

Unit-1: Fundamentals of Cloud Computing

Introduction to Cloud Computing

A modern cloud computing visual showcasing a central cloud connected to devices like laptops, smartphones, and servers. This symbolizes the key concept of cloud computing—seamless access to shared resources over the internet.



What is Cloud Computing?

Definition:

Cloud computing is the delivery of computing services—including servers, storage, databases, networking, software, analytics, and intelligence—over the internet (“the cloud”) to offer faster innovation, flexible resources, and economies of scale. Instead of owning and maintaining physical infrastructure, users access resources on-demand, paying only for what they use.

Example:

A widely known example of cloud computing is Google Drive, a cloud-based storage service. Users can upload, access, and share files like documents, photos, and videos from anywhere with an internet connection. The data is stored on Google's servers, eliminating the need for users to manage physical storage devices.

Key Attributes of Cloud Computing**On-Demand Self-Service**

Users can provision computing resources like servers or storage as needed without requiring human interaction with the service provider.

Example: A startup can instantly deploy virtual machines on Amazon Web Services (AWS) to handle increased traffic during a product launch.

Broad Network Access

Services are available over the internet and can be accessed from a variety of devices such as smartphones, tablets, and laptops.

Example: Employees use Microsoft OneDrive to access shared documents from any location using their devices.

Resource Pooling

Resources are pooled to serve multiple users, dynamically allocated based on demand. Users are abstracted from the physical location of the resources.

Example: Google Cloud Platform hosts services for numerous businesses while optimizing resource allocation.

Rapid Elasticity

Resources can be scaled up or down dynamically based on demand, ensuring efficient utilization.

Example: During seasonal sales, e-commerce platforms like Flipkart leverage cloud services to handle the surge in user traffic, scaling back once the demand decreases.

Benefits of Cloud Computing

Cost-Efficiency

- Cloud computing eliminates the need for upfront capital investments in physical hardware.
- Users pay only for the resources they consume, reducing operational costs.
- It lowers expenses associated with maintenance, power, and IT staffing.

Accessibility from Anywhere

- Cloud services can be accessed via the internet, enabling remote work and global collaboration.
- This ensures data and applications are available across devices like laptops, tablets, or phones.
- Teams can work in real-time, enhancing productivity and flexibility.

Scalability and Flexibility

- Resources can be scaled up or down dynamically based on demand.
- Businesses can handle sudden traffic surges without over-investing in infrastructure.
- This adaptability makes cloud computing ideal for startups and growing enterprises.

Reliability and Disaster Recovery

- Cloud providers offer robust infrastructure with minimal downtime and automatic backups.
- Data recovery is faster due to redundant systems spread across multiple locations.
- It ensures business continuity even during hardware failures or natural disasters.

Real-World Applications

Examples: Google Drive, Dropbox, AWS, Microsoft Azure.

Industries using cloud: Healthcare, Banking, IT, Education. provide image for each point

Evolution of Cloud Computing

1960s: Mainframe Computing

- Origin of shared computing resources in centralized mainframes.
- Time-sharing models allowed multiple users to access a single system.
- **Example:** IBM Mainframe systems.

1970s: Virtualization

- Introduction of virtual machines (VMs) enabled running multiple OS on a single physical machine.
- Laid the foundation for modern cloud infrastructures.
- **Example:** IBM's VM/370 operating system.

1990s: The Internet and SaaS

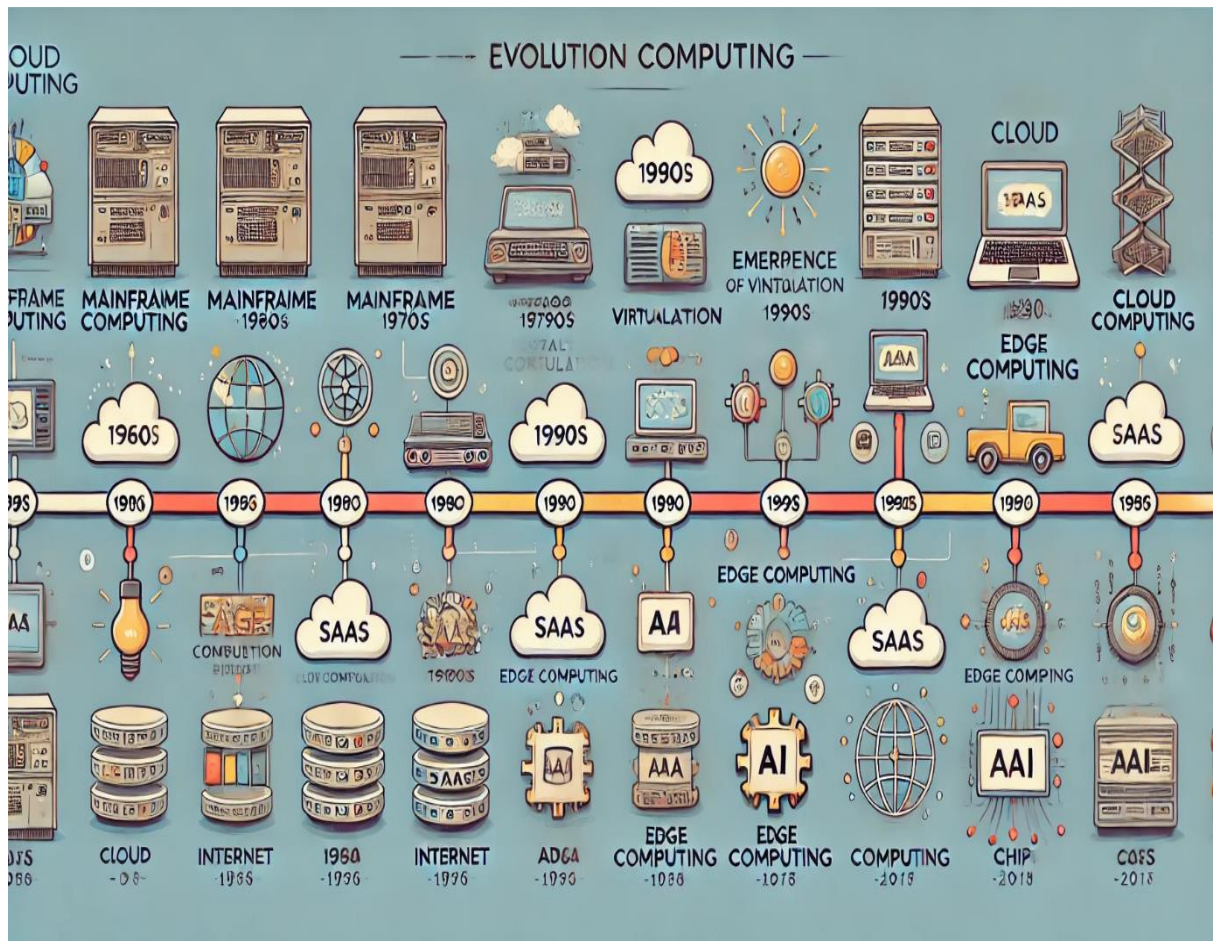
- Expansion of the internet allowed remote access to shared resources.
- Rise of Software-as-a-Service (SaaS) like Salesforce.
- **Example:** Introduction of web-based email services such as Hotmail.

2000s: Cloud Service Models

- Emergence of cloud providers like AWS (2006) offering IaaS, PaaS, and SaaS.
- Democratized access to scalable computing resources.
- **Example:** AWS Elastic Compute Cloud (EC2).

2010s and Beyond: AI and Edge Computing

- Integration of AI and machine learning for smarter cloud services.
- Edge computing minimized latency by processing data closer to users.
- **Example:** Microsoft Azure's AI-powered services and Google Cloud's IoT solutions.



Cloud Generations

1st Generation: Infrastructure-Centric Cloud

- Focused on providing virtualized hardware and storage resources.
- Enabled organizations to replace physical infrastructure with cloud-based solutions.
- Offered basic services like virtual machines and data storage.
- **Examples:** Amazon Web Services (AWS) EC2, early Google Cloud storage services.

2nd Generation: Service-Oriented Cloud

- Introduced service models like SaaS, PaaS, and IaaS, revolutionizing software delivery.
- Facilitated businesses with ready-to-use platforms for app development and deployment.
- Enhanced user accessibility with subscription-based models.
- **Examples:** Salesforce (SaaS), Microsoft Azure (PaaS), AWS (IaaS).

3rd Generation: AI-Driven and Edge-Computing-Enabled Cloud

- Integrated artificial intelligence for automation and smarter decision-making.
- Supported edge computing for real-time data processing closer to the source.
- Ideal for IoT, smart cities, and industries requiring low-latency solutions.
- **Examples:** Google Cloud's AI-powered tools, Azure IoT Hub.



Characteristics of Cloud Computing with Examples

On-Demand Self-Service

- Users can provision computing resources like storage and processing power as needed, without human intervention.
- **Example:** A developer using Amazon Web Services (AWS) can launch a virtual machine instantly via the AWS console.

Broad Network Access

- Cloud services are accessible over the internet and available across multiple devices like smartphones, laptops, and tablets.
- **Example:** Employees using Google Workspace can collaborate on shared documents from any location and device.

Resource Pooling

- Providers pool resources to serve multiple customers, dynamically assigning resources as required, while maintaining data isolation.
- **Example:** Microsoft Azure uses resource pooling to allocate servers and storage for different organizations securely.

Rapid Elasticity

- Cloud resources can scale up or down automatically to handle workload changes, ensuring optimal performance.
- **Example:** During a sale event, an e-commerce site like Flipkart scales its cloud infrastructure to handle increased traffic.

Measured Service

- Cloud systems monitor and measure resource usage, enabling a pay-as-you-go model for users.
- **Example:** Netflix pays for AWS resources based on the number of streaming sessions served.

Multi-Tenancy

- Multiple customers share the same physical infrastructure while keeping their data and resources isolated.
- **Example:** CRM software like Salesforce provides secure multi-tenant environments for businesses.

High Availability and Reliability

- Cloud systems are designed with redundancy to ensure services remain available even during failures.
- **Example:** Google Cloud Platform offers 99.99% uptime, ensuring minimal service disruptions.