# Understanding Cloud Computing

Def$^n$ " a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies".

- Forrester Research provided its def$^n$ of cloud computing as:
" a standardized IT capability (services, software, or infrastructure) delivered via Internet technologies in a pay-per-use, self-service way".

· National Institute of Standards & Technology (NIST) defined cloud in Sept 2011 as:
"Cloud computing is a model for enabling ubiquitos, convenient, on demand network access to a shared pool of configurable resources (eg, networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal managment effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models".

"Cloud computing is a specialized form of distributed computing that introduces utilization models for remotely provisioning scalable and measured resources".

## Business Drivers

The primary business drivers that promoted the growth of modern cloud-based technology are:

### i) Capacity Planning

- Capacity planning is the process of determining & fulfilling future demands of an organization's IT resources, products and services.
- Capacity represents the maximum amount of work that an IT resource is capable of delivering in a given period of time.,
- Capacity planning is focused on minimizing the differences between the capacity of an IT resource and its demand to achieve predictable effeciency & performance.

Different capacity planning strategies are:

- Lead Strategy - adding capacity to an IT resource in anticipation of demand.
- Lag Strategy - adding capacity to an IT resource reaches its full capacity.
- Match Strategy - adding IT resource capacity in small increments, as demand increases.

ii) Cost Reduction

Two costs need to be accounted for: the cost of acquiring new infrastructure & the cost of its ongoing ownership.

Infrastructure related operating overhead includes:

- technical personnel required to keep the environmen operational
- upgrades & patches that introduce additional testing & deployment cycles.
- utility bills and capital expense investments for power and cooling.
- security and access control measures that need to be maintained & enforced to protect infracture resources.

- administrative & accounts staff to keep track of licenses & support arrangements

iii) Organizational Agility

- Organizational agility is the measure of an organization's responsiveness to change.
- An IT enterprise often needs to respond to business change by scaling its IT resources beyond the scope of what was previously predicted or planned for.
- Changing business needs & priorities may require IT resources to be more available & reliable than before.
- Due to a lack of reliability controls within the infrastructure, responsiveness to consumer or customer requirements may be reduced to a point whereby a business's overall continuity is threatened.
- The up-front investments & infrastructure ownership cost.
- The business may decide against proceeding with an automation solution altogether upon review of its infrastructure of budget.

## Technology Innovations

Established technologies are often used an inspiration & the actual foundations upon which new technology innovations are derived & built.

### i) Clustering

- A cluster is a group of independent IT resources that are interconnected & work as a single system.
- Availability & reliability are increased, since redundancy and failover features are inherent to the cluster.
- A general prerequisite of hardware clustering is that its component systems have reasonably identical hardware & operating system to provide similar performance when one failed component is to be replaced by another.
- Component devices that form a cluster are kept in synchronization through dedicated', high-speed communication links.
- The basic concept of built-in redundancy & failover is core to cloud platforms.

## ii) Grid Computing

- A computing grid provides a platform in which computing resources are organized into one or more logical pools.
- These pools are collectively coordinated to provide a high performance distributed grid, refered as "super virtual computer".
- Grid computing systems involve computing resources that are heterogeneous and geographically dispersed, which is not possible with cluster systems.
- The technological advancements achieved by grid computings such as networked access, resource pooling, & scalability & resiliency. These features influenced various aspects of cloud computing platforms & mechanisms.
- Grid computing is based on a middleware layer that is deployed on computing resources.
- IT resources participate in a grid pool that implements a series of workload distribution and coordination functions.

- The middle tier can contain load balancing logic, failover controls and autonomic configuration management.

iii) Virtualization

- Virtualization represents a technology platform used for the creation of virtual instances of IT resource
- A layer of virtualization software allows physical IT resource to provide multiple virtual images of themselves so that their underlying processing capabilities can be shared by multiple users.
- The virtualization process severs the software-hardw dependency, as hardware requirements can be simulated by emulation software running in virtualized environments.
- Modern virtualization technologies emerged to overcome the performance, reliability and scalabil limitations of traditional virtualization platforms.

## iv) Technology Innovations vs Enabling Technologies

The cloud enabling technologies existed in some form prior to the formal advent of cloud computing. such cloud - enabling technologies, are:

- Broadband Networks & Internet Architecture
- Data Center Technology
- (Modern) Virtualization Technology
- Web Technology
- Multitenant Technology
- Service Technology.

es

e

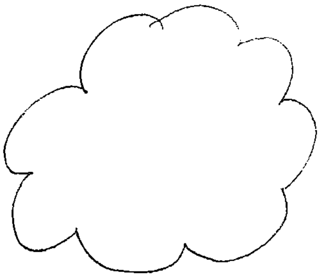
ting,

## Basic Concepts and Terminology

### Cloud

- A cloud refers to distinct IT environment tha is designed for the purpose of remotely provision scalable & measured IT resources.

- Cloud originated as a metaphor for Internet, a network of networks providing remote access to a set of decentralized IT resources.



The symbol used to denote the boundary of cloud environment.

- As a specific environment ~~has~~ used to remotely provision IT resources, a cloud has a finite, boundary.

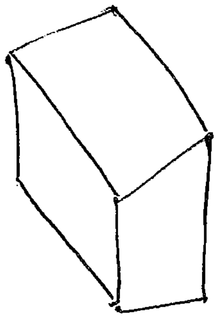
- Internet provides open access to many Web-based IT resources, a cloud is typically privately owned & offers access to IT resource that is metered.

- IT resources provided by cloud environments are dedicated to supplying back-end processing

- A cloud can be based on the use of any protocols that allow for the remote access to its IT resources.

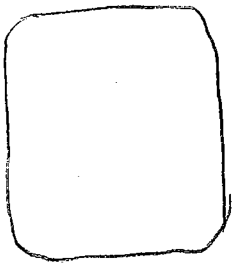
## IT Resource

An IT resource is a physical or Virtual IT-related artifact that can be either software-based, such as a virtual server or a custom software program, or hardware-based, such as a physical server or a network device.

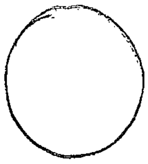


Physical Server

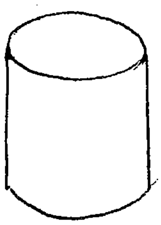
Virtual Server



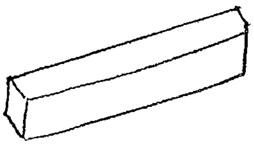
Software program



Service



Storage device



Network device.

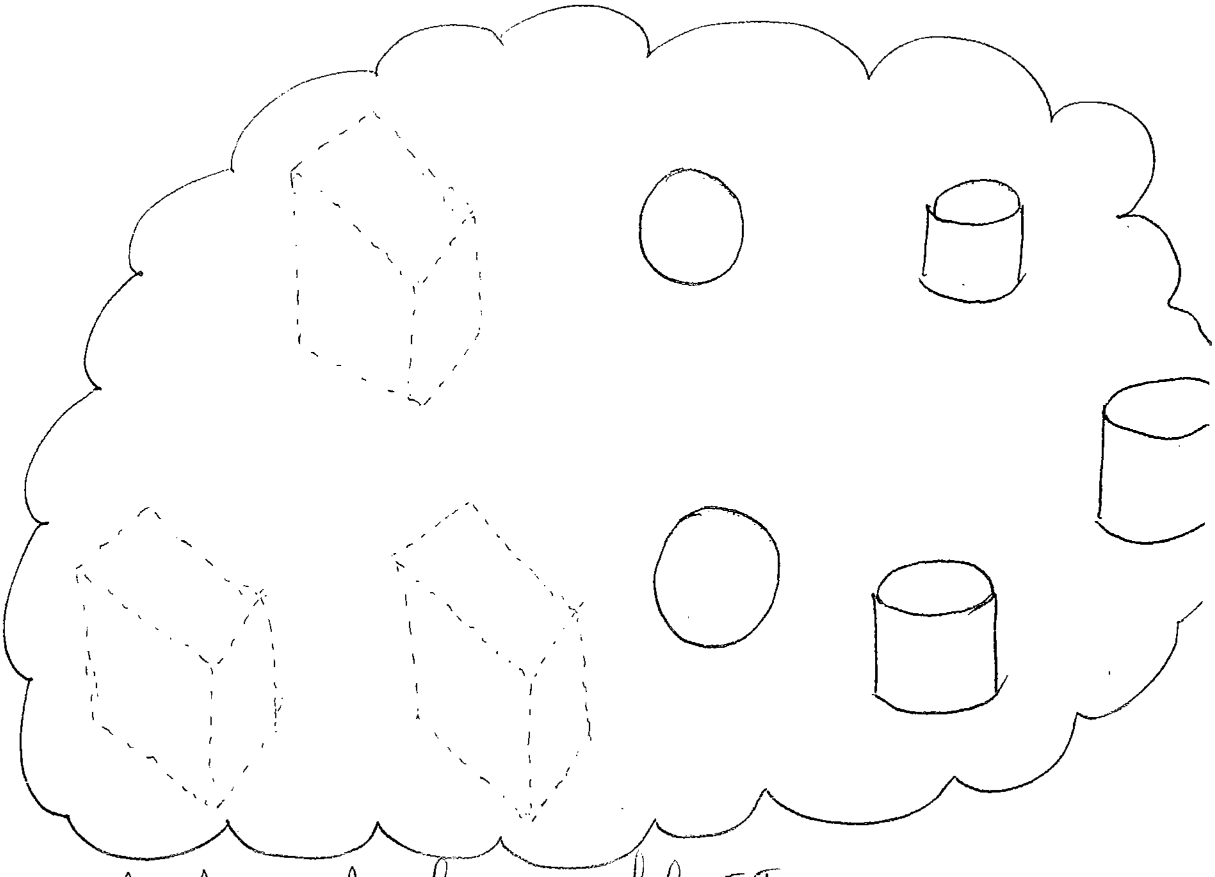


Fig: A cloud is hosting eight IT resources: three virtual servers, two cloud services, & three storage devices.

These diagrams provide abstracted views of the underlying technology architectures. This means that only a portion of the actual technical details are shown.

## On Premise

- An IT resource that is hosted in a conventiona IT enterprise within an organizational boundary considered to be located on the premises of the IT enterprise, or on-premise for short.

- on the premises of a controlled IT environment that is not cloud based.
- An IT resource that is on-premise cannot be cloud-based, and vice-versa.

- An on-premise IT resource can access and interact with a cloud-based IT resource.
- An on-premise IT resource can be moved to a cloud, thereby changing it to a cloud-based IT resource.
- Redundant deployments of an IT resource can exist in both on-premise and cloud-based environments.

## Cloud Consumers and Cloud Providers

- The party that provides IT resources is the cloud provider.
- The party that uses cloud-based IT resources is the cloud consumer.

## Scaling

Scaling, from an IT resource perspective, represents the ability of the IT resource to handle increased or decreased usage demands.

Types of scaling:

- Horizontal Scaling - scaling out and scaling in
- Vertical Scaling - scaling up and scaling down.

### Horizontal Scaling

- The allocating or releasing of IT resources that are of the same type is referred as horizontal scali
- The horizontal allocation of resources is referred to as scaling out.
- The horizontal releasing of resources is referred as scaling in.
- It is a common form of scaling within cloud environments.

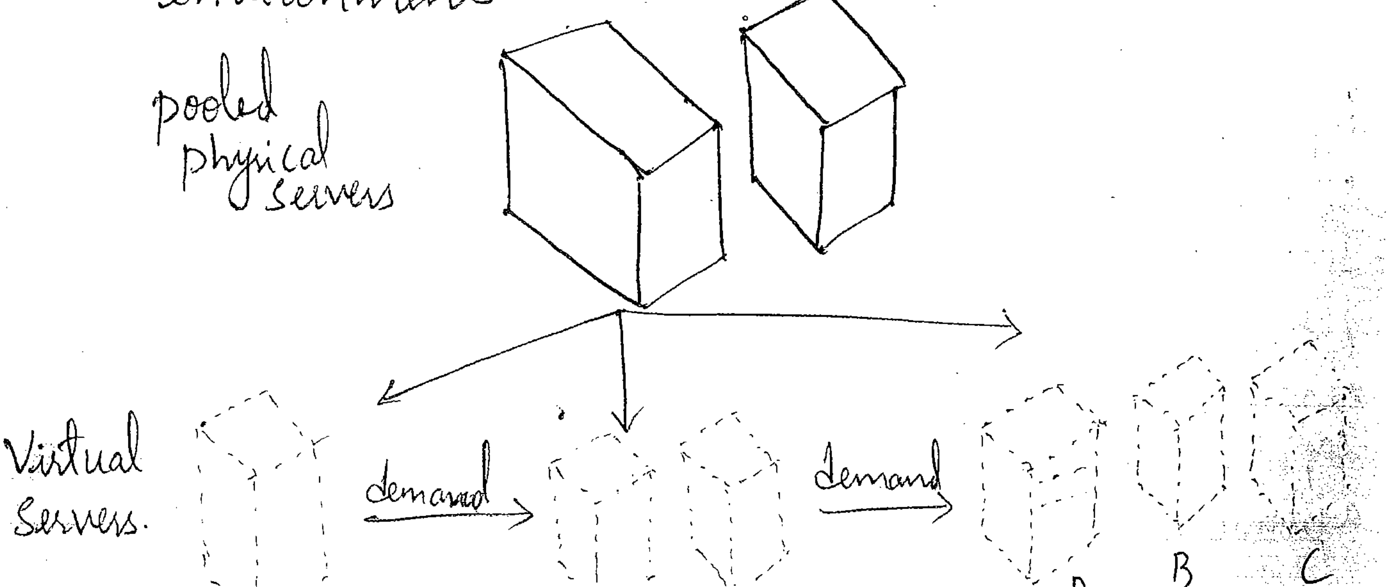


Fig: An I resource, Virtual Server A, is scaled out by adding more of the same IT resources, Virtual servers B & C.

## Vertical Scaling.

- When an existing IT resource is replaced by another with higher or lower capacity, is called vertical scaling.
- The replacing of an IT resource with another that has a higher capacity is referred to as scaling up.
- The replacing an IT resource with another that has a lower capacity is considered scaling down.
- Vertical scaling is less common in cloud environment due to the downtime required while replacement is taking place.

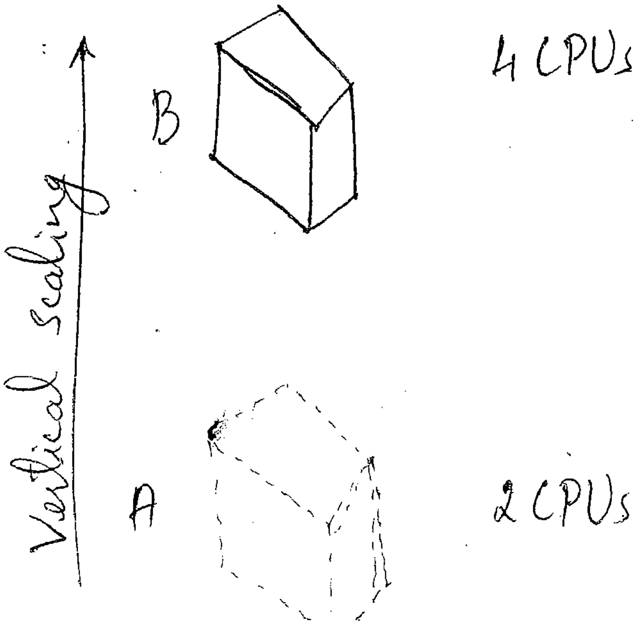


Fig: An IT resource, a virtual server with two CPUs, is scaled up by replacing it with a more powerful IT resource with increased capacity for data storage, a physical server with 4CPUs.

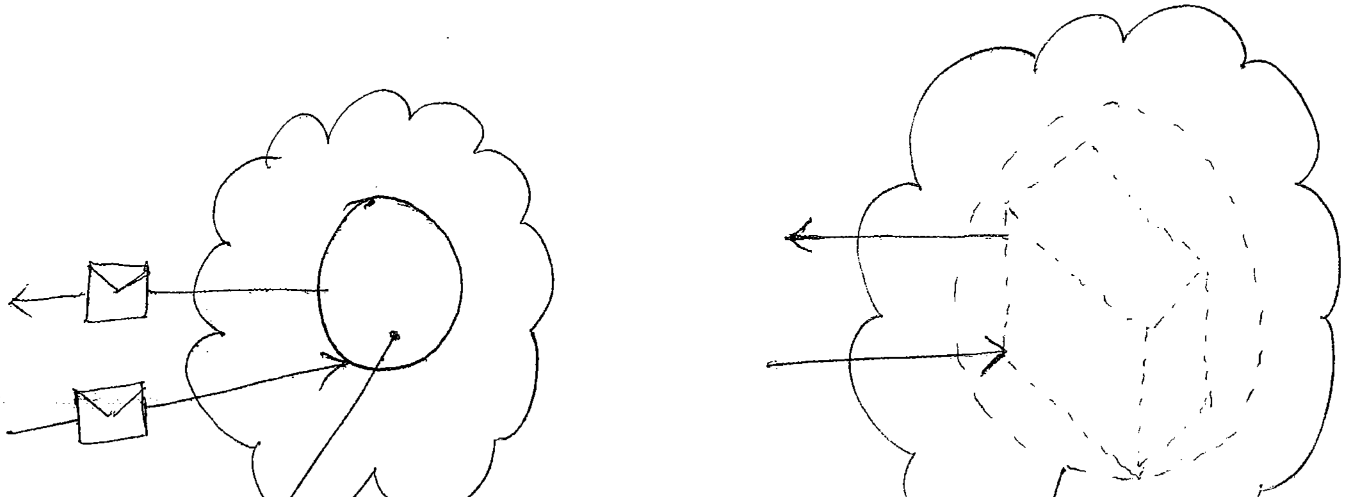
## Comparison of horizontal and Vertical scaling

| Horizontal Scaling | Vertical Scaling |
|---|---|
| 1) less expensive (through commodity hardware components) | - more expensive (specialized servers) |
| 2) IT resources instantly available | - IT resources normally instantly available |
| 3) resource replication & automated scaling | - additional setup is normally needed |
| 4) additional IT resources needed | - no additional IT resources needed |
| 5) not limited by hardware capacity | - limited by maximum hardware capacity. |

## Cloud Service

- Not all IT resources residing within a cloud can be made available for remote access. For Ex, a database or a physical server deployed within a cloud may only be accessible by other IT resourc that are within the same cloud.
- A software program with a published API may be deployed specifically to enable access by remote clients.

- A cloud service is any IT resource that is made remotely accessible via a cloud.
- A cloud service can exist as a simple Web-based software program with a technical interface invoked via the use of a messaging protocol or as a remote access point for administrative tools or larger environments & other IT resources.



remotely accessed Web Service acting as a cloud service

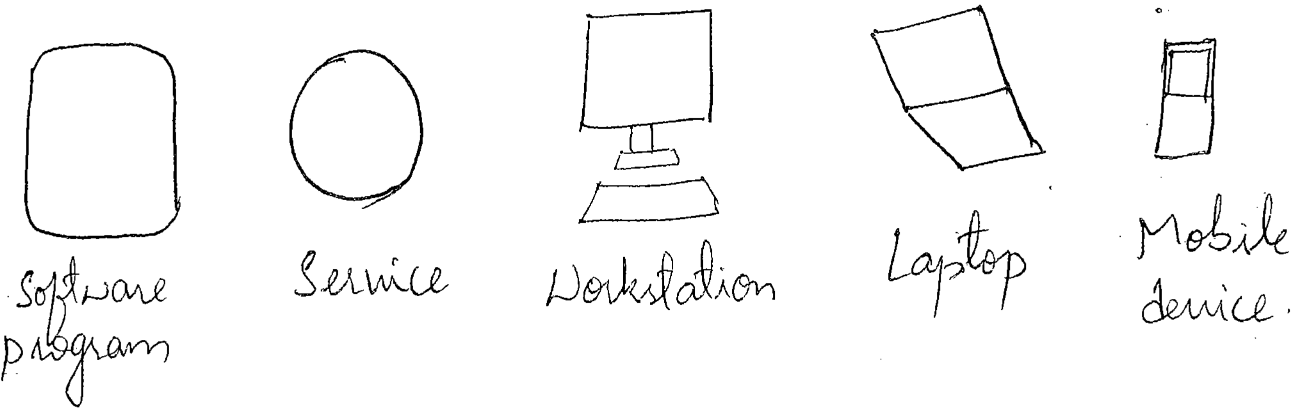
remotely accessed virtual Server acting as a cloud service.

Fig: A cloud service with a published technical interface is being accessed by a consumer outside of the cloud (left). A cloud service that exists as a virtual server is also being accessed from outside of the cloud's boundary (right). The cloud service on the left is likely being invoked by a consumer program that was designed to access the cloud service's published technical interface. The cloud service on the right may be accessed by a human user that has remotely logged on to the virtual server.

## Cloud Service Consumer

Cloud service consumer is a temporary runtim role assumed by a software program when it accesse a cloud service.

As shown in fig below, common types of cloud service consumers can include software programs & service capable of remotely accessing cloud services with published service contracts, as well as workstations, laptops & mobile devices running software capable of remotely accessing other IT resou positioned as cloud services.



Common measurable benefits to cloud consumers are:

- On-demand access to pay-as-you-go computing resources on a short-term basis & the ability to release these computing resources when they are no longer needed.
- The perception of having unlimited computing resources that are available on demand, thereby reducing the need to prepare for provisioning.
- The ability to add or remove IT resources at a fine-grained level, such as modifying available storage disk space by single gigabyte increments
- Abstraction of the infrastructure so applications are not locked into devices or locations & can be easily moved if needed.
- The financial benefits of dynamic scaling and the risk transference of both over-provisioning and under-provisioning must also be accounted for.

## 2) Increased Scalability

- By providing pools of IT resources, along with tools and technologies, cloud can instantly and dynamically allocate IT resources to cloud consumers, on-demand or via the cloud consumer's direct configuration.
- This empowers cloud consumers to scale their cloud-based IT resources to accommodate processing fluctuations and peaks automatically or manually.
- The inherent, built-in feature of cloud to provide flexible levels of scalability to IT resources is directly related to proportional costs benefit.
- The ability of IT resources to always meet & fulfill unpredictable usage demands avoids potential loss of business that can occur when usage thresholds are met.
- The lag and Match Strategies are generally more applicable due to a cloud's ability to scale IT resources on demand.

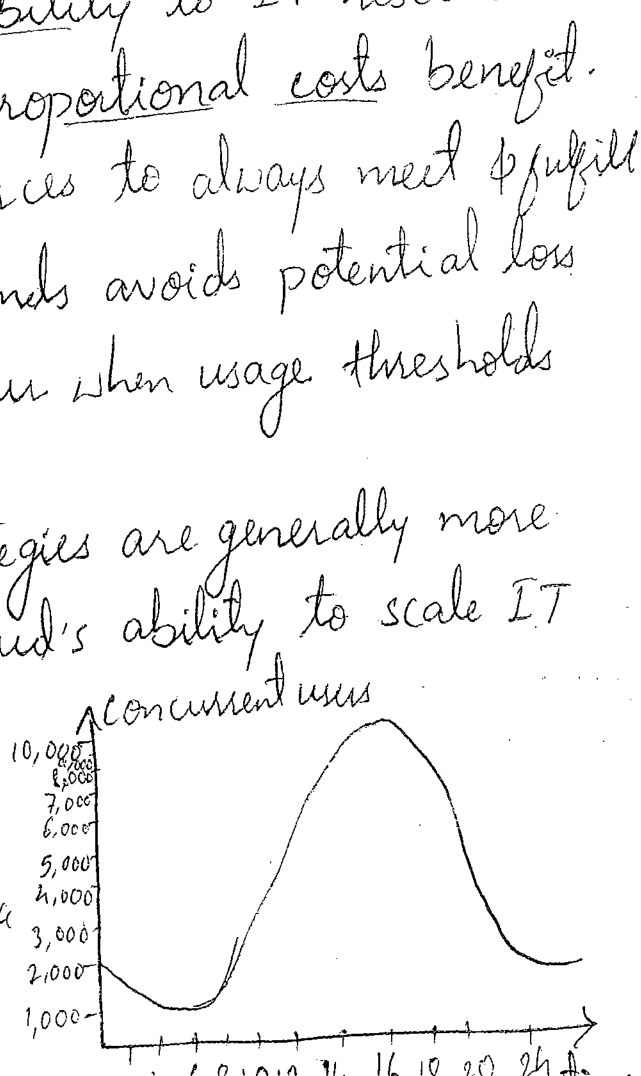


Fig: An example of an organization's changing demand for an IT resource over the course of a day.

## 3) Increased Availability and Reliability

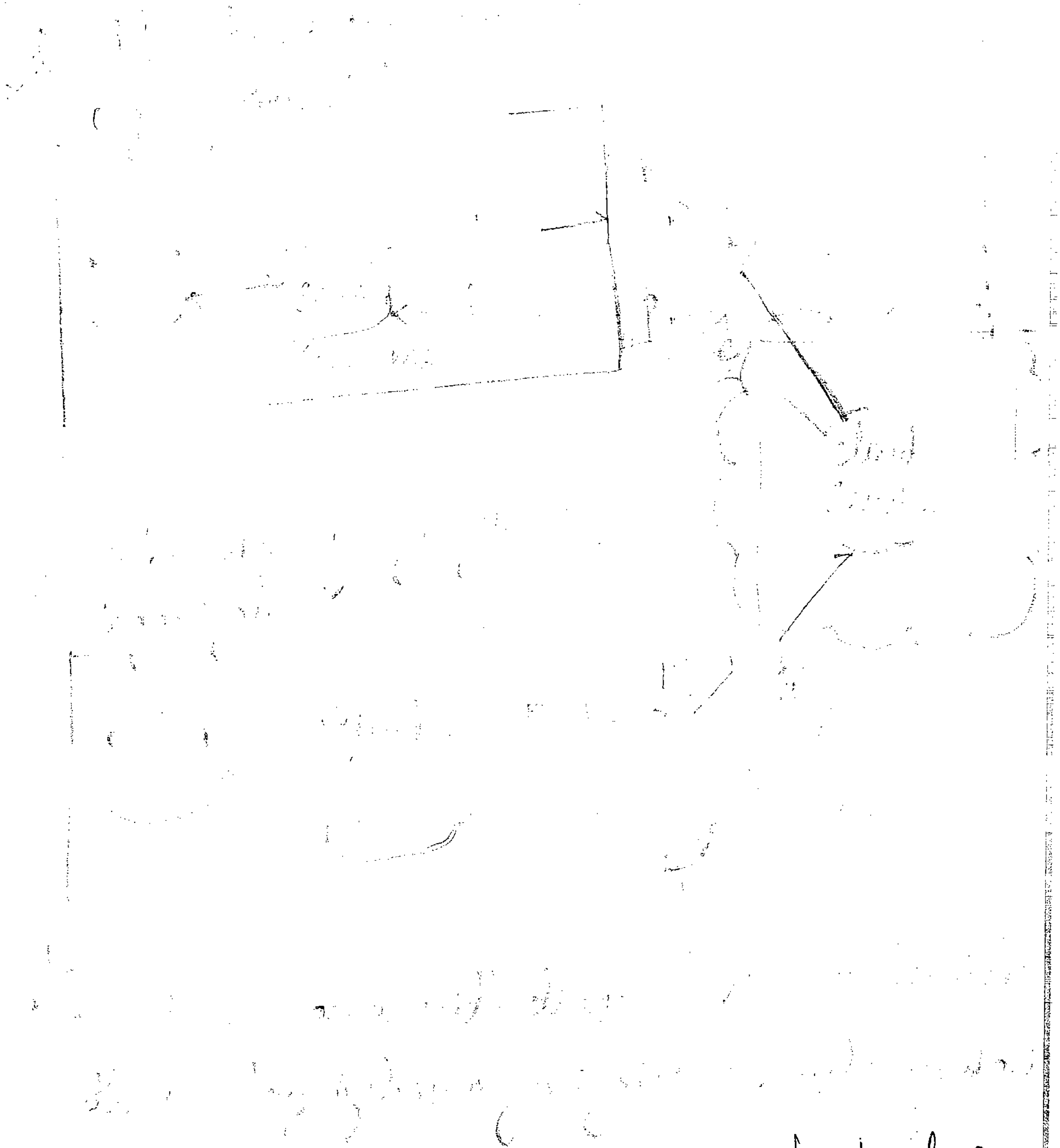
- An IT resource with increased availability is accessible for longer periods of time.
- Cloud providers generally offer "resilient" IT resourc for which they are able to guarantee high levels of availability.
- An IT resource with increased reliability is able to better avoid & recover from exception conditions.
- The modular architecture of cloud environments provides extensive failover support that increase reliability.

## Risks and Challenges

### 1) Increased Security Vulnerabilities

- The moving of business data to the cloud means that the responsibility over data security becomes shared with the cloud provider.

- The extent to which the data is secure is now limited to the security controls and policies applied by both the cloud consumer and cloud provider.
- There can be overlapping trust boundaries from different cloud consumers due to the fact that cloud-based IT resources are commonly shared.
- The overlapping of trust boundaries & the increased exposure of data can provide malicious cloud consumers with greater opportunities to attack IT resources & steal or damage business data.



The above fig illustrates a scenario whereby two organizations accessing the same cloud service are required to extend their respective trust boundaries to the cloud, resulting in overlapping trust boundaries.

2) <u>Reduced Operational Governance Control</u>

- Cloud Consumers are usually allotted a level of governance control that is lower than that over on-premise I.T resources.

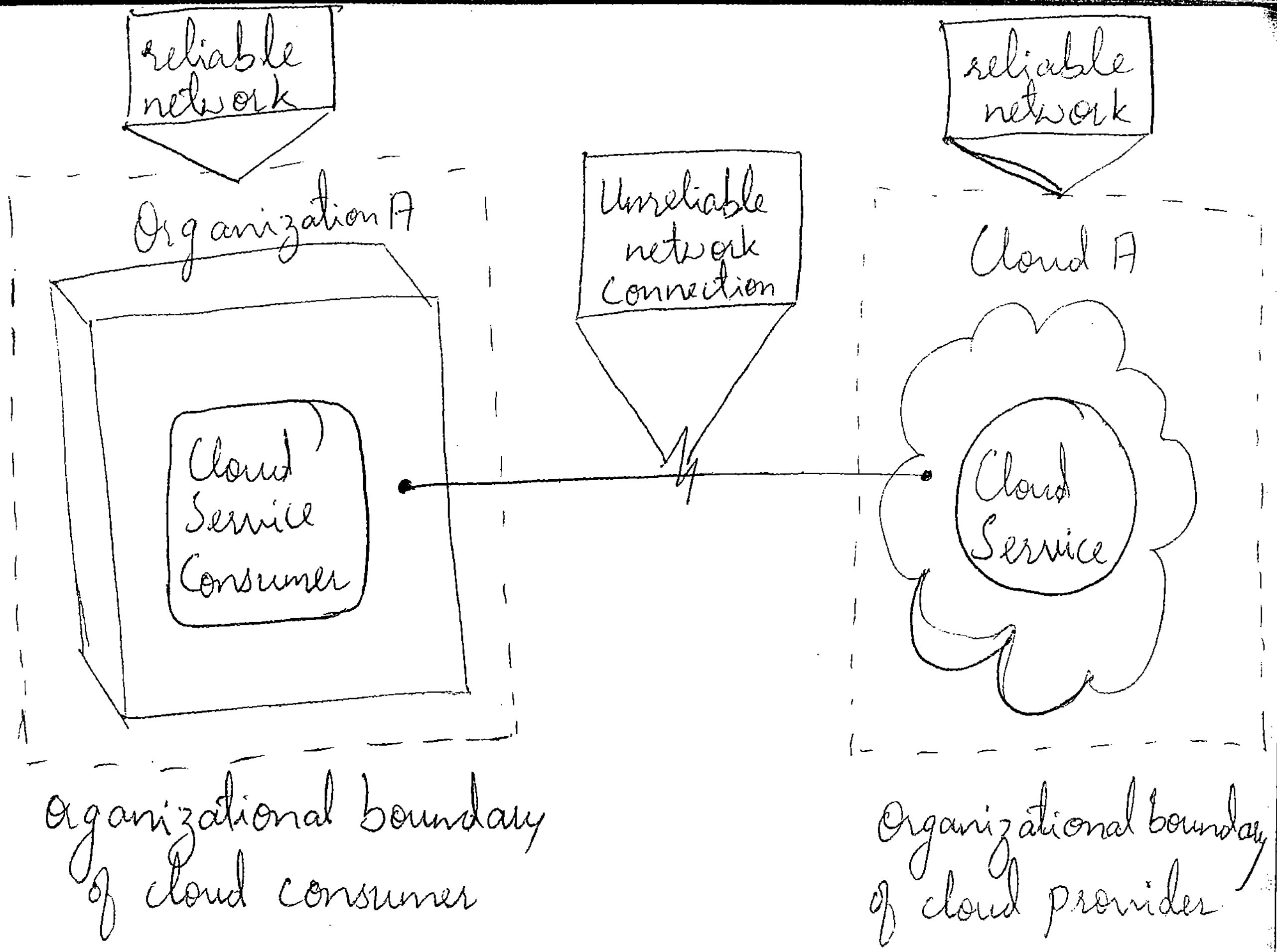
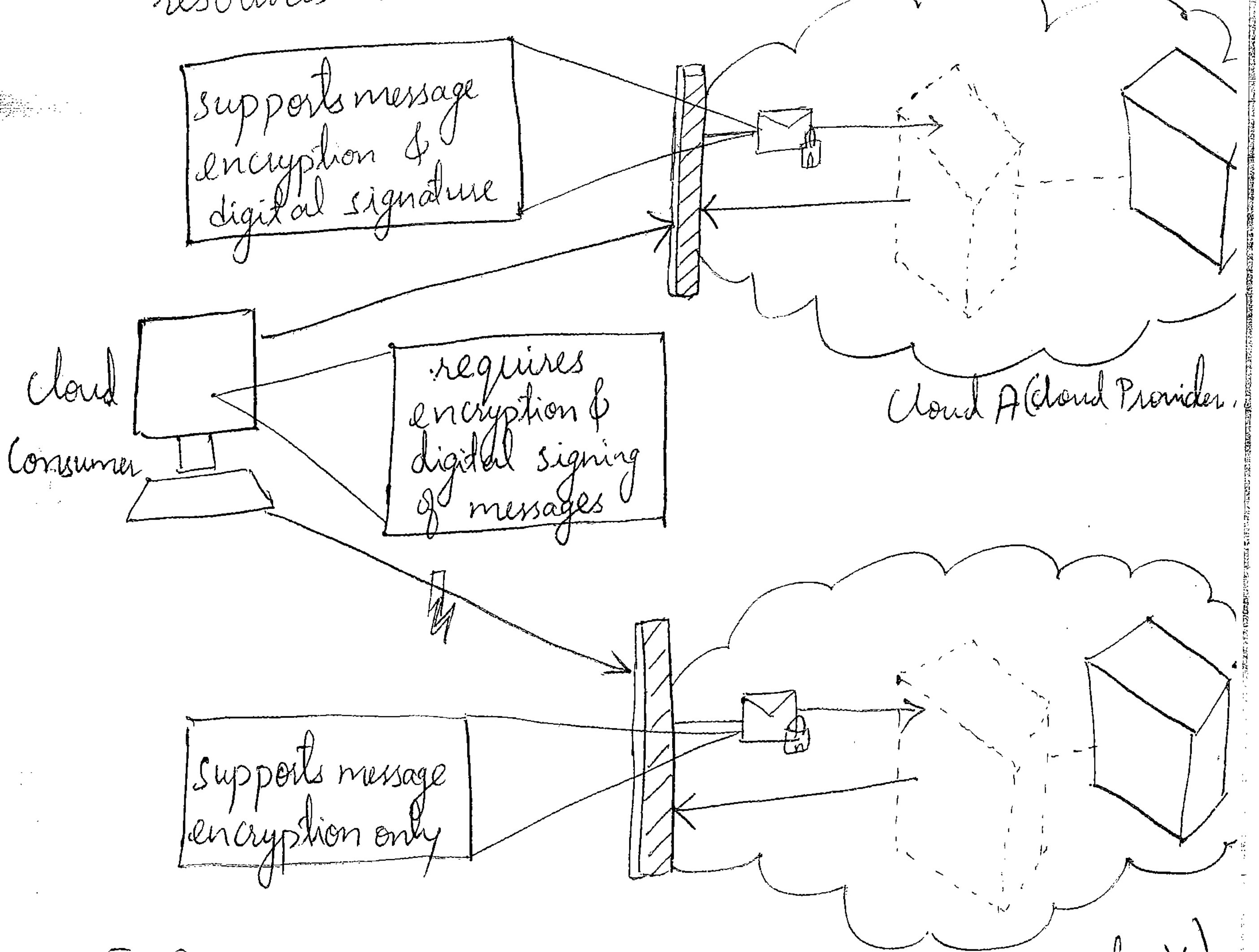

reliable network
Organization A
Unreliable network connection
reliable network
Cloud A
Cloud Service Consumer
Cloud Service
Organizational boundary of cloud consumer
Organizational boundary of cloud provider

Fig: An unreliable network connection compromises the quality of communication between cloud consumer and cloud provider environments.

- Longer geographic distances between the cloud consumer and cloud provider can require additional network hops that introduce fluctuating latency and potential bandwidth constraints.
- Legal contracts, when combined with SLA's, technology inspections, and monitoring, can lessen the risks and issues.

3) Limited Portability Between Cloud Providers

- Due to a lack of established industry standards within the cloud computing industry, public clouds are commonly proprietary to various extents.
- For cloud consumers that have custom-built solutions with dependencies on these proprietary environments, it can be challenging to move from one cloud provider to another.
- Portability is a measure used to determine the impact of moving cloud consumer IT resources & data between clouds. Fig below:


supports message encryption & digital signature
Cloud Consumer
requires encryption & digital signing of messages
Cloud A Cloud Provider.
supports message encryption only

FigA: A cloud consumers' application has a decreased level of portability when assessing a potential migration from cloud A to cloud B, because the cloud provider of cloud B does not support the same security technologies as cloud A.

4) Multi-Regional Compliance and Legal Issues

- Third-party cloud providers will frequently establish data centers in affordable or convenient geographical locations.

- Cloud consumers will often not be aware of the physical location of their IT resources & data when hosted by public clouds.

- For some organizations, this can pose serious legal concerns pertaing to industry or government regulations that specify data privacy & storage policies.

- Another potential legal issue pertains to the accessibility & disclousure of data.

# FUNDAMENTAL CONCEPTS & MODELS

## Roles and Boundaries

1. **Cloud Provider**

- The organization that provides cloud-based IT resource is the cloud provider
- An organization is responsible for making cloud services available to cloud consumers, as per agreed upon SLA guarantees.
- Provider has to take care of any required managemen and administrative duties to ensure the on-going operation of the overall cloud infrastructure.
- Cloud providers normally own the IT resources that are made available for lease by cloud consumers.
- Some cloud providers also "<u>resell</u>" IT resources leased from other cloud providers.

**Cloud Consumer**

- A cloud Consumer is an organization (or a human) that has a formal contract or arrangement with a cloud provider to use IT resources made available by the cloud provider.

## Cloud Service Owner

- The person or organization that legally owns a cloud service is called a cloud service owner.
- The cloud service owner can be the cloud consumer, or, the cloud provider that owns the cloud within which the cloud service resides.
- A cloud consumer that owns a cloud service hosted by a third-party cloud does not necessarily need to be the user of the cloud service.
- Several cloud consumer organizations develop and deploy cloud services in clouds owned by other parties for the purpose of making the cloud services available to the general public.
- The reason a cloud service owner is not called a cloud resource owner is because the cloud service owner role only applies to cloud services.

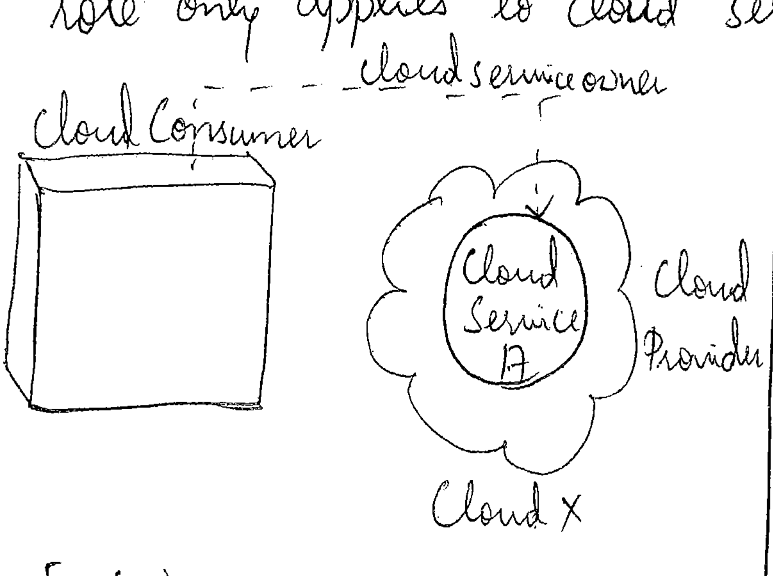
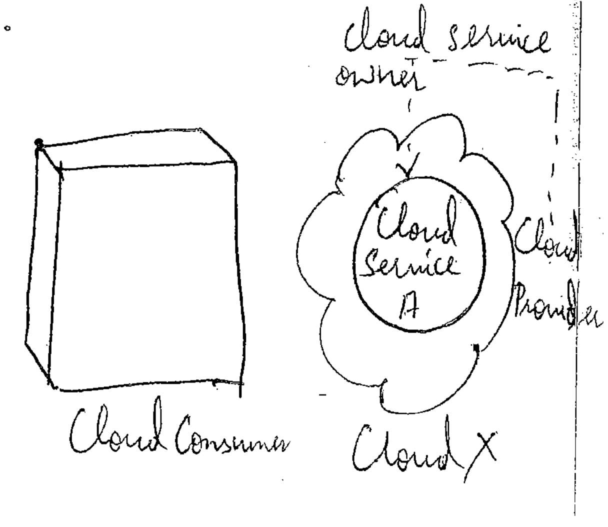


Fig:(a)



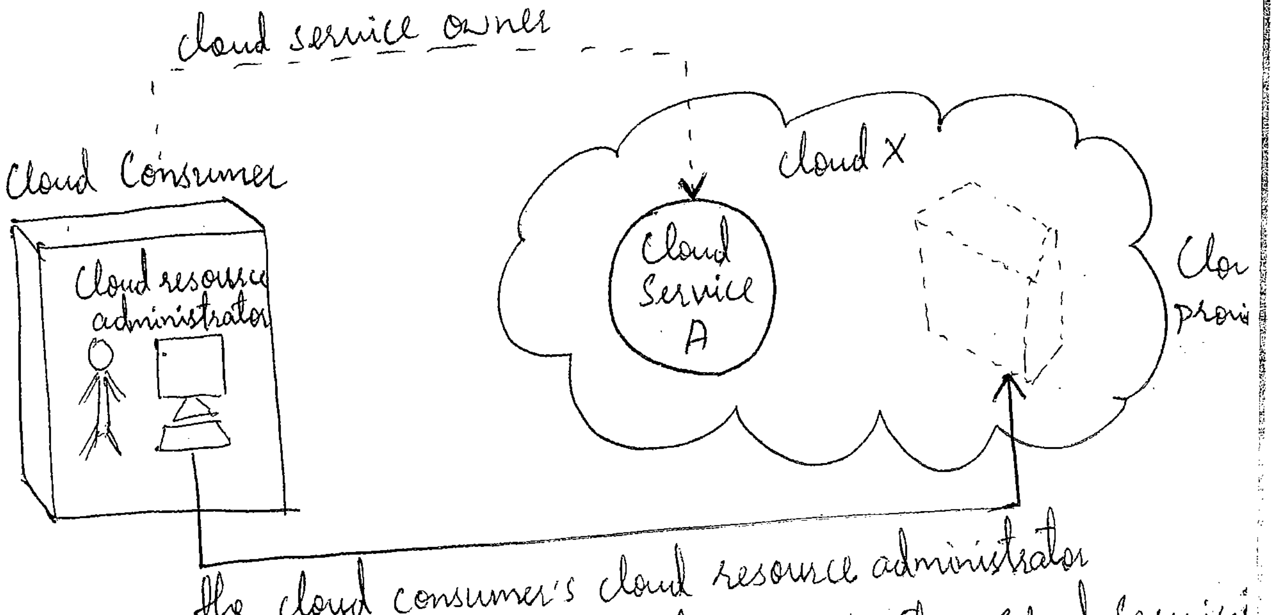
Fig(b)

Fig(a): A cloud Consumer can be a cloud service owner when it deploys its own service in a cloud.

Fig(b): A cloud provider becomes a cloud service owner if it deploys its own cloud service, typically for other cloud consumer to use.

## Cloud Resource Administrator

- A cloud resource administrator is the person or organization responsible for administering a cloud-based IT resource.
- The cloud resource administrator can be the cloud consumer or cloud provider of the cloud withi which the cloud service resides. Alternatively, it cc be a third-party organization contracted to administer the cloud-based IT resource.



The cloud consumer's cloud resource administrator remotely accesses the virtual server hosting Cloud Service

## Additional Roles

- Cloud Auditor
- Cloud Broker
- Cloud Carrier.

- Cloud Auditor - A third party that conducts independent assessments of cloud environment, evaluation of security controls, privacy impacts, & perfamance is called cloud Auditor.
  - The main purpose of the cloud auditor is to provide an unbiased assement of cloud to help strengthen the trust relationship between consumers & providers.
- Cloud Broker - party that assumes the responsibility of managing and negotiating the usage of cloud services like intermediation, aggregation & arbitrage, between cloud consumer & provider.
- Cloud Carrier - party responsible for providing the wire-level connectivity bet$^{n}$ cloud consumer & provider.

Organizational Boundary

- An organizational boundary represents the physical perimeter that surrounds a set of IT resources that are owned & governed by an organization.

Trust Boundary

A trust boundary is a logical perimeter that typically spans beyond physical boundaries to represents the extent to which IT resources are trusted.

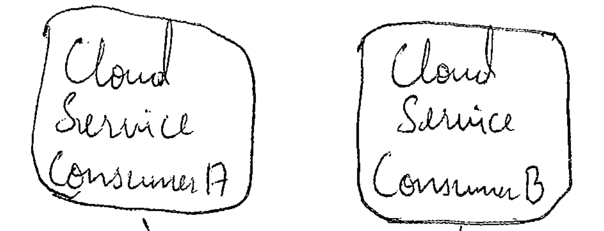
Cloud Characteristics

1. On-demand usage - A cloud consumer can unilaterally access cloud-based IT resources giving the cloud consumer the freedom to self-provisioned IT resource that can be automated, requiring no further human involvement

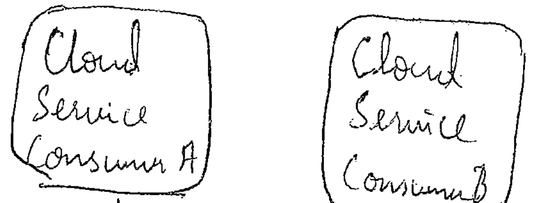
2. Ubiquitous Access - It is ability for a cloud service to be widely accessible.
   - Ubiquitous access for a cloud serive can require support for a range of devices, transport protocols, interfaces, & security technologies that is required to provide particular needs of different cloud service consumer.

## 3) Multitenancy (and Resource Pooling)

- The characteristic of a software program that enables an instance of the program to serve different consumers whereby each is isolated from the other, is referred to as multitenancy.
- Multitenancy rely on the use of virtualization.
- Different physical and virtual IT resources are dynamically assigned & reassigned according to cloud consumer demand.



Fig(a): Single tenant environment, each cloud consumer has a separate IT resource instance.



Fig(b): Multitenant environment, a single instance of an IT resource, such as cloud storage device, serves multiple consumer

4) <u>Elasticity</u>

- Elasticity is the automated ability of a cloud to transparently scale IT resources.
- Reduced Investment & Proportional costs benefit.

5) <u>Measured Usage</u>

- The measured usage characteristic keep track of the usage of its IT resources & used by consumers.
- The cloud provider can charge a cloud consumer only for the IT resources actually used and/or for the timeframe.
- Helps in tracking statistics for billing purposes.
- General monitoring of IT resources & related usage.

6) <u>Resiliency</u>

- Resilient computing is a form of failover that distributes redundant implementations of IT resources across physical locations.
- Resiliency can refer to redundant IT resources within the same cloud or across multiple clouds.

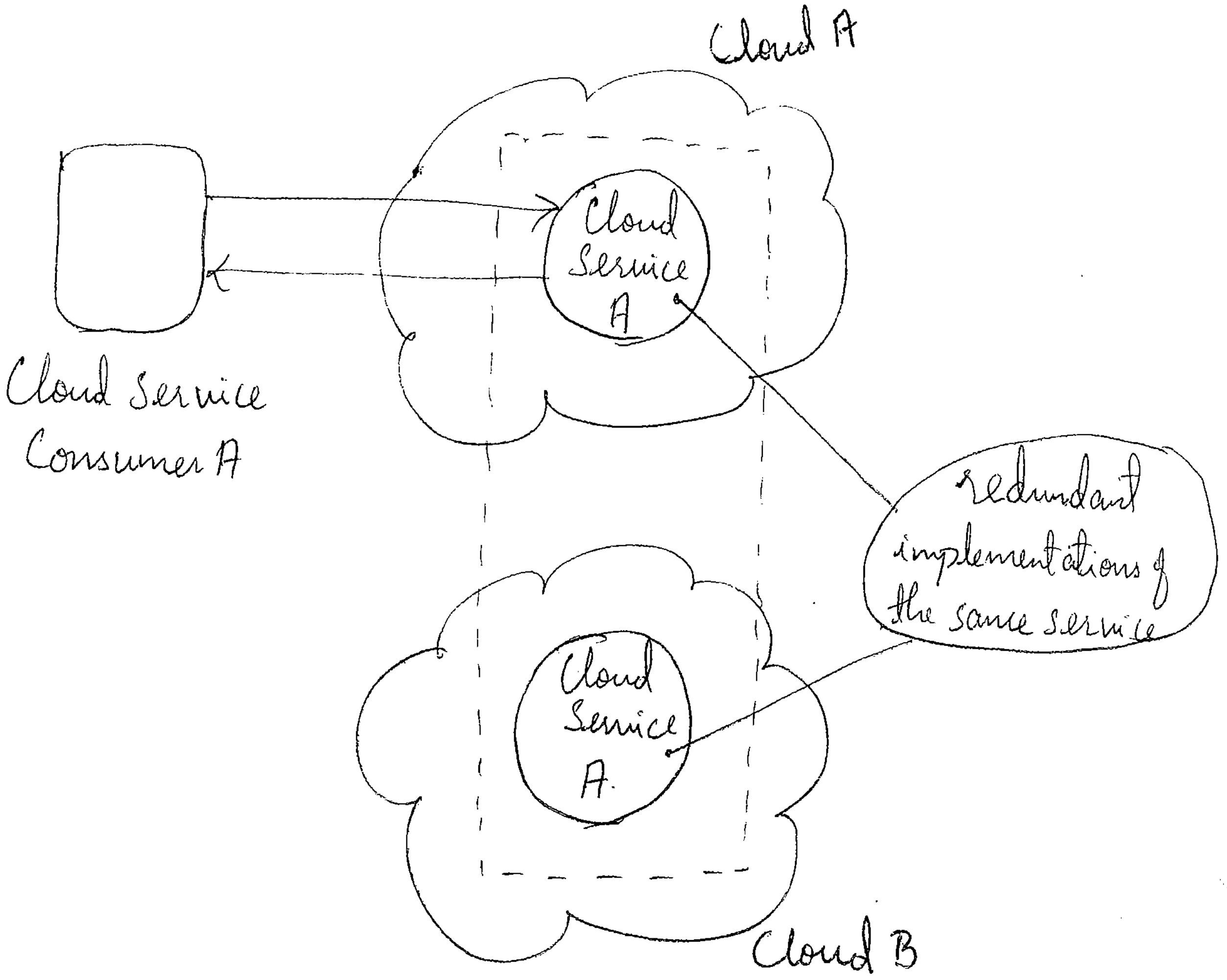


Fig: A resilient system in which cloud B hosts a reduntant implementation of cloud Service A to provide failover in case cloud service A on cloud A becomes unavailable.

# Cloud Delivery Models

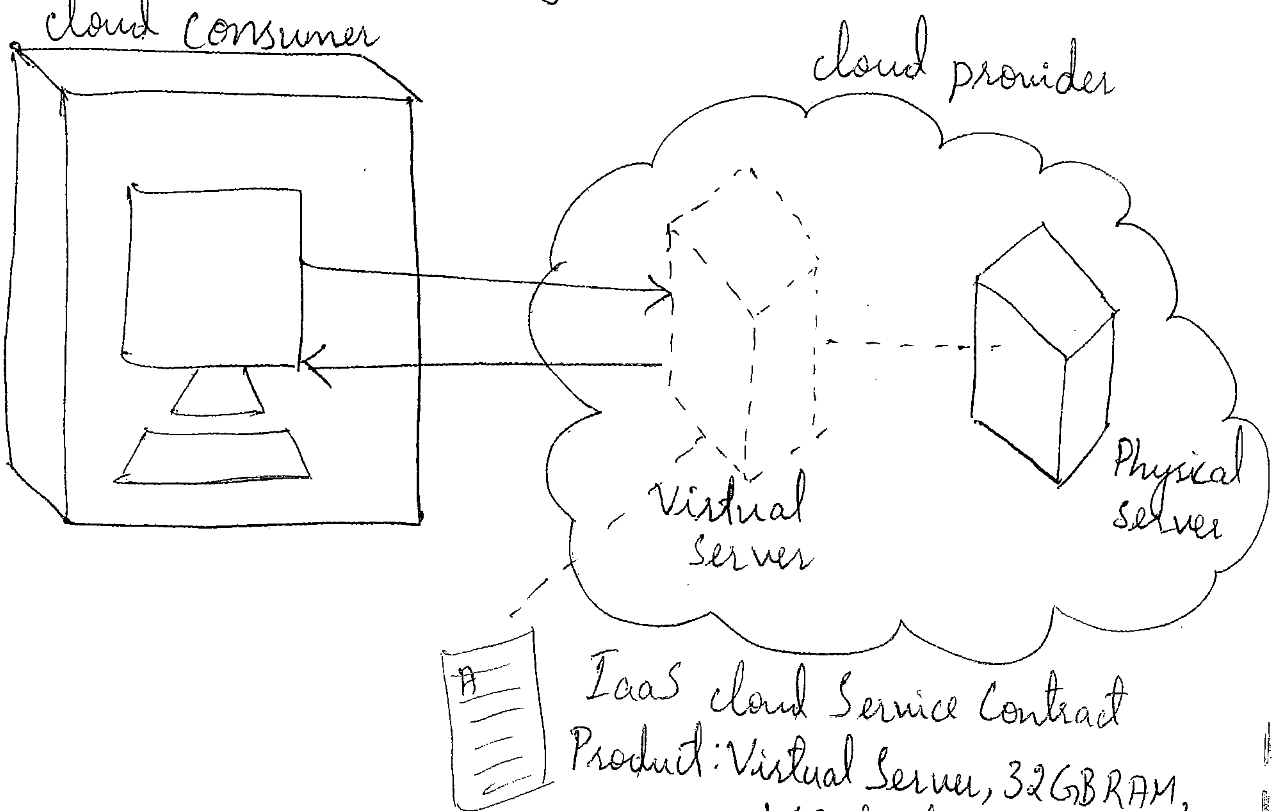
A cloud delivery model represents a specific, pre-packaged combination of IT resources offered by a cloud provider.

- Infrastructure-as-a-Service (IaaS)
- Platform-as-a-Service (PaaS)
- Software-as-a-Service (SaaS)

## Infrastructure-as-a-Service (IaaS)

- IaaS delivery model represents a self-contained IT environment comprised of infrastructure-centric IT resources that can be accessed & managed via cloud service-based interfaces & tool
- This environment can include hardware, network, connectivity, operating sytems, & other 'raw' IT resource
- Cloud consumers will have high level of control and responsibility over IaaS environment configuration & utilization.
- IT resources available through IaaS environments are generally offered as freshly initialized virtual instances.

- Virtual servers are leased by specifying server hardware requirements, such as processor capacity, memory, & local storage space as shown below:



IaaS cloud Service Contract
Product: Virtual Server, 32GB RAM, 4GB local storage
SLA: availability = 99.5%, no failover
Price: $0.95 per hour, $0.05 per GB transfered out of cloud.

Fig: A cloud consumer is using a virtual server within an IaaS environment. Cloud consumers are provided with a range of contractual guarantees by cloud provider, pertaining to characteristics such as capacity, performance, & availability.

## Platform-as-a-Service (PaaS)

- PaaS delivery model represents a pre-defined "ready-to-use environment typically comprised of already deployed & configured IT resources.

- Common reasons a cloud consumers would use & invest in a PaaS environment include:
  - The cloud consumer wants to extend on-premise environments into the cloud for scalability & economic purposes.

Paper : DDoS attacks in cloud computing : Collateral Damage to non-targets

Intro : Survey by Kaspersky in 2014, 1/5 business have been attacked by DDoS.

- harmed business losses, reputation, downtime.
- SaaS are most attacked.
- 4 major attacks
  - Attack on Microsoft & Sony gaming servers by Lizard Squad were first.
  - Attack on Rackspace
  - Amazon EC2 server
  - Linode, for over 1 week, the attack

DDoS/EDoS attack in the cloud

9