**GOVERNMENT OF TAMILNADU**
**DIRECTORATE OF TECHNICAL EDUCATION**
**CHENNAI – 600 025**

## STATE PROJECT COORDINATION UNIT

**Diploma in** Computer Engineering

**Course Code:** 1052
**M – Scheme**

**e-TEXTBOOK**
**on**

# Cloud Computing

**for**
**V Semester** Dip. In Comp. Engg.


**Convener for Computer Engineering Discipline:**

**Mrs.A.Ghousia Jabeen,**
Principal,
Thanthai Periyar E.V.Ramasamy Govt. Polytechnic College for Women,
Vellore – 632 002

**Team Members :**

**Mr. B Krishnakumar**
Head of Dept. / Computer Engineering
228, Arasan Ganesan Polytechnic College
Sivakasi – 626 130.

**Mrs.P. Bhavani,**
Lecturer / Computer Engg.
Government Polytechnic College,
Purasaiwalkam
Chennai – 6000 12

**Mr.N. Jegan,**
Lecturer / Computer Engineering
228, Arasan Ganesan Polytechnic College
Sivakasi – 626 130.

**Validated By**
**Mrs. S.Sharmila**
HOD / IT
P.S.G. Polytechnic College
Coimbatore

# UNIT I

CLOUD COMPUTING BASICS

- To learn the concepts of cloud computing.
- To understand the essential characteristics of Cloud Computing.
- To study the architectural influences.
- To understand the benefits of cloud computing.
- To learn the concept of security concerns.

**Introduction:**

Cloud computing is the delivery of computing services like servers, storage, databases, networking, software etc over the Internet Companies offering these computing services are called cloud providers and typically charge for cloud computing services based on usage, similar to how you are billed for water or electricity at home.

## 1.1 Cloud computing overview

Cloud computing is the process of delivering or providing computational resources like software and/or hardware as a service over the cloud (internet).

The name "cloud computing" comes from the most common usage of cloud shaped symbol for internet.

It is also known as internet computing. With cloud computing the users can access the database resources through internet from anywhere in the world. The databases in the cloud are very dynamic.

### 1.1.1 Origin of cloud computing

- This concept of providing resources via global networks starts from 1960's.

- This idea was introduced byJ.C.R.Licklider, who was responsible for ARPANET in 1969.
- Since then the cloud computing has developed a lot. After the wide internet usage in 1990's, large bandwidth was offered. Thus this concept came to use by public.
- One of the first milestones for cloud-computing was the introduction of salesforce.com in 1999. It introduced the concept of delivering enterprise applications through a simple website. This helped software companies to deliver applications over the cloud.
- Amazon Web Services in 2002 was the next major development in cloud computing, which provided a group of cloud based services.
- Amazon launched its Elastic Compute Cloud (EC2) in 2002 as a commercial web service which allows individuals and small scale industries to hire/rent computers on which they can run their own applications.
- Another big milestone was the introduction of web2.0 in 2009. Web2.0 is a website that allows the users to interact with each other in a social media creator. Examples of Web2.0 include social networking websites, blogs etc.
- Google and others started to offer browser based enterprise applications, through services such as Google apps.
- Other important factors that have enabled cloud computing to evolve are virtualization technology, development of universal high-speed bandwidth and universal software standards.
- Most of the IT professionals use cloud computing as it offers increased storage space, high flexibility and very low cost.

Thus cloud computing has brought enormous benefits for users.

### 1.1.2 Cloud Components

There are three cloud components, they are

1. Clients

2. Data centers
3. Distributed servers

Each element has an explicit role in delivering a application which is cloud based.

**Clients**

Clients are the end user devices where the users interact with cloud to manage their information. They are usually computers and also laptops,notebook computers, tablets, mobile phones, PDAs etc. Clients usually filtered in three categories, they are

- Mobile clients.
- Thin clients.
- Thick clients.

**Mobile clients:**This refer to mobile devices including PDAs or smart phones like iPhone.

**Thin Clients:**This refer to computers that do not have internal hard drives. They allow the server do all the work, but then display the information on the screen.

**Thick clients:**This refers to regular computers, that uses a web browser like Chrome, Firefox, Internet Explorer to connect to the internet.

Now a days thin clients are becoming more popular due to the following reasons,

- Lower hardware costs
- Data security
- Less power consumption
- Easy to repair or replacement
- Less noise

**Data Centers**

The data center is the collection of numerous servers. It could be a large room with full of servers located anywhere in the world. The clients can access these servers through the cloud.

**Distributed Servers**

Servers are in geographically separate locations in the world. For the cloud subscriber, these servers act as if they are very near. This gives the service provider more flexibility in security and options.

For example,Amazon has servers all over the world. If there is a failure at one site, the service would still be accessed through another site.

### 1.1.3 Essential characteristics:

Nowadays the term cloud is often used but still confused by nontechnical crowd. Read about cloud. Now coming to essential characteristics of cloud computing there are five most essential characteristics, they are :

1) On-Demand self service
2) Broad network access
3) Location independent resource pooling
4) Rapid elasticity
5) Measured service

#### 1.1.3.1 On-Demand self service

It one of the essential characteristic of cloud that allows user to receive the services such as computing resources, server time, network storage automatically without direct interaction with the service provider.

The applications and resources can be assigned and removed within minutes using cloud catalogs. Some of the popular on demand self service providers are AWS (Amazon Web Services), Google, Microsoft, IBM, Salseforce.com.

#### 1.1.3.2 Broad network access

This is another essential aspect that is available over the network. They are accessed by using standard mechanisms in thick or thin client platforms.

#### 1.1.3.3 Location independent resource pooling

The service providers resources are pooled in order to serve multiple consumers. There is a sense of location independence as the customer has no control over location where the resources are provided. Consumers need not worry about how the cloud allocates the provided resources.

#### 1.1.3.4 Rapid elasticity

The definition of elasticity is the ability to scale the resources up and down as required. The storage on cloud seems to be unlimited for the client. The consumer can use as much as he needs at any time.

#### 1.1.3.5 Measured services

Another essential attribute is that the resources can be measured, controlled and reported. This provides transparency for both provider and consumer of the used service. Metering capability is used to control and optimize resource use.

## 1.2 Architectural influences

### 1.2.1 High Performance Computing

1. High-performance computing (HPC) is the use of super computers and parallel processing techniques for solving complex computational problems.
2. HPC technology focuses on developing parallel processing algorithms and systems by incorporating both administration and parallel computational techniques.

   High-performance computing is typically used for solving advanced

problems and performing research activities through computer modelling, simulation and analysis. HPC systems have the ability to deliver sustained performance through the concurrent use of computing resources.

3. The terms high-performance computing and supercomputing are sometimes used interchangeably. High-performance computing (HPC) evolved due to meet increasing demands for processing speed.
4. HPC brings together several technologies such as computer architecture, algorithms, programs and electronics, and system software under a single canopy to solve advanced problems effectively and quickly.A highly efficient HPC system requires a high-bandwidth, low-latency network to connect multiple nodes and clusters.

HPC technology is implemented in multidisciplinary areas including:

1. Geographical data
2. Scientific research
3. Oil and gas industry modelling
4. Electronic design automation
5. Climate modelling
6. Media and entertainment

### 1.2.2 Utility and enterprise grid computing

Utility computing is a service providing model in which a service provider makes computer resources and infrastructure management available to the customer whenever needed.

The consumer is charged for specific usage of the resources. It is an On-Demand computing method.

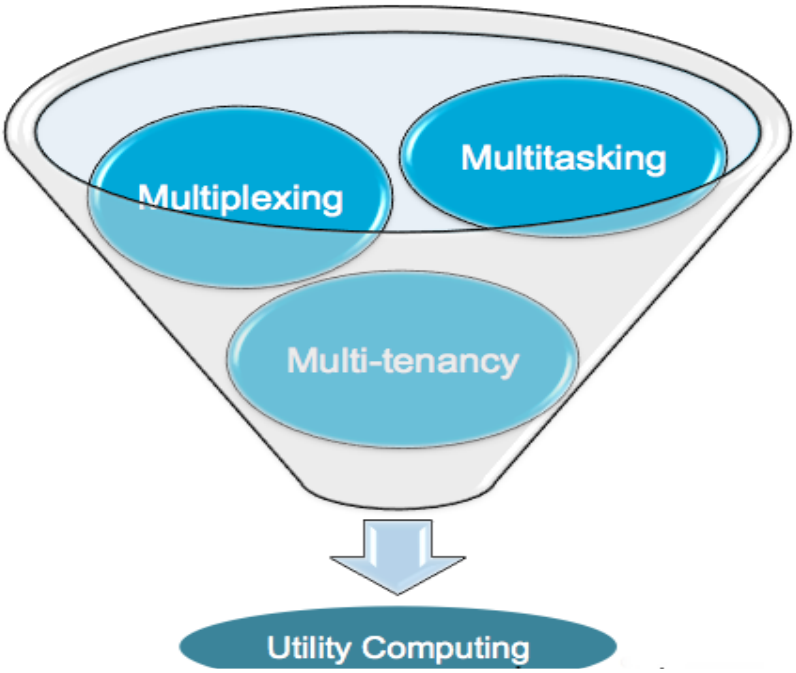



Fig 1.1 Utility Computing

Renting the computer resources such as hardware, network bandwidth, software to the customer on demand basis is called on demand computing. This type of approach is also called as Pay per use model or metered service.

The method used in utility computing is becoming popular in enterprise computing. It is sometimes used for consumer market, internet service, web site access, file sharing and other applications.The main advantage of this model is there is no initial cost or at-least low cost to get the required computer resources.

Another version of utility computing, which is carried out within an enterprise centralizes its computer resources to serve a large number of users. This will save time and money.

Early leaders in utility computing are IBM, HP, Microsoft and then industries like Google, Amazon etc took the lead in 2008, as they started to establish their own utility services for computing, storage and applications.

### 1.2.3 Enterprise Grid Computing

Enterprise-grid computing is a form of computing particularly inside an enterprise. It is a collection of network components under the control of a grid management entity, represented in Fig 1.2. This management entity manages the assignment of resources to services to meet the specific business goals.

Enterprise computing is used in single data center or multiple data centers. Enterprise- grid supports various types of workloads such as transactions.

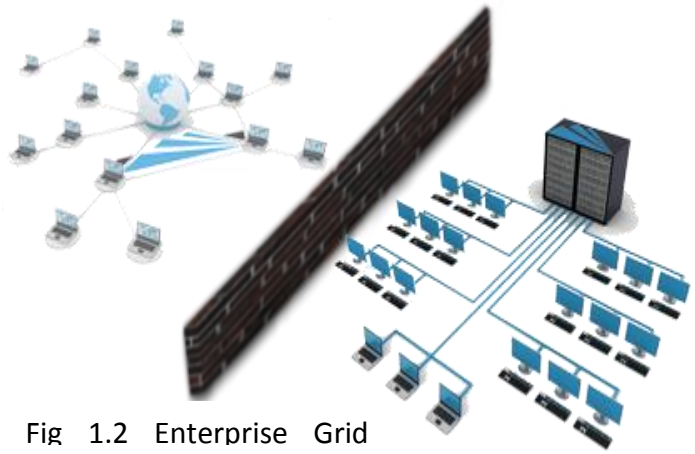


Fig 1.2 Enterprise Grid

These types of grids are used to meet the specific set of business goals.

The services that run on an enterprise-grid may range from the traditional commercial enterprise applications such as ERP to new concepts such as distributed applications. Enterprise computing enables the companies to the following

- To dynamically provide resources.
- To simplify tasks.
- To consolidate computing components.
- To set standards across enterprise.
- To scale the resources and workload.

Some of the advantages of enterprise-grid computing are,

- It reduces hardware, software costs
- It also reduces employee costs.
- It improves the quality of service through quick response time.

### 1.2.4 Autonomic computing

Autonomic computing refers to the self managing characteristics of distributed computing resources, adapting to unpredictable changes.

It controls the functioning computer applications and systems without input from the user. This computing model has systems that run themselves, capable of doing high level functions.

The complexity of the system is invisible to the users. The concept of autonomic computing was first introduced by IBM in 2001. This model aims to develop computer systems capable of self management. This overcomes the rapidly growing complexity of computing systems management.

These systems take decisions on their own by using high level policies.

It will constantly check and optimize the status. Thus it automatically adapts itself to changing conditions.

An autonomic computing framework is composed of autonomic components (AC) interacting with each other.

The following are the main components of autonomic computing:

- Two main control loops ( local and global )
- Sensors ( for self monitoring )
- Effectors ( for self adjustments )
- Knowledge
- Planner ( for analyzing policies ).

### 1.2.4 Service Consolidation

In computing, consolidation refers to when data storage or server resources are shared among multiple users and accessed by multiple applications.

Consolidation aims to make more efficient use of computer resources and prevent servers and storage equipment from being under-utilized and taking too much space.

The two main types of consolidation are server consolidation and storage consolidation.

**Server consolidation** involves reducing the number of servers and server locations within an organization. The intended result is more efficient use of server resources and occupied space. However, this also increases the complexity of the servers, data and applications, which may be challenging for users.

Server virtualization attempts to address this problem by masking that complexity from users.

Another option is to use blade servers, which are actual servers in the form of modular circuit boards on a card.

They occupy less rack space and consume less power.

**Storage consolidation**, or storage convergence, is a method of centralizing data storage through any one of three architectures:

- Network Attached Storage (NAS): Dedicated storage hard drives do not have to compete with other computers for processing resources.
- Redundant Array of Independent Disks (RAID): Data is located on multiple disks but appears as a single logical drive.
- Storage Area Network (SAN): Fiber channel technology is used to provide high throughput, data sharing, data migration and service to many

clients (also called subscribers) over a large geographical area. SANS is the most sophisticated storage consolidation method of the three.

### 1.2.5 Horizontal Scaling

Horizontal scaling is the capability of an operation to be scaled up to meet the demand through request and the distribution of request across the servers as in Fig 1.3 as demand increases the servers are scaled up. Scalability is the capacity of a model to be enlarged to handle the expanding volume of production in an effective way.

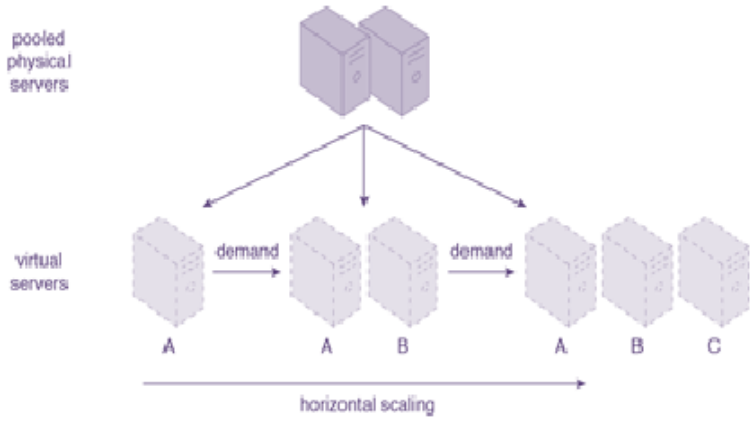



Fig 1.3 Horizontal Scaling

Methods of adding more resources for a particular application falls into two broad categories, they are

- Horizontal-scaling
- Vertical-scaling

This ( horizontal-scaling) is achieved by adding more servers by using clustering and load balancing solution.

Numerous number of computers may be configured in a cluster to obtain combined computing power. Horizontal scaling is also called as **scale out**.

An example of the scaling method is scaling out from one web server system to three servers. This type of scaling is used by the cloud computing as it

is very cheaper. In a cloud the load will be distributed across various additional servers.

This model created the demand in sharing data with enormous I/O performance and particularly for the processing of huge volume of data.

### 1.2.6 Web services

Web services are a set of services over the web or the technical term cloud. This a service which is “always on”, same as in the concept of utility computing. It is a standard way for integrating the web applications, It is considered as the next evolution of web. Web services converts your applications into Web application which can be published, found and used over the internet.

Web services communicate using open protocols and can be used by other applications.

These services are hardware independent, operating system independent, programming language independent.

The basic platform for these services are XML and HTTP.

Fig 1. 4 Web Services Management

The components of web services are web service server code, web service consumer code, SOAP, XML, WSDL and UDDI.SOAP is a protocol for accessing web service. It stands for **Simple Object Access Protocol**. SOAP is a communication protocol for sending messages.Fig 1.4

XML is a markup language. It stands for **eXtensible Markup Language.** The contents are encoded in xml codes. WSDL is an XML based language. It stands for **Web Services Description Language**. It is used to describe and locate the services.

UDDI is a directory service. It stands for **Universal Description, Discovery and Integration**. UDDI is used for storing the information about the web services.

### 1.2.6 High Scalability Architecture

Scalability is the ability of a model to be enlarged to handle the growing amount of work in an effective manner. The cloud resources can be rapidly scaled ( i.e., increased ) based on the demand which includes storage, CPU time, memory, web service requests.

Customers buy the services given by service provider based on scalability, availability and performance.

High scalability architecture is an important parameter to select a service provider. In cloud computing, it is important to have a High scalability architecture that is capable of handling numerous users on an on-demand basis.

To achieve this, Google introduced Bigtable which has high scalability architecture. Bigtable was developed with very high speed, flexibility and has very high scalability. A Bigtable database is petabytes(PB) in size and spans thousands of distributed servers.

- Each bigtable is a multi dimensional table.
- The table is made up of rows and columns.

- Each cell has a time stamp. With this stamp, you can select certain versions of a web page or delete cells that are older than a given date and time.

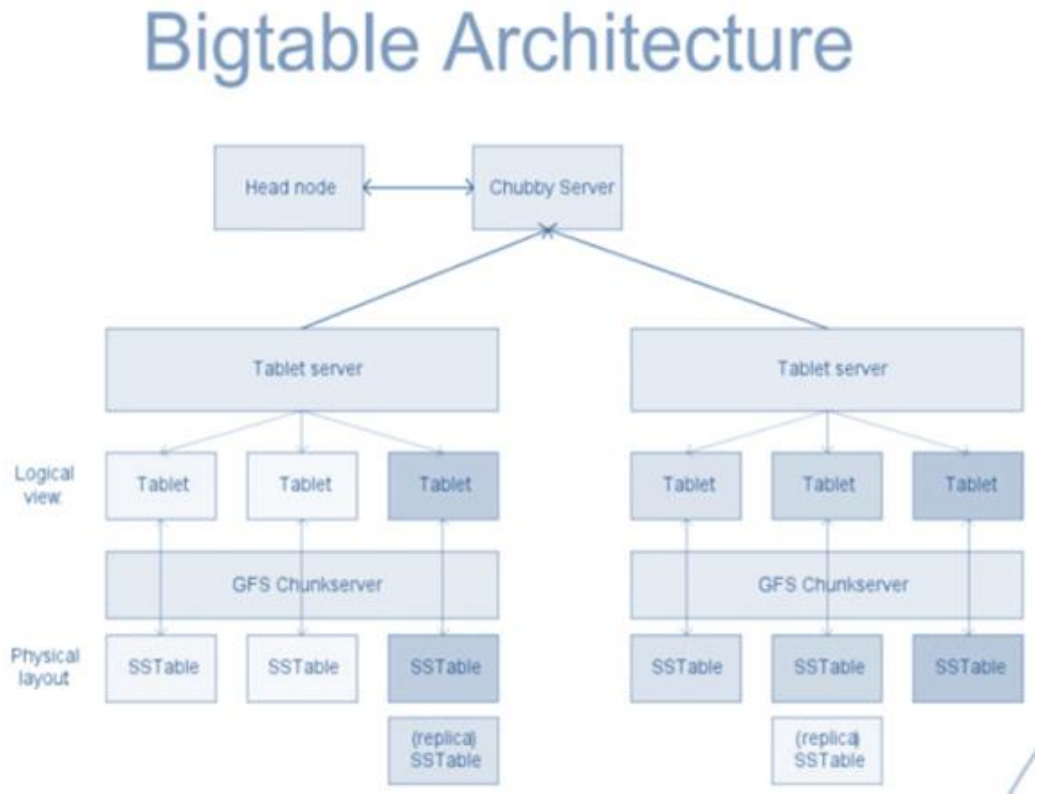



Fig 1.5 Bigtable Architecture

- Since the tables are very large, Bigtable splits them at row boundaries and saves them as tables.
- Each tables is about 200 MB and each server stores about 100 tables.
- The data from a database is likely to be stored in many different servers.
- This architecture also allows for load balancing ( i.e.,) If one table is getting a lot of queries, it can move the busy table to another machine that is not busy.
- Also, when a machine fills up, it can compress some tablets freeing more drive space.

## 1.3 Cloud Scenario

### 1.3.1 Benefits of cloud computing

It offers various advantages in which the main components are listed below:

- Scalability
- Simplicity
- Vendors
- Security


Fig 1.6 Cloud Benefits

**Scalability:**

The ability of a model to be extended to manage the amount of work growth in an effective manner is called scalability.Cloud-computing resources can be rapidly scaled according to subscribers convenience. If there is a sudden necessity for more computer resources, instead of buying new equipment we can buy additional resources form cloud providers. After the endeavor is over we can stop using those services.

**Simplicity:**

In most cases cloud-computing is free to use. It is very simple that users can easily understand which is the biggest advantage of cloud-computing. It is possible to get our application started instantly.

**Vendors:**

The service providers are called vendors. Some of the well known vendors are Google, Amazon,Microsoft, IBM. These providers offer reliable services to their customers.

**Security:**

There are also some risks when using a cloud vendor. But the reputed firms work hard to keep their consumers data safe and secure. They use complex cryptographic algorithms to authenticate users. To make it even more secure we can encrypt our information before storing it in cloud.

### 1.3.2 Limitations in cloud

There is not any product without a few flaws and so is cloud computing. There are some cases where the cloud computing may not be the best solution for computational requirements. Such cases are called as limitations,there are two main limitations:

#### 1.3.2.1 Sensitive Information :

Storing sensitive information on the cloud is always dangerous. Any important information about a person or an organization is called sensitive information. Once the data leaves our hands, the control over the data is lost but that does not mean we cannot manage the data on cloud. It needs to be kept safe. Some of the popularly known limitations in the issue of sensitive information are:

- Government can get the information from service providers easily.

- In few cases, the service providers itself share our data with marketing companies.

The best way is to encrypt the data before storing it in cloud or sending it to third party. Programs like PGP ( Pretty Good Privacy ) or open source True Crypt can encrypt the file so that only the one who owns the password can view the details stored in the uploaded file.

**1.3.2.2 Application development :**

This is the other important limitation in cloud computing. In some cases the applications we need may not be available on cloud or not work as expected. And sometimes some applications will not be securely communicated over the internet. In that case our data will be at risk. Thus there are only two ways to get the desired product. One is, to develop their own application and other is to approach application developer to build the desired product for you.

### 1.3.3 Security concerns

In cloud computing world, security is a two sided coin. The security is very important particularly when moving critical applications and sensitive data to public and shared environments.

### 1.3.4 Privacy concern with a third party

The important security concern is for privacy considerations. That is, if third party is hosting all our data, we do not know if it is safe or not. Everything that is placed on cloud can be accessed by anyone.

There are also other privacy concerns because government can get the data that is placed on cloud easily from organization's servers. Though there are popular companies who provide good security to keep the data safe, it can be hacked. The best procedure is not to perform critical tasks on a cloud platform without extensive security.

If it cannot be managed then it is advisable to have less critical data on cloud.

### 1.3.4.1 Security level of third party

Service providers are doing all they can to protect their customer's data. As a matter of fact, the vendors will have to make sure that the subscriber has been fully satisfied from their service or else the firm will not be gaining customers.

Most of the security problems are due to loss of control, lack of trust and multi-tenancy.

Multi-tenancy – it is an architecture in which single instance of a software application serves multiple customers. Each customer is called tenant.

These problems exist mainly in third party management models. So there should be strong protection measures in order to prevent the hacking of data.

### 1.3.4.2 Security Benefits

Providers do endeavor to ensure security. Cloud provide some of the security measures ensuring the customers data are safe:

#### *1.1.1.1* Centralised Data

There are some good security traits that come with centralizing your data, making your system more inherently secure.

**Reduced Data Leakage**:

**If the data is centralized and the various devices used like laptop, notebook computers can access the data, no need to backup the data**.

There is threat for theft of the handheld devices. If the data are lost and although any security measures like encryption is applied and it may be compromised and the entire data may be in the hands of the thief.

Moreover by maintaining data on the cloud, employing strong access control, limiting the employee downloading to only what they need to perform a

task, computing can limit the amount of information that could be potentially be lost.

**Monitoring benefits**:

Central storage is easier to control and monitor. The flipside is the nightmare scenario of comprehensive data theft. If your data is maintained on a cloud, it is easier to monitor security than have to worry about the security of numerous servers and clients. The security professional figuring out smart ways to protect and monitor access to data stored in one place (with the benefit of situational advantage) than trying to figure out all the places where the company data resides. You can get the benefits of Thin Clients today but Cloud Storage provides a way to centralize the data faster and potentially cheaper. The logistical challenge today is getting Terabytes of data to the Cloud in the first place.

*1.1.1.2* **Instant Swapover** - if a server in the Cloud gets compromised (i.e. broken into), then clone that server at the click of a mouse and make the cloned disks instantly available to the Cloud Forensics server. When the swapover is performed its seamless to the users. No need to spend time to replicate the data or fix the breach. Abstracting the hardware allows to do it instantly.

***1.1.1.3*** **Logging**

In cloud logging is improved. Logging is often an afterthought, to solve the issues insufficient disk space is allocated. Cloud Storage changes all this - no more ‘guessing’ how much storage you need for standard logs. With your logs in the Cloud you can leverage Cloud Compute to index those logs in real-time and get the benefit of instant search results. This help to Compute instances and to measure in and scale as needed based on the logging load - meaning a true real-time view.

Most modern operating systems offer extended logging in the form of a C2 audit trail. This is rarely enabled for fear of performance degradation and log size. Now you can ‘opt-in’ easily - if you are willing to pay for the enhanced logging, you can do so. Granular logging makes compliance and investigations easier.

### *1.1.1.4* Secure builds

When you developed your own network and you have to buy third-party security software to get the level of protection you want. With the cloud solution, those tools can be bundled in and available to you and you can develop your system with whatever level of security you desire.

**Easier to test impact of security changes**: this is a big one. Spin up a copy of your production environment, implement a security change and test the impact at low cost, with minimal startup time. This is a big deal and removes a major barrier to ‘doing’ security in production environments.

**Improve the state of security software (performance**)

**Drive vendors to create more efficient security software**: Billable CPU cycles get noticed. More attention will be paid to inefficient processes; e.g. poorly tuned security agents. Process accounting will make a comeback as customers target ‘expensive’ processes. Security vendors that understand how to squeeze the most performance from their software will win.

*1.1.1.5*

### *1.1.1.6* Security Testing

**Reduce cost of testing security:** a SaaS provider only passes on a portion of their security testing costs. It is shared among the cloud users. The end results is that because you are in a pool with others but you never see the other users but you realize the lower cost for testing.

Even with Platform as a Service (PaaS) where your developers get to write code, but the cloud code –scanning tools check for security weakness.

### 1.3.4 Regulatory Issues

In the case of cloud computing, regulation might be exactly what we need. Without some rules in place there are chances for unsecure with service or even shifty enough to make off with your data.

'Sensitive Data' is defined as personal information that relates to:

(a)passwords;

(b)financial information such as Bank account or credit card or debit card or other payment instrument details;

(c)physical, psychological and mental health condition;

(d)sexual orientation;

(e)medical records and history;

(f)biometric information;

any detail relating to the above received by the body corporate for provision of services; or

8. any information relating to (a) – (g) that is received, stored or processed by the body corporate under a lawful contract or otherwise.

No existing regulation:

Currently there is no existing regulation.

While comparing cloud service providers to banks there are similarities, Banks deal with money whereas cloud service providers deal with data, both are immense value to consumers and organizations alike.

**Location of Stored Data** – Service providers generally do not disclose the location where the service subscriber's data are stored. It leaves users in the dark regarding the extent of protection applied to their critical information.

Although security certifications could lessen the user's anxiety, the matter of determining if the provider's compliance with legal and regulatory laws

includes those that cover the geographical location where data is stored, aside from the laws of the areas where the data was collected.

Government to the rescue?

Is the Government's place to regulate cloud computing? If government can figure out a way to safeguard data, either from loss or theft of any company facing such a loss would applaud the regulation. One such example is the greatest bank failure in American History. In 2008 the United States government took control of Washington Mutual.

On the other hand, there are those who think the government should stay out of it and let competition and market forces guide cloud computing.

Who owns the Data?

To overcome the various issues in cloud, the Government has to work out on important questions , like

Who owns the data?

Should law enforcement agencies have easier access to personal information on cloud data than that stored on a personal computer?

Also the big problem is that people using cloud services are not aware of the privacy and security implication on their online email accounts, their LinkedIn account, their MySpace page, and so forth. While these are popular sites for individuals, they are still considered cloud services and their regulation may affect other cloud services.

Government Procurement

There are also questions about whether government agencies will store their data on the cloud.

Procurement regulations will have to change for government agencies to be keen on jumping on the cloud.

The General Service Administration (GSA) is making a push toward cloud computing, in an effort to reduce the amount of energy their computers

consume. The GSA is working with a vendor to develop an application that will calculate how much energy government agencies consume.

**Government Policies**:

Government Policies

The aim of the cloud policy of government is to realise a comprehensive vision of a government cloud (GI Cloud) environment available for use by central and state government line departments, districts and municipalities to accelerate their ICT-enabled service improvements. As per the guidelines, both cloud service provider (CSP) and government department will have to share responsibility for the managing services provisioned using cloud computing facility.

To implement the policy, Government of India has made an initial step "GI Cloud" which has been coined as 'Meghraj'. The focus of this initiative is to accelerate delivery of e-services in the country while optimizing the expenditure of the Government.

The ministry of electronics and IT has issued on important guideline regarding the location of data as follows : "The terms and conditions of the Empanelment of the Cloud Service Provider has taken care of this requirement by stating that all services including data will be guaranteed to reside in India".

The cloud computing service enables its user to hire or use software, storage, servers as per requirement instead of purchasing the whole system.

Meity( Ministry of Electronics and Information Technology) has empanelled the following companies for providing cloud computing services to government departments :

1. Microsoft Corp.,
2. Hewlett Packard,

3. IBM India ,
4. Tata Communications,
5. Bharat Sanchar Nigam Ltd (BSNL),
6. Net Magic IT Services,
7. Sify Technologies and
8. CtrlS Data Centers.

The architectural vision of GI Cloud as mentioned above consists of a set of discrete cloud computing environments spread across multiple locations, built on existing or new (augmented) infrastructure as given below :

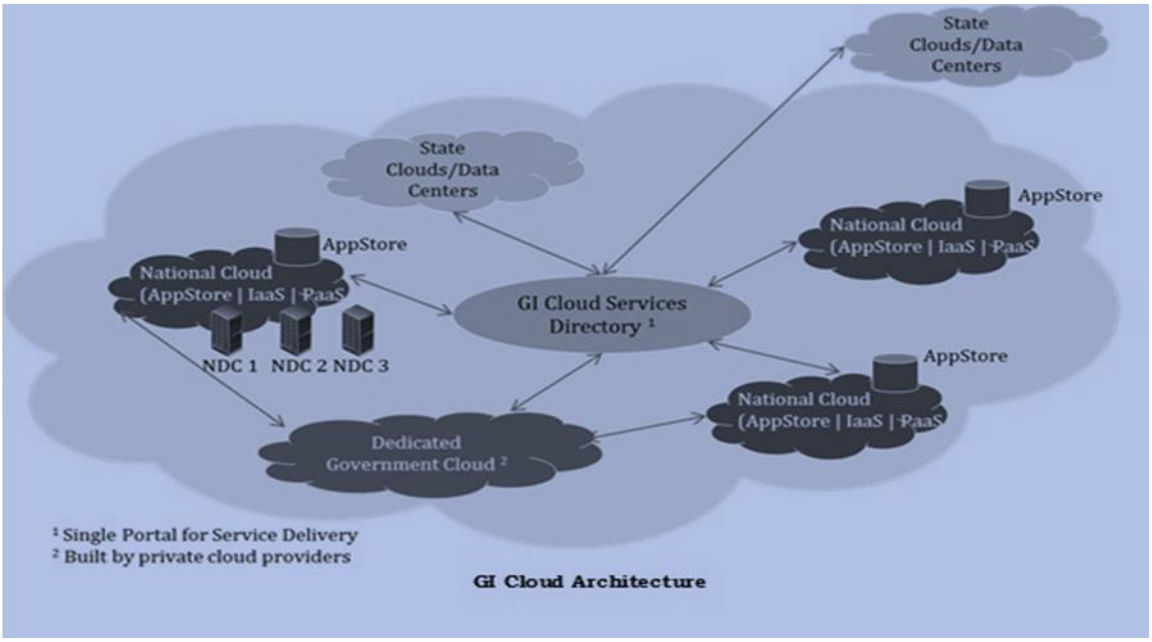



**Fig 1. 7 GI Cloud Architecture**

## 2. Components of Meghraj

1. Setting up of State and National Clouds
2. Set up an e-Gov Appstore
3. Empanelment of Cloud Service Providers

4. Empanelment of Cloud Auditors
5. Setting up of Cloud Management Office
    1. Setting up an eco-system for Cloud proliferation (Policies, Guidelines, templates, security norms, certification, business models for applications, tariff & revenue models for private sector Cloud services)
    2. Awareness workshops, training programs and migration support for cloud adoption by departments
6. MeghRaj (GI-Cloud) service Directory
7. Setting up of Clouds by other Government entities

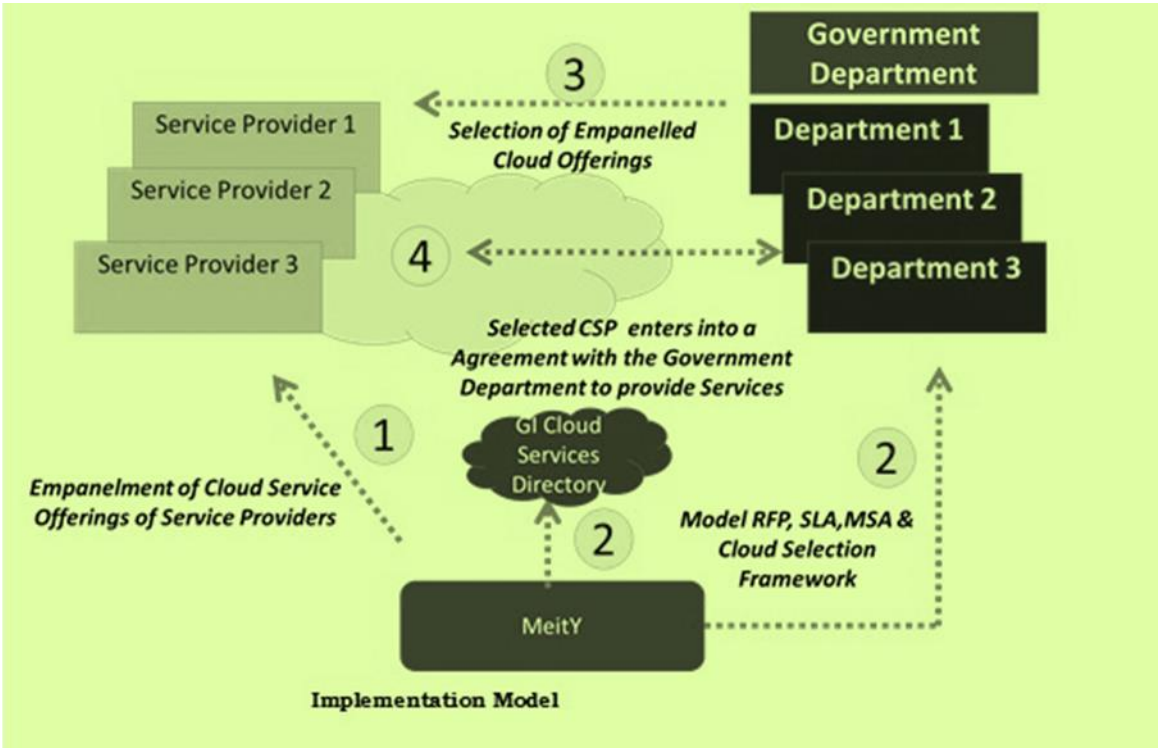



Fig 1.8 Implementation model

Cloud Deployment Models:

The empanelment of the Cloud service offerings of CSPs has been done for a combination of the Cloud Deployment models and Service models as mentioned below:

Public Cloud :

A shared multi-tenant IT infrastructure is made available over the internet. It is owned and operated by a Cloud Service Provider delivering cloud services to the Government Department.

Virtual Private Cloud :

A logically separated Cloud Infrastructure (Servers, Storage, Network infrastructure and Networks) to protect data, applications and servers and provide robust virtual isolation for the Government Department.

Government Community Cloud

A cloud with IT infrastructure resources which will be dedicated for two or more Government Departments that have common privacy, security and regulatory considerations.

* * *

SUMMARY

- ✓ Cloud computing is the process of delivering or providing computational resources like software and/or hardware as a service over the cloud (internet).
- ✓ Cloud Components-There are three cloud components, They are Clients,Data centers,Distributed servers
- ✓ Essential characteristics: 1) On-Demand self service,2) Broad network access,3) Location independent resource pooling,4) Rapid elasticity,5) Measured service

- ✓ High-performance computing (HPC) is the use of super computers and parallel processing techniques for solving complex computational problems.
- ✓ Server consolidation involves reducing the number of servers and server locations within an organization.
- ✓ Benefits of cloud computing Scalability,Simplicity,Vendors,Security
- ✓ Limitations in cloud are Sensitive Informationand application development**.**

## Questions

### Part-A

1. Define Cloud Computing.
2. Define Utility Computing.
3. What is grid computing?
4. What is autonomic computing?
5. Mention the benefits of Cloud Computing.
6. List out any two limitations of cloud computing.
7. What is meant by on-demand self service?
8. What is Data center?
9. What is horizontal scaling?
10. Mention two essential characteristics of cloud computing.

### Part-B

1. What are the components of cloud computing?
2. Write the essential characteristics of cloud computing.
3. What are the types of clients?
4. State any 3 security benefits.
5. What is High Scalability Architecture?

## Part-C

1. What is cloud computing? Discuss the origin of cloud computing.
2. Explain about components in detail.
3. Write short notes on high performance computing.
4. Write short notes on :
   (a) Web Services.
   (b) High scalability architecture.
5. Discuss the limitations of cloud computing.
6. List out the benefits of cloud computing and explain.
7. Explain the architectural influences of cloud computing.
8. Write about security concerns in cloud.
9. Explain about utility computing and enterprise grid computing.
10. Explain about regulatory issues and government policies in cloud computing.

* * *

# UNIT – II

# CLOUD COMPUTING ARCHITECTURE AND SERVICES

**Objectives**

**To learn the cloud delivery model**

**To understand the difference between SPI vs traditional IT model**

**To learn the services of the SPI framework**

**To learn the cloud deployment model and its types.**

**Introduction:**

**The Cloud computing architecture** refers to the components and subcomponents required for cloud computing. These components typically consist of a front end platform (fat client, thin client, mobile device), back end platforms (servers, storage), a cloud based delivery, and a network (Internet, Intranet, Intercloud). Combined, these components make up cloud computing architecture. The cloud services are categorized into SaaS,PaaS,IaaS.

The evolution of SPI when compared with the traditional IT model,helps to know how the cloud is deployed based on the requirement of the user/organization is discussed in detail.

## 2.1 Cloud Architecture

Cloud architecture is the design of software application that uses internet access and on-demand service. Cloud architecture is used when it is needed to retrieve the resources on demand and perform the specific job. It will dispose the resources once the job is finished.

These services can be accessed anywhere from the world.

In Fig 2.1 the cloud computing architecture is depicts the various components :

The major components are defined and explained below.

- Cloud consumer
- Cloud provider
- Cloud auditor

- Cloud broker
- Cloud carrier

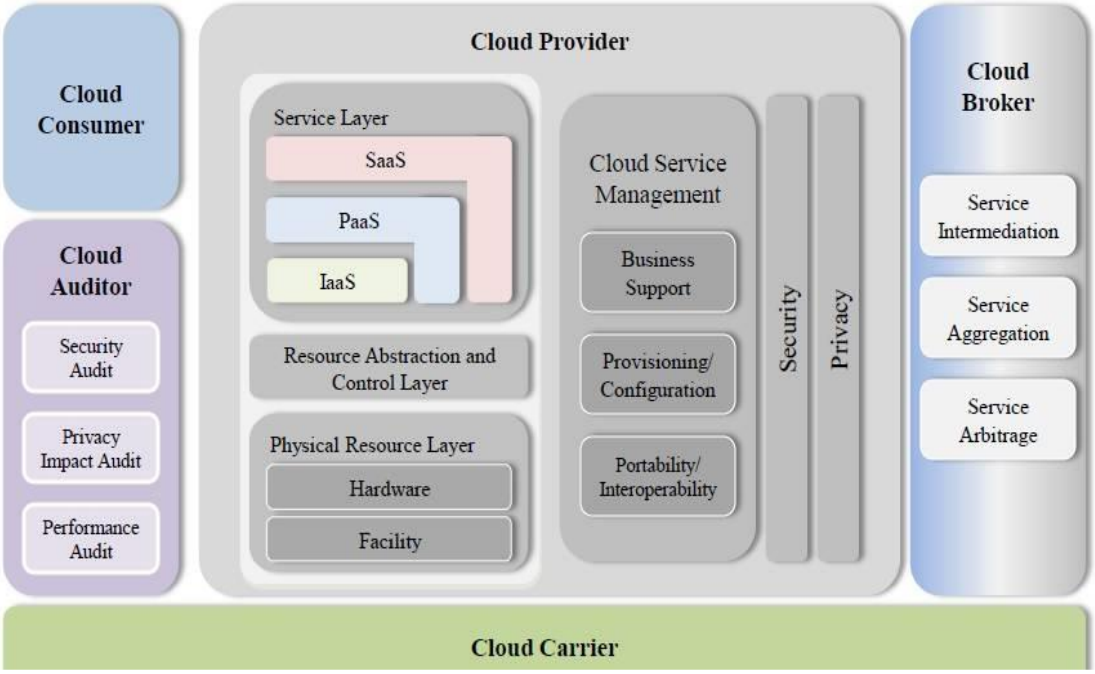



Fig 2.1 Cloud Architecture

Cloud computing architecture is categorized into two sections. They are front end and back end. Front end is the part seen by the client. Back end is the place where it has the server and databases.

Cloud has centralized server administration system and adjusts demands, avoids congestion, monitors traffic and client supply.

**Cloud consumer** is the person or organization that maintains a business relationship between the user and services from cloud providers.

**Cloud provider** is the person or organization that is responsible for making the service available for cloud consumers.

**Cloud auditor** is the person who can conduct assessment of cloud services, performance, security of cloud, information system operations.

**Cloud broker** is used to manage the performance, usage and delivery of services. It is also used to negotiate between cloud consumers and cloud providers.

**Cloud carrier** is the intermediate level that provides connectivity of cloud services.

### 2.1.1 Cloud delivery model

Cloud computing services can be delivered to customers(users) in different ways and depicted in Fig 2.2. The cloud computing delivery models include infrastructure, platform and software.

These services are provided and used over the internet.

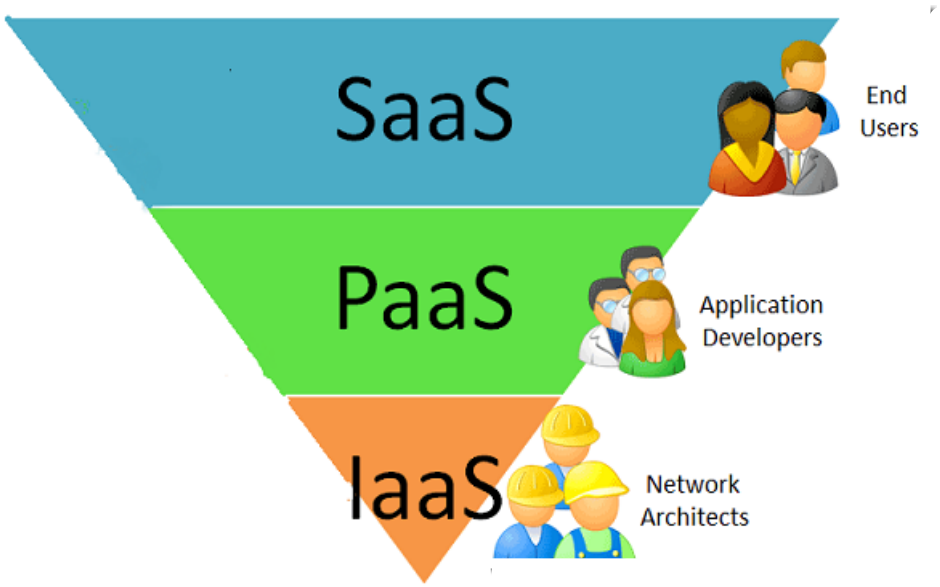



**Fig 2.2 Cloud Delivery Model**

**Infrastructure as a Service ( IaaS )**

It is one of the cloud delivery model which provides computer infrastructure or hardware like servers, networking technology, storage as a service. It may also include the delivery of operating system and virtualization technologies to manage these resources.

**Platform as a Service ( PaaS )**

PaaS is another cloud delivery model which delivers more than just infrastructure. It provides solution stack which is an integrated set of software. This model provides everything a developer needs to build an application for software development and run time.

**Software as a Service ( SaaS )**

SaaS is a delivery model which delivers business application designed for a specific purpose. This service is provided over the internet which eliminates the need to install and run the applications on consumer's own computers. It simplifies maintenance and support.

Some of the properties of SaaS are:

- Network or online access
- Centralized management
- Powerful communication features

SaaS works on two distinct modes, they are

- Single multi tenancy
- Fine-grain multi tenancy

### 2.1.2 SPI Framework

The acronym for SPI stands for three major services provided through the cloud. They are as follows,

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

The Fig 2.3 gives overall view about the various resources are available to the end-user.

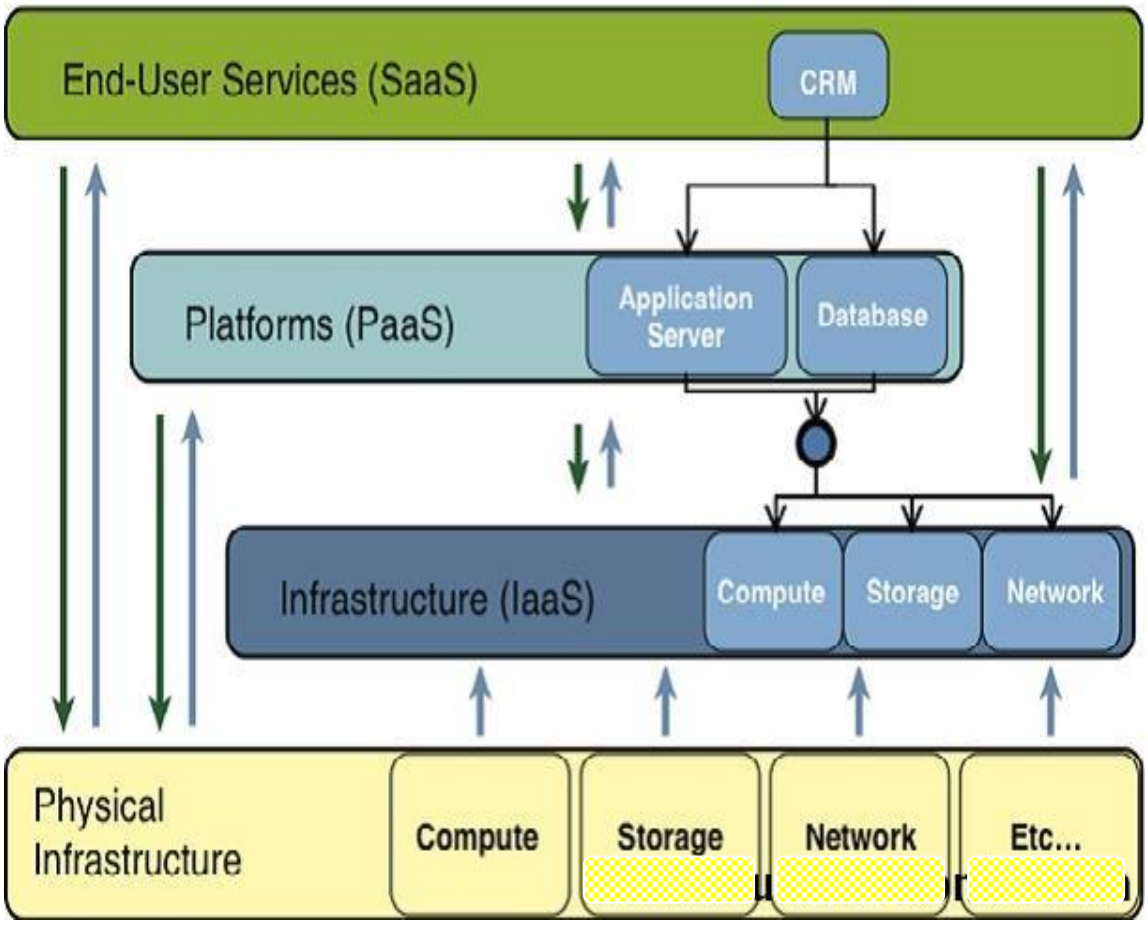



Fig 2.3 SPI Framework

### 2.1.3 SPI evolution

To understand how the SPI framework evolved, perhaps it's helpful to place it in context with the development of Internet Service Providers (ISPs).

The various version and services of ISPs are:

Version 1.0 — As ISPs originally began to provide Internet services, dial-up modem service for homes and organizations grew, making the Internet a commercial commodity.

Version 2.0 — During a period of merging and consolidation, ISPs began offering other services, such as e-mail and off-site data storage.

Version 3.0 — The increasing demand for infrastructure to host their customers applications and data led to the creation of data centers known as collocation facilities, where multiple customers could centralize their servers, storage, and communications systems on the ISPs premises.

Version 4.0 — The commoditization of collocation facilities led to the development of application service providers (ASPs). ASPs provided software applications tailored to an organization, owning both the application and the infrastructure.

Version 5.0 — The ASP model eventually evolved into cloud computing, which brought new delivery models, such as the SPI framework, with its SaaS, PaaS, and IaaS service models, and various deployment models, such as private, community, public, and hybrid cloud models.

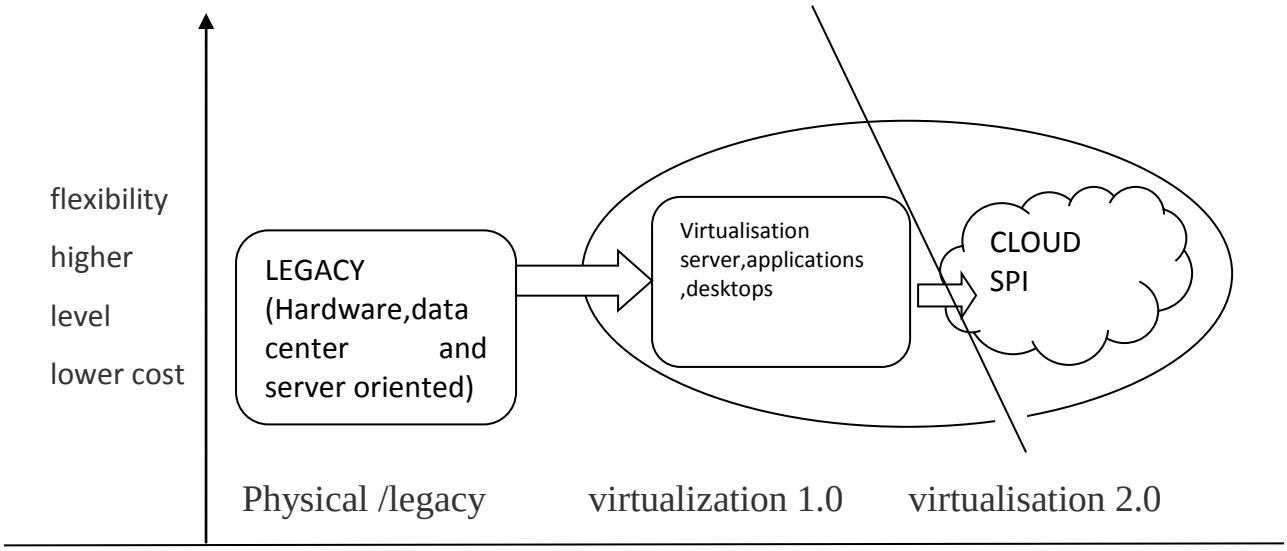



**Fig 2.4 SPI evolution through virtualization**

### 2.1.4 SPI vs Traditional IT Model

**Traditional IT Model**

The traditional enterprise-wide application rollout requires resources and coordination from many parts of the organization. This rollout may require numerous new hardware (servers, perimeter network devices, workstations, backup systems), operating systems, communication link provisioning, and user and management training, for example.

One advantage of the traditional model is that software applications are more customizable, but even this advantage often comes at a high cost in resources and effort.

In the traditional IT model, software applications may require substantial licensing and support costs. These licensing costs may be based on formulae that do not translate well to the actual intended use of the application, such as hardware requirements (number of servers, processors, communication links) or other company characteristics unrelated to the original intent of the application (total number of employees in the organization, total number of remote offices, etc.).

In addition, changes in the original licensing structure due to usage increases (additional per-seat needs) may create substantial costs down the line, such as additional hardware; support SLAs, and IT resources.

In the traditional IT model, security is often owned "in-house," with security professionals and supporting security infrastructure (firewalls, intrusion detection/prevention systems, e-mail and web monitoring systems, etc.) under the direct control of the organization.

This may make it easier to provide regulatory compliance for auditing purposes using the traditional model. However, the drawback of this security ownership is the infrastructure overhead, which requires considerable resources of manpower to secure properly.

Typically, organizations employing the SPI framework do not own the infrastructure that hosts the software application. They instead license application usage from the cloud provider, by employing either a subscription-based license or a consumption-oriented model.

**SPI Framework**

This enables companies to pay for only the resources they need and use, and to avoid paying for resources they do not need, thus helping them avoid a large capital expenditure for infrastructure. The cloud service provider that delivers some or all of the SPI elements to the organization can also share infrastructure between multiple clients.

This helps improve utilization rates dramatically by eliminating a lot of wasted server idle time. Also, the shared use of very high-speed bandwidth distributes costs, enables easier peak load management, often improves response times, and increases the pace of application development.

Another benefit of adopting the SPI framework for organizational computing is reduced startup costs. Eliminating the resource requirements mentioned above lowers the barrier to entry, and in many cases provides an organization much quicker access to computing power and software development than the traditional IT model did.

The table 2.1 consolidates the SPI vs Traditional model.

| | Traditional IT model | SPI model |
|---|---|---|
| Data access | While in traditional computing, the user can access data only on the system in which data as to be stores | Is the ability to access the data anywhere and anytime. |
| Automation & maintenance | **No automatic updates**.Fully trained IT personnel need to regular monitoring and maintenance of server. | **Automatic updates**.no need to monitored. |
| Cost | Start up cost is high. | **Low cost : pay for what you used** |
| Security | **Low security**: the organization takes the responsibility to protect the data | 1. **high security :** Data is stored in 256 bit AES ( Advanced Encryption Standard ) and customer account information is encrypted before storing in database. |
| Easy to use | Training is needed to maintain the resources. Difficult to use | To work with these apps, no training is needed. Users can easily understand without having any technical knowledge. |
| Infrastructure | Advantage of traditional IT model is software applications. It is more customizable. It may require substantial licensing and support cost. | SPI frameworks does not own the infrastructure. It employed by the subscription based license. |

**Table 2.1 SPI model Vs traditional IT model**

## 2.2 Saas (Software as a Service)

SaaS in one of the SPI models.

In this, the customer does not need to purchase software. They rent it for use on a subscription or pay per use plan. Some of these services could be free for limited usage.

Some of the benefits of SaaS are,

- Easier administration
- Automatic updates

SaaS is a software delivery model where the applications are hosted by the service provider or the dealer. This architecture made the services available to customers over a network typically the internet. The Fig 2.5 steps involved in using SaaS.

Fig 2. 5 SaaS


Benefits of Software as a service model :

- Easier administration
- Automatic updates
- Patch management
- Compatibility : All users will have the same version of software
- Easier collaboration as all users have same version.
- Global accessibility.

### 2.2.1 Software as a service providers :

Some of the service providers are

1. Amazon Web services
2. Google Apps
3. icloud
4. Oracle
5. Salesforce.com
6. Windows Azure

### 2.2.2 Web Services

A **web service** is a service offered by an electronic device to another electronic device, communicating with each other via the World Wide Web. In a Web service, Web technology such as HTTP, originally designed for human-to-machine communication, is utilized for machine-to-machine communication, more specifically for transferring machine readable file formats such as XML and JSON, overall concept is depicted in Fig . 2.6.

In practice, the web service typically provides an object-oriented web-based interface to a database server, utilized for example by another web server, or by a mobile application, that provides a user interface to the end user.

Another common application offered to the end user may be a mashup, where a web server consumes several web services at different machines, and compiles the content into one user interface.

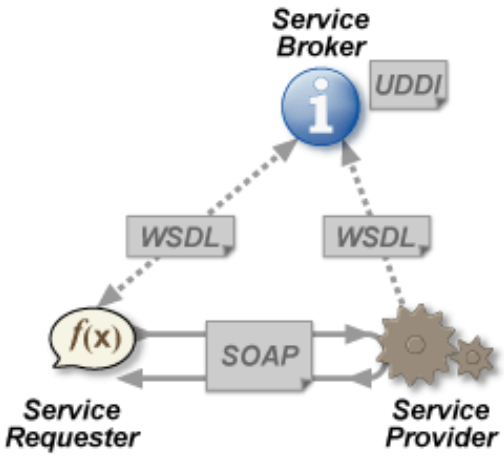



**Fig 2.6 Web Services**

### 2.2.3 Web 2.0

**Web 2.0** describes World Wide Web websites that emphasize user-generated content, usability (ease of use, even by non-experts), and interoperability (this means that a website can work well with other products, systems and devices) for end users.

A Web 2.0 website may allow users to interact and collaborate with each other in a social media dialogue as creators of user-generated content in a virtual community, in contrast to the first generation of Web 1.0-era websites where people were limited to the passive viewing of content. Examples of Web 2.0 include social networking sites and social media sites (e.g., Facebook), blogs, wikis, folksonomies ("tagging" keywords on websites and links), video sharing sites (e.g., YouTube), hosted services, Web applications , collaborative consumption platforms, and mashup applications.

### 2.2.4 Web Operating System:

In cloud computing, users work with Web-based, rather than local, storage and software. These applications are accessible via a browser and look and act like desktop programs.

With this approach, users can work with their applications from multiple computers. In addition, organizations can more easily control corporate data and reduce malware infections.

Also, cloud computing makes collaboration easier and can reduce platform-incompatibility problems. Now, a growing number of organizations are adding to the cloud concept by releasing commercial and open source Web-based operating systems. While the idea is not new, the proliferation of users and applications distributed over the Web, including those at scattered corporate sites, has made it more interesting, relevant, and, vendors hope, commercially viable.

It also includes many of a traditional operating system capabilities, including a file system, file management, and productivity and communications applications. As is the case with Web-based applications, the Web OS functions across platforms from any device with Internet access.

**Web operating systems** are interfaces to distributed computing systems, particularly cloud or utility computing systems. Web OS can be referred as a virtual Desktop accessible through a web browser with multiple integrated inbuilt applications that allow the users to manage their data from any location and is independent of the traditional operating systems. These are also termed as Internet Operating Systems

### 2.2.4.1 Google App Engine:

Google app engine is a SaaS provider which was introduced in 2008. It was quite unique cloud system compared to other systems. It provides platform to create applications. It provides infrastructure for hosting.

Many high level services which needs to be build are available when using an App Engine.

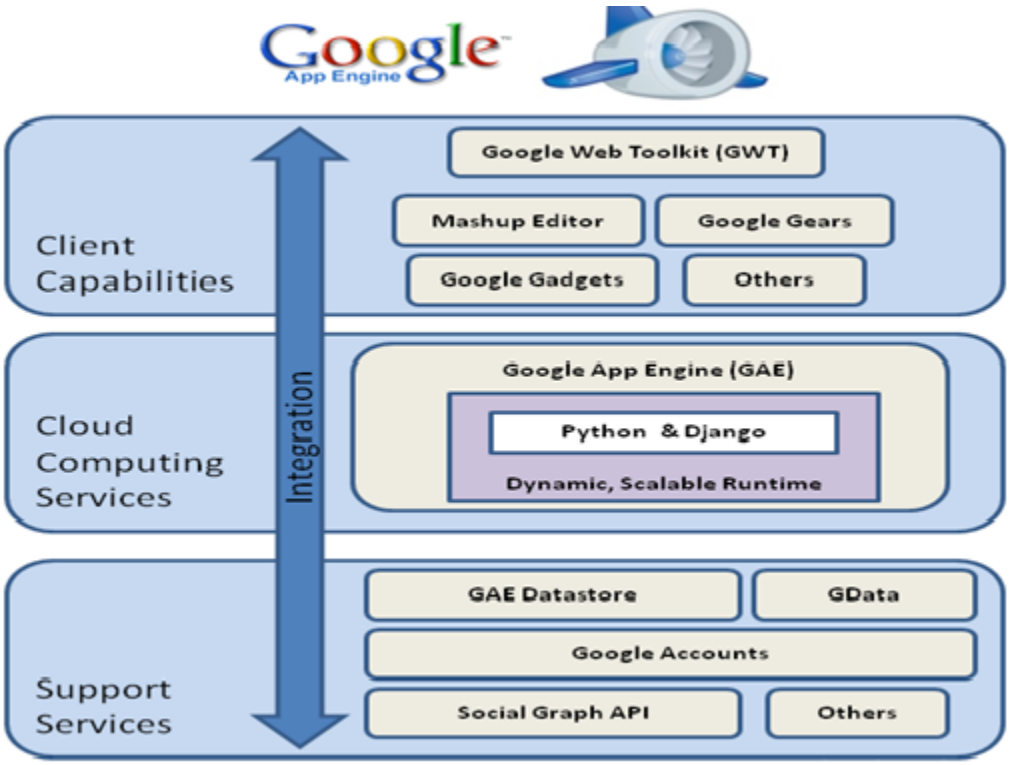



Fig 2.7 Google App Engine

A development server is provided to allow users to test their code before running in actual production. Deployment is very simple as Google handles it.

Google App Engine enables developers to build web applications on some scalable systems that power their own applications.

Scalable applications are easily designed that attracts millions of users without any infrastructure risk. Every Google app engine application will have enough CPU, bandwidth and storage to service around five million monthly page views for free.

With this app engine we pay only for what we use.

There are no setup costs and fees. All applications can use up to one gigabyte of storage. It supports various applications written in different programming languages like Java, JVM, JavaScript, Ruby.

With the help of Google App Engine we can create apps that runs under heavy load, large data and very much reliable.

**Features :**

Dynamic web serving with full support for common web technologies.

Automatic scaling.

Load balancing.

API's for authentication users and sending email using Google Account.

The application can run in any one of the following environment.

They are GoGRID Environment, the Java environment and the Python Environment.

Standard protocols and common technologies for web application are provided.

### 2.2.4.2 Salesforce.com and google platform :

One of the service of Salesforce, is enabling customers to use Salesforce, one of the CRM, with Google Apps.

Together with Google have delivered seamless integrations between Salesforce and Google Drive, Gmail and Calendar.

With Google on its new Sheets and Slides APIs, which were announced today at Google I/O. With these new APIs, you can seamlessly connect your favorite Salesforce apps to Sheets and Slides for increased ease of use and real-time collaboration.

And now, with new integrations between Salesforce and the APIs, data and reports will flow seamlessly between these solutions so you always have access to Salesforce data directly within the Google apps you use every day.

In addition, you will be able to connect Salesforce to Google Sheets with the new Sheets Lightning Component from Salesforce Labs help to quickly update Salesforce CRM data, such as opportunities, directly from Sheets. Since Salesforce data and records are synchronized with Sheets in real time and always working with latest information.

### 2.2.5 Benefits of SaaS

1. **Low cost :** There are no software installation and also no need to maintain the resources. It helps to decrease the resource cost.
2. **Quick deployment :** It can be set up and ready to work in few minutes.
3. **Easy to use :** To work with these apps, no training is needed. Users can easily understand without having any technical knowledge.
4. **Increased collaboration :** Web platform allows the solution to be used across the industries world wide.
5. **Scalable :** This model provides unlimited scalability and very quick processing time.
6. **Geo specific hosting :** It ensures that the data is kept in the specific location.
7. **On demand :** The solution is self served and is made available to use anytime we need.
8. **Secure access :** Data is stored in 256 bit AES ( Advanced Encryption Standard ) and customer account information is encrypted before storing in database.
9. **On going benefits :** It is not just one time solution. The organization can enjoy

### 2.2.5.1 Operational benefits of SaaS :

SaaS can improve the consumer's organization effectiveness based on the following benefits,

1. **Managing business driven IT project:**

   A SaaS model provide the necessary infrastructure and thus leads to technology projects that address true business needs.

2. **Increasing consumer demand :**

   SaaS model provides reliability to deliver near perfect 99.99% system availability. So any number of users can access the system at anytime from anywhere.

3. **Addressing growth :**

   This model provides scalability that is easily supported by an increasing number of consumers to meet their own objectives.

4. **Serving new markets quickly and easily :**

   SaaS allows the organization to quickly and easily adds programs so as to adapt the changes based on the demand at a faster rate.

5. **On demand :** The solution is self serve and available for use as needed.
6. **Scalable :** It allows for the infinite scalability and quick processing time.

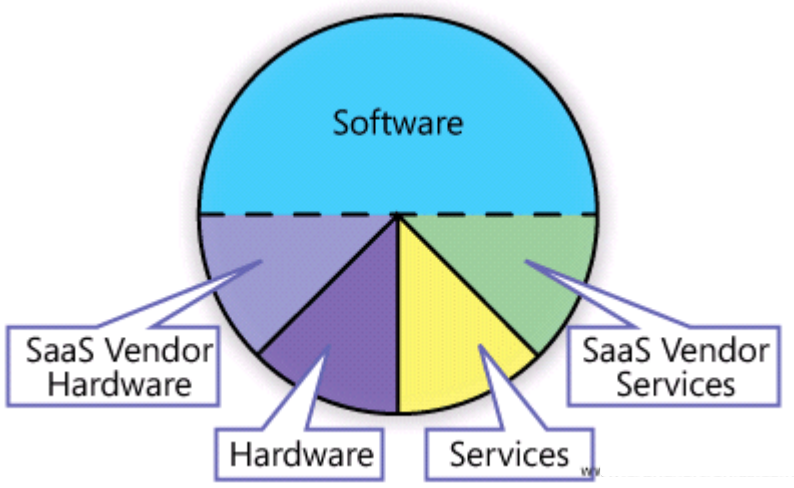



Fig 2.8 SaaS utilization

### 2.2.5.2 Economic benefits of SaaS :

SaaS not only saves time but also has greater financial benefits.

1. It reduces IT expenses.
2. The implementation cost of SaaS is much lower than the traditional software.
3. It redirects savings expenses towards business improvements.
4. It strengthens the financial capability.
5. By utilizing SaaS, we are free to use as much of any software as we need. This gives you easy and economical access to many programs.
6. SaaS vendors release upgrades for their software, thus users need not put any effort into installing and upgrading the software.
7. Another main benefit in SaaS is that it can quickly and easily be accessed from anywhere by using a web browser.

### 2.2.6 Evaluating SaaS

Consumer should consider few features for evaluating SaaS, and they are as follows

1. Disaster recovery
2. Security
3. Flexibility and quality of service
4. Service level agreements ( SLA )
5. Global reach

Evaluating SaaS is divided into two criteria. They are

- Business criteria
- Technical criteria

The business criteria includes

- Pricing and billing
- Service levels and agreements
- Support and communication

The technical criteria includes

- Identity
- Integration
- Managements
- Security
- Storage
- Network

By placing all these criteria together, it becomes important to provide organization or a customer needs more reliability and assurance. SaaS should provide highly secured data over internet. It should have the security measures.

It should be flexible in involving number of things. SaaS provider should provide the business for long time. Finally it should meet consumer needs with availability, scalability and security. Thus the things discussed above must be considered for evaluating SaaS.

### 2.3. Platform as a Service – PaaS

PaaS is one of the SPI models. In PaaS, the vendor offers a development environment to application developers. They develop applications and offer those services through provider's platform. It helps to use web based applications without the cost and complexity of buying servers and setting them up. Some of the benefits of PaaS are,

- Each platform component is provided as a service
- Reduces total cost of ownership

PaaS – Platform as a Service is a service model in cloud computing. It provides a solution stack and computing platform for customers as a service.

In this standard, the customer develops the application using libraries or tools from the service provider. The consumer also controls the software deployment and configuration settings. The PaaS vendor provides the networks , servers and other required services.


Fig 2.9 Paas users

PaaS provides the users with easy deployment of applications without much cost and complexity of getting the needed resources. It manages the underlying hardware and software, and gives hosting capabilities. There are various types of platform as a service vendors who offer application hosting and a development environment along with various integrated services. These services have varying levels of scalability and maintenance.

### 2.3.1 Cloud Platforms and Management

**Google Cloud Platform** is a cloud computing service by Google that offers hosting on the same supporting infrastructure that Google uses internally for end-user products like Google Search and YouTube. Cloud Platform provides developer products to build a range of programs from simple websites to complex applications.

Google Cloud Platform is a part of a suite of enterprise services from Google Cloud and provides a set of modular cloud-based services with a host of development tools. For example, hosting and computing, cloud storage, data storage, translations APIs and prediction APIs.

**IBM Cloud Orchestrator** is a cloud management platform for automating provisioning of cloud services using policy-based tools. It enables you to configure, provision, deploy development environments, integrate service management—and add management, monitoring, back-up and security—in minutes. And do it again—in minutes—whenever you need. All from a single, self-service interface.

Some of the cloud platforms from various service providers are:
Google Cloud Platform are: Google Compute Engine,Google App Engine, Bigtable,Big Query,Google Cloud Functions, Google Cloud Datastore.

Amazon Web Services: Amazon EC2,AWS Elastic Beanstalk, Amazon DynamoDB,Amazon Redshift.

Microsoft Azure: Azure Virtual Machines

### 2.3.2 Computation and storage

Computation and Storage in the Cloud is the first comprehensive and systematic work investigating the issue of computation and storage trade-off in the cloud in order to reduce the overall application cost.

Scientific applications are usually computation and data intensive, where complex computation tasks take a long time for execution and the generated datasets are often terabytes or petabytes in size.

Storing valuable generated application datasets can save their regeneration cost when they are reused, not to mention the waiting time caused by regeneration.

However, the large size of the scientific datasets is a big challenge for their storage. By proposing innovative concepts, theorems and algorithms, this book will help bring the cost down dramatically for both cloud users and service

providers to run computation and data intensive scientific applications in the cloud.

Computation and Storage in the Cloud is the first comprehensive and systematic work investigating the issue of computation and storage trade-off in the cloud in order to reduce the overall application cost. Scientific applications are usually computation and data intensive, where complex computation tasks take a long time for execution and the generated datasets are often terabytes or petabytes in size.

Storing valuable generated application datasets can save their regeneration cost when they are reused, not to mention the waiting time caused by regeneration. However, the large size of the scientific datasets is a big challenge for their storage.

Of course, for the average computer user, some of the most important SaaS offerings are cloud storage services. These allow file hosting, file sharing, and remote data backup.

Basically, after signing up for a cloud storage service, you get a certain amount of free storage space (usually 2–5 GB) for hosting whatever data you would like, along with the option to pay for an upgrade to access more storage space. Just like with other SaaS software, we can increase or decrease your use of the service very quickly, without having to interact with any of the related computing infrastructure.

### 2.3.3 PaaS service providers :

Some of the platform as a service providers are,

- Force.com
- Microsoft Azure
- Google Engine
- Rack space cloud
- Site cloud

- Right Scale
- Salesforce.com

### 2.3.4 Right Scale

Right Scale is one of the PaaS service providers. Right Scale is a web based cloud computing management solution for managing cloud infrastructure from multiple providers. It enables firms to easily deploy and manage business critical applications across public, private and hybrid clouds. This also enables customers to manage hybrid cloud infrastructure by migrating workloads between their private clouds and public clouds operated by Amazon Web Services ( AWS ), Rack Space and Tata.

**Salesforce.com**

Salesforce.com brings the trust and speed to build and deploy the developed applications on cloud. The process is faster than any other paas model. It is the trusted leader in cloud computing and customer relationship management and also it is simple, scalable and most importantly reliable service.

**Rack Space**

Rack space is a world's leading specialist in hosting and cloud computing industry. Rack space hosting started its service in 1998. It provides three different types of services such as managed hosting, cloud hosting and hybrid hosting.

**Force.com**

Force.com is a standalone platform which delivers paas service to customers in a new way to build and deploy apps that makes developers and companies to concentrate on their applications rather than the software and

infrastructure. There is no need to buy software or server. Force.com gives app programmers or creators, the quickest way to transform innovative ideas into business. It creates the business app which is very simple and at the same time sophisticated. It allows consumers to run multiple applications within the same instance.

### 2.3.6.1 Services of PaaS

PaaS includes various services for application design, application development, testing and deployment.

The main services of PaaS are :

- Team collaboration
- Web service
- Integration
- Managing databases
- Security
- Scalability
- Storage
- Presistance
- State management
- Application version
- Application instrumentation
- Developer community facilitation

### 2.3.6.2 Benefits of PaaS

Some of the main benefits of PaaS are given below.

- Each platform components are provided as a services.
- Provides services required to complete the building and deploying services and web applications through internet.
- Service for deploying, testing and maintaining application in same IDE.

- It follows pay per use model.
- PaaS reduces the total cost of ownership since there is no need to buy all the system software, platforms, tools needed to build the application. User can only rent them for a certain period of time till the resources are needed.
- It has elasticity and scalability to afford same efficiency and experience irrespective of load and usage.
- It helps to build applications rapidly with PaaS. System features can be changed and upgraded frequently.

### 2.4 Infrastructure as a Service ( IaaS )

IaaS is a SPI model which is based on the idea of offering computing services. The vendor charges for amount of processing power and disk space. Few benefits of IaaS are,

- Easy to use
- Metered service

Infrastructure as a Service is a provision model in which an organization outsources the equipment to support the operations like storage, hardware, servers and networking components. The service provider will own the equipment and he is responsible for running and maintaining it.

Hardware as a Service ( HaaS ) is an another term for IaaS.


Fig 2.6 HaaS

**Characteristics and components of IaaS**

- Utility computing service.
- Billing model.
- Automation of administrative tasks.
- Dynamic scaling.
- Desktop virtualization.
- Policy based services.
- Internet connectivity.

### 2.4.1 IaaS service providers

IaaS providers are the organizations who provide the most basic needs in IT like servers, networking and storage based on payment models. They typically make a big investment in their business like data centers and other hardware for renting to users.

Some of the IaaS providers are :

- Amazon
- At & T
- Blue Lock
- Cloud Scaling
- Data pipe
- Go GRID
- Verizon

### 2.4.2 Amazon EC2

Amazon EC2 is an acronym for Amazon Elastic Compute Cloud. It is a central part of Amazon.com's cloud computing platform where it is an American multinational electronic e-commerce company which is a major service provider in cloud computing services.

EC2 allows the users to rent virtual computers on which run their applications. Scalable deployment of applications are provided by EC2 for web services by which a subscriber can create a virtual machine.A user can create, launch and terminate server instances as needed by paying. That is why it is defined as “elastic”. EC2 provides users with control over geographical location of instances that allows high levels of redundancy.


Fig 2.7 Amazon EC2 services.

**EC2 Functionality**

EC2 presents a true virtual computing environment. This allows the users to use web service interface to launch instances with a variety of operating systems.

It can be loaded with the customized application environment, manage the access permission and run the image for any number of systems.

**Features of EC2**

Some of the features of EC2 are :

- Amazon elastic book store
- EBS optimized instances
- Multiple locations
- Elastic load balancing
- Amazon cloud watch.

### 2.4.3 GoGRID

GoGRID is a cloud infrastructure service. It hosts Windows and Linux virtual machines manages by a multi server control panel and a RESTful API. Representation State Transfer ( REST ) has emerged as a important web API design model. It is privately held service. It is used to provide and scale virtual and physical servers, storage, networking, load balancing and firewall in real time across the multiple data centers using Go GRID's API.

Its infrastructure is used when we need instant access to highly available multi server environments. It can be accessed and operated by the standard network protocols and IP address.

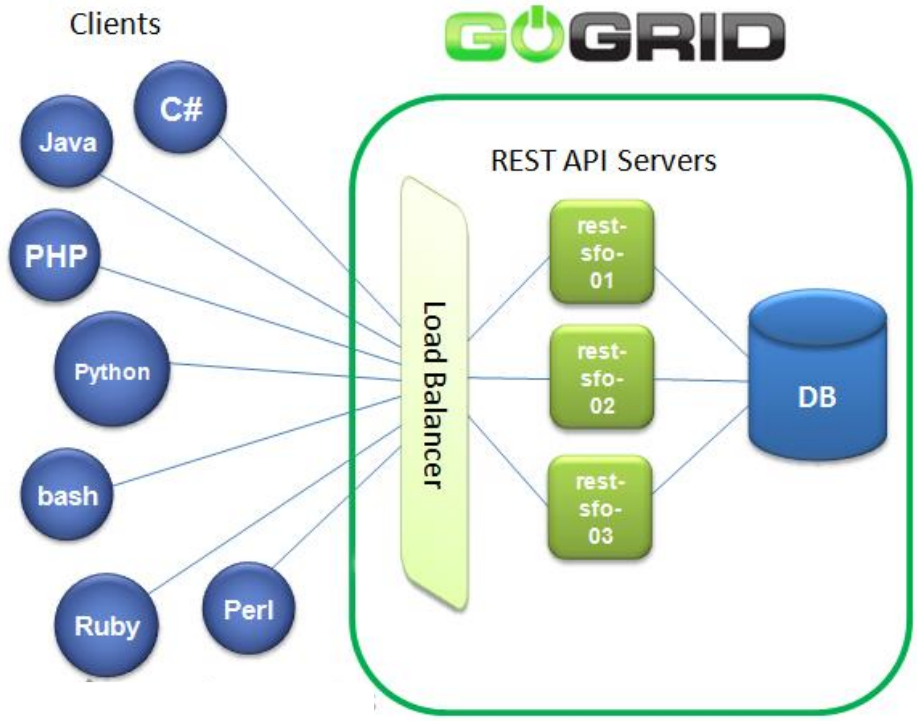



Fig 2.7 GoGrid

**Advantages of Go Grid**

- Large volume of data
- High availability
- Scalability
- E – commerce
- Development
- Testing

### 2.4.4 Microsoft implementation and support

Microsoft has a very extensive cloud computing portfolio under active development. Efforts to extend Microsoft products and third-party applications into the cloud are centered around adding more capabilities to existing Microsoft tools.

Microsoft's approach is to view cloud applications as software plus service. In this model, the cloud is another platform and applications can run locally and access cloud services or run entirely in the cloud and be accessed by browsers using standard Service Oriented Architecture (SOA) protocols.

Microsoft calls their cloud operating system the Windows Azure Platform. You can think of Azure as a combination of virtualized infrastructure to which the .NET Framework has been added as a set of .NET Services.

The Windows Azure service itself is a hosted environment of virtual machines enabled by a fabric called Windows Azure AppFabric. You can host your application on Azure and provision it with storage, growing it as you need it. Windows Azure service is an Infrastructure as a Service offering.

A number of services interoperate with Windows Azure, including SQL Azure (a version of SQL Server), SharePoint Services, Azure Dynamic CRM, and

many of Windows Live Services comprising what is the Windows Azure Platform, which is a Platform as a Service cloud computing model.

Eventually, many more services will be added, encompassing the whole range of Microsoft's offerings. This architecture positions Microsoft to either extend its product into the Web or to license its products, whichever way the cloud computing marketplace develops. From Microsoft's position and that of its developers, Windows Azure has lots of advantages.

Windows Live Services is a collection of applications and services that run on the Web. Some of these applications called Windows Live Essentials are add-ons to Windows and downloadable as applications. Other Windows Live Services are standalone Web applications viewable in a browser.

An important subset of these Windows Live Services is available to Windows Azure applications through the Windows Live Messenger Connect API. A set of Windows Live for Mobile applications also exists.

## Exploring Microsoft Cloud Services

Microsoft has a vast array of cloud computing products and initiatives, and a number of industry-leading Web applications.

Going forward, Microsoft sees its future as providing the best Web experience for any type of device, which means that it structures its development environment so the application alters its behavior depending upon the device. For a mobile device, that would mean adjusting the user interface to accommodate the small screen, while for a PC the Web application would take advantage of the PC hardware to accelerate the application and add richer graphics and other features.

That means Microsoft is pushing cloud development in terms of applications serving as both a service and an application. This duality—like light, both a particle and a wave—manifests itself in the way Microsoft is currently structuring its Windows Live Web products. Eventually, the company intends to create a Microsoft app store to sell cloud applications to users.

Microsoft Live is only one part of the Microsoft cloud strategy. The second part of the strategy is the extension of the .NET Framework and related development tools to the cloud. To enable .NET developers to extend their applications into the cloud, or to build .NET style applications that run completely in the cloud, Microsoft has created a set of .NET services, which it now refers to as the Windows Azure Platform. .NET Services itself had as its origin the work Microsoft did to create its BizTalk products.

Azure and its related services were built to allow developers to extend their applications into the cloud. Azure is a virtualized infrastructure to which a set of additional enterprise services has been layered on top, including:

• A virtualization service called Azure AppFabric that creates an application hosting environment. AppFabric (formerly .NET Services) is a cloud-enabled version of the .NET Framework.

• A high capacity non-relational storage facility called Storage.

• A set of virtual machine instances called Compute.

• A cloud-enabled version of SQL Server called SQL Azure Database.

• A database marketplace based on SQL Azure Database code-named “Dallas.”

• An xRM (Anything Relations Management) service called Dynamics CRM based on Microsoft Dynamics.

• A document and collaboration service based on SharePoint called SharePoint Services.

• Windows Live Services, a collection of services that runs on Windows Live, which can be used in applications that run in the Azure cloud.

Eventually the entire Microsoft server portfolio will be available as a cloud-based application or service, including Exchange. So the Windows Azure Platform can be viewed in a sense as the next Microsoft operating system, the first one that is a cloud OS.

The applications developed in Visual Studio or through PHP and other languages deployed to the cloud, existing local (on-premises) applications interacting with Azure with standard SOA protocols (SOAP, REST, and XML), all running on the Windows Azure virtualized infrastructure.

### 2.4.5 Amazon EC Service Level Agreement

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers.

An SLA serves as both the blueprint and warranty for cloud computing.

This Amazon EC2 Service Level Agreement ("SLA") is a policy governing the use of Amazon Elastic Compute Cloud ("Amazon EC2") and Amazon Elastic Block Store ("Amazon EBS") under the terms of the Amazon Web Services

Customer Agreement (the “AWS Agreement”) between Amazon Web Services, Inc. and its affiliates (“AWS”, “us” or “we”) and users of AWS’ services (“you”).

This SLA applies separately to each account using Amazon EC2 or Amazon EBS.

### 2.4.6 Recent developments

Recent development of Infrastructure as a Service ( IaaS ) is Dynamic Load Balancer by Go GRID. It is a cloud based load balancing solution to manage the high availability infrastructure. Using this, customers can deploy and scale network services dynamically to support essential infrastructure. This highly available solution helps in business control spending and easy maintenance.This provides the elasticity and on demand control needed to efficiently manage cloud.

Recent development in IBM is Advanced enterprise ready IaaS ( Infrastructure as a Service ). This delivers a range of security rich, enterprise cloud services. Based on open standards, IBM IaaS features advanced cloud management.

Recent development in Microsoft is Windows Azure. This enables us to extend our data centers into the cloud while using our current windows resources. We can build applications using any language, tool or framework.

### 2.4.7 Benefits

Some of the advantages of Infrastructure as a Service ( IaaS ) are,

- Rapid scaling
- Completely managed cloud solutions
- Security
- Easy to use
- Metered services
- Flexibility

**Rapid Scaling**

It can be scaled both horizontally and vertically in a very short period of time. Horizontal scaling provides unique separate environment to private cloud. Vertical scaling provides additional hardware resources. IaaS services can be easily scaled depending upon the user's need.

**Completely managed cloud solution**

It does not need costlier infrastructure investment. This model provides more advantages for companies with limited investment in the computing resources.

**Security**

It provides high level security for every environment.

**Easy to use**

It is very much easy to use the web portal and API. It provides plug and play integration with existing infrastructure and networks.

**Metered services**

The service usages are measured and charged on the number of units consumed ( used ). It follows Pay for what you use and when you use.

**Flexibility**

It can be accessed from anywhere, anytime and on any device.

## 2.5 Cloud deployment model

Cloud computing is a fast improving technology which offers an new way to provide software, data storage and computing services.

The National Institute for Science and Technology ( NIST ) has define four different cloud computing deployment models. They are,

1. Private cloud
2. Community cloud

3. Public cloud
4. Hybrid cloud

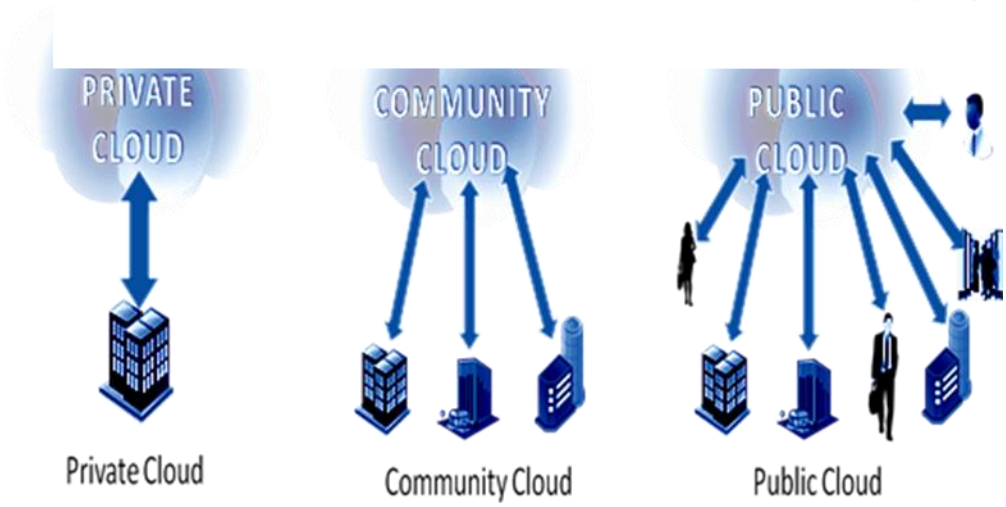



Fig 2.7 Types of cloud

**2.5.1 Public cloud**

In public cloud hosting apps, storage and other resources are made available to the public by service provider. These services are provided as pay per use model or for free. Most popular public cloud providers are Amazon, Google and Microsoft. Its services can be accessed through internet. Direct contact with the service provided is not offered in public cloud.

**2.5.2 Private cloud**

Private cloud is a cloud infrastructure that is operated particularly for a single organization. It can be managed internally or by a third party. It can be hosted internally or externally.

It delivers the services to single organization. This model shares many characteristics of traditional client server architecture.Like any other cloud model services are delivered on demand .

In this the resources can be managed inside the organization or by third party. It provides more security and privacy.Hosting in public, community and

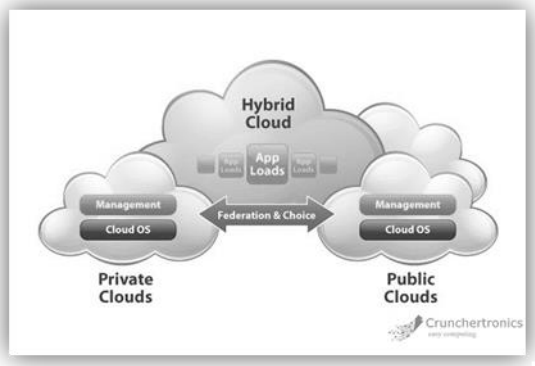
hybrid cloud.As we seen earlier about private cloud, we are now moving on to other types of cloud hosting.

### 2.5.3 Community cloud

In community cloud hosting, the infrastructure is shared among the number of organizations with similar interests and requirements. The cost for the services are spread over few users.

The number of subscribers for community cloud is less than public cloud but more than private cloud. This can be managed third party or by thyself. Community cloud can be hosted internally or externally.

### 2.5.4 Hybrid cloud:




The hybrid cloud is a combination of two or more cloud models such as private, public or community. It has unique features of every cloud hosting models bound together. It offers benefits of multiple deployment models. Hybrid clouds use cloud bursting for scaling across clouds.

The main advantage of cloud bursting and hybrid cloud, is that an organization only pays for extra resources when needed. Hybrid cloud lacks in flexibility, security and certainty of in-house applications.

### 2.5.5 Advantages of cloud computing

**Scalability:** The ability of a model to be extended to manage the amount

of work growth in an effective manner is called scalability. Cloud-computing resources can be rapidly scaled according to subscriber's convenience. If there is a sudden necessity for more computer resources, instead of buying new equipment we can buy additional resources form cloud providers.

**Simplicity:** In most cases cloud-computing is free to use. It is very simple that users can easily understand which is the biggest advantage of cloud-computing. It is possible to get our application started instantly.

**Vendors:** The service providers are called vendors. Some of the well known vendors are Google, Amazon, Microsoft, IBM. These providers offer reliable services to their customers.

**Security:** There are also some risks when using a cloud vendor. But the reputed firms work hard to keep their consumers data safe and secure. They use complex cryptographic algorithms to authenticate users. To make it even more secure we can encrypt our information before storing it in cloud.

**Flexibility and Resiliency :** A major benefit of cloud computing is the flexibility that is provided by the following:

- Freedom from concerns about updating servers
- Freedom from having to install software patches
- Automated provisioning of new services and technologies
- Acquiring increased resources on an as-needed basis
- Ability to focus on innovation instead of maintenance details
- Device independence

**Reduced Costs :**The cloud paradigm, in general, is a basis for cost savings because capability and resources can be paid for incrementally without

the need for large investments in computing infrastructure. This model is especially true for adding storage costs for large database applications. Therefore, capital costs are reduced and replaced by manageable, scalable operating expenses. Pay per use model

**Reduced Time to Deployment:** In a competitive environment where rapid evaluation and development of new approaches is critical, the cloud offers the means to use powerful computational resources in a short time frame and large amounts of storage without requiring sizeable initial investments in hardware, software, and personnel. This rapid provisioning can be accomplished at relatively small cost and offers the client access to advanced technologies that are constantly being acquired by the cloud provider. Improved delivery of services obtained by rapid cloud provisioning improves time to market and market growth.

* * *

**Summary**

- ✓ Cloud architecture is the design of software application that uses internet access and on-demand service.
- ✓ Cloud delivery model includes Iaas,Paas,Saas.
- ✓ Web operating systems are interfaces to distributed computing systems, particularly cloud or utility computing systems.
- ✓ An SLA serves as both the blueprint and warranty for cloud computing.
- ✓ Cloud Deployment Model – the different types of clouds are Public cloud, private cloud, Community cloud and Hybrid Cloud.
- ✓ Advantages of Cloud Computing: Scalability, Simplicity, Vendors, Security, Flexibility and Resiliency ,Reduced Costs ,Reduced Time to Deployment.

# Review Questions

## Part-A (2 marks)

1. What is Cloud Architecture?
2. What is SPI?
3. What is IaaS?
4. Expand PaaS and SaaS.
5. What are the benefits of SaaS?
6. State any 2 service providers of SaaS.
7. State any 2 service providers of PaaS.
8. State any 2 service providers of IaaS.
9. What is Public Cloud?
10. What is Private Cloud?
11. What is SPI?
12. What is salesforce.com?
13. What is cloud delivery model?
14. What is other name for IaaS?
15. Define Hybrid cloud.
16. Give examples of PaaS.

## Part-B (3 marks)

1. What is Web Operating System?
2. What is Amazon EC2?
3. What isSalesforce.com?
4. What is Google App Engine?
5. Mention the services of PaaS.
6. What is community clouds?

Part-C (10 marks)

1. Explain the SPI framework in detail.
2. What are the benefits of SaaS?
3. Explain the criteria in evaluating SaaS.
4. Write short notes on
   (a) Web Services.
   (b) Web Operating System.
5. Write short notes on GoGrid.
6. Explain the types of Cloud.
7. Differentiate the SPI vs Traditional IT model.
8. Describe the types of cloud deployment model.
9. Explain in detail about cloud delivery model.
10. Discuss the operational and economic benefits of SaaS.

* * *

# Unit – III

# Virtualization

**OBJECTIVES:**

At end the this unit, students can

- Define Virtualization
- Explain Virtualization and Cloud Computing
- State the need and limitations of Virtualization
- State the types of Hardware Virtualization
- Explain Full, partial and para Virtualization
- Explain Desktop, Software, Memory, Storage, Data and Network Virtualization
- State Microsoft Implementation
- Explain Microsoft Hyper V
- Explain VMWare features and infrastructure
- Explain Virtual Box and Thin Client

**INTRODUCTION:**

Virtualization is the creation of a virtual (rather than actual) version of something, such as an operating system, a server, a storage device or network resources. Imagine, Google has built a huge DataCentre and they reach a point where they cannot add more resources but, they do not want to loose their customers, then they may choose to virtualize their DataCentre (physical resources) to overcome this issue or even accommodate more users.

## 3.1 Virtualization

Virtualization is the key component of cloud computing for providing computing and storage services. Virtualization is the ability to run multiple operating systems on a single physical system and share the underlying hardware resources. It is the process by which one computer hosts the appearance of many computers.

### 3.1.1 Virtualization and cloud computing

Virtualization is very important for cloud computing. It benefits the cloud computing for scalability. Because each virtual server is allocated only enough computing power and storage capacity that the client needs more virtual servers can be created. Without Virtualization, cloud computing would not exist.

Virtualization and cloud computing were both developed to maximize the use of computing resources while streamlining processes and increasing efficiencies to reduce the total cost. But virtualization and cloud computing are truly very different approaches.

Virtualization software allows one physical server to run several individual computing environments. This technology is fundamental to cloud computing. Cloud providers have large data centers full of servers to power the cloud offerings but they are not able to devote a single server to each customer. Thus they virtually partition the data on the server, enabling each client to work with the separate virtual instance of some software.

Cloud computing encompasses virtualization. It gives access to complex applications and large computing resources via internet.

A virtualized server makes better use of the server's available capacity than a non-virtualized server. Each virtual machine can run its own operating system as well as any business applications as needed. It can also be applied to storage hardware.

With cloud computing we can implement enterprise grade application. In order to get the service, it can be chosen from a variety of cloud computing providers and cloud based services. It is used for big applications.

Thus both virtualization and cloud computing operate on a one to many model. Virtualization can make one computer to perform like many separate computers and cloud computing allows many different companies to access one application. The virtualization is employed locally while cloud computing is accessed as a service.

### 3.1.2 Need for Virtualization

Virtualization is needed for the following reasons

#### Costs

With virtualization, administration becomes a lot easier, faster and cost effective. Virtualization lowers the existing cost.

It dramatically simplifies the ownership and administration of their existing IT servers. The operational overhead of staffing, backup, hardware and software maintenance has become verify significant in IT budgets and business. In such case virtualization reduces these costs beneficially. By using virtualization we can save the operational costs.

Virtualization concentrates on increasing utilization and consolidation of equipment thereby reducing capital costs, cabling, operational costs such as power, cooling, maintenance cost of hardware and software.

Thus Virtualization is cost effective.

#### Administration

Administrating virtualization has to be done in efficient manner since all the resources are centralized security issues has to be categorized more sensitively. The users access the resources like data storage, hardware or software has to be allocated properly. Since more users will utilize the resources, the sharing of needed resources is complicated.

Administration of virtual server is done through virtual server administration website. By using this virtual server is assigned to application for access.

In this, virtual IP addresses are configured on the load balancer. When a request is sent from the user from certain port on a virtual IP load balancer, it distribute the incoming request among multiple server the needed service will be provided to particular user.

## Fast Deployment

Deployment of consolidated virtual servers, migrating physical, servers has to be done.

Virtualization deployment involves several phases and planning. Both server and client systems can support several operating systems simultaneously, virtualization providers offer reliable and easily manageable platform to large companies.

It can be built with independent, isolated units which work together without being tied to physical equipment. Virtualization provides much faster and efficient way of deployment of services by some third party software like VMware, Oracle etc. Thus it provides the fastest service to the users.

## Reducing Infrastructure Cost

Virtualization essentially allows one computer to do the job of the multiple computers by means of sharing the resources of a single computer across multiple environments.

Virtual servers and virtual desktops allow hosting multiple operating systems and multiple applications locally and in remote locations. It lowers the expense by efficient use of the hardware resources.

It increases utilization rate for server and cost savings efficiently by altering the physical resources by virtual sharing.

## Some other reasons are

1. To run old Apps
2. To access virus infected Data
3. To safely browse
4. Test software, upgrades or new configurations
5. To run Linux on top of Windows
6. To backup a entire operating system
7. To create a personal cloud computer
8. To reuse old hardware

## Limitations

Some of the limitations of virtualization are

1. If the CPU does not allow for hardware virtualization we can run some operating system in software virtualization but it is generally slower. Some operating system will not run in software virtualization and require to have CPU with hardware virtualization so it would cost more if CPU with hardware virtualization is not possible.

2. If we want a own server and intend to resell a virtual server then it cost high. This mean purchase of 64 bit hardware with multiple CPU's and multiple hard drives.
3. Some of the limitations are in analysis and planning which problems can be divided into three types they are
    a. Technical limitation
    b. Marketing strategies
    c. Political strategies
4. It has a high risk in physical fault.
5. It is more complicated to set up and manage virtual environment with high critical servers in a production environment. It is not easy as managing physical servers.
6. It does not support all applications.

These are some of the limitations of virtualization in cloud computing.

### Types of virtualization are:

- Hardware virtualization
- Desktop virtualization
- Software virtualization
- Memory virtualization
- Storage virtualization
- Data virtualization
- Network virtualization

## 3.2 Types of hardware virtualization

Hardware virtualization or platform virtualization refers to the creation of a virtual machine that acts like a real computer with an operating system. Software executed on these virtual machines is separated from the underlying hardware resources.

In hardware virtualization, the host machine is the actual machine on which the virtualization takes place and the guest machine is the virtual machine. The words host and guest are used to distinguish the software that runs on the physical machine from the software that runs on the virtual machine. The software that creates a virtual machine on the host hardware is called hypervisor or virtual machine manager.

Different types of hardware virtualization are :

1. Full virtualization
2. Partial virtualization
3. Para virtualization

### 3.2.1 Full virtualization

It is a virtualization technique used to provide a virtual machine environment. Full virtualization requires that every features of the hardware can be reflected into one of several virtual machines which include the full instruction set, I/O operations, interrupts,

memory access and all the elements used by the software that runs on the base machine which will run in the virtual machine. In such case any software capable of executing on that hardware can be run in mutual machine.

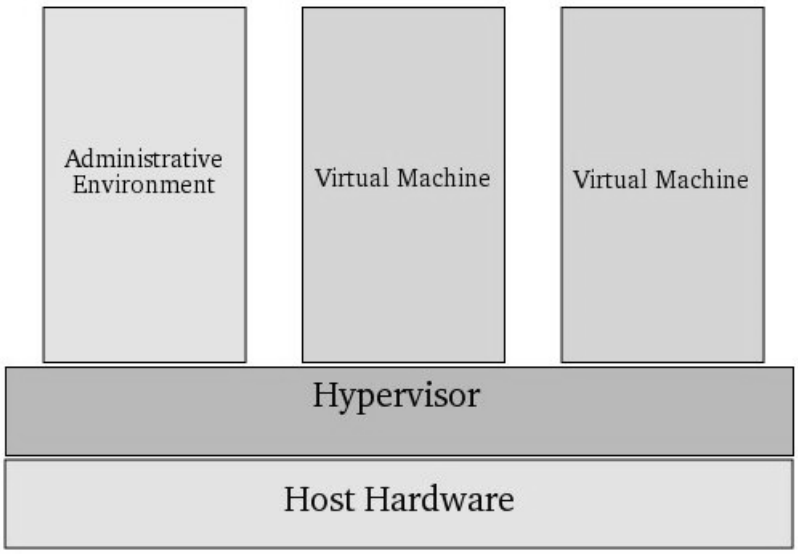
Full virtualization is possible only with the right combination of hardware and software.

The obvious test of virtualization is whether an operating system intended for stand-alone use can successfully run in a virtual machine.

The effects of the every operation performed within a given virtual machine must be kept within that virtual machine. Virtual operation cannot be allowed to alter the state of any other virtual machine, the control program or the hardware.

Full virtualization is successful for

- Sharing a computer system among multiple users.
- Isolating users from each other
- Emulating new hardware to achieve improved reliability, security and productivity.




Virtualization features built into the latest generations of CPU's from the technologies, known as Intel VT and AMD-V respectively, provide extensions necessary to run unmodified guest virtual machines without the internet in full virtualization. CPU emulation Hypervisor can operate essentially leaving ring 0 available for unmodified guest operating systems. Hypervisor based virtualization solutions include Xen, VMWare ESX server and Microsoft's Hyper-V technology.

### 3.2.2 Partial virtualization

Partial virtualization is a virtualization technique used to implement a certain kind of virtual environment. One that provides a partial simulation of the underlying hardware environment particularly address spaces. Using this entire operating system cannot run in the virtual machine which would be the sign of full virtualization but that many applications

can run. A key form of partial virtualization is address space virtualization in which each virtual machine consists of an independent address space. This capability address relocation hardware is present in most of partial virtualization. It is the milestone to full virtualization.

### 3.2.3 Para virtualization

Under para virtualization the kernel of the guest operating system is modified to run on the hypervisor. Hypervisor performs the task instead of guest kernel.

It is very difficult to build the more sophisticated binary translation support necessary for full virtualization.

It involves modifying the OS kernel to replace non-virtualizable instructions with hypercalls that communicate directly with the virtualization layer hypervisor.

The hypervisor also provides hypercall interfaces for other critical kernel operations such as memory management, interrupt handling and time keeping.

Para virtualization is different from full virtualization, where the unmodified OS does not know it is virtualized and sensitive OS calls are trapped using binary translation.

Para virtualization cannot support unmodified OS.

Example: Xen-modified Linux Kernel and a version of Windows XP.

## 3.3 Desktop virtualization

It is a software technology that separates the desktop environment and associated application software from the physical client device that is used to access it.

### 3.3.1 Software virtualization

It is the virtualization of applications or computer programs. One of the most widely used software virtualization is Software Virtualization Solution (SVS) which is developed by Altris.

It is similar to hardware which is simulated as virtual machines. Software virtualization involves creating a virtual layer or virtual hard drive space where applications can be installed. From this virtual space, the application can be run as they have been installed onto host OS.

Once user finished using application, they can switch it off. When a application is switched off, any changes that the application made to the host OS will be completely reversed. This means that registry entries and installation directories will have no trace of the application being installed, executed at all.

Benefits of software virtualization are,

- The ability to run applications without making permanent registry or library changes.
- The ability to run multiple versions of the same application.
- The ability to install applications that would otherwise conflict with each other.

- The ability to test new applications in an isolated environment.
- It is easy to implement.

### 3.3.2 Memory virtualization

Virtual memory is a feature of an operating system that enables a process to use a memory (RAM) address space that is independent of other processes running in the same system. It uses the space that is larger than the actual amount of RAM. Virtual memory enables each process to act as it has the whole memory space to itself, since the address that it uses to reference memory are translated by the virtual memory mechanism into different addresses in physical memory. This allows different processes to use same memory address. The purpose of virtual memory is to enlarge the address.

Memory virtualization decouples volatile random access memory (RAM) resources form the individual systems in the data center, and then aggregates those resources into a virtualized memory in the cluster. The memory pool is accessed by the operating system or applications running on the top of the operating system.

Memory virtualization allows networked and therefore it can be distributed. So that servers to share a pool of memory to overcome physical memory limitations. It is integrated into the network, applications of memory to improve overall performance, system utilization, increases memory usage efficiency and enables new uses cases.

Memory virtualization is different from shared memory. Shared memory systems do not permit abstraction of memory resources, thus it requires implementation with a single operating system instance.

#### Benefits:

- Improves memory utilization via the sharing of scarce resources.
- Increases efficiency and decreases run time for data and I/O bound application.
- Allows applications on multiple servers to share data without replication; decreasing total memory needs.
- Lowers the latency and provides faster access than other solutions.

### 3.3.3 Storage Virtualization

Storage virtualization is the pooling of physical storage from multiple network storage devices into what appears to be single storage device that is managed from a central console. Storage virtualization is commonly used in storage area networks. Storage Area Network is a high speed sub network of shared storage devices and makes tasks such as archiving, back-up and recovery easier and faster. Storage virtualization is usually implemented via software applications.

Storage systems typically use special hardware and software along with disk drives in order to provide very fast and reliable storage for computing and data processing.

Storage systems can provide either block discussed storage or file accessed storage. In storage system there are two primary types of virtualization. They are, Block virtualization and file virtualization. Block virtualization is used to separate the logical storage from

physical storage. So that it may be accessed without regard to physical storage. The separation allows the administrator of the storage system with greater flexibility in how they manage the storage of end users.

File virtualization eliminates the dependences between the data accessed at the file level and the location where the files are physically stored. This provides opportunities to optimize storage use and server consolidation.

### Benefits

- Ability of data migration or data transfer
- Improved utilization
- Central management

### 3.3.4 Data virtualization

It describes that process of abstracting desperate systems like database, application, the repositories websites, data services etc. Through a single data access layer which may be any of several data access mechanism.

This abstraction enables data access clients to target a single data access layer, serialization, formal structure etc. rather than making each client tool handle multiples of any of these. This data virtualization is often used in data integration, business intelligence, service oriented architecture, data services etc.

Data virtualization is a technology which provides some of these capabilities.

- Abstraction
- Virtualized Data Access
- Transformation
- Data federation
- Flexible data delivery

### 3.3.5 Network virtualization

Network virtualization is the process of combining hardware and software network resources and network functionality into a single, software based administrative entity which is said to be virtual network. Network virtualization involves platform virtualization.

Network virtualization is categories into external network virtualization and internal network virtualization.

External network virtualization is combining of many networks into a virtual unit.

Internal network virtualization is providing network like functionality to the software containers on a single system.

Network virtualization enables connections between applications, services, dependencies and end users to be accurately emulated in the test environment.

## 3.4 Microsoft Implementation

### 3.4.1 Microsoft Hyper – V

Hyper-V is code named viridian and formerly known as windows server virtualization. It is a native hypervisor that enables platform virtualization. Hyper-V has been released in a free standalone version. Hyper-V exists in two variants. They are standalone product called Microsoft Hyper-V 2012 and Microsoft Hyper-V Server 2008.

With any virtualization platform, Hyper-V makes for a more efficient data center, maximizing resources and reducing costs. Hyper-V provides end to end functionality for an enterprise grade virtualization product. It provides the basic functionality to create a virtualization layer over the physical layer of the host server machine and enables guest operating systems to be installed and managed through an integrated management console.

Hyper-V isolates part of a physical machine into child partitions and allocates them to different guest operating systems with Windows Server 2008 which act as a primary host/parent.

Hyper-V also assigns appropriate hardware and software resources for each of the guest operating system it's hosting because they don't have direct access to the computer hardware resources.

#### Benefits

- It is cost effective
- It improves scalability and performance
- It has a better performance

Hyper-V has three major components. They are,

- Hyper-V cloud deployment guide
- Hyper-v cloud Fast Tract
- Hyper-V cloud service providers

### 3.4.2 VMWare features

VMWare is an American software company that provides cloud and virtualization software and services. It was founded in 1998 and based in Palo Alto, California, USA. The company was acquired by EMC Corporation in 2004 and now operates as a separate software subsidiary.

VMWare's desktop software runs on Microsoft Windows, Linux and Mac OS X and VMWare ESX are embedded hypervisors that run directly on server hardware without requiring and additional operating system.

VMWare virtualization lets us to

- Run multiple operating systems and applications on a single computer.
- It consolidates hardware to get higher productivity of new applications.

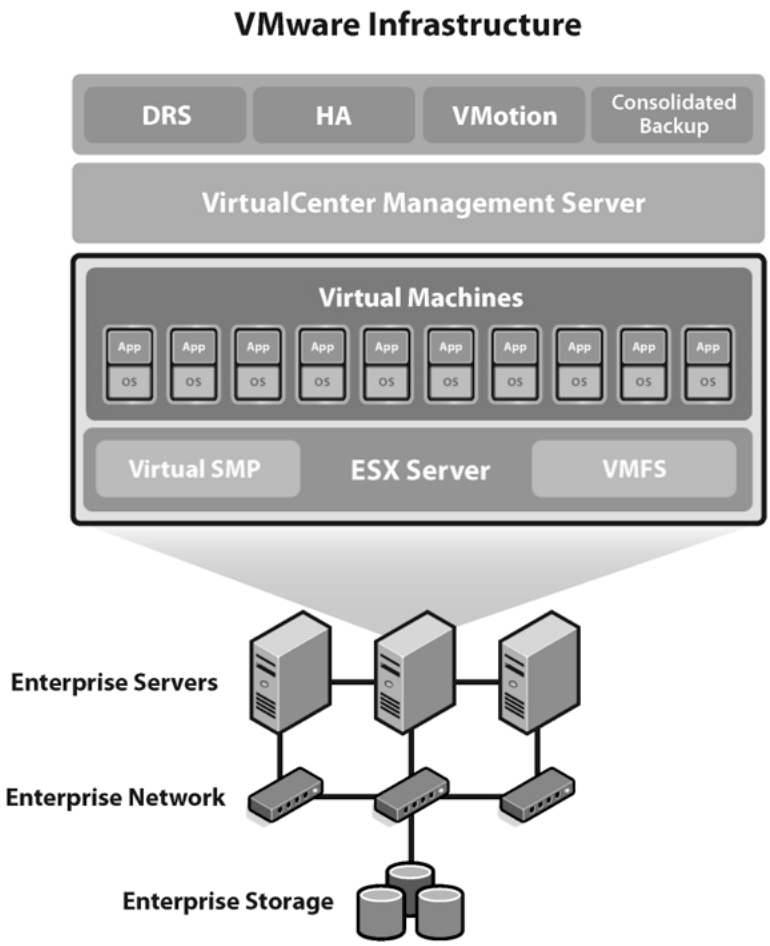
- It saves 50% or more on overall IT costs.
- It speeds and simplify IT management, maintenance and deployment of new applications.

## VMWare infrastructure

VMWare infrastructure includes the following components are

- **VMWare ESX Server**

A production-proven virtualization layer run on physical servers that abstract processor, memory storage and networking resources to be provided to multiple virtual machines.




- **VMWare virtual Machine File System (VMFS)**

A high performance cluster file system for virtual machine.

- **VMWare virtual symmetric Multi-processing (SMP)**

Enables a single virtual machine to use multiple physical processors simultaneously.

- **Virtual Center Management Server**

The central point for configuring, provisioning and managing virtualized IT infrastructure.

- **Virtual Infrastructure Client (VI Client)**

The central point for configuring, provisioning and managing virtualized IT infrastructure.

- **Virtual Infrastructure Web Access**

A web interface for virtual machine management and remote consoles access.

- **VMWare VMotion**

Enables the live migration of running virtual machines form one physical server to another with zero downtime continuous service availability and complete transaction integrity.

- **VMWare High Availability (HA)**

It provides easy to use, cost-effective high availability for applications running in mutual machines. In the event of server failure, affected mutual machines are automatically restarted on the other server.

- **VMWare Distributed Resource Scheduler (DRS)**

Intelligently allocates and balances computing capacity dynamically across collections of hardware resources for virtual machines.

- **VMWare Consolidated Backup**

It provides easy usage, centralized facility for agent feedback of virtual machines. It simplifies backup administration and reduces the load on ESX server installations.

- **VMWare Infrastructures (SDK)**

It provides a standard interface for VMWare and third-party solutions to access VMWare structure.

### 3.4.3 Virtual Box

Virtual box is originality developed by Innotek Gambit and related in 2007 as an open source software package. It is purchased by Sun Microsystem.

A virtual box is a software virtualization package that installs on an operating system as an application. Virtual box allows additional operating systems to be installed on it, as a guest OS and run in a virtual environment.

It will share the RAM and CPU power of other OS. If Linux is running and install virtual box, then it is possible to install Windows and run it.

Virtual box was the most popular, virtualization software application. Separating systems which are supported are Windows XP, Vista, Windows 7, Mac OS X, Linux, Solaris and Open Solaris.

Oracle Corporation is developing the software package with little Oracle VM virtual box.

### 3.4.4 Thin Client

It is a computer or a computer program which depends heavily on some other computer to fulfill its computational roles. Thin client is designed to be especially small processing occurs on the server.

Thin client is increasingly used for computers such as network computers which are designed to serve as the client for Client/Server architecture. A thin client is a network computer without a hard disk drive, whereas a fat client includes a disk drive.

### Types of Thin client

- Thin client – Micro range
- Citrus Thin client
- PC Thin client
- Windows Server 2012
- Wyse Thin client

### Uses

Thin clients are used where a lot of people need to use computers. This includes public places like libraries, airports and schools. Thin client setup is also popular in places where people need to be able to save and access information from a central location like an office, a call center.

### Summary:

- ✓ Virtualization is the key component of cloud computing.
- ✓ Virtualization and cloud computing were both developed to maximize the use of computing resources.
- ✓ Virtualization is needed for cost, administration, fast deployment, reduce infrastructure cost.
- ✓ Virtualization has some limitations also.
- ✓ Three types of Hardware Full virtualization: Full, Partial and Para virtualization
- ✓ Desktop virtualization is a software technology that separates the desktop environment and associated application software.
- ✓ Software virtualization is the virtualization of applications or computer programs.
- ✓ Memory virtualization is Virtual memory and it is a feature of an operating system.
- ✓ Storage virtualization is the pooling of physical storage from multiple network storage devices into a single storage device.
- ✓ Data virtualization describes the process of abstracting desperate systems.
- ✓ Network virtualization is the process of combining hardware and software network resources.
- ✓ Microsoft Hyper-V is code named viridian and formerly known as windows server virtualization.

- ✓ VMWare is an American software company that provides cloud and virtualization software and services.
- ✓ VMWare infrastructure includes the many components: VMWare ESX Server, VMWare virtual Machine File System (VMFS), VMWare virtual symmetric Multi-processing (SMP), Virtual Center Management Server, Virtual Infrastructure Client (VI Client), Virtual Infrastructure Web Access, VMWare VMotion, VMWare High Availability (HA), VMWare Distributed Resource Scheduler (DRS), VMWare Consolidated Backup, VMWare Infrastructures (SDK)
- ✓ A virtual box is a software virtualization package that installs on an operating system as an application.
- ✓ Thin Client is a computer or a computer program which depends heavily on some other computer to fulfill its computational roles.

---

## Review Questions

### Part – A

1. Define virtualization
2. Write any two needs for virtualization
3. List any two limitations of virtualization
4. List out the types of hardware virtualization
5. Define Desktop virtualization
6. Write any two benefits of software virtualization
7. Define memory virtualization
8. Define storage virtualization
9. What is Data virtualization
10. What is Microsoft Hyper-V ?
11. Define virtual box
12. What is Thin client ?

### Part – B

1. Write about the limitations of virtualization
2. Write short notes on Partial virtualization
3. Write short notes on Para virtualization
4. Write short notes on Network virtualization
5. Write short notes on Virtual Box
6. Write short notes on Thin Client and its types

### Part – C

1. Differentiate virtualization and cloud computing
2. Explain the types of hardware virtualization with its benefits

3. Explain software and memory virtualization with its benefits
4. List out the VMWare features
5. Explain VMWare infrastructure with a neat diagram

---

# Unit – IV

# Storage Management

**OBJECTIVES:**

At end the this unit, students can

- Define Storage Network
- Explain the Architecture of storage Network
- Explain the analysis and planning of storage Network
- State the Storage network design considerations
- Explain NAS and FC SANs
- Explain hybrid storage networking technologies
- Explain ISCSI, FCIP, FCoE
- Explain design for storage virtualization in cloud computing
- Explain File systems or object storage

**INTRODUCTION:**

Many storage management technologies, like storage virtualization, deduplication and compression, allow companies to better utilize their existing storage. The benefits of these approaches include lower costs both the one-time capital expenses associated with storage devices and the ongoing operational costs for maintaining those devices.

## 4.1 Storage Network

Storage networking is the practice of linking together storage devices and connecting them to other IT networks. Storage networks provide a centralized repository for digital data that can be accessed by many users, and they use high-speed connections to provide fast performance.

The phrase "storage networking" is commonly used in reference to storage area networks (SANs). A SAN links together multiple storage devices and provides block-level storage that can be accessed by servers.

### 4.1.1 Architecture of storage, analysis and planning

Cloud storage is a model of data storage in which the digital data is stored in logical pools, the physical storage spans multiple servers (and often locations), and the physical environment is typically owned and managed by a hosting company. These cloud storage providers are responsible for keeping the data available and accessible, and the physical environment protected and running. People and organizations buy or lease storage capacity from the providers to store user, organization, or application data.

Cloud storage services may be accessed through a co-located cloud computer service, a web service application programming interface (API) or by applications that utilize

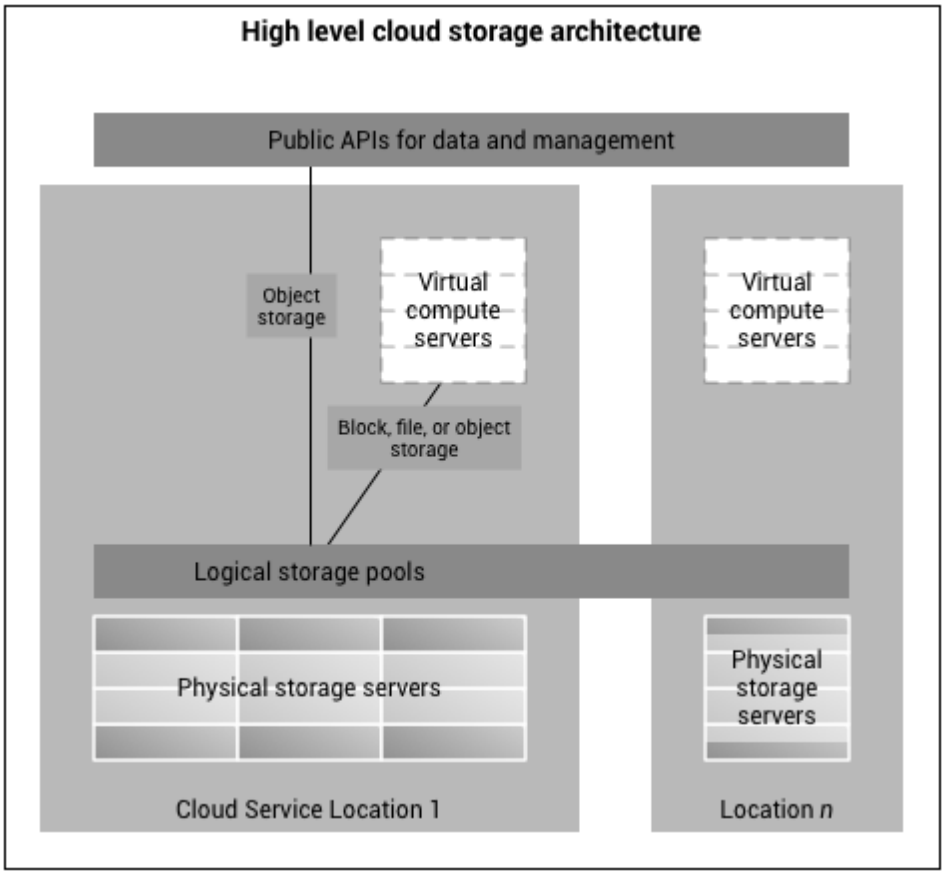
the API, such as cloud desktop storage, a cloud storage gateway or Web-based content management systems.

## Architecture

Cloud storage is based on highly virtualized infrastructure and is like broader cloud computing in terms of accessible interfaces, near-instant elasticity and scalability, multi-tenancy, and metered resources. Cloud storage services can be utilized from an off-premises service (Amazon S3) or deployed on-premises (ViON Capacity Services).

Cloud storage typically refers to a hosted object storage service, but the term has broadened to include other types of data storage that are now available as a service, like block storage.

Object storage services like Amazon S3 and Microsoft Azure Storage, object storage software like Openstack Swift, object storage systems like EMC Atmos, EMC ECS and Hitachi Content Platform, and distributed storage research projects like OceanStore and VISION Cloud are all examples of storage that can be hosted and deployed with cloud storage characteristics.




Cloud storage is:

- Made up of many distributed resources, but still acts as one, either in a federated or a cooperative storage cloud architecture
- Highly fault tolerant through redundancy and distribution of data

- Highly durable through the creation of versioned copies
- Typically eventually consistent with regard to data replicas

## Analysis and Planning

In cloud computing, the business requirements are mandatory to consider before deploying the applications to cloud. Following are the things to consider while planning:

- Data Security and privacy requirement
- Budget requirements
- Data backup, training
- Type of cloud i.e public, private or hybrid
- Dashboard and reporting requirements
- Client access requirements
- Data export requirements

To achieve these, well-compiled planning is required.

## Phases of Planning

Following are the phases to migrate the entire business to cloud.

6. Strategy phase
7. Planning phase
8. Deployment phase

## 1. Strategy phase

This phase analyze the strategy problems that customer might face. There are two steps to perform this analysis:

- **Cloud Computing Value Proposition**

  In this, examine the factors affecting the customers while applying the cloud computing mode. Target the key problems they want to solve. The key factors are:

  - IT management simplification
  - Operation and maintenance cost reduction
  - Business mode innovation
  - Low cost outsourcing hosting
  - High service quality outsourcing hosting

  All of the above analysis helps in decision making for future development.

- **Cloud Computing Strategy Planning**

  The strategy establishment is based on the analysis result of the above step. In this step, a strategy document is prepared according to the conditions a customer might face when applying cloud computing mode.

## 2. Planning Phase

This step performs analysis of problems and risks in the cloud application to ensure the customers that the cloud computing is successfully meeting their business goals. This phase involves the following planning steps:

- **Business Architecture Development**

  In this step, we recognize the risks that might be caused by cloud computing application from a business perspective.

- **IT Architecture development**

  In this step, we identify the applications that support the business processes and the technologies required to support enterprise applications and data systems.

- **Requirements on Quality of Service Development**

  Quality of service refers to the non-functional requirements such as reliability, security, disaster recovery, etc. The success of applying cloud computing mode depends on these non-functional factors.

- **Transformation Plan development**

  In this step, we formulate all kinds of plans that are required to transform current business to cloud computing modes.

## 3. Deployment Phase

This phase focuses on both of the above two phases. It involves the following two steps:

- **Selecting Cloud Computing Provider**

  This step includes selecting a cloud provider on basis of Service Level Agreement (SLA), which defines the level of service the provider will meet.

- **Maintenance and Technical Service**

  Maintenance and Technical services are provided by the cloud provider. They need to ensure the quality of services.

## 4.1.2 Storage network design considerations

The best storage area network design for a customer will take into consideration a number of critical issues:

9. Uptime and availability
10. Capacity and scalability
11. Security
12. Replication and disaster recovery

Find out how each of these factors will influence storage area network design choices.

- **Uptime and availability**

Because several servers will rely on a SAN for all of their data, it's important to make the system very reliable and eliminate any single points of failure. Most SAN hardware vendors offer redundancy within each unit like dual power supplies, internal controllers and emergency batteries.

In a typical storage area network design, each storage device connects to a switch that then connects to the servers that need to access the data. To make sure this path isn't a point of failure, the client should buy two switches for the SAN network. Each storage unit should connect to both switches, as should each server. If either path fails, software can failover to the other. Some programs will handle that failover automatically, but cheaper software may require you to enable the failover manually.

- **Capacity and scalability**

A good storage area network design should not only accommodate the client's current storage needs, but it should also be scalable so that the client can upgrade the SAN as needed throughout the expected lifespan of the system.

Because a SAN's switch connects storage devices on one side and servers on the other, its number of ports can affect both storage capacity and speed. By allowing enough ports to support multiple, simultaneous connections to each server, switches can multiply the bandwidth to servers. On the storage device side, enough ports for redundant connections to existing storage units, as well as units to add later should be present.

- **Security**

With several servers able to share the same physical hardware, the security plays an important role in a storage area network design.

Most of this security work is done at the SAN's switch level. Zoning allows giving only specific Servers access to certain LUNs, like a firewall allows communication on specific ports for a given IP address. If any outward-facing application needs to access the SAN, like a website, the switch should be configured so that only the server's IP address can access it.

- **Replication and disaster recovery**

With so much data stored on a SAN, the client wants to build disaster recovery into the system. SANs can be set up to automatically mirror data to another site, which could be a failsafe SAN a few meters away or a disaster recovery (DR) site hundreds or thousands of miles away.

If client wants to build mirroring into the storage area network design, one of the first considerations is whether to replicate synchronously or asynchronously. Synchronous mirroring means that as data is written to the primary SAN, each change is sent to the secondary and must be acknowledged before the next write can happen.

The alternative is to asynchronously mirror changes to the secondary site. This replication can be configured to happen as quickly as every second, or every few

minutes or hours. This means that the client could permanently lose some data, if the primary SAN goes down before it has a chance to copy its data to the secondary.

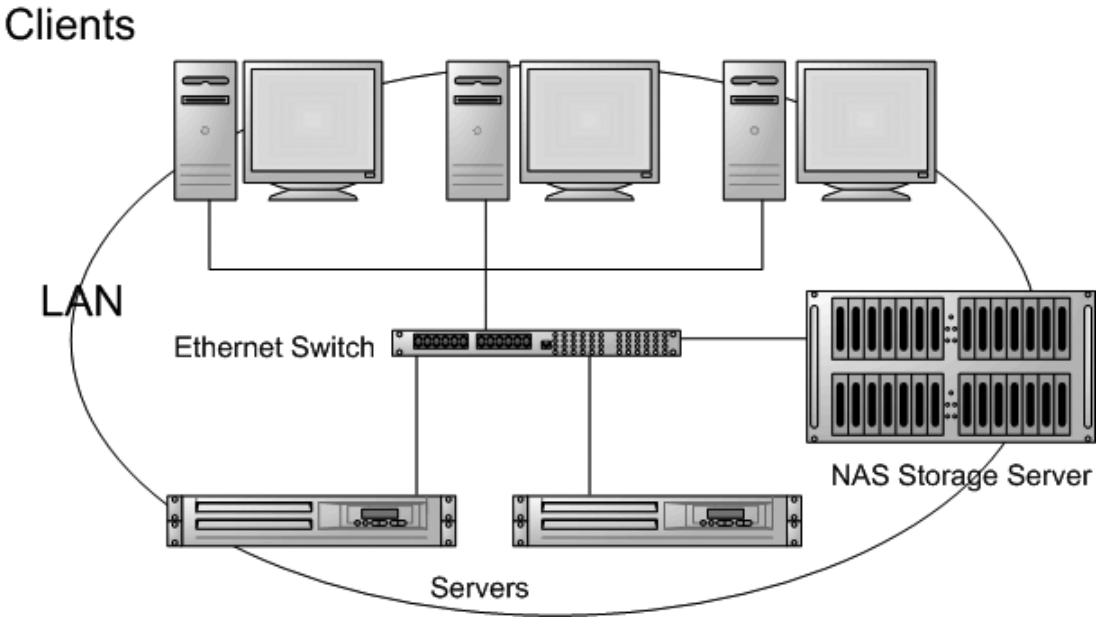
## 4.2 NAS and FC SANs

### 4.2.1 Network-attached storage (NAS)

Network-attached storage (NAS) is a file-level computer data storage server connected to a computer network providing data access to a heterogeneous group of clients. NAS is specialized for serving files either by its hardware, software, or configuration. It is often manufactured as a computer appliance – a purpose-built specialized computer.

Network-attached storage removes the responsibility of file serving from other servers on the network. Potential benefits of dedicated network-attached storage, compared to general-purpose servers also serving files, include faster data access, easier administration and simple configuration.

Network-attached storage devices are flexible and scale-out, meaning that if we need additional storage, we can add on to what we have. A Network-attached storage is like having a private cloud in the office. It's faster, less expensive and provides all the benefits of a public cloud onsite, giving complete control.




Network-attached storage devices are perfect for small businesses because they are:

13. Simple to operate, a dedicated IT professional is often not required
14. Lower cost
15. Easy to use for back up of data, so it's always accessible when you need it
16. Good at centralizing data storage in a safe, reliable way

With a Network-attached storage device, data is continually accessible, making it easy for employees to collaborate, respond to customers in a timely fashion, and promptly

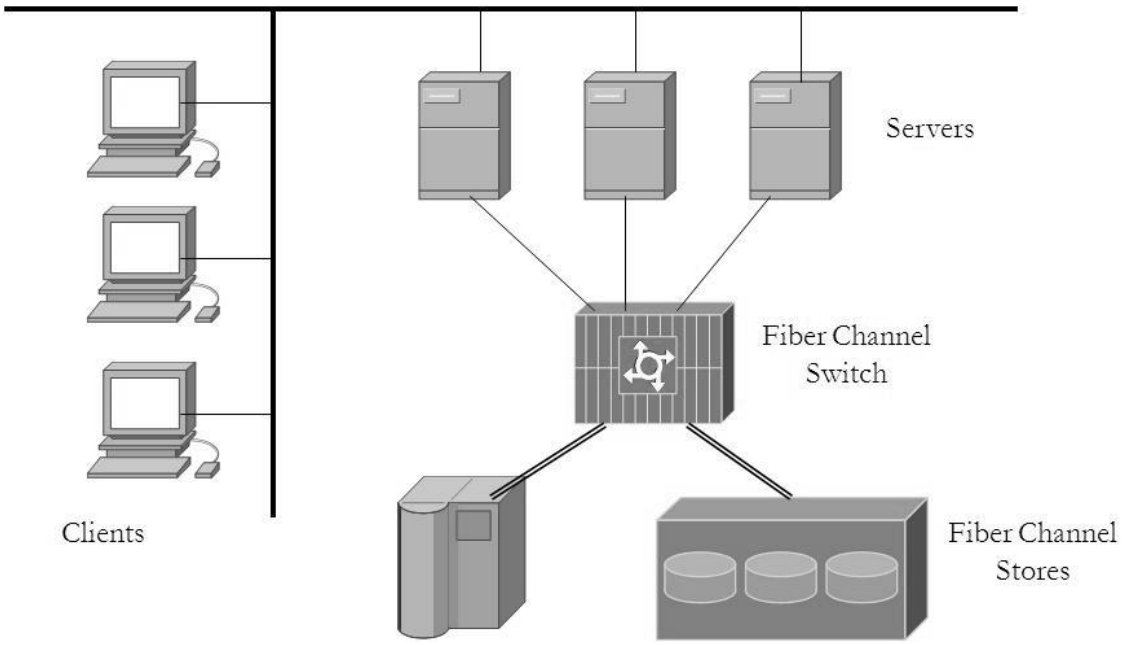
follow up on sales or other issues because information is in one place. Because a Network-attached storage device is like a private cloud, data may be accessed remotely using a network connection, meaning employees can work anywhere and anytime.

### 4.2.2 FC SANs

A storage-area network (SAN) is a dedicated high-speed network (or sub-network) that interconnects and presents shared pools of storage devices to multiple servers.

Fiber Channel (FC) is a high speed serial interface for connecting computers and storage systems. A fiber channel storage area network (FC SAN) is a system that enables multiple servers to access network storage devices. A storage area network enables high-performance data transmission between multiple storage devices and servers.

Native FC is a standards-based SAN interconnection technology within and between data centers limited by geography. It is an open, high-speed serial interface for interconnecting servers to storage devices (discs, tape libraries or CD jukeboxes) or servers to servers. FC has large addressing capabilities. Similar to SCSI, each device receives a number on the channel. It is the dominant storage networking interface today. The Fibre Channel can be fully meshed providing excellent redundancy. FC can operate at the following speeds: 1, 2, 4, 8, 16 and 32 Gb/s with 8Gb/s to 16 Gb/s currently being predominant. The transmission distances vary with the speed and media.




### 4.2.3 Hybrid Storage Networking Technologies

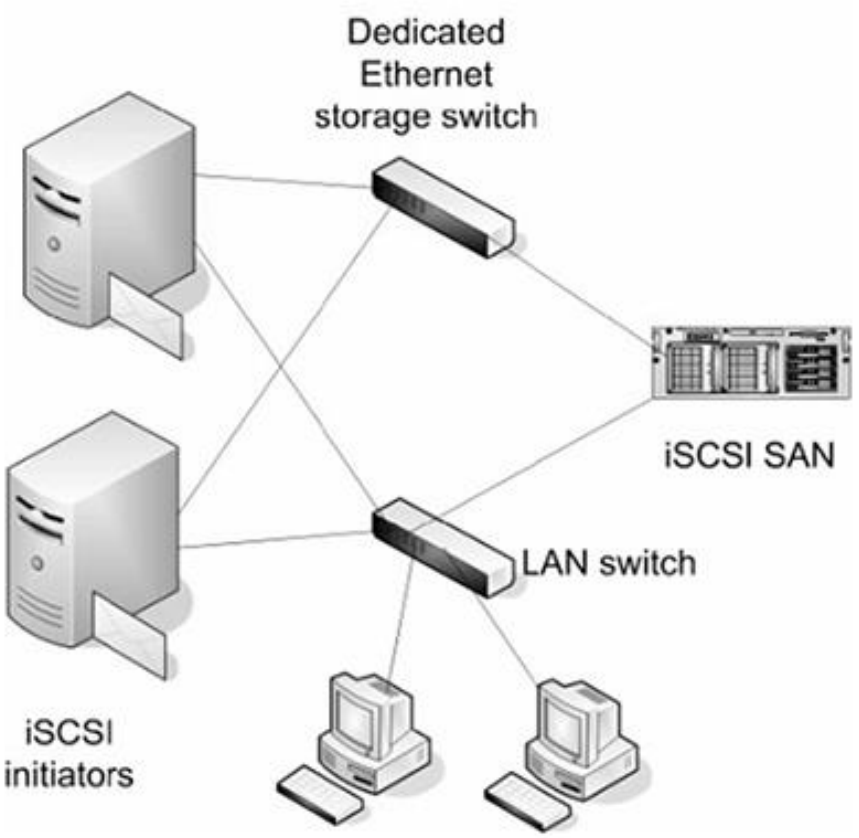
Hybrid cloud storage is an approach to managing storage that uses both local and off-site resources. Hybrid cloud storage is often used to supplement internal storage with public cloud storage.

Ideally, a hybrid cloud implementation behaves as if it is homogeneous storage. Hybrid cloud storage is most often implemented by using proprietary commercial storage

software, by using a cloud storage appliance that serves as a gateway between on-premise and public cloud storage or by using an application program interface (API) to access the cloud storage.

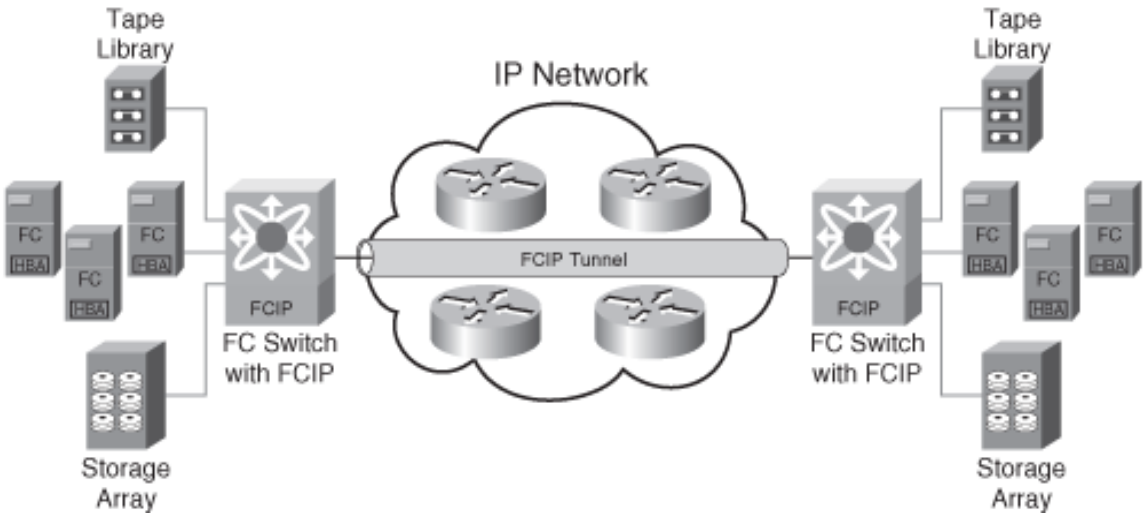
## iSCSI

Internet Small Computer Systems Interface is an Internet Protocol (IP)-based storage networking standard for linking data storage facilities. It provides block-level access to storage devices by carrying SCSI commands over a TCP/IP network. iSCSI is used to facilitate data transfers over intranets and to manage storage over long distances. It can be used to transmit data over local area networks (LANs), wide area networks (WANs), or the Internet and can enable location-independent data storage and retrieval.



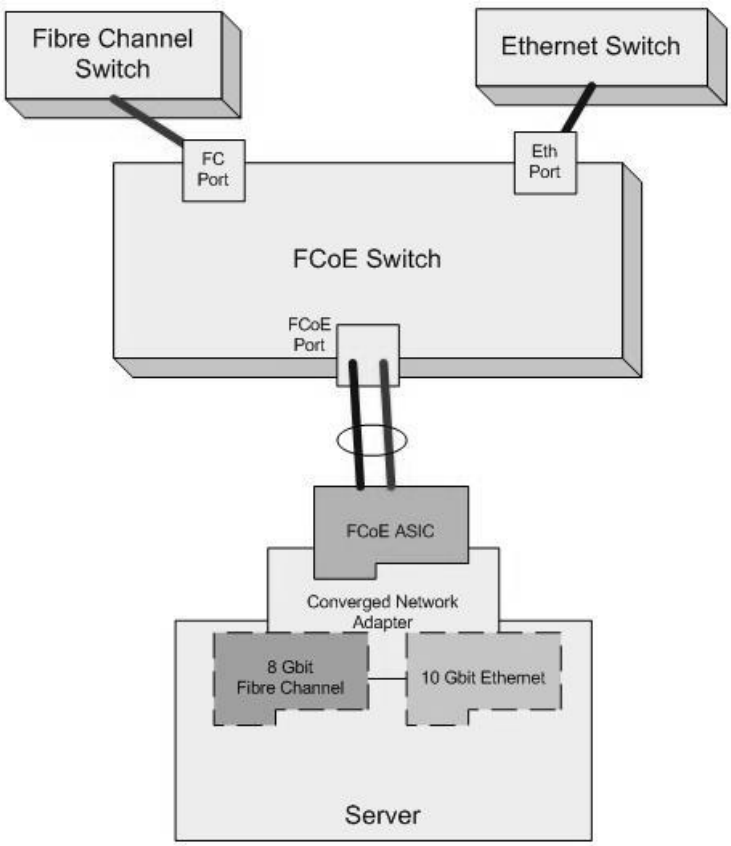

## FCIP

Fiber Channel over IP created by the Internet Engineering Task Force (IETF) for storage technology. An FCIP entity functions to encapsulate Fiber Channel frames and forward them over an IP network. FCIP entities are peers that communicate using TCP/IP. FCIP technology overcomes the distance limitations of native Fiber Channel, enabling geographically distributed storage area networks to be connected using existing IP infrastructure, while keeping fabric services intact. The Fiber Channel Fabric and its devices remain unaware of the presence of the IP Network.




## FCoE

Fiber Channel is accomplished on a separate network than the Ethernet network. With Fiber Channel over Ethernet, Converged Network Adapters are used in place of Ethernet adapters and allow a single channel to pass both Ethernet and Fiber Channel encapsulated packets across a standard IP network extending distance over an entire enterprise, regardless of geography via Ethernet routers and bridges. For replication between storage systems over a wide area network, FCoE provides a mechanism to interconnect islands of FC SAN or FCoE SANs over the IP infrastructure (LANs/MANs/WANs) to form a single, unified FC SAN fabric.



## 4.2.4 Design for storage virtualization in cloud computing

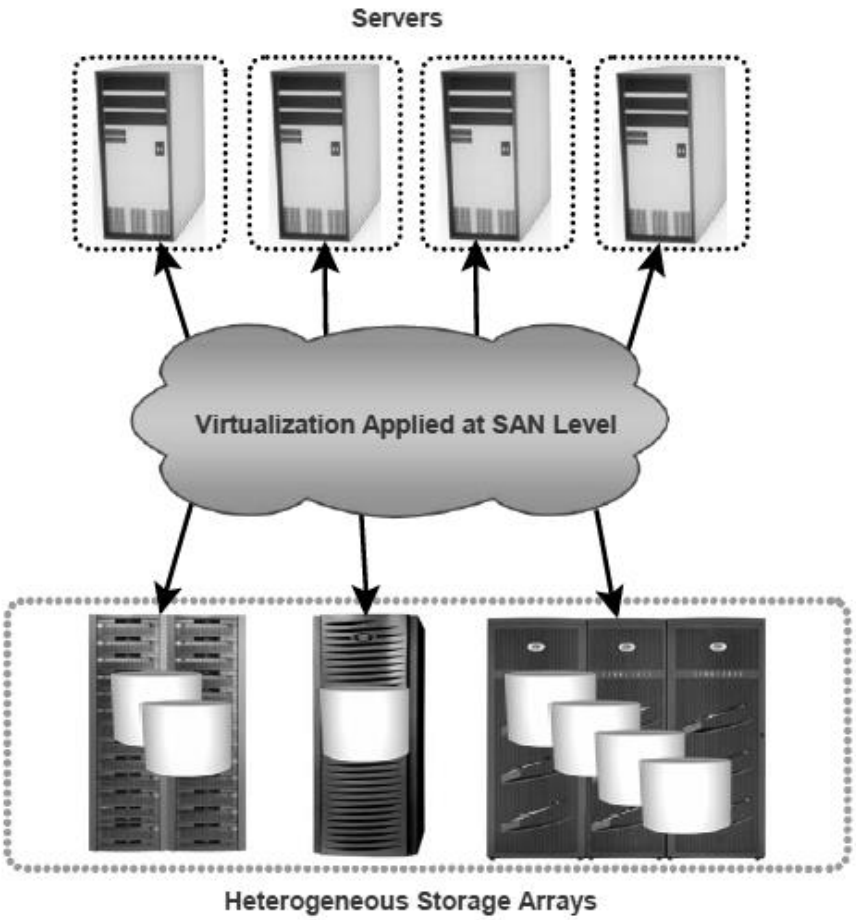
One of the most popular storage virtualization techniques is the pooling of physical storage from multiple network storage devices into what appears to be a single logical storage device that can be managed from a central point of control (console). Storage virtualization techniques are commonly used in a storage area network (SAN), but are also applicable to large-scale NAS environments where there are multiple NAS filers.

There are two primary types of virtualization that can occur:

17. Block level storage virtualization
18. File level storage virtualization

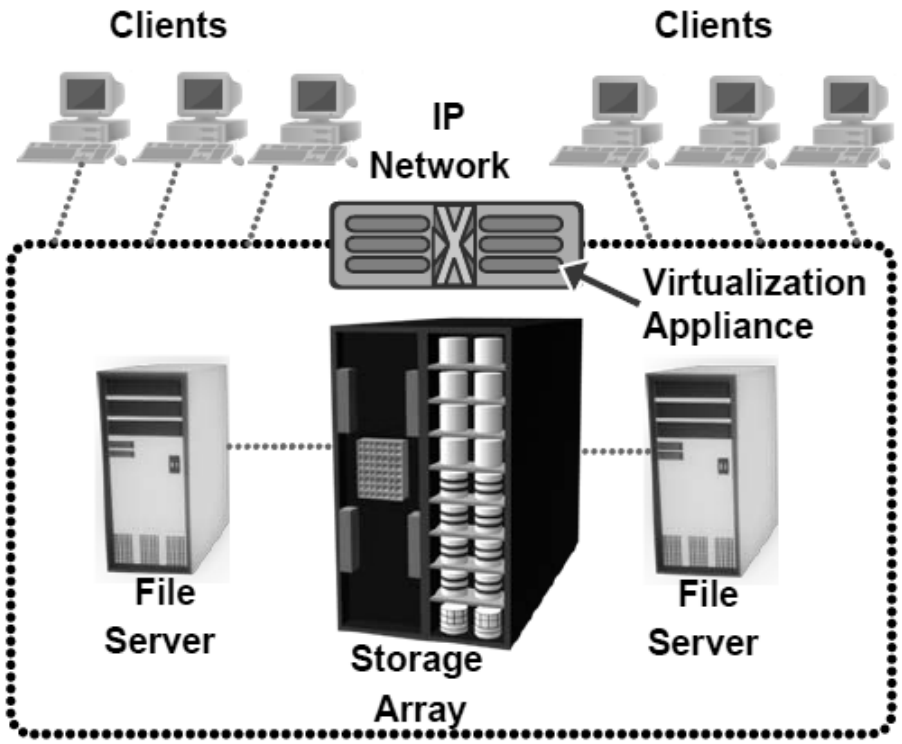
### Block level storage virtualization

Block level storage virtualization is a storage service that provides a flexible, logical arrangement of storage capacity to applications and users while abstracting its physical location. As a software layer, it intercepts I/O requests to that logical capacity and maps them to the appropriate physical locations. In this way virtualization enables administrators to provide the storage capacity when and where it's needed while isolating users from the potentially disruptive details of expansion, data protection and system maintenance.



## File level storage virtualization

File level storage can be defined as a centralized location, to store files and folders. These files systems are Network attached and so form a platform for Network Attached Storage, This level of storage requires file level protocols (computer communication language) like NFS presented by Linux and VMware and SMB/CIFS which is presented by Windows.




## 4.3 File systems or object storage

Object storage (also known as object-based storage) is a computer data storage architecture that manages data as objects, as opposed to other storage architectures like file systems which manage data as a file hierarchy and block storage which manages data as blocks within sectors and tracks. Each object typically includes the data itself, a variable amount of metadata, and a globally unique identifier.

Object storage can be implemented at multiple levels, including the device level (object storage device), the system level, and the interface level. In each case, object storage seeks to enable capabilities not addressed by other storage architectures, like interfaces that can be directly programmable by the application, a namespace that can span multiple instances of physical hardware, and data management functions like data replication and data distribution at object-level granularity.

Object storage systems allow retention of massive amounts of unstructured data. Object storage is used for purposes such as storing photos on Facebook, songs on Spotify, or files in online collaboration services, such as Dropbox.

The majority of cloud storage available in the market uses the object storage architecture. Two notable examples are Amazon Web Services S3, which debuted in 2005,

and Rackspace Files. Other major cloud storage services include IBM Bluemix, Microsoft Azure, Google Cloud Storage, Alibaba Cloud OSS, Oracle Elastic Storage Service and DreamHost based on Ceph.

## Characteristics of Object Storage

- Performs best for big content and high storage throughput
- Data can be stored across multiple regions
- Scales infinitely to Petabytes (bigger than terabyte) and beyond
- Customizable metadata, not limited to number of tags

## Advantages

- Scalable capacity
- Scalable performance
- Durable
- Low cost
- Simplified management
- Single Access Point
- No volumes to manage/resize/etc.

## Disadvantages

- No random access to files
- The Application Programming Interface (API), along with command line shells and utility interfaces (POSIX utilities) do not work directly with object-storage
- Integration may require modification of application and workflow logic
- Typically, lower performance on a per-object basis than block storage

## The Object Storage is suited for the following:

- Unstructured data
    - Media (images, music, video)
    - Web Content
    - Documents
    - Backups/Archives
- Archival and storage of structured and semi-structured data
    - Databases
    - Sensor data
    - Log files

## The Object Storage is not suited for the following:

- Relational Databases
- Data requiring random access/updates within objects

---

Summary:

- ✓ Storage networking is the practice of linking together storage devices and connecting them to other IT networks.
- ✓ Cloud storage is a model of data storage in which the digital data is stored in logical pools, the physical storage spans multiple servers (and often locations), and the physical environment is typically owned and managed by a hosting company.
- ✓ Cloud storage is based on highly virtualized infrastructure and is like broader cloud computing in terms of accessible interfaces, near-instant elasticity and scalability, multi-tenancy, and metered resources.
- ✓ In cloud computing, the business requirements are mandatory to consider before deploying the applications to cloud.
- ✓ The phases to migrate the entire business to cloud are as follows: Strategy phase, Planning phase, Deployment phase.
- ✓ Strategy phase analyze the strategy problems that customer might face. There are two steps to perform this analysis: Cloud Computing Value Proposition, Cloud Computing Strategy Planning.
- ✓ Planning Phase performs analysis of problems and risks in the cloud application to ensure the customers that the cloud computing is successfully meeting their business goals.
- ✓ Deployment Phase focuses on both of the above two phases.
- ✓ The best storage area network design for a customer will take into consideration a number of critical issues: Uptime and availability, Capacity and scalability, Security, Replication and disaster recovery.
- ✓ Network-attached storage (NAS) is a file-level computer data storage server connected to a computer network providing data access to a heterogeneous group of clients.
- ✓ A fiber channel storage area network (FC SAN) is a system that enables multiple servers to access network storage devices.
- ✓ Hybrid cloud storage is an approach to managing storage that uses both local and off-site resources.
- ✓ Internet Small Computer Systems Interface (iSCSI) is an Internet Protocol (IP) based storage networking standard for linking data storage facilities.
- ✓ A Fiber Channel over IP (FCIP) entity functions to encapsulate Fiber Channel frames and forward them over an IP network.
- ✓ With Fiber Channel over Ethernet (FCoE), Converged Network Adapters are used in place of Ethernet adapters and allow a single channel to pass both Ethernet and Fiber Channel encapsulated packets across a standard IP network extending distance over an entire enterprise, regardless of geography via Ethernet routers and bridges.
- ✓ One of the most popular storage virtualization techniques is the pooling of physical storage from multiple network storage devices into what appears to be a

single logical storage device that can be managed from a central point of control (console).

- ✓ There are two primary types of virtualization that can occur: Block level storage virtualization and File level storage virtualization.
- ✓ Block level storage virtualization is a storage service that provides a flexible, logical arrangement of storage capacity to applications and users while abstracting its physical location.
- ✓ File level storage can be defined as a centralized location, to store files and folders.
- ✓ Object storage (also known as object-based storage) is a computer data storage architecture that manages data as objects, as opposed to other storage architectures like file systems which manage data as a file hierarchy and block storage which manages data as blocks within sectors and tracks.
- ✓ Characteristics of Object Storage are, 1. Performs best for big content and high storage throughput, 2. Data can be stored across multiple regions, 3. Scales infinitely to Petabytes (bigger than terabyte) and beyond and 4. Customizable metadata, not limited to number of tags
- ✓ Advantages are Scalable capacity, Scalable performance, Durable, Low cost, simplified management, Single Access Point, No volumes to manage/resize/etc.
- ✓ Disadvantages are, 1. No random access to files, 2. The Application Programming Interface (API), along with command line shells and utility interfaces (POSIX utilities) do not work directly with object-storage, 3. Integration may require modification of application and workflow logic, 4. Typically, lower performance on a per-object basis than block storage.
- ✓ The Object Storage is suited for the following: 1. Unstructured data, 2. Archival and storage of structured and semi-structured data.
- ✓ The Object Storage is not suited for the following: 1. Relational Databases, 2. Data requiring random access/updates within objects.

---

## Review Questions

### Part – A

1. What is Storage Network ?
2. Define : Architecture of storage
3. What is cloud storage ?
4. Write the phases to migrate business to cloud.
5. What is strategy phase ?
6. What is Planning phase ?
7. What is deployment phase ?

8. What are the considerations for storage network design ?
9. Define : NAS
10. Define : FC SAN
11. What is iSCSI ?
12. What is FCIP ?
13. What is the use of FCoE ?
14. What are two types of virtualization in cloud storage ?
15. What is block level virtualization ?
16. What is file level virtualization ?
17. Define : Object storage.

## Part – B

1. Write on architecture of storage in cloud.
2. Write on analysis and planning phase in cloud.
3. Write on Strategy phase of planning
4. Write on Planning phase.
5. Write on Deployment phase.
6. Write on Replication and disaster recovery.
7. Write on FC SANs.
8. What are Hybrid storage networking technologies? Write on it.
9. Write on iSCSI.
10. Write on Block level storage virtualization.
11. Write on File level storage virtualization.

## Part – C

1. Write on Architecture of storage, analysis and planning.
2. Explain storage network design consideration.
3. Describe on NAS.
4. Write on FC SAN and Hybrid storage technologies.
5. Write on FCIP and FCoE
6. Explain design for storage virtualization in cloud computing.
7. Write on File systems or object storage.

---

## UNIT - V : SECURITY IN THE CLOUD

### Objectives

At end the this unit, students can

- Define cloud security.
- Define data security in the cloud.
- Explain cloud users and providers on security service boundary.
- Mention the CSA cloud reference model.
- Describes the cloud consumers using brokered cloud storage access.
- Mention the benefits of cloud broker and categorized.
- Define Service Level Agreements.
- Describe Encryption.
- Define cloud computing security challenges.
- Describe implementations of security policy.
- State the types of policy.
- Describes the virtualization security.
- Explain the virtual threats.

### 5.1. Introduction:

Cloud computing and storage provides users with capabilities to store and process their data in third-party data centers. Organizations use the cloud in a variety of different service models (with acronyms such as SaaS, PaaS, and IaaS) and deployment models (private, public, hybrid, and community). Security concerns associated with cloud computing fall into two broad categories: security issues faced by cloud providers (organizations providing software, platform, or infrastructure-as-a-service via the cloud) and security issues faced by their customers (companies or organizations who host applications or store data on the cloud).

When an organization elects to store data or host applications on the public cloud, it loses its ability to have physical access to the servers hosting its information. As a result, potentially sensitive data is at risk from insider attacks. According to a recent Cloud Security Alliance Report, insider attacks are the sixth biggest threat in cloud computing.

Virtualization in implementing cloud infrastructure brings unique security concerns for customers or tenants of a public cloud service. Virtualization alters the relationship between the OS and underlying hardware - be it computing, storage or even networking. This introduces an additional layer that itself must be properly configured, managed and secured.

Specific concerns include the potential to compromise the virtualization software, or "hypervisor".

**Definition cloud security:**

Cloud security refers to a set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing. It is a sub-domain of computer security, network security, and more information security.

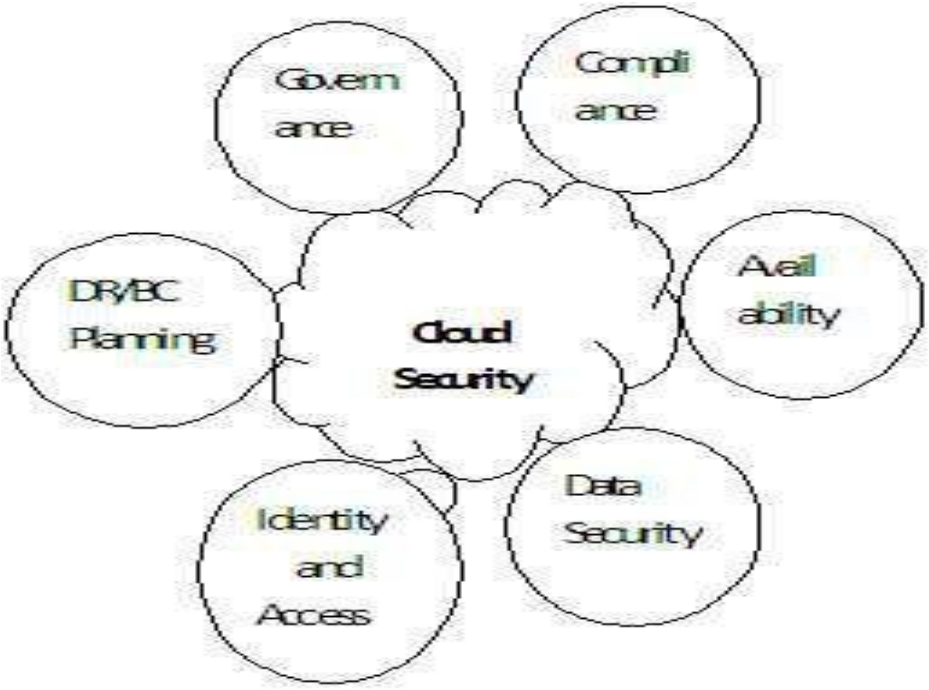



Fig 5.1 Cloud Security

## 5.1.1. Understanding Cloud Security

Many of the tools and techniques that used to protect the data, regulations and maintain the integrity of the systems are shared by others and many of them are outsourced. Cloud computing service providers are well aware of these concerns and have developed new technologies to use them.

Different types of cloud computing service models provide different levels of security services. The least amount of built in security with an Infrastructure as a Service provider (IaaS) and most with a Software as a Service provider (SaaS). The concept of security boundary is separating the clients and vendors responsibilities. The Data stored in the cloud must be transferred and stored in an encrypt format. It uses proxy and brokerage services to separate clients from direct access to shared cloud storage.

### 5.1.2. Securing the Cloud:

Cloud computing has all the vulnerabilities associated with Internet applications, and additional vulnerabilities arise from pooled, virtualized, and outsourced resources. The following areas of cloud computing that they felt were uniquely troublesome:

- Auditing
- Data integrity
- e-Discovery for legal compliance
- Privacy
- Recovery
- Regulatory compliance

In order to evaluate risks, the following analyses are needed to perform:

1. Determine which resources (data, services, or applications) are planning to move to the cloud.
2. Determine the sensitivity of the resource to risk. Risks that need to be evaluated are loss of privacy, unauthorized access by others, loss of data, and interruptions in availability.
3. Determine the risk associated with the particular cloud type for a resource. Cloud types include public, private (both external and internal), hybrid, and shared community types. In each type need to consider the data and functionality will be maintained.
4. Take into account the particular cloud service model that we will be using.
5. Different models such as IaaS, SaaS, and PaaS require their customers to be responsible for security at different levels of the service stack.
6. The cloud service provider need to evaluate the system to understand how data is transferred, where it is stored, and how to move data both in and out of the cloud.

### 5.1.3. Security service boundary

The Security-as-a-Service Working Group of the Cloud Security Alliance, a not-for-profit association formed by cloud-computing stakeholders, issued a report. The report is aimed at providing cloud users and providers on security as a service in order to ease its.

- **Identity and Access Management:** It should provide controls for assured identities and access management. Identity and access management includes people, processes and systems that are used to manage access to enterprise resources.
- **Data Loss Prevention:** It is the monitoring, protecting and verifying the security of data at rest, in motion and in use in the cloud and on-premises.
- **Web Security:** This is real-time protection offered either on-premise through software/appliance installation or via the cloud by proxy or redirecting web traffic to the cloud provider.
- **E-mail Security:** It should provide control over inbound and outbound e-mail and protecting the organization from malicious attachments. Digital signatures enabling

identification and non-repudiation are features of many cloud e-mail security solutions.

- **Intrusion Management:** This is the process of using pattern recognition to detect and react to statistically unusual events.
- **Security Information and Event Management:** The systems accept log and event information.
- **Encryption:** The systems typically consist of algorithms that are computationally difficult or infeasible to break, along with the processes and procedures to manage encryption and decryption, hashing, digital signatures, certificate generation and renewal and key exchange are also used.
- **Network Security:** It consists of security services that allocate access, distribute, monitor and protect the underlying resource services.

### 5.1.4. CSA Cloud Reference Model

CSA is an industry working group that studies security issues in cloud computing and offers recommendations to its members. It gives the cloud computing stack model, which shows how different functional units in a network stack relate to one another. The CSA functional cloud computing hardware/software stack is the Cloud Reference Model i.e. The CSA functional cloud computing hardware/software stack is the Cloud Reference model.

IaaS(Infrastructure as a Service) is the lowest level service, with PaaS(Platform as a Service) and SaaS(Software as a Service) the next two services above in Iaas. It moves upward in the stack, each service model inherits the capabilities.

IaaS supplies the infrastructure, PaaS use added application development frameworks, transactions and control structures and SaaS is an operating environment with applications, management and the user interface. Each different type of cloud service delivery model creates a security boundary, at which the service provider's responsibilities end and customer's responsibilities begin. Any security mechanism below the security boundary must be built into the system, and any security mechanism about must be maintained by the customer. It moves up the stack, it becomes the type and level of security is part of Service Level Agreement.

In the SaaS model, the vendor provides security like compliance, governance, and liability levels for the entire stack. The PaaS model, the security boundary may include the software framework and middleware layer. In Ieast built-in security IaaS model, software of any kind is the customer's problem. A private cloud may be internal or external of the organization and a public cloud is most often external only.

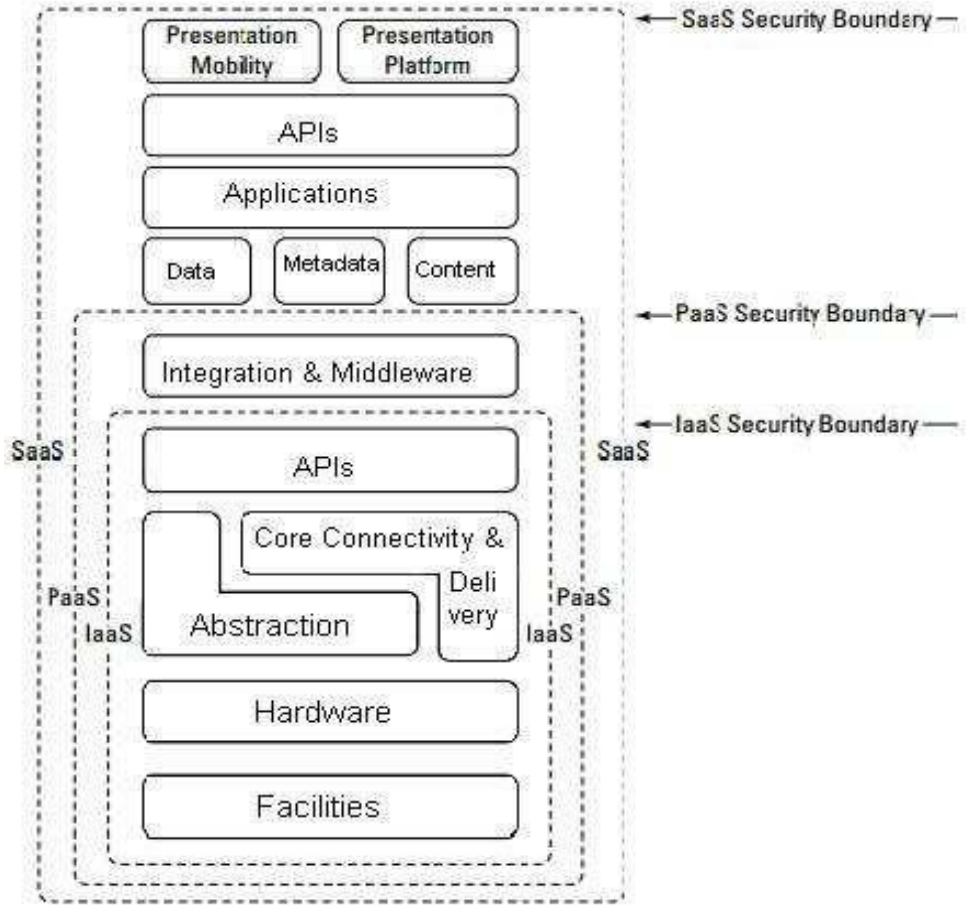



**Fig 5.2:** CSA Cloud Reference Model

### 5.1.5. Securing Data

Securing data sent to received from stored in the cloud is the single largest security concern of most organizations should have with cloud computing. Such as any WAN traffic, the data can be intercepted and modified. The traffic to a cloud service provider and stored-off premises is encrypted. It is for general data as it is for any account ID and passwords.

These are the key mechanism for protecting data mechanisms:

- Access control
- Auditing
- Authentication
- Authorization

### 5.1.6. Brokered cloud storage access

Cloud Broker is an entity that manages the use, performance and delivery of cloud services, and relationships between cloud providers and cloud consumers.

All the data stored in the cloud. It can be located in the cloud service provider's system used to transfer data from sent and received. The cloud computing has no physical system that

serves this purpose. To protect the cloud storage is the way to isolate data from client direct access. They are two services are created. One service for a broker with full access to storage but no access to the client, and another service for a proxy with no access to storage but access to both the client and broker. These important two services are in the direct data path between the client and data stored in the cloud.

Under this system, when a client makes a request for data, here's what happens:

1. The request goes to the external service interface of the proxy.
2. The proxy using internal interface, forwards the request to the broker.
3. The broker requests the data from the cloud storage system.
4. The storage system returns the results to the broker.
5. The broker returns the results to the proxy.

The proxy completes the response by sending the data requested to the client.

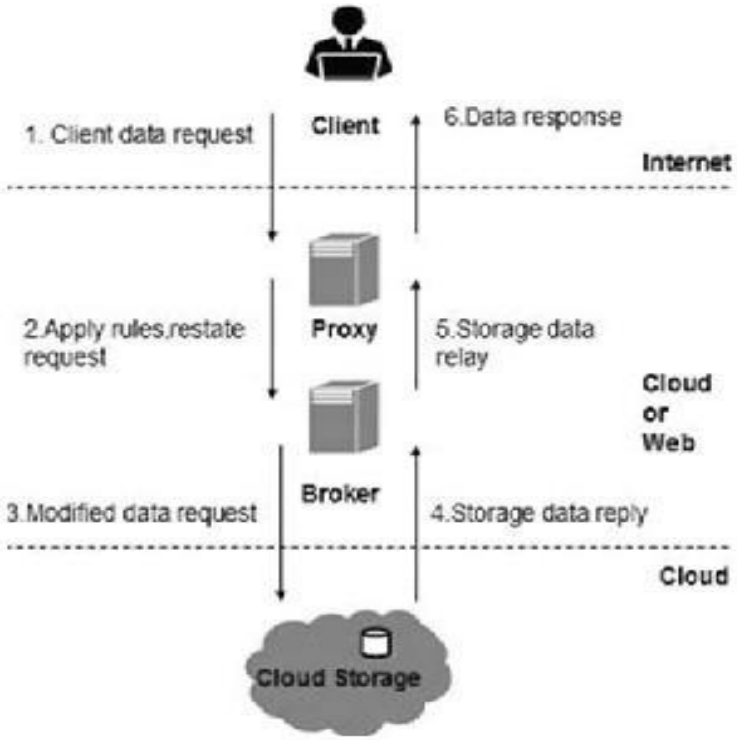



**Fig 5.3 Cloud storage with proxy/broker service**

Even if the proxy service is compromised, that service does not have access to the trusted key that is necessary to access the cloud storage. In the **multi-key solution**, not eliminated all internal service endpoints, but proxy service run at a reduced trust level is eliminated. The creation of storage zones with associated encryption keys can further protect cloud storage from unauthorized access.

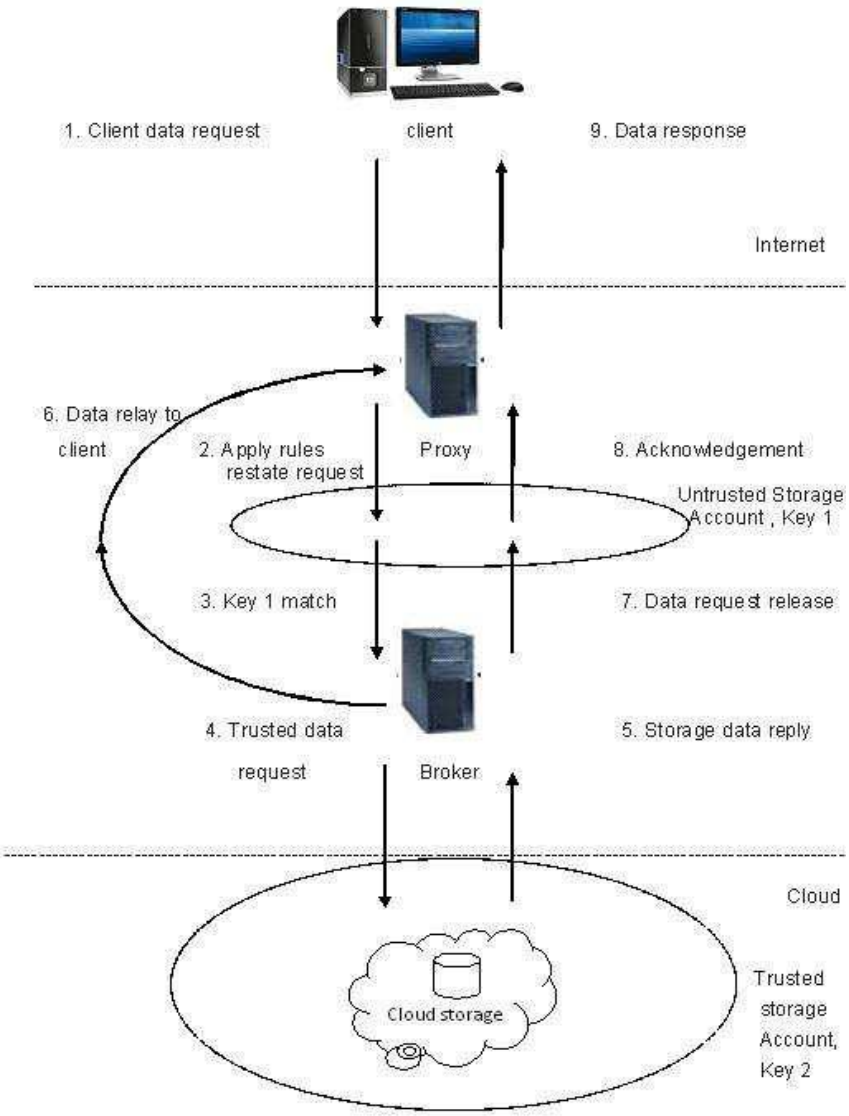



**Fig 5.4** storage zone with encrypted keys

Cloud brokers provide services in three categories:

**Aggregation:** A cloud broker combines and integrates multiple services into one or more new services.

**Arbitrage:** This is similar to service aggregation, except that the services being aggregated are not fixed.

**Intermediation:** The cloud broker give service by improving capability and providing value-added services to cloud consumers. The improvement can be managing access to cloud services, identity management, performance reporting, enhanced security, etc.

**Benefits of using a cloud broker**

Benefits of using a cloud broker for a business or technical purpose include the following:

- Cloud interoperability - Integration between several cloud offerings.
- Cloud portability - Move application between different cloud vendors.
- Increase business continuity by reducing dependency from one cloud provider.
- Cost savings.

### 5.1.7. Storage location and tenancy:

Cloud service providers as per their Service Level Agreements, need to contractually store and process data in locations that are predetermined by their contract. It gets the commitment for specific data site storage the cloud vendor is under contract to conform to privacy laws.

Because data stored in the cloud is usually stored from multiple tenants the each vendor has its own unique method for segregating one customer's data from another. It's important to understand how the specific service provider maintains data segregation. Cloud storage provider provides privileged access to storage. Most cloud service providers store data in an encrypted form to protect the data used in security mechanism. Hence, data cannot be accessed by the unauthorized user.

It is important to know what impact a disaster or interruption occur on the stored data. Since data are stored across multiples sites, it may not be possible to recover data in a timely manner.

### 5.1.8. Encryption

Cloud encryption is the transformation of a cloud service customer's data into cipher text. Cloud encryption is commonly used to prevent unauthorized access to private information, protect sensitive data stored in the cloud. Cloud customer must take time to learn about the provider's policies and procedures for encryption are called encryption key management. The cloud encryption capabilities of the service provider need to match the level of sensitivity of the data being hosted.

Strong encryption technology is a core technology for protecting data in transit to and from the cloud as well as data stored in the cloud. The goal of encrypted cloud storage is to create a virtual private storage system that maintains confidentiality and data integrity. Encryption should separate stored data (data at rest) from data in transit. Depending upon the particular cloud provider, such as Microsoft allows up to five security accounts per client, and can use these different accounts to create different zones. On Amazon Web Service, we can create multiple keys and rotate those keys during different sessions.

Although encryption protects data from unauthorized access, it does nothing to prevent data loss. The losing encrypted data is to lose the keys that provide access to the data. Therefore, need to approach key management seriously. This schemes used to protect keys are the creation of secure key stores that restricted role-based access, automated key stores backup, and recovery techniques.

### 5.2.1. Cloud Computing Security Challenges:

The cloud security challenges are classified into five categories:

1. **User authentication:** Only authorized persons are allowed to access the data which rests in the cloud. To ensure the integrity of user authentication and to confirm that only the authenticated users are accessing the data.
2. **Data protection:** So protecting the data is to be considered in two aspects. Such as data at rest and in transit.
3. **Contingency planning:** The data is being secured and also measures the Cloud Service Provider (CSP) is implementing to assure the integrity and availability of the data.
4. **Interoperability**: It applies to cloud computing is at its simplest, the requirement for the components of a processing system to work together to achieve their intended result.
5. **Portability:** It applies to the cloud provides for application and data components to continue to work the same way when moved from one cloud environment to another without having to be changed.

.

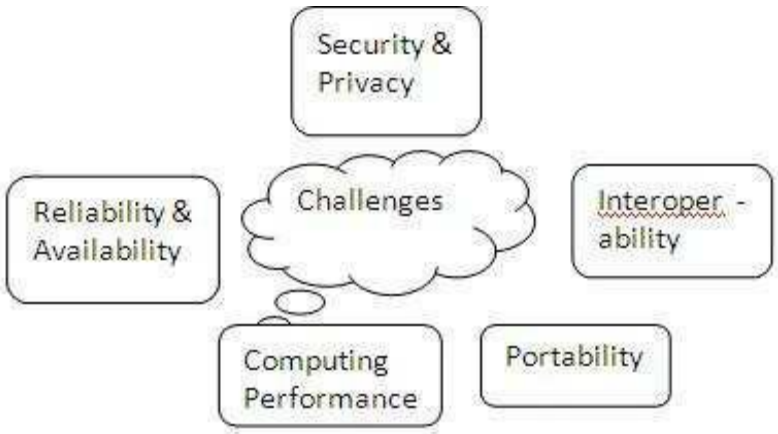



**Figure** 5.5: Security Challenges

### 5.2.2. Security Policy Implementation:

Security policies are the foundation of a sound security implementation. The organizations will implement technical security solutions without first creating this foundation of policies, standards, guidelines, and procedures, unintentionally creating unfocused and ineffective security controls.

A policy is one of those terms that can mean several things. For example, there are security policies on firewalls, which refer to the access control and routing list information. Standards, procedures, and guidelines are also referred to as policies in the larger sense of a global information security policy.

A policy, for example, can literally be a lifesaver during a disaster, or it might be a requirement of a governmental or regulatory function. A policy can also provide protection from liability due to an employee's actions, or it can control access to trade secrets. A policy can also provide protection from liability due to an employee's actions, or it can control access to trade secrets.

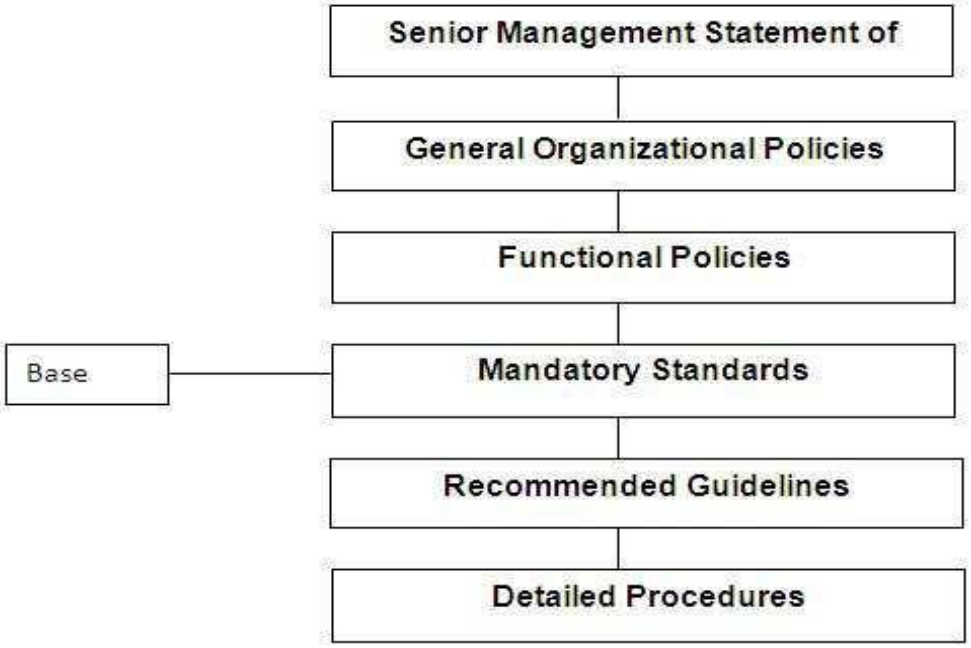



**Figure 5.6:** Security policy hierarchy

### 5.2.3. Policy Types:

Policies are the first and highest level of documentation, from which the lower-level elements of standards, procedures, and guidelines flow. These higher-level policies, which reflect the more general policies and statements (process, strategic reasons, tactical element can follow). Management should ensure the high visibility of a formal security policy. This is all employees at all levels will in some way be affected, major organizational resources will be addressed, and many new terms, procedures, and activities will be introduced. There are four types of Policy.

**Senior Management Statement of Policy**

The first policy of any policy creation process is the senior management statement of policy. This is high-level policy that acknowledges the importance of the computing resources to the business model. Such as support for information security throughout the enterprise, and commits to authorizing and managing the definition (lower-level standards, procedures, and guidelines).

**Regulatory Policies**

Regulatory policies are security policies that an organization must implement due to compliance, regulation, or other legal requirements. Such as companies might be financial institutions, public utilities, or some other type of organization that operates in the public interest.

**Advisory Policies**

Advisory policies are security policies that are not mandated but strongly suggested, perhaps with serious consequences defined for failure to follow them (such as termination, a job action warning, and so forth).

**Informative Policies**

Informative policies are policies that exist simply to inform the reader. There are not implied or specified requirements, and the audience for this information could be certain internal (within the organization) or external parties.

### 5.2.4. Virtualization Security Management

Historically, the development and implementation of new technology has preceded the full understanding of its inherent security risks, and virtualized systems are no different. The global adoption of virtualization is a relatively recent event, threats to the virtualized infrastructure.

A virtual machine (VM) is an operating system (OS) or application environment that is installed on software, which imitates dedicated hardware. The Virtual Machine (VM), Virtual Memory Manager (VMM), and hypervisor or host OS are the minimum set of components needed in a virtual environment.

### Virtualization Types:

Based on the minimum set of components, we classify the Virtual Environments in the following distinct ways.

- Type 1 virtual environments are considered “full virtualization” environments and have VMs running on a hypervisor that interacts with the hardware.

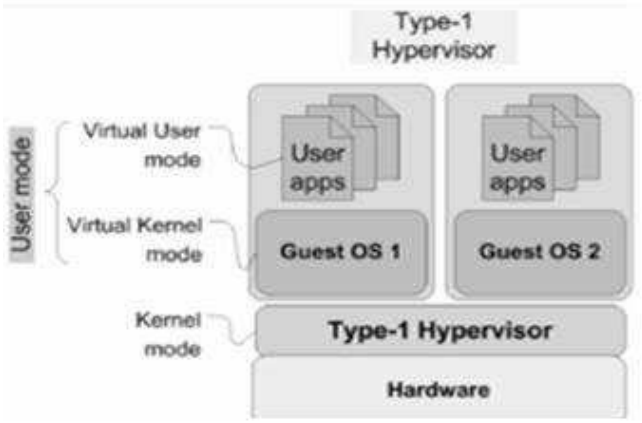



**Figure 5.6a**: Type 1

- Type 2 virtual environments are also considered “full virtualization” but work with a host OS instead of a hypervisor.

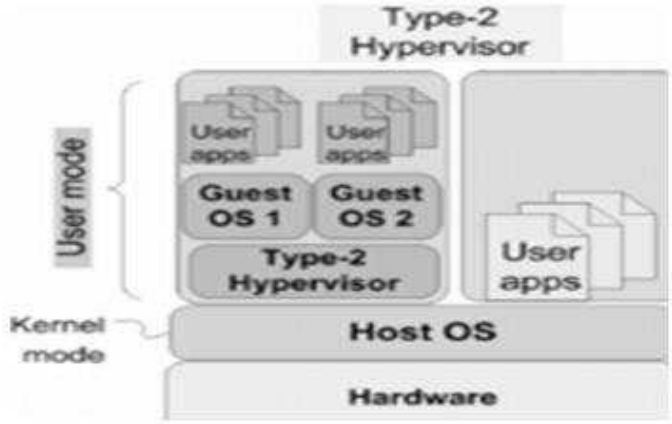



**Figure 5.6b**: Type 2

- Para virtualized environments offer performance gains by eliminating some of the emulation that occurs in full virtualization environments.
- Other type designations include hybrid virtual machines (HVMs) and hardware-assisted techniques.

These classifications are somewhat ambiguous in the IT community at large. The most important thing to remember from a security perspective is that there is a more significant impact when a host OS with user applications and interfaces is running outside of a VM at a level lower than the other VMs (i.e., a Type 2 architecture). Because of its architecture, the Type 2 environment increases the potential risk of attacks against the host OS. For example, a laptop running VMware with a Linux VM on a Windows XP system inherits the attack surface of both OSs, plus the virtualization code (VMM).

**Virtualization Management Roles:**

The roles assumed by administrators are the Virtualization Server Administrator, Virtual Machine Administrator, and Guest Administrator. The roles assumed by administrators are configured in VMS and are defined to provide role responsibilities.

1. Virtual Server Administrator — This role is resp onsible for installing and configuring the ESX Server hardware, storage, physical and virtual networks, service console, and management applications.
2. Virtual Machine Administrator — This role is res ponsible for creating and configuring virtual machines, virtual networks, virtual machine resources, and security policies. The Virtual Machine Administrator creates, maintains, and provisions virtual machines.
3. Guest Administrator — This role is responsib le for managing a guest virtual machine Tasks typically performed by Guest Administrators include connecting virtual devices, adding system updates, and managing applications that may reside on the operating system.

### 5.2.5 Virtual Threats:

Some threats to virtualized systems are general in nature, as they are inherent threats to all computerized systems (such as denial-of-service, or DoS, attacks). Other threats and vulnerabilities, however, are unique to virtual machines. Many VM vulnerabilities stem from the fact that vulnerability in one VM system can be exploited to attack other VM systems or the host systems, as multiple virtual machines share the same physical hardware.

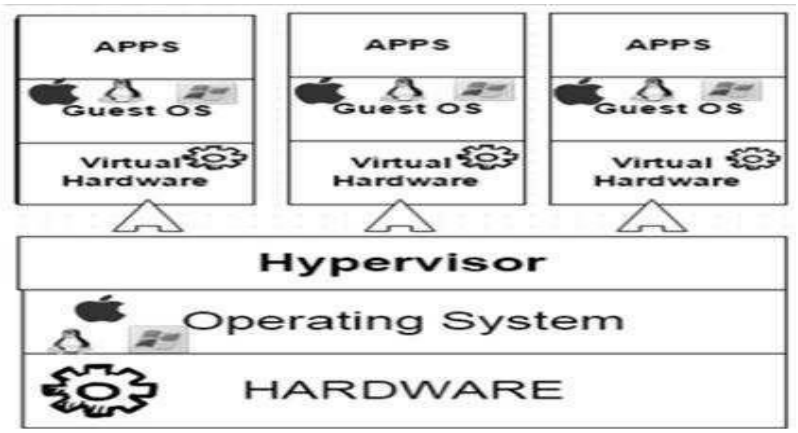



**Figure 5.7:** Virtual Threats

Some of the vulnerabilities exposed to any malicious-minded individuals regarding security in virtual environments:

- **Shared clipboard** — Shared clipboard technology allows data to be tra nsferred between VMs and the host, providing a means of moving data between malicious programs in VMs of different security realms.
- **Keystroke logging** — Some VM technologies enable the logging of keystr okes and screen updates to be passed across virtual terminals in the virtual machine, writing to host files and permitting the monitoring of encrypted terminal connections inside the VM.
- **VM monitoring from the host** — Because all network packets coming from or going to a VM pass through the host, the host may be able to affect the VM by the following:
    1. Starting, stopping, pausing, and restart VMs.
    2. Monitoring and configuring resources available to the VMs, including CPU, memory, disk, and network usage of VMs.
    3. Adjusting the number of CPUs, amount of memory, amount and number of virtual disks and number of virtual network interfaces available to a VM.
    4. Monitoring the applications running inside the VM.
    5. Viewing, copying, and modifying data stored on the VM's virtual disks.
- **Virtual machine monitoring from another VM** — Usually, VMs should not be able to directly access one another's virtual disks on the host.

- **Virtual machine backdoors** — A backdoor, covert communications channel between the guest and host could allow intruders to perform potentially dangerous operations.

## SUMMARY

Cloud security refers to a set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand.

Cloud Security Alliance (CSA) is a nonprofit organization that promotes research into best practices for securing cloud computing and the ability of cloud technologies to secure other forms of computing.

Cloud security controls: Deterrent controls, Preventive controls, Detective controls, Corrective controls.

Data storing in the cloud is called cloud storage.

Different types of models and services used in cloud security. Such as IaaS, PaaS, SaaS. Cloud reference models are used in various services. Such as IaaS is supplies the infrastructure, PaaS use adds application development frameworks, transactions and control structures and SaaS is an operating environment with applications, management and the user interface.

Cloud brokers provide three categories of services. They are aggregation, Arbitrage, Intermediation.

Aggregation means cloud broker combines and integrates multiple services into one or more new services.

Arbitrage means a broker has the flexibility to choose services from multiple Providers, depending upon the characteristics of the data or the context of the service.

Intermediation means cloud broker enhances a given service by improving some specific capability and providing value-added services to cloud consumers.

Policies are four types. They are Senior Management Statement of Policy, Regulatory Policies, Advisory Policies, and Informative Policies.

Virtual machine (VM) is an operating system (OS) or application environment that is installed on software, which imitates dedicated hardware.

Virtualization security is the collective measures, procedures and processes that ensure the protection of a virtualization infrastructure / environment.

A virtual security is a computer appliance that runs inside virtual environments.

* * *

## Review Questions

### Part-A

1. Define cloud security.
2. Expand: IaaS and SaaS.
3. What is securing the cloud?
4. List the security service boundary.
5. What is cloud encryption?
6. Define network security in cloud.
7. Expand: CSA.
8. Expand: SLA.
9. What are the various securing the data mechanism?
10. What are the benefits of cloud broker?
11. Define: Data Protection.
12. What is interoperability?
13. What is the policy?
14. What are the types of policies?
15. What is virtualization?
16. Define contingency planning.
17. What is virtual machine?
18. What is virtual security?
19. What is the SPI model?
20. How secure is cloud computing?
21. What is Cloud Security Alliance?

### Part – B

1. How to understand the cloud security?
2. What are to be performed in risk analysis in securing the cloud?
3. Difference between web security and e-mail security.
4. Draw the CSA reference model.
5. What is the cloud broker?
6. Define service level agreement.
7. What is portability?
8. Define security policy.
9. Compare the regulatory policy and informative policy.
10. What is virtual machine?
11. What is proxy?
12. What are the types of virtualization?

13. List the virtual management roles.
14. What is the virtual threat?
15. Define shared clipboard.
16. Define Advisory policies.
17. What is cloud computing?
18. What are the cloud provider's Encryption policies?
19. How do protect virtual machine?
20. Which is the key mechanism for protecting data?
21. Which is the standard for interoperable cloud-based key management?
22. What is privacy policy?
23. What are the cloud security controls?

## Part –C

1. Explain on securing the cloud.
2. Explain about the various securities service boundaries.
3. Describes the CSA cloud reference model.
4. Explain the brokered cloud storage access.
5. Write notes on storage location and tenancy.
6. Explain the Encryption in the cloud.
7. Explain about the cloud computing security challenges.
8. Explain security policy implementation.
9. Explain the types of policies.
10. Describes the virtualization security management.
11. Write notes on: virtual threats.
12. Describes the Virtualization Management Roles.
13. Explain Cloud Provider's Disaster Recovery Plan?
14. How do manage the encryption keys?
15. What are the security measures you use to authenticate users?
16. Explain the top security threats facing the company?

* * *