INTRODUCTION TO CLOUD COMPUTING

Cloud computing is the delivery of computing services, including servers, storage, databases, networking, software, and analytics, over the internet or "the cloud." Cloud computing enables companies to rent IT infrastructure rather than build and maintain their own.

Cloud computing offers numerous benefits, including scalability, flexibility, cost savings, and increased efficiency. Companies can quickly scale up or down their computing resources to meet changing demands, only pay for what they use, and eliminate the need for upfront infrastructure investments.

Types of Cloud Computing

- Private Cloud
- Public Cloud
- Hybrid Cloud
- Multi-Clouds

Types of Cloud Computing Services

- **SaaS** provides users access to software applications over the internet, eliminating the need for organizations to install and run software on their own computers or data centers. E.g Google Workspace, Microsoft Office 365, Salesforce, Dropbox, Zoom

- **PaaS** provides a platform for developers to build and deploy their own applications without having to worry about the underlying infrastructure. E.g. Heroku, Google App Engine, AWS Elastic Beanstalk, Microsoft Azure

- **IaaS** provides virtualized computing resources over the internet, such as virtual machines, storage, and networking, allowing organizations to run their own software and applications on a cloud provider's infrastructure. E.g. Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, DigitalOcean, Linode

Cloud Providers

- Amazon Web Services (AWS)
- Microsoft Azure.
- Google Cloud Platform (GCP)
- Alibaba Cloud.
- Digital Ocean
- Oracle Cloud.
- IBM Cloud (Kyndryl)
- Tencent Cloud.
- OVHcloud.

Amazon Web Services (AWS) is a cloud computing platform offered by Amazon. It provides a wide range of cloud-based computing services, including Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

AWS offers a massive global infrastructure of data centers and services that can be used to build and deploy applications, store and process data, and handle a variety of other computing needs.

- **IAM (Identity and Access Management)**: IAM is a web service that helps you securely control access to AWS resources. IAM allows you to create and manage users, groups, and permissions to access AWS services and resources. With IAM, you can grant permissions to users to perform specific actions on specific resources, and also manage authentication and authorization for AWS resources.

- **EC2 Instances, EBS Volumes, ELB (Elastic Compute Cloud)**: EC2 is a web service that provides resizable compute capacity in the cloud. EC2 instances are virtual machines that can be launched and run in the cloud. EBS (Elastic Block Store) volumes provide persistent block storage for EC2 instances. ELB (Elastic Load Balancer) provides a load balancing service for EC2 instances, distributing incoming traffic across multiple instances.

- **S3 (Simple Storage Service)**: S3 is a web service that provides scalable object storage for data. S3 allows you to store and retrieve any amount of data, at any time, from anywhere on the web. S3 provides a highly durable and available storage infrastructure, and can be used for a wide range of use cases, including backup and recovery, archiving, and data analytics.

- **CloudWatch**: CloudWatch is a monitoring and management service that provides real-time monitoring of AWS resources and applications. CloudWatch allows you to collect and track metrics, collect and monitor log files, and set alarms. You can use CloudWatch to monitor resources such as EC2 instances, RDS databases, and ELB load balancers, and also to monitor custom metrics and logs.

- **RDS (Relational Database Service)**: RDS is a web service that provides managed relational database services in the cloud. RDS allows you to launch and run popular relational database engines such as MySQL, PostgreSQL, and Oracle. RDS provides automatic backups, automatic software patching, and automatic scaling of compute and storage resources.

- **Autoscaling**: Autoscaling is a web service that allows you to automatically scale your EC2 instances based on demand. Autoscaling allows you to set scaling policies that automatically add or remove EC2 instances based on predefined criteria. This allows you to automatically adjust your capacity to meet demand, while minimizing costs.

- **Route53**: Route53 is a web service that provides scalable domain name system (DNS) services. Route53 allows you to register domain names and route traffic to AWS resources such as EC2 instances, S3 buckets, and load balancers. Route53 provides a highly available and scalable DNS infrastructure, and can also be used for health checking and failover of resources.

- **Lambda**: Lambda is a serverless computing service that allows you to run code without having to manage infrastructure. With Lambda, you can run code in response to events, such as changes to data in an S3 bucket or incoming messages in an SQS queue. Lambda automatically scales your code based on demand, so you only pay for the resources you use.

- **DynamoDB**: DynamoDB is a NoSQL database service that provides scalable and high-performance document and key-value storage. DynamoDB allows you to store and retrieve any amount of data, at any time, from anywhere on the web. DynamoDB provides a fully managed, highly available, and durable database infrastructure, and can be used for a wide range of use cases, including gaming, mobile, and web applications.

- **CloudFront**: CloudFront is a content delivery network (CDN) service that allows you to distribute content to users worldwide with low latency and high data transfer speeds. CloudFront can be used to deliver static and dynamic content, including video streaming, APIs, and web applications. CloudFront also provides security features, such as SSL/TLS encryption and DDoS protection.

- **Amazon SNS**: Amazon SNS is a fully managed messaging service that enables you to send notifications from the cloud to multiple recipients or subscribers. SNS provides a flexible, cost-effective, and highly scalable way to send notifications to a large number of subscribers or endpoints, such as mobile devices, email addresses, or AWS Lambda functions.

- **Elastic Beanstalk**: Elastic Beanstalk is a platform as a service (PaaS) that allows you to deploy and manage web applications in the cloud. Elastic Beanstalk provides an easy-to-use interface for deploying and managing applications, and automatically handles the underlying infrastructure, including EC2 instances, load balancers, and databases. Elastic Beanstalk supports a wide range of web technologies, including Java, PHP, Node.js, Python, and Ruby.

Q & A SESSION