**\*\*\*\* AWS Cloud \*\*\***

**What is cloud computing** ?

Cloud computing is the **on-demand** delivery of compute power, database, storage, applications, and other IT resources through a cloud services platform via the internet with **pay-as-you-go pricing.**

**Features:**

1. Broad Network Access
2. Resource Pooling
3. Rapid Elasticity
4. Measured Service
5. Scalability
6. Fault Tolerance and Reliability
7. Security
8. On demand self-service

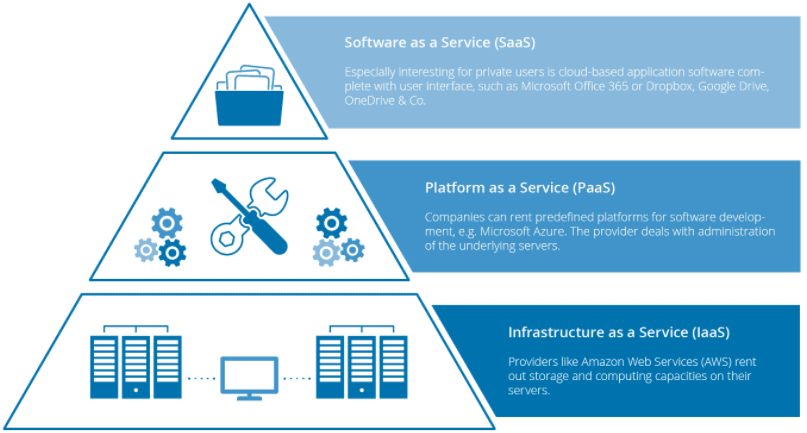
## **Cloud Computing Service Models**

There are three main types of service models of cloud computing. Each type of cloud computing provides different levels of control, flexibility, and management so that you’ll select the proper set of services for your needs.

## **Three Common Cloud Service Models**

The **three Cloud Service Models** are as follows:

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



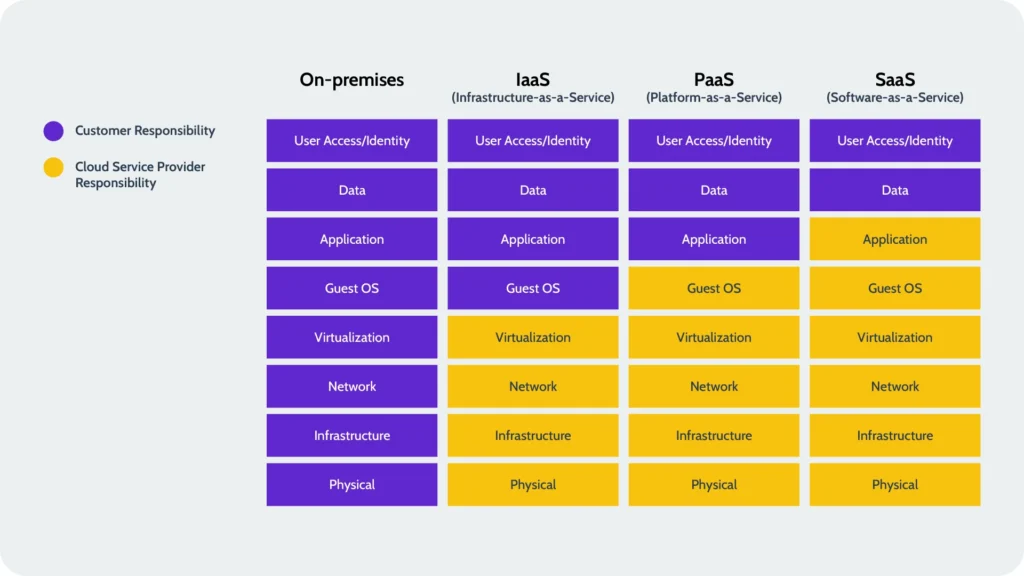
## 

Types of cloud:

1. Public Cloud: Services offered over the internet by third-party providers like aws ,azure,gcp .A cloud-based application is fully deployed in the cloud and all parts of the application run in the cloud.
2. Private Cloud(on-premise): Cloud infrastructure exclusively used by a single organization.
3. Hybrid Cloud: Integration of private and public cloud environments.
4. Multi-Cloud: Utilizing services from multiple cloud providers.

## **Cloud Shared Responsibility Model**

The shared responsibility model defines cloud security, but it changes for IaaS, PaaS, and SaaS.



# **AWS global infrastructure: Regions, Availability Zones, and Edge Locations**

**AWS infrastructure includes Regions (geographic areas with clustered data centers), Availability Zones (isolated locations within a Region with redundant data centers), and Edge Locations (endpoints for caching content closer to users). These components ensure high availability, fault tolerance, and low-latency access to AWS services globally.**

Amazon Web Services (AWS) operates a global infrastructure that consists of Regions, Availability Zones, and Edge Locations, which collectively ensure high availability, fault tolerance, and low-latency access for their services. Let me break down each component:

1. Regions:
   * Regions are physical locations around the world where AWS clusters data centers. These regions are geographically distributed and are designed to provide low-latency access to AWS services for users in various parts of the world.
   * Each AWS Region is a separate geographic area that contains multiple Availability Zones. AWS typically names its regions based on the geographical location, such as us-east-1 (North Virginia), eu-west-1 (Ireland), ap-southeast-1 (Singapore), etc.
   * AWS Regions are completely isolated from each other to provide fault tolerance and high availability.
2. Availability Zones (AZs):
   * Availability Zones are distinct locations within a Region that are engineered to be isolated from failures in other Availability Zones.
   * Each Availability Zone typically consists of one or more data centers, with redundant power, networking, and connectivity.
   * By spreading resources across multiple Availability Zones, AWS ensures that if one zone goes down due to a failure, other zones can continue to operate, thus providing high availability and fault tolerance.
   * Typically, AWS Regions consist of at least two Availability Zones, though some regions have more.
3. Edge Locations:
   * Edge Locations are endpoints for AWS services that are used for caching content closer to end-users to improve latency and data transfer speeds.
   * Unlike Regions and Availability Zones, Edge Locations are not used for storing data or running compute instances. Instead, they primarily support services like Amazon CloudFront (AWS's content delivery network) and Amazon Route 53 (AWS's DNS service).
   * Edge Locations are spread across the globe in major cities and are strategically placed to reduce the latency of delivering content to end-users.

# The 6 Pillars of the AWS Well-Architected Framework

The [AWS Well-Architected Framework](https://aws.amazon.com/architecture/well-architected/) helps cloud architects build the most secure, high-performing, resilient, and efficient infrastructure possible for their applications.

* **Cost Optimization**: Maximizing the value of IT spending by optimizing resource usage, selecting the right pricing models, and continuously monitoring costs.
* **Operational Excellence**: Implementing best practices for operations, including automation, monitoring, and incident response, to deliver reliable and efficient services.
* **Reliability**: Ensuring that systems can recover from failures and meet service level agreements (SLAs) by designing for fault tolerance and implementing scalable architectures.
* **Performance Efficiency**: Optimizing the performance of systems and applications by selecting appropriate resources, scaling as needed, and tuning configurations.
* **Security**: Protecting data, systems, and infrastructure by implementing comprehensive security measures, including encryption, access controls, and compliance frameworks.
* **Sustainability**: Minimizing the environmental impact of cloud operations through energy efficiency, resource optimization, and sustainable practices.

Virtualization vs Hypervisor :