changes to the network topology and the manner in which the existing network topology interacts with the cloud provider's network topology should be taken into account . There are four significant risk factors in this use case:

Ensuring the confidentiality and integrity of organization's data-in-transit to and from a public cloud provider;

- Ensuring proper access control (authentication, authorization, and auditing) to whatever resources are used at the public cloud provider;
- Ensuring the availability of the Internet-facing resources in a public cloud that are being used by an organization, or have been assigned to an organi-zation by public cloud providers;
- Replacing the established model of network zones and tiers with domains.

5.4.2Infrastructure Security – The Host Level

When reviewing host security and assessing risks, the context of cloud services delivery models (SaaS, PaaS, and IaaS) and deployment models public, private, and hybrid) should be considered [7]. The host security responsibilities in SaaS and PaaS services are transferred to the provider of cloud services. IaaS cus-tomers are primarily responsible for securing the hosts provisioned in the cloud (virtualization software security, customer guest OS or virtual server security).

5.4.3Infrastructure Security – The Application Level

Application or software security should be a critical element of a security pro-gram. Most enterprises with information security programs have yet to institute an application security program to address this realm. Designing and implement-ing applications aims at deployment on a cloud platform will require existing ap-plication security programs to reevaluate current practices and standards. The application security spectrum ranges from standalone single-user applications to sophisticated multiuser e-commerce applications used by many users. The level is responsible for managing [7], [9], [10]:

- Application-level security threats;
- End user security;
- SaaS application security;
- PaaS application security;
- Customer-deployed application security
- IaaS application security
- Public cloud security limitations

It can be summarized that the issues of infrastructure security and cloud com-puting lie in the area of definition and provision of security specified aspects each party delivers.

5.5 Legal Issues

While far from an exhaustive list, some of the key legal issues that need to be agreed upon by the customer and the service provider are as follows:

- Governing Law and Jurisdiction - Virtually without exception, the service provider will outline that it is liable and governed within its own country, and that all disputes that arise from the contract are under the jurisdiction of the courts of the service provider's country. Many customers may want to have this amended to move any legal jurisdiction to their home country and in some cases when he service provider is a large multi-national, this may be possible. It may also be possible to remove such a provision from a contract and allow legal debate to decide, when or if a situation should arise
- Data Location - Many service provider contracts explicitly outline the right to maintain customer data on any of their sites, regardless of the origin of the data. While some service providers do not address the issue directly, most follow a similar policy on the grounds that not explicitly prohibiting the practice legitimizes it. Although maintaining data across multiple geographical locations provides a greater level of security, it does raise issues in relation to export control and needs to be addressed directly within the contract, legislating against extraterritorial storage
- Privacy and Confidentiality - In many cases, data collected for a specific purpose may only be used for that specific purpose. For example, student information stored in college databases typically may only be outsourced to designated vendors with legitimate interests in the data. Contracts governing data outsourcing need to ensure data usage specifically for the required service, and non-disclosure of data by the third party without authorization. Without being expressed explicitly within a contract, enforcement may be compromised
- Data Security - In the case of contracts that address data security, most limit their provision to a "reasonable" level of security, or to implement "industry standard" security practices. Despite providing a level of confidence in the service provision, these terms are widely open to argument and interpretation. To ensure a greater level of security, this needs to be replaced by independent specific security standards, and updated and audited periodically. Also, any contract should place a requirement on the service provider to give notice of data or security breaches
- Data Access for E-Discovery - While it is not an absolute necessity, an understanding of the architecture of the service being provided is important. In order to prepare for any e-discovery requirements that could arise, knowledge of the format used for data storage and available tools for data access is required. Some services fail to provide such tools, turning e-discovery into a cumbersome and time consuming task
- End User Responsibility - Service provider contracts may require the customer to ensure that the end users of the service abide by the service providers usage terms and conditions. While this is an understandable condition on the service provider - customer relationship, it also places the liability of the third party usage of the system with the customer. An alternative would be to enforce agreement between third parties and the service provider
- Inappropriate and Unauthorized Usage - Some service providers may place the responsibility of preventing inappropriate and unauthorized usage of the provided service with the customer. Considering that the service provided resides in the cloud, and is by and large outside of the control of the customer, it is recommended that the contract limits the

liability to the customer not authorizing or knowingly allowing prohibited usage of the service.

- End User Account Suspension - Occasionally service providers may specify the right to suspend the accounts of an end user on the violation of the service provider’s terms and conditions. With a broad statement of right, service providers can suspend the customer’s end users at will. It is preferable for the customer to restrict the service provider’s right of suspension to material or significant violations that compromise the security of the vendors system
- Emergency Security Issues - Service providers may have legislation inserted to suspend – without notice – a service provision, should an offending use of the service cause an emergency issue. It is in the best interest of the customer to clearly define the constitution of an emergency issue, thus limiting the flexibility or discretion of the service provider, and ideally should only incorporate a significant violation of the service provider’s terms and conditions
- Service Suspension and Termination - Typically service providers reserve the right to suspend a service, or to even terminate a service, in the event of specified events. Although such conditions are practical and legitimate for the service provider’s point of view, they too need to be limited to a strict set of events without any ambiguity. Such clauses need to provide the customer with an opportunity to remedy the situation, rather than an instant denial of service (with the exception of extreme emergencies), and to provide the customer to make alternative arrangements for service provision. It is also essential that, in the occurrence of such an event, the customer's data is available in a usable format for a specified amount of time after service termination. Finally, the service provider needs to be obliged to return or destroy any customer data once the service termination is complete
- Data Ownership - It is essential that the contract between the service provider and the customer explicitly states that all data is the property of the customer, and that the service provider does not acquire any licenses or rights to the customer’s data based on the transaction. The restriction of any security interest in the customer’s data by the service provider should also be noted
- Publicity - Occasionally, the service provider may be permitted to the use of the customer’s name, trademarks or logos for the service providers own publicity. If such stipulations cannot be removed, a modification should take place that requires the customers approval for any use of the customer brand, or at the very least to limit the use to the customer name without implying an endorsement
- Service Level Agreements - Guarantees for the service provision need to be detailed to provide for the minimum amount of uptime, the process, and the timescale associated with correcting downtime. Consequences for falling outside the agreed SLAs need to be precise and detailed
- Disclaimer of Warranty - Typically, a service provider contract will disclaim all warranties, occasionally explicitly including any guarantee that the service providers offering is not in breach of the intellectual property rights of a third party. As a minimum requirement, the contract should guarantee that the provided service functions according to its specifications, and that it is not in breach of the rights of any third party. In the absence of such warranties, an enforceable assurance of the service functionalities is not possible,

or that the service provider even has the authority to provide the service. In the event of service failure, or liable action being taken against the

- Customer Indemnification - Some service provider contracts require indemnification for the service provider in the event of illicit third party actions, along with customer actions. While this does not constitute adopting an extra liability as the customer may face legal action over third party content, it is in the best interest of the customer to avoid accepting this liability voluntarily
- Vendor Indemnification - It is rare for service provider contracts to outline any indemnification that benefits the customer, despite legal protection being essential in a minimum of two scenarios - third party intellectual property rights infringement and a breach or unauthorized disclosure of sensitive customer data. In both scenarios, the responsibility lies solely with the service provider, and defending or remedying either situation can prove extremely costly. By refusing to accept liability in either scenario, the service provider is displaying a lack of confidence in their provision, and careful consideration needs to be taken by the customer before making a decision to adopt the service
- Contract Modifications - In many cases, the service provider will reserve the right to modify their services as they deem appropriate. Given the nature of the industry, such modification rights are necessary to provide upgrades and patches to services. However, specifying the rights in a vague manner once again exposes the customer to the possibility of a deterioration of the service provided. It is within the customer's interests to limit such modifications to commercially reasonable ones that are not materially detrimental to the service provided

VERY SHORT QUESTIONS

1. What is cloud security?
2. What does privacy protection mean?
3. What does data integrity mean?
4. What does encryption mean?
5. What is cryptojacking?

SHORT QUESTIONS

1. What is data secutiy?
2. What is infrastructure security?

LONG QUESTIONS

1. What are the legal issues in cloud security??

UNIT 6 STORAGE AS A SERVICE

6.1 Learning objectives

- To know about SAAS
- To discuss benefits of cloud storage
- To elaborate challenges of cloud storage
- To discuss SAN

6.2 What is SAAS(Storage as a Service)

Storage as a service (SaaS) is a business model in which a company leases or rents its storage infrastructure to another company or individuals to store data. Small companies and individuals often find this to be a convenient methodology for managing backups, and providing cost savings in personnel, hardware and physical space.

6.3 Benefits of Storage as a service

Storing data in the cloud lets IT departments transform three areas:

1. Total Cost of Ownership. With cloud storage, there is no hardware to purchase, storage to provision, or capital being used for "someday" scenarios. You can add or remove capacity on demand, quickly change performance and retention characteristics, and only pay for storage that you actually use. Less frequently accessed data can even be automatically moved to lower cost tiers in accordance with auditable rules, driving economies of scale.

2. Time to Deployment. When development teams are ready to execute, infrastructure should never slow them down. Cloud storage allows IT to quickly deliver the exact amount of storage needed, right when it's needed. This allows IT to focus on solving complex application problems instead of having to manage storage systems.

3. Information Management. Centralizing storage in the cloud creates a tremendous leverage point for new use cases. By using cloud storage lifecycle management policies, you can perform powerful information management tasks including automated tiering or locking down data in support of compliance requirements.

6.3 challenges of cloud storage

- *Cloud storage problem #1: Not choosing the right cloud storage provider*

The old adage is that no-one got fired for choosing IBM, and when it comes to cloud storage it's tempting to choose one of the two biggest cloud providers: AWS or Microsoft Azure.

But while they may well be the best choice for many companies, they may not be the best choice for yours. Depending on the size of your organizations it may make sense to look at smaller storage providers who will be able to give you more attention.

The things to look for with other storage providers include:

- Downtime history, to get an idea of how reliable they have been in the past – and therefore an indication of how reliable they may be in the future.
- Data accessibility, including what bandwidth they have within their data center, between their data centers and to the Internet.
- Their pricing structure, including fixed charges and bandwidth charges to move data in and out. A common cloud storage problem is to neglect to establish how easily you can scale your requirements up and down. For example, are you committed to a certain amount of storage every month, or can pay only for what you use each day, week or month?

Familiarity with your industry vertical. Choosing a storage provider that understands your business and your likely data requirements can make life much easier for you, and failing to choose a good provider that specializes in your industry is a clouds storage pitfall that could put you at a disadvantage compared to your competitors. That's because service providers familiar with your industry may be better equipped to accommodate your industry's usage patterns and performance requirements and to demonstrate compliance with relevant industry regulations.

- *Cloud storage problem #2: Neglecting connectivity*

You may have a state of the art network in your data center running at 100Gbps or 10Gbps, with perhaps 10Gbps, 1Gbps or even 100Mbps in the rest of the organization. But when it comes to connectivity with the Internet your bandwidth will likely be much slower – perhaps as low as 10Mbps – and it may well be asymmetric (meaning uploads to a cloud storage provider will be much slower than downloads from it.)

Cloud storage gateways and other WAN optimization appliances can help alleviate the problem, but if the connectivity to your cloud storage provider is not sufficient then a move to cloud storage is unlikely to enable high enough storage performance to get many of the potential benefits.

- *Cloud storage problem #3: Not getting the service level agreement (SLA) right*

Most cloud storage providers will offer you a boilerplate SLA outlining their obligations to you and what they will do if things go wrong. But there is no reason why you have to accept it – IDC estimates that about 80% of cloud customers accept the boilerplate SLA they are offered, but 20% negotiate alterations to this boilerplate to ensure that it more closely meets their needs.

For example, a provider may offer you "four nines" (i.e 99.99%) uptime guarantee, allowing 50 minutes downtime per year. But this may be calculated on an annual basis, so the service could be down for 50 minutes on the first day of the contract and you would have to wait until the end of the year to find out if the SLA had been breached and you were therefore entitled to any compensation.

In the meantime you would have to bear any resultant losses yourself. To avoid this cloud storage problem it may be possible to negotiate that while 50 minutes per year is permissible, there should be no more than (say) 15 minutes per month if that suits your business needs better.

- *Cloud storage problem #4: Overestimating the compensation you might get if the provider* breaches the SLA

It's tempting to think of an SLA as some kind of insurance policy: your business can survive as long as the terms of the SLA are met, and if they are not you'll be OK because your cloud storage provider will provide compensation that is tied to the impact on your business of the breach.

But that is simply not the case and it's a common cloud storage problem. In most cases breach penalties come in the form of service credits (i.e. free storage for a few months), and in the case of a serious breach – such as all your data being lost – the most you should hope for is a monetary payment of three or four times your annual contract value. In many cases that will be nothing like the cost to your business of losing so much data.

Of course it may be possible to negotiate higher compensation payments from your cloud storage provider, but then it's likely you will have to pay much more for your storage. In most cases it would work out cheaper to buy insurance cover from a third party.

- *Cloud storage problem #5: Failing to monitor your SLA effectively*

Working with a cloud storage provider adds another layer of complexity between the business users who use corporate data and the data itself. The IT department, which monitors the SLA, is somewhere in the middle.

A common cloud storage pitfall when it comes to data access problems is that users or business units may bypass the IT department and go directly to the cloud storage provider's help desk to resolve issues when they occur. If that happens then you can't necessarily rely on the provider to record every problem that occurs, and that means accurate monitoring of the SLA is effectively impossible. Avoiding this cloud storage pitfall comes down to educating users that your IT helpdesk should be their first point of contact in all cases.

- *Cloud storage problem #6: Failing to get a clear understanding of how to get your data back or move it to another provider*

Cloud storage providers may fall over themselves to make it easy for you to give them your data in the first place – perhaps by collecting physical media such as hard disk drives from your data center or offering free data ingress over a network connection. But if you decide that you no longer want to use the provider's services it can often prove unexpectedly difficult or expensive to get it back.

To avoid this cloud storage pitfall it's important to get satisfactory answers to the following questions:

* How will your data be made available – over a network connection or can it be placed on physical storage media for collection?

* How soon will it be available – will you be expected to wait for days or weeks?

* How much bandwidth will be available if you plan to download your data? That's important because even with a 1Gbps link, it would take almost two weeks to get 150TB of data back from a cloud storage provider to your data center.

* What bandwidth costs will be involved if you move your data back over a network, and what are the costs of having it put on physical media?

* How long will it take for copies and backups of your data to be deleted, and what formal confirmation can you expect that all copies have been deleted?

* In what format will data be made available – will it be provided in a .csv file or in some other more closed format?

- *Cloud storage problem #7: Assuming that using a cloud storage provider absolves you of all* security responsibilities

Cloud providers are meant to be experts at what they do, including keeping their clouds and the data within it secure. But if there is a data security breach then it is you that your customers will hold responsible and seek compensation from, and it is you that will suffer the embarrassment, loss of reputation and possible loss of business.

That means that to avoid this cloud storage problem it is up to you to do due diligence and satisfy yourself that the security offered by the cloud storage provider is good enough for your needs. To do this you will need to find out as much as possible about the security arrangements that are in place, and what guidelines and regulations (think HIPPA, PCI-DSS, SSAE 16) it has been certified to comply with.

- *Cloud storage problem #8: Fixating on costs without considering other factors*

For many companies one of the key drivers for moving to the cloud is reduced costs, or at the very least a switch from a single large capital expenditure to small regular operating expenditures. While that may be beneficial from a business point of view it's important to remember that as well as changing how you pay, you are also paying for something fundamentally different.

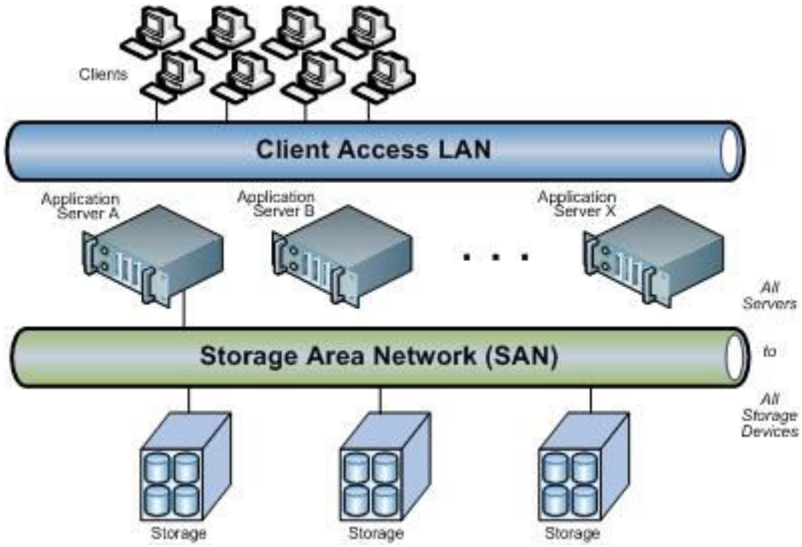
Cloud storage, in other words, is not the same as your existing data center storage, and as well as new security, compliance and accessibility challenges there are also new performance characteristics to consider. What this boils down to is that some applications that you run in your data center aren't performance sensitive and are well suited to being used in conjunction with cloud storage. For other applications that's not the case.

That means that if you decide to use cloud storage for these latter applications then the applications themselves may also have to run in the cloud, close to the cloud storage. And that in turn means that moving your data to cloud storage may need to be part of a far larger consideration of the viability of moving some or all of your applications to the cloud.

6.4 Storage area network

A Storage Area Network (SAN) is a specialized, high-speed network that provides block-level network access to storage. SANs are typically composed of hosts, switches, storage elements, and storage devices that are interconnected using a variety of technologies, topologies, and protocols. SANs may also span multiple sites.

A SAN presents storage devices to a host such that the storage appears to be locally attached. This simplified presentation of storage to a host is accomplished through the use of different types of virtualization.



SANs are often used to:

- Improve application availability (e.g., multiple data paths)
- Enhance application performance (e.g., off-load storage functions, segregate networks, etc.)
- Increase storage utilization and effectiveness (e.g., consolidate storage resources, provide tiered storage, etc.), and improve data protection and security.
- SANs also typically play an important role in an organization's Business Continuity Management (BCM) activities.

6.4.1 Types of SAN

SAN solutions are available as two types:

- Fiber Channel (FC): Storage and servers are connected via a high-speed network of interconnected fiber channel switches. This is used for mission-critical applications where uninterrupted data access is required.
- Internet Small Computer System Interface (iSCSI) Protocol: This infrastructure gives the flexibility of a low-cost IP network.

6.4.2 Advantages of SAN

The advantages of SAN include:

- Storage Virtualization: Server capacity is no longer linked to single storage devices, as large and consolidated storage pools are now available for software applications.
- High-Speed Disk Technologies: An example is FC, which offers data retrieval speeds that exceed 5 Gbps. Storage-to-storage data transfer is also available via direct data transmission from the source to the target device with minimal or no server intervention.
- Centralized Backup: Servers view stored data on local disks, rather than multiple disk and server connections. Advanced backup features, such as block level and incremental backups, streamline IT system administrator responsibilities.
- Dynamic Failover Protection: Provides continuous network operation, even if a server fails or goes offline for maintenance, which enables built-in redundancy and automatic traffic rerouting.

VERY SHORT QUESTIONS

1. What does SAAS stand for?
2. What does SAN stand for?
3. What does ISCSI stand for?
4. What are the two advantages of SAN?
5. Name the types of SAN.

SHORT QUESTIONS

1. What are the benefits of storage as a service?
2. What are the challenges faced by storage as a service?
3. What are the advantages of SAN?
4. What is storage area networks?

LONG QUESTIONS

1. What are the challenges faced by cloud storage? Explain.

UNIT 7 SCHEDULING IN CLOUD

7.1 Learning objectives

- To know about cloud scheduling.
- Elaborate different cloud scheduling algorithms.
- Issues in cloud scheduling.

7.2 Cloud scheduling

Cloud computing is known as a provider of dynamic services using very large scalable and virtualized resources over the Internet. Various definitions and interpretations of “clouds” and / or “cloud computing” exist. With particular respect to the various usage scopes the term is employed to, we will try to give a representative (as opposed to complete) set of definitions as recommendation towards future usage in the cloud computing related research space. We try to capture an abstract term in a way that best represents the technological aspects and issues related to it. In its broadest form, we can define a 'cloud' is an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality of service. To be more specific, a cloud is a platform or infrastructure that enables execution of code (services, applications etc.), in a managed and elastic fashion, whereas “managed” means that reliability according to pre defined quality parameters is automatically ensured and “elastic” implies that the resources are put to use according to actual current requirements observing overarching requirement definitions – implicitly, elasticity includes both up- and downward scalability of resources and data, but also load-balancing of data throughput.

Job scheduling is one of the major activities performed in all the computing environments. Cloud computing is one the upcoming latest technology which is developing drastically. To efficiently increase the working of cloud computing environments, job scheduling is one the tasks performed in order to gain maximum profit. The goal of scheduling algorithms in distributed systems is spreading the load on processors and maximizing their utilization while minimizing the total task execution time Job scheduling, one of the most famous optimization problems, plays a key role to improve flexible and reliable systems. The main purpose is to schedule jobs to the adaptable resources in accordance with adaptable time, which involves finding out a proper sequence in which jobs can be executed under transaction logic constraints. There are main two categories of scheduling algorithm. 1) Static scheduling algorithm and 2) Dynamic scheduling algorithm. Both have their own advantage and limitation. Dynamic scheduling algorithm has higher performance than static algorithm but has a lot of overhead compare to it.

7.3 Different Types of Scheduling

There has been various types of scheduling algorithm exist in distributed computing system. Most of them can be applied in the cloud environment with suitable verifications. The main advantage of job scheduling algorithm is to achieve a high performance computing and the best system throughput. Traditional job scheduling algorithms are not able to provide scheduling in the cloud environments. According to a simple classification, job scheduling algorithms in cloud computing can be categorized into two main groups; Batch Mode Heuristic scheduling Algorithms (BMHA) and online mode heuristic algorithms. In BMHA, Jobs are queued and collected into a set when they arrive in the system. The scheduling algorithm will start after a fixed period of time. The main examples of BMHA based algorithms are; First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Min–Min algorithm and Max–Min algorithm.

By On-line mode heuristic scheduling algorithm, Jobs are scheduled when they arrive in the system. Since the cloud environment is a heterogeneous system and the speed of each processor varies quickly, the on-line mode heuristic scheduling algorithms are more appropriate for a cloud environment. Most Fit Task scheduling algorithm is suitable example of On-line mode heuristic scheduling algorithm.

a. First Come First Serve Algorithm:

Job in the queue which comes first is served. This algorithm is simple and fast.

b. Round Robin Algorithm:

In the round robin scheduling, processes are dispatched in a FIFO manner but are given a limited amount of CPU time called a time-slice or a quantum. If a process does not complete before its CPU-time expires, the CPU is preempted and given to the next process waiting in a queue. The preempted process is then placed at the back of the ready list.

c. Min–Min Algorithm:

This algorithm chooses small tasks to be executed first, which in turn delays large tasks for long time.

d. Max–Min Algorithm:

This algorithm chooses large tasks to be executed first, which in turn delays small tasks for long time.

e. Most Fit Task Scheduling Algorithm:

In this algorithm task which fit best in queue are executed first. This algorithm has high failure ratio.

f. Priority Scheduling Algorithm:

The basic idea is straightforward: each process is assigned a priority, and priority is allowed to run. Equal-Priority processes are scheduled in FCFS order. The Shortest-Job-First (SJF) algorithm is a special case of general priority scheduling algorithm. An SJF algorithm is simply a priority algorithm where the priority is the inverse of the (predicted) next CPU burst. That is, the longer the CPU burst, the lower the priority and vice versa. Priority can be defined either internally or externally. Internally defined priorities use some measurable quantities or qualities to compute priority of a process.

Scheduling Process

Scheduling process in cloud can be generalized into three stages namely:

- Resource discovering and filtering: Data center Broker discovers the resources present in the network system and collects status information related to them.
- Resource selection: Target resource is selected based on certain parameters of task and resource. This is deciding stage.
- Task submission: Task is submitted to resource selected.

Existing Scheduling Algorithms

The following scheduling algorithms are currently prevalent in clouds.

- *Resource-Aware-Scheduling Algorithm (RASA)*

Saeed Parsa and Reza Entezari-Maleki [1] proposed a new task scheduling algorithm RASA. It is composed of two traditional scheduling algorithms; Max-min and Min-min. RASA uses the advantages of Max-min and Min-min algorithms and covers their disadvantages. Though the deadline of each task, arriving rate of the tasks, cost of the task execution on each of the resource, cost of the communication are not considered. The experimental results show that RASA is outperforms the existing scheduling algorithms in large scale distributed systems.

- *An Optimal Model for Priority based Service Scheduling Policy for Cloud Computing Environment*

Dr. M. Dakshayini, Dr. H. S. Guruprasad [2] proposed a new scheduling algorithm based on priority and admission control scheme. In this algorithm priority is assigned to each admitted queue. Admission of each queue is decided by calculating tolerable delay and service cost. Advantage of this algorithm is that this policy with the proposed cloud architecture has achieved very high (99%) service completion rate with guaranteed QoS. As this policy provides the highest precedence for highly paid user service-requests, overall servicing cost for the cloud also increases.

- *Extended Max-Min Scheduling Using Petri Net and Load Balancing*

El-Sayed T. El-kenawy, Ali Ibraheem El-Desoky, Mohamed F. Al-rahamawy [3] has proposed a new algorithm based on impact of RASA algorithm. Improved Max-min algorithm is based on the expected execution time instead of complete time as a selection basis. Petri nets are used to model the concurrent behaviour of distributed systems. Max-min demonstrates achieving schedules with comparable lower makespan rather than RASA and original Max-min.

- *Reliable Scheduling Distributed in Cloud computing (RSDC)*

Arash Ghorbannia Delavar, Mahdi Javanmard, Mehrdad Barzegar Shabestari and Marjan Khosravi Talebi [4] proposed a reliable scheduling algorithm in cloud computing environment. In this algorithm, major job is divided to sub jobs. In order to balance the jobs the request and acknowledge time are calculated separately. The scheduling of each job is done by calculating the request and acknowledges time in the form of a shared job. So that efficiency of the system is increased.

- *Improved Cost-Based Algorithm for Task Scheduling*

Mrs. S.Selvarani, Dr. G. Sudha Sadhasivam [5] proposed an improved cost-based scheduling algorithm for making efficient mapping of tasks to available resources in cloud. The improvisation of traditional activity based costing is proposed by new task scheduling strategy for cloud environment where there may be no relation between the overhead application base and the way that different tasks cause overhead cost of resources in cloud. This scheduling algorithm divides all user tasks depending on priority of each task into three different lists. This scheduling algorithm measures both resource cost and computation performance, it also Improves the computation/communication ratio.

- *An Optimistic Differentiated Job Scheduling System for Cloud Computing*

Shalmali Ambike, Dipti Bhansali, Jaee Kshirsagar, Juhi Bansiwal [6] has proposed a differentiated scheduling algorithm with non-preemptive priority queuing model for activities performed by cloud user in the cloud computing environment. In this approach one web application is created to do some activity like one of the file uploading and downloading then there is need of efficient job scheduling algorithm. The Qos requirements of the cloud computing user and the maximum profits of the cloud computing service provider are achieved with this algorithm.

- *A Priority based Job Scheduling Algorithm in Cloud Computing*

Shamsollah Ghanbari, Mohamed Othman proposed a new scheduling algorithm based on multi–criteria and multi-decision priority driven scheduling algorithm [7]. This scheduling algorithm consist of three level of scheduling: object level, attribute level and alternate level. In this algorithm priority can be set by job resource ratio. Then priority vector can be compared with each queue. This algorithm has higher throughput and less finish time.

- *Performance and Cost Evaluation of Gang Scheduling in a Cloud Computing System with Job Migrations and Starvation Handling*

Ioannis A. Moschakis and Helen D. Karatza has proposed a gang scheduling algorithm with job migration and starvation handling in which scheduling parallel jobs, already applied in the areas of Grid and Cluster computing. The number of Virtual Machines (VMs) available at any moment is dynamic and scales according to the demands of the jobs being serviced. The aforementioned model is studied through simulation in order to analyze the performance and overall cost of Gang Scheduling with migrations and starvation handling. Results highlight that this scheduling strategy can be effectively deployed on Clouds, and that cloud platforms can be viable for HPC or high performance enterprise applications.

7.4 Issues in cloud scheduling

- Execution time: In which program is running and single instruction, such as addition or multiplication is carried out in the computer instruction.
- Response time: The response time is the sum of the service time and wait time. Technically response time is the time of system takes to react to a given input.
  Make span Difference between the depart and close of a sequence of jobs.
- Energy Consumption: It is the consumption of energy or power. It is also defined in some quarters as the use of energy as a raw material in the process of manufacturing utilities.
- Throughput: It refers to how much data can be transferred from one location to another in a given amount of time.
- Scalability: The increasing demands and growing amount of the work is known as scalability.
- Resource utilization: Resource utilization is the use of a resource in such a way that increases through output. The sources used to perform a particular task.
- Load Balancing: Load balancing is the most above-board method of surmounting away an application host base. As application need growths, new hosts can be easily summed to the imagination pond and the load balancer will instantly start posting traffic to new host.

VERY SHORT QUESTIONS

1. What is First Come First Serve Algorithm?
2. What is Round Robin Algorithm?
3. What is Min–Min Algorithm?
4. What is Max–Min Algorithm?
5. What is cloud security?

SHORT QUESTIONS

1. What are the issues in cloud computing?

LONG QUESTIONS

2. What are the different scheduling algorithms in cloud computing?