Unit-1	**Introduction to Cloud Computing**	Contact Hours: 15
	Chapter1: Introduction and history of cloud computing, Characteristics and advantages of cloud computing, Cloud service providers and market overview, Virtualization and its role in cloud computing. Chapter 2: Cloud Deployment Models: Public, private, and hybrid cloud models, Comparisons and selection criteria for deployment models, Case studies and examples of cloud deployments. .	
Unit-2	**Cloud Service Models and Security**	Contact Hours: 15
	Chapter1: Cloud Service Models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), Comparison and use cases of service models. Cloud Platform and Tools: Introduction to popular cloud platforms (AWS, Azure, Google Cloud). Chapter2: Security in Cloud Computing, Security challenges and threats in the cloud, Identity and access management, Data privacy and protection, Compliance and legal issues in the cloud	
Unit-3	**Scaling Techniques and Emerging Trends in Cloud Computing**	Contact Hours: 15
	Chapter1: Scalability and Performance Optimization, Scaling techniques for cloud-based applications. Load balancing and auto-scaling, Performance monitoring and optimization. Chapter2: Emerging Trends in Cloud Computing, Server less computing, Edge computing, Internet of Things (IoT) and cloud integration, Future directions and innovations in cloud computing.	

Cloud Computing →

- Storing and accessing the data and programs on remote servers that are hosted on internet rather than on local server.
- On demand service with pay as you go pricing.

1961 – John MacCharty
1999 – salesforce.com
2010 – Microsoft launch
2002 – Amazon (web services) 2006 – Elastic compute cloud
2008 – Google Apps

Before Cloud Computing →

- Client Server Architecture
  where all the data and control of client resides in server side.
- Destributed Computing
  all computers are networked together, user can share their resource when needed.

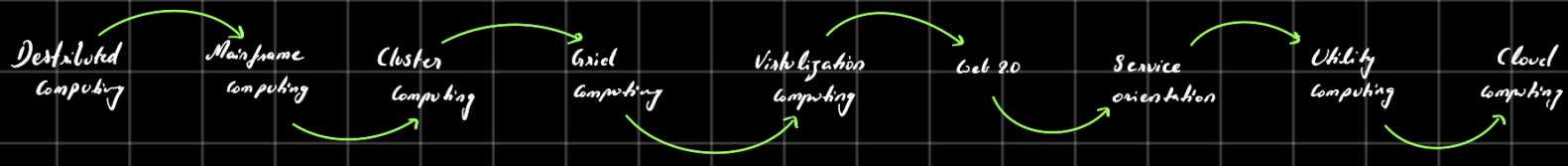
Client server Computing
↓
Distributed Computing
↓
Cloud Computing

Characteristic of Cloud Computing →

- Agility
  - works in distributed computing environment.
  - shares resources among users and works very fast.
- High availability and reliability
  - chances of failure less
  - service availability
- High Scalibility
  - On demand

- Multisharing
  - Multiple users and applications can work more efficiently.
  - Sharing common Infrastructure.
- Device and Location Independence
  - can be accessed with internet connectivity
- Maintainance
  - No maintainance required
- Low Cost
  - Rental Based

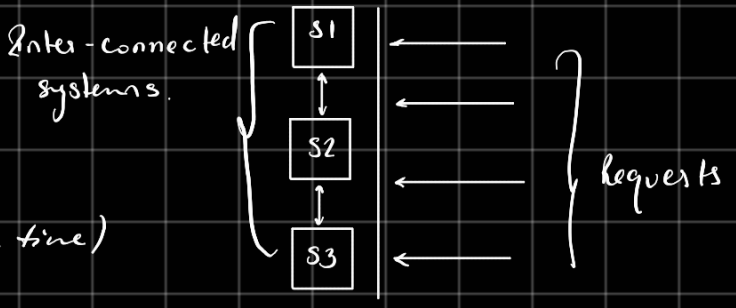
- Application of Cloud Computing in real world →
  1) Online Data Storage
  2) Backup and Recovery
  3) Testing and Development
  4) Anti-Virus Application



Destributed →

- A collection of Independent computers that appears to its users as one computer.
- Purpose of distributed system is to share resources and use them effectively and efficiently.

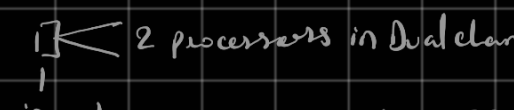
Characteristics →
1) Scalable
2) Availability
3) Concurrency (happening at the same time)



Mainframe System →

- Can handled a huge amount of request and responds to them at the same time.
- Has close to no delay in processing.
- Highly reliable

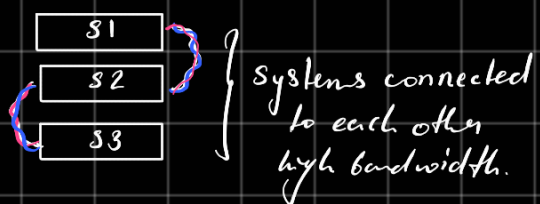
Con → Expencive



1 Main frame IBM have
600 processors
+
SAP
System Assistant Processor

Cluster Computing →

- Each machine is connected to each other with high bandwidth.



- Equally capable of high computation.

Con → Geographic lock

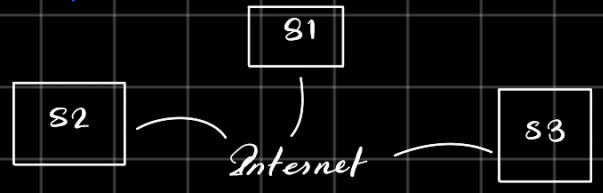
Middleware → (Manager)
- Administer
- Overload
- Security

Homogeneous Network
- System with same configuration.

Heterogeneous Network
- System with non-same configuration.

Grid Computing →

- Different systems are at different geo location and are connected to each other with internet.



eg :- ATM's.

nodes → Controller, Provider, User

switching of nodes also happens

Virtualization →

- Creating a virtual layer over the hardware allowing to run multiple instances simultaniously on the hardware.
- Physical resources available as virtual resources.

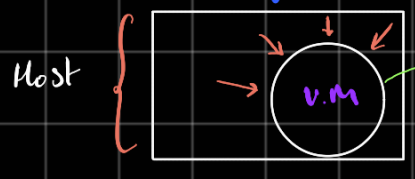
Virtualization allows multiple environment to run on a single physical computer or server.
It involves creating a virtual representative of physical hardware.

Hypervisor manages and controll the allocation of physical resources to VMs.

Concept behind the Virtualization →
- Host machine
- Guest machine
- Hypervisor

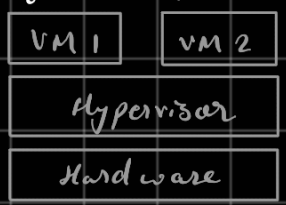
Hardware Virtualization →

Creating virtual machine over existing O.S. and H/w.



V.M is managed by S/w "Hypervisor"

Architecture of Virtualization →



Type - 1

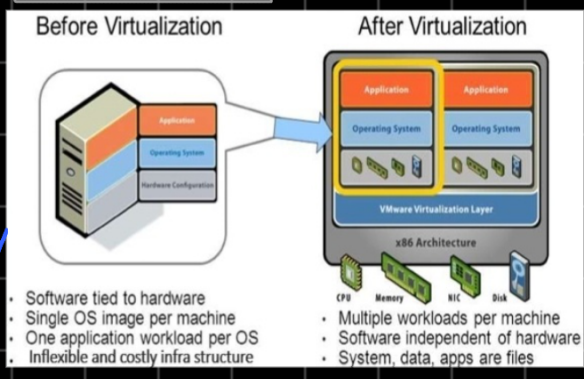
- Runs directly on the hardware
- High performance
- More expensive
- More secure
- More complex
- Costly / Corporate use

Bare metal hypervisor

Type - 2

Hosted hypervisor

- Runs over an existing OS
- Less performant
- Less expensive
- Less secure
- Less complex
- Preferred / small industry use.



Types of Virtualization →

- **Hardware Virtualization**

  Virtual Machine Manager (VMM) is directly installed on the hardware.

  The main job of hypervisor is to control and monitor the procesr, storage, other hardware resources.

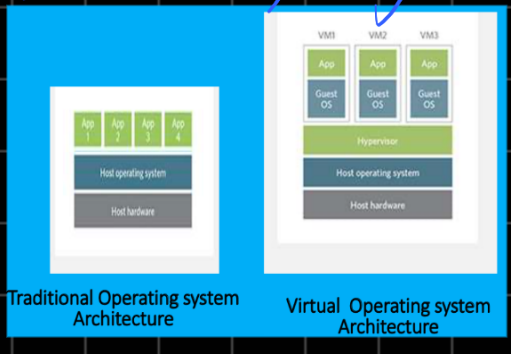
- **Operating System Virtualization**

  When Virtual Machine Manager (VMM) is installed on the host operating system

- **Server Virtualization**

  When Virtual Machine Manager is is directly installed on the Server System.

  Is done to divide single physical server into multiple servers.



- **Storage Virtualization**

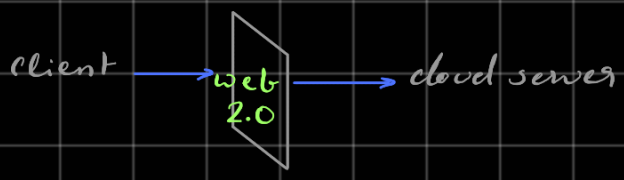
  Grouping up of all physical storage through out the network and looking like one single.

Types of Server Virtualization →

Hypervisor	Paravirtualization	Full Virtualization
• VMM between OS and hardware.	• Improves VM by modifing guest OS.	• Single host per multiple guests.
• Setups VM on the system.		• Comple simulation of hardware
• Share host's resources.		

Web 2.0 →

Interface through which the cloud computing services interacts with the client.



EC2 RDS S3

AWS – 2006

MAzure – 2010

GC - fastest

Service Orientation →

SaaS

PaaS

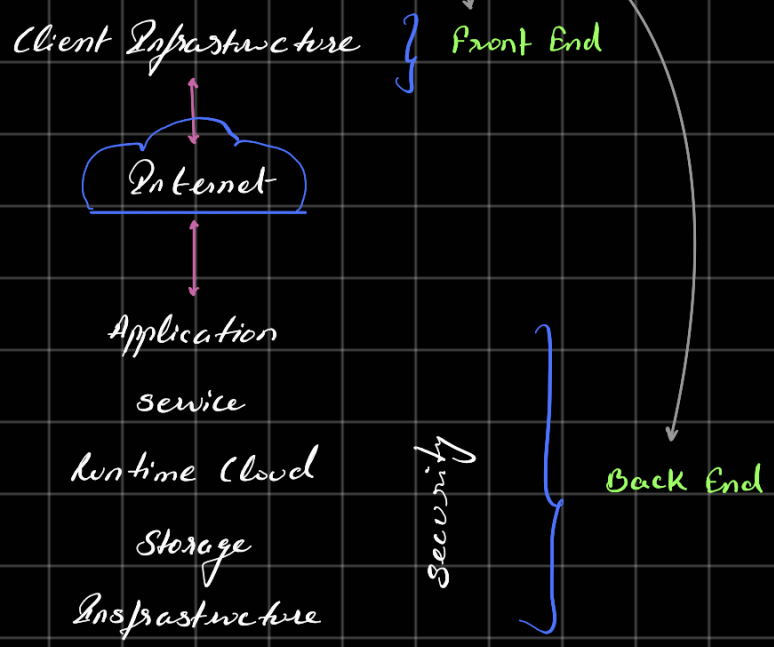
PaaS

XaaS

Utility Computing →

is a cloud computing service model.

Cloud Computing Architecture



- Client Infrastructure →

Client Infrastructure is a front end provides GUI to interact with the cloud.

- Application

  Platform that a client wants to access.

- Services

  cloud computing provides different types of services →

  1. SaaS (Software as a Service) ex → Google maps (service over the internet access using web.)
  2. PaaS (Platform as a Service) ex → Window Azure

     PaaS provides a platform for software creation, by using SaaS we can access software over the Internet without the need of any platform.
  3. IaaS (Interface as a Service) ex → AWS, GCE

     It is responsible for managing application data, middleware and runtime environment.
  4. XaaS (Anything as a Service) delivery of anything as a service.

- Runtime Cloud

  Provides execution and runtime environment to the virtual machines.

- Storage

  It provides huge amount of storage capacity in the cloud storage and manage data.

- Infrastructure

  Provides services on the host level, application level and network level.
  cloud infrastructure includes hardware and software.

- Management

  Used to manage - application, services, runtime cloud, storage, infrastructure.

- Security

  Inbuilt back end component of the cloud computing.

- Internet

Internet is the medium through which front end and backend can interact and communicate with each other.

Cloud computing	Grid Computing
• Client server	• Distributed
• Centralised Management System	• Distributed Management System
• Resources are used in centralized manner.	• Resources are used in colaborative manner.
• High Accessibility	• Low Accessibility
• Pay as you go	•
• AaaS, PaaS, IaaS	• Destributed Computing, Destributed Information.

Utility Computing →

- User only needs to pay for what he is using for how long.
- User can customize as per his needs.

How utility computing works →

- One company pays another company for computing services.
- Service might include : hardware, software, storage or computing power.
- Many companies offers bundle of resources.

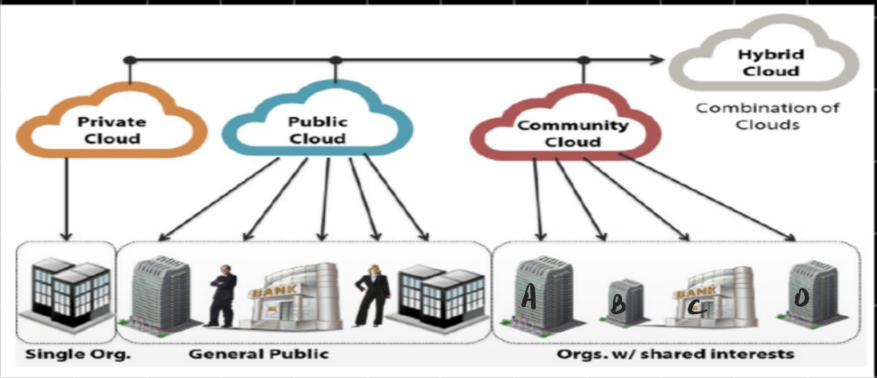
Advantages →

- No need to buy any hardware, software or licence to do business
- Large company with many department, problem can arise with computing software. Each department might depend on different software suits. One file might be incompatible for other department.

Disadvantages →

- If the provider faces financial issue then the service could get cut off
- Attractive target for hackers.
- Reliability.

Deployment Models of Cloud computing →



- **Private Cloud**

Advantages →

Owned by single organization.

High security and privacy.

Storage and network components are customizable.

Disadvantages →

Skilled person is required.

Area of operation is limited.

Maintainance is required.

- **Public Cloud**

Advantages →

3rd party cloud provider.

The provider will look after maintainence.

Scalability.

Cost effective.

24/7 uptime.

Disadvantages →

Data security.

Reliability.

- **Hybrid Cloud**

Advantages →

Cost effective.

Compose of Private and Public.

Scalable.

Critical activities are executed in Private environment.

Disadvantages →

Complex to manage.

Reliability depends upon the provider.

- Community Cloud

Advantages →

Multiple organization uses for joint-projects.

Cost effective.

Secure than Public Cloud.

Collabrative and distributed environment.

Resource Sharing with in limits

Disadvantages →

Not useful for non-collabrative work.

Not as secure as Private.

Fixed data storage and bandwidth is shared.

Benifits of Cloud computing →

- Cost effective (Saving money and resources)
- High speed (Customise your virtual machine ad run programs in few and easy clicks)
- Accessibility (can access your resources from any location with internet connectivity)
- Backup (cloud can be used as an data backup)
- Scalable
- No maintainance

Some of these as same in the case of risk of cloud computing

Limitations of Cloud computing →

Dependent on internet connectivity

Vinsuble to attacks

Latency

Limited control

Stuck with single operator

Risks of Cloud Computing →

- Privacy

Data can be accessed by the host company with or without permission.

Service providers may access the data that is on the cloud at any point in time.

- Security

Hard computing →

Its is a traditial way of approach of solving problem using logic and mathematics

Soft Computing →

Set of computational techniques that are uncertainty, partial truth and approximation.

Inspired by how human mind solve problems.

Soft computing	Hard Computing
◦ Less accurate	◦ More Accurate
◦ Less precision	◦ More Precise
◦ Slow	◦ Fast
◦ More complex	◦ Less complex
◦ More Flexible	◦ Less flexible

Cloud Service model →

1. IaaS also Known as Hardware as a Service (HaaS)

Advantages –

Allows multiple users to share the same physical infrastructure.

Allows access to resources over the internet

Provides Pay-As-Per-Use model

Focus on the core business

On demand scalability

- Provides Infrastructure –
  - Servers
  - Networking
  - Storage
  - VM
  - Operating system
  - Databases
- Website hosting
- Storing and analyzing data
- Development and testing environment

Disadvantages –

Security

Maintenance and Upgrade

Difficult to migrate VM to another IaaS provider.

2. PaaS

- Provides Run-time environment
- Developers to create / run / test and deploy.
- Database
- Storage file

Advantages –

Simplified Development

No hardware or software purchase required

Instant community

Scalability

- Development and Deployment of website
- Running Mobile application
- Running IOT application

Disadvantages –

Vendor lock in

Data Privacy

Integration with the rest of the systems applications

3. SaaS → Provides access to software application over the internet.

- Affordable
- Easy to use
- Scalable

Advantages –

No server setup is required, as software is hosted remotely

No maintainance

Pay-As-Per-Use

No hardware or software purchase required.

API integration

2 modes of SaaS →

Simple multi-tendency model → each user own resource

Fine grain multi-tendency mode → resources are shared between multiple users.

Disadvantages –

Security

Latency

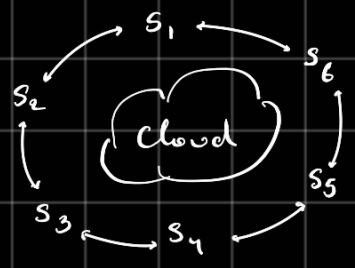
Total dependency on internet

Switching between vendor

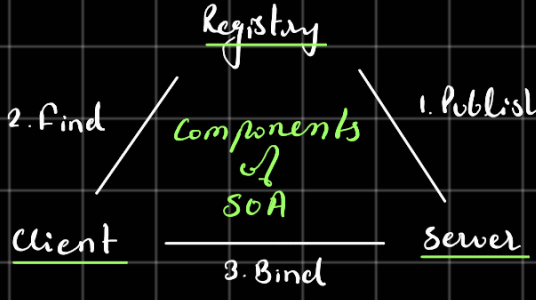
Service Oriented Architecture (SOA)

Is a software design pattern that structures an application as a collection of loosely coupled services.

- Each service is a stand alone and provides a specific functionality..



- Cloud with multiple individual and independent services.
- SOA provides an architecture and protocol so other services can communicate with each other.
- User asks S_6 for multiple services out of which two services are already provided by others service provider so S_6 askes other provider for the particular service.
- Saving time and cost.



Creating a Cab App → Map service from Google → Payment method Paytm →

- Code reuseability.
- Cost and Time efficient.
- Easy development
- SOA application are more flexible as SOA application's are combination of multiple loosely packed service.
- Easily scalable
- Maintainability — SOA app as combination of independent loosely packed services that can be maintained independently.

Principle of SOA →

- Loose coupling having as low as possible dependency on external resources.
- Abstraction hiding up of service data from client or user.
- Interoperability any system can run a service, regardless of platform or programming language.

Platform / Tools →

SaaS
Amazon WorkMail : Manage Mail and calander.

PaaS
AWS App Runner : full service for managing, develop and deploy of application

AWS →

- Comprehensive and Brodly adopted cloud platform.
- Destributed
- Data centers in over 200 countries

IaaS
EC2 (computation)
RDS (Relation Data Ser)
S3 (Simple Storage service)

Pros :- • Service Offering (IaaS, PaaS, SaaS)
- Scalability / flexibility
  Consumers can up or down scale its infrastructure need.
- Global Reach
  Multiple servers around the world improves its latency.
- Security
  Robust security feature, including Identity and Access Management.
- Reliable

Cons :-
- Complexity
  Wide range of services
- Cost Management
  Need to continously monitor the amount.
- Vendor lock-in

IoT
• Azure IoT Hub monitoring and managing IoT devices.

IaaS
• Azure SQL Database
• Azure Virtual Machine

Azure →

Pros :-
- Services offered
  IaaS, PaaS, SaaS and AI of IoT
- Hybrid cloud capabilities
- Global Reach
- Integration with Microsoft products

PaaS
• Azure App Service, for building, deploying and scaling.

SaaS
• Microsoft 365

AI/ML learning Services
• Azure Machine Learning

- Developer-Friendly
  Provides SDK and tools for developers

Cons :-

- Complex, multiple services
- Pricing Complexity
  Multiple tiers and payment methods and models.
- Vendor lock-in
- Few data centers compare to "AWS"
- Support not as good, criticized by many users.

Google Cloud Platform →

Pros :-

- Global Infrastructure
- Services Related to AI ML and Analytics
- Open Source tools, making easy integration
- Serverless computing like Google Cloud function and App Engine
- The fastest of the bunch

Cons :-

- Lowest market share
- Complex Pricing Model Requires regular monitoring for cost efficiency.
- Less Mature ecosystem
- Learning Curve, for AWS and Azure users.

IaaS
- Google Kubernetes Engine
- Google Virtual Private Cloud

PaaS
- App Engine building, running, deploying and scaling web and Mobile application.

SaaS
- G Suite (G Mail, G Photo, Docs, Sheet...)

Databases and Data services
- Google Cloud SQL

Google Cloud IoT core.

AI/ML services
- Google cloud AI Platform

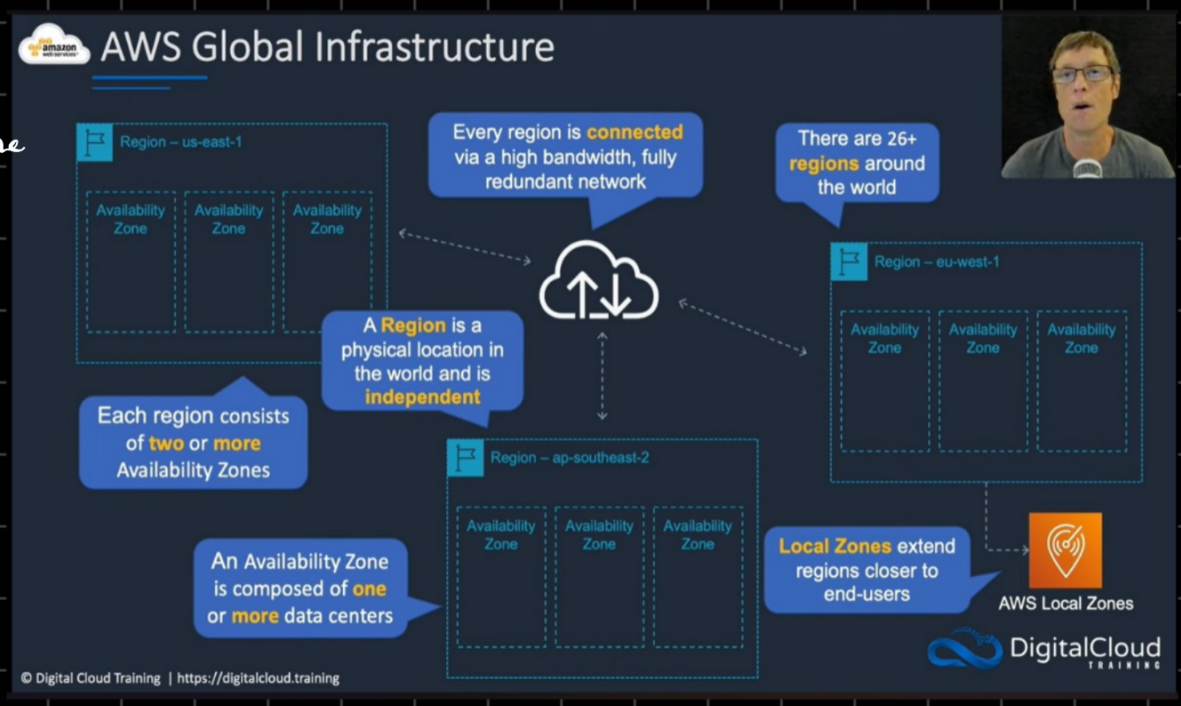
Factor	AWS	Azure	GCP
Service Offerings	Vast array of services; comprehensive ecosystem; IaaS, PaaS, SaaS	Wide range of services; strong integration with Microsoft products; IaaS, PaaS, SaaS	Strong data analytics and machine learning; IaaS, PaaS, SaaS
Market Share	Largest market share	Second-largest and growing, especially among enterprises	Third-largest but gaining traction
Global Reach	Data centers in numerous regions worldwide	Data centers in multiple regions	Network of data centers, fewer regions
Pricing	Pay-as-you-go model; various pricing options; requires diligent cost management	Competitive pricing; various pricing models; complex pricing	Competitive pricing; discounts available; complex pricing
Ecosystem	Mature ecosystem with third-party integrations; wide range of tools	Strong integration with Microsoft products; growing ecosystem	Strong support for open-source; growing ecosystem
Strengths	Extensive service offerings, strong security, mature ecosystem	Integration with Microsoft products, hybrid cloud capabilities, enterprise-focused solutions	Advanced AI, data analytics, global network, open-source support
Weaknesses	Complex pricing; competition-driven	Complex pricing; competition-driven	Smaller market share; fewer regions; less enterprise-focused

Users

- Netflix
- Coursera
- Fedex
- Audi
- Spotify
- Toyota

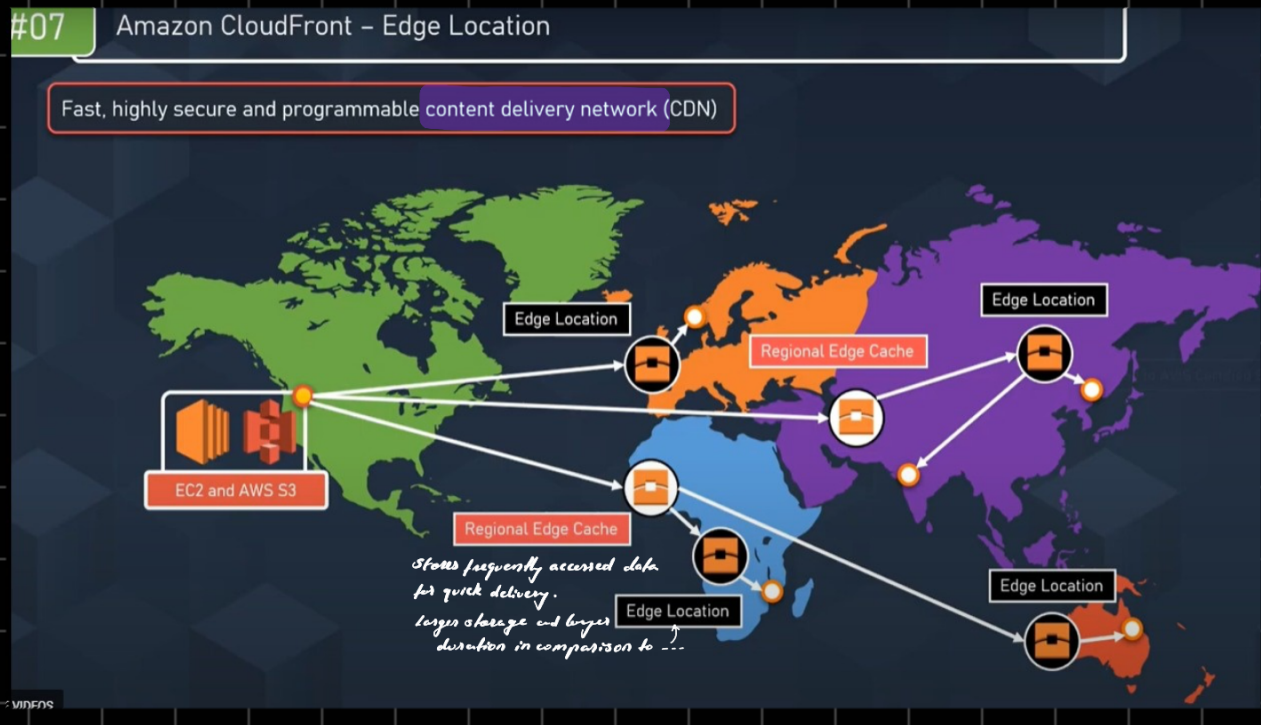
AWS Global Infrastructure →

- Region
- Availability Zone



AWS CloudFront →

- Regional edge cache
- Edge location



Tools →

- AWS →
  1. AI and ML →
     - AWS has a library Gluon
     - Open source, deep learning library for developer and non-developer
     - For creating Neural network
     - Using AI powered camera

2. Sage Maker to Serverless →
   - Service used to train and deploy ML models.
   - Also integrates lex conversational interface.

- Azure →
  1. Congnitive Server →
     - Bing AI search
     - Face API
     - Computer Vision API
     - Serverless computing

  2. Support MSFT Software →
     - Provides support for Microsoft product installed on premises.
     - Windows system backup linked to Azure Backup

- Google Cloud →
  1. IOT to serverless →
     - Google clude includes API for natural language, speech, translation and other complex technology.

  2. Big On AI →
     - Google cloud fortuner of AI development.
     - Tensorflow, open source library for ML.

What is cloud security →
- Why required
- Benifits of Cloud security system
- Denial of Service Attack (DDoS)

Cloud security consists of
- Policies
- Technology use
- Procedure
- Controls

Security challenges and threats in Cloud computing →

Security Concerns →

- Cloud service provider can't garentee 100% data security.
- Cyber attacks
- Insider Threath
- Government Interuption
- Compromise in security for competative pricing.
- Anything stored even if in private cloud if found agint contry or service provider policy, legal action can be taken.
- Lack of support
- Lack of standardization

Threats →

- DOS (Denial of service) — Active Attack / Passive Attack
- Man In The Middle
- Data Traffic Monitoring
- Alter the data traffic

Cloud Security Mechanism →

- IAM (Identity and Access Management)
  - Identity Verification
    - Verifying the identity accessing the cloud, by user id and password, two factor authentication.
  - Access Control
    - IAM to determine what the authority user and entities are allowed to do within the cloud.

- Identity lifecycle management
  - The process of creating and maintaining digital identity for human and non human entity.

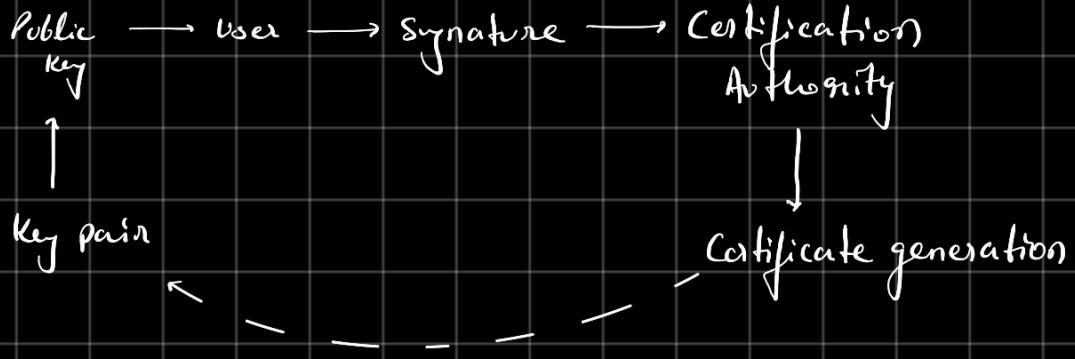
- Multifactor Authentication
  - Implementation of two or more authentication factors

- Identity Governance
  - IAM system monitors what users do with their resource access, wheather doing any prohibited task or not.

IAM Main Components:

Authentication :- username + pwd Combination. ↳ Digital Signatures ↳ Digital certificates ↳ Biometric	**Authorization** Access Control U1 → r1 (✓) → r2 (✗)
User management ↳ Creating new user ↳ adding user to the required access grp ↳ maintaining pwd policies, privileges	**Credential management** ↳ establishes identities & access control rules for users.

- Public Key Infrastructure →
  - Asymetric key cryptography
  - Digital certificate generation



- Verification of sender.

- Encryption →
  - Mentain confidentiality and Entigrity of data
  - Symmetric key and Asymmetric key encryption.
    - Confidentiality
    - Confidentiality
    - Authentication

Cloud Scalability →

It refers to the ability to increase or decrease IT resources as need to meet changes.

Elasticity →

It refers to the system's ability to grow or shrink dynamically as per work load.

- It can be automated, to perform capacity planning in advance to the occasion

Types of scaling Technique for cloud based application →

1) Vertical (Scale up)
2) Horizontal (Scale Out)
3) Diagonal
4) Auto Scaling
5) Load balancing
6) Serverless computing
7) Caching
8) Content Delivary Network (CDNs)

Vertical →

- Upgradation of systems hardware
- No load balancing is required
- Quick implementation

Pro
- easy maintain
- less complex
- Quick implement

Con
- More downtime
- less flexible

Horizonal →

- Purchasing of new system's/servers.
- Load balancing is required.

Con
- Expensive
- Takes time

- Chances of system failure decreases.

Diagonal →

- Linear incrementation of both.

Scalability requirement →

- Manage load on the server
- Flexibility and Speed
- Cost saving
- Disaster recovery

Load Balancing →

Is the method of distributing network traffic equally accross a pool of resources that support an application.

Modern application must process milions of users simultaniously and return the correct text videos images and other data to each user in a fast and reliable manner.

Benifits →

- Application availability
- Application Scalability
- security
- Performance

Auto scaling →

Automated approach to dynamically adjust resources based on demand.

- Monitoring application metrics like CPU, Memory, network traffic and automatically scale resources.

Serverless Computing →

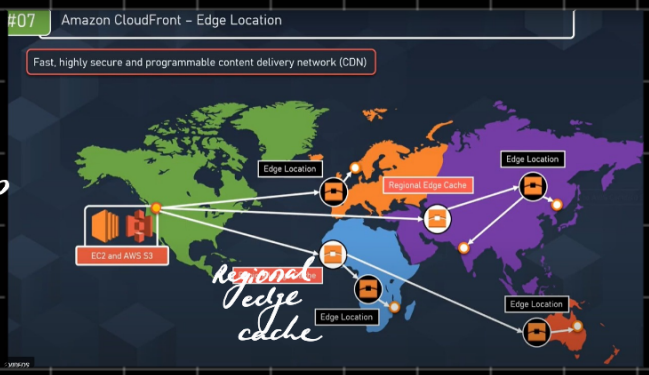
All the underlying infrastructure and scaling is automatically maintained by the cloud service provider, allowing developers to purely concentrate on the code writing.

Caching →

Is a set of fast accessable memory storing frequently used data/resource helps in fast execution of task.

CDNs →

Storing data near edge location for lower latency.



" ASG (Auto scaling Group) "

Auto scaling (Automatic scaling) is a cloud computing feature that allows system to automatically adjust its capacity based on demand.

The goal is maintain performance and insure that resources are efficiently utilized.

Auto scaling is used in cloud environment where resourse can be provisioned and deprovisioned dianamically.

Cloud Elasticity and Cloud Scalability →

Cloud Elasticity	Cloud Scalability
Elasticity is used just to meet the sudden up and down in the workload for a small period of time.	Scalability is used to meet the static increase in the workload.
Elasticity is used to meet dynamic changes, where the resources need can increase or decrease.	Scalability is always used to address the increase in workload in an organization.
Elasticity is commonly used by small companies whose workload and demand increases only for a specific period of time.	Scalability is used by giant companies whose customer circle persistently grows in order to do the operations efficiently.
It is a short term planning and adopted just to deal with an unexpected increase in demand or seasonal demands.	Scalability is a long term planning and adopted just to deal with an expected increase in demand.

What is cloud Elasticity ?

- Property of cloud to grow or shrink for CPU, memory and storage resources to adapt to changing demands.

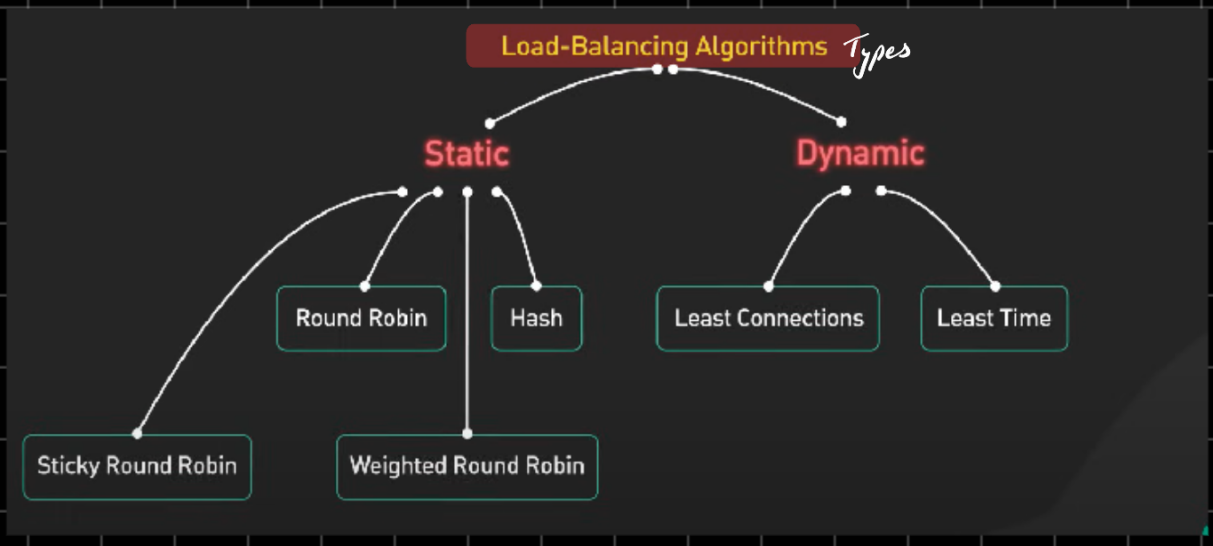
Overcoming the problem of overloading on the servers →

- Single - Server →

  The server is upgraded to a higher performing server.

- Multiple - Service →

  A scalable service system is build on cluster of servers.



1) Static Algorithm

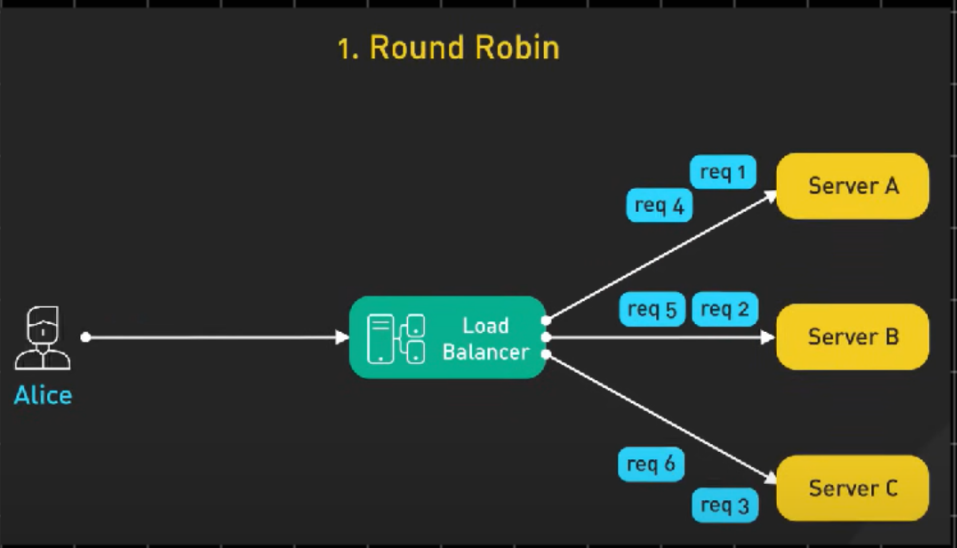
- For systems with very little load.
- Traffic devided equally between servers.
- Load shifting not depend on current state

2) Dynamic Algorithm

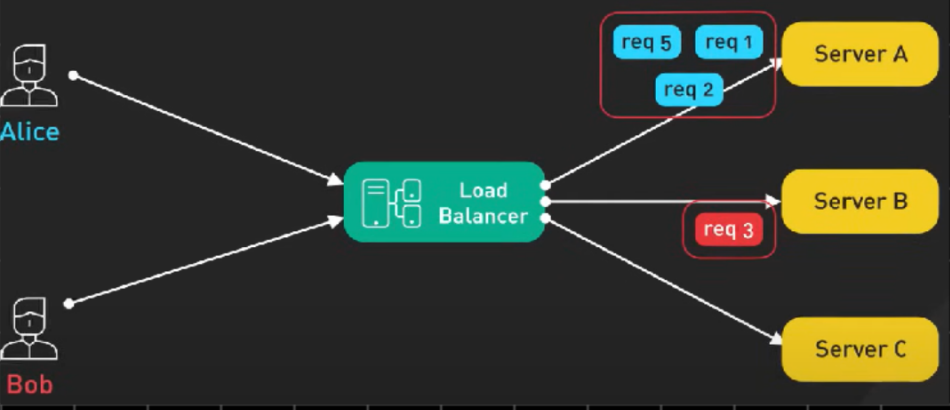
- Auto scaly as per load.
- Every action depends on current state.
- Complex implementation

  High performance

- sequential task destribution
- Doesn't check the complexity of task
- Doesn't check the present status of the server.

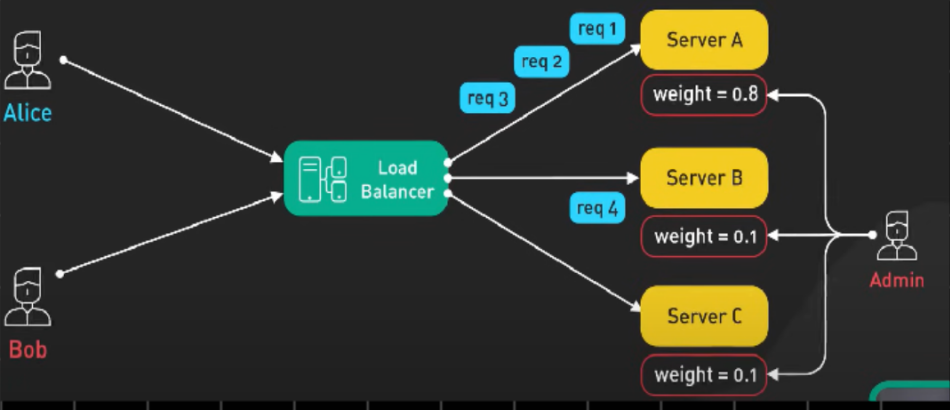


2. Sticky Round Robin



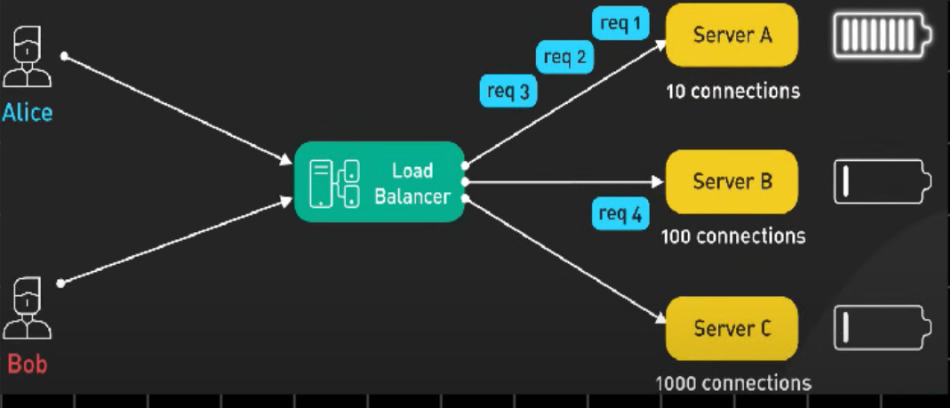
- Server assigned to a user.
- All the data is present on a single server, computation faster
- No task distribution
- Overloading of servers still problem
- Random assignment of new user.

3. Weighted Round Robin



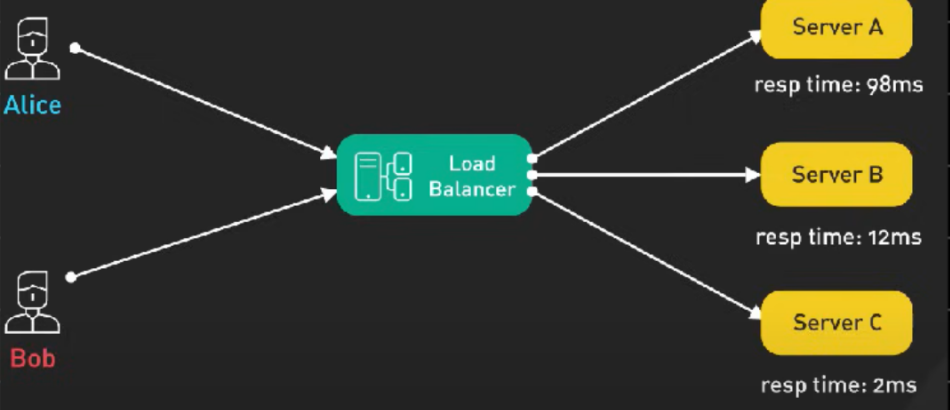
- Servers with higher weight gets get higer number of requests.
- Need to be maintained / regulated (weights manually)

5. Least Connections



- Server with least congestion gets higher number of request.
- Active system tracking
- Compilation of multiple request on a single server

6. Least Time



- Active system tracking
- Requests distributed based on latency.
- Complex
- Regular Monitoring increases overhead.

Type of load balancing →

1) Network

Utilizes the layers in the network protocol for load balancing.
Use of 4th layer in TCP/UDP layer, diversion of network traffic.

2) Application

Diversion / distribution of network over different target on application.

3) HTTP

The oldest method of code balancing.
Utilizes the 7th layer operation for code balancing.

4) Internal load balancing

consists of multiple methods of load balancing →

Hardware based

Utilization of physical hardware for management of traffic over network and application.

Software based

Installed before use.
Cheap, easy to use, economical

Vertical load balancer

Deployment of software over hardware - load balancer device on the vertical machine.

Cloud based Open-Source Databases →

1) AWS Relation Database
2) Azure SQL
3) Google SQL

4) Oracle

Measured Services in Cloud Computing Mean →

Utility provides monitors how much of a particular service each customer consumes within a designated time period.

Advance Security Challanges →

Enlarge Surfaces
Lack of visibility and Tracking
Changing workload
Complex Environment
Cloud Governance

Emerging trends in Cloud Computing →

1) AI and ML
2) Serverless Computing
3) IoT
4) Block chain Technology
5) Edge Computing

↳ Serverless Computing

Is a cloud computing execution needed in which the cloud server provides / allocates resources on demand taking care of the server on behalf of their

customer.

Aserver less architecture is a way to build and run application and services without having to manage infrastructure.

Characteristics →

- No infrastructure management
  Infrastructure is managed by cloud providers
- Auto scaling based on incoming request
- Reduces cost or cost effective
- It enables developers to focus on code as they have no need to manage infrastructure.

Big Data 5V →

- Volume
- Velocity (Increase of data at
- Variety (Offerent types of data → structured
  → unstructured
  → semi structure
- Veracity (Trust worthy "coding legal or not")
- Value

Edge Computing →

- Enable data to be analyzed process and transfered at the edge of the network.
- The idea is to analyze data locally closer to where it stored in real time without latency.
- To leverage 5G wireless technology and AR to enable faster response time, lower latency (ability to process very high volume of data with minimal delay) and simplified maintenance.