**CLOUD BASICS (PREREQUISITE CONCEPTS)**

**VIRTUALIZATION:**

Virtualization is a technology that allows you to create multiple virtual environments or virtual machines (VMs) on a single physical computer or server. It enables the abstraction and isolation of computer resources, such as processors, memory, storage, and network, from the underlying hardware.

In virtualization, a software layer called a hypervisor or virtual machine monitor (VMM) is installed on the host machine. The hypervisor manages the virtual machines and provides a means for them to access the physical resources of the host system. Each virtual machine operates as an independent entity, running its own operating system and applications.

**Virtualization provides several benefits:**

**1. Server consolidation:** Virtualization allows you to run multiple virtual machines on a single physical server, which helps in consolidating server infrastructure and reducing hardware costs.

**2. Resource optimization:** Virtualization enables efficient utilization of hardware resources by dynamically allocating them to virtual machines as needed. It allows for better resource management and flexibility.

**3. Isolation:** Each virtual machine operates in its own isolated environment, providing security and stability. If one virtual machine crashes or encounters issues, it does not affect the others.

**4. Testing and development:** Virtualization is commonly used in software testing and development environments. Developers can quickly set up virtual machines to test different configurations, software versions, or to replicate production environments.

**5. Disaster recovery and high availability:** Virtual machines can be easily backed up, replicated, and migrated between physical servers. This allows for efficient disaster recovery strategies and high availability solutions.

**CLOUD NATIVE:**

Cloud native refers to an approach in software development and deployment that leverages cloud computing principles and technologies to build and run applications. It is characterized by designing applications specifically for deployment in cloud environments and taking advantage of cloud-native services and platforms.

Key characteristics of cloud native applications include:

**1. Microservices architecture:** Cloud native applications are typically built using a Microservices architecture, where an application is composed of loosely coupled and independently deployable services. Each service focuses on a specific function and can be developed, deployed, and scaled independently.

**2. Containers:** Containers provide lightweight and portable packaging of applications and their dependencies. Cloud native applications are often packaged as containers, allowing them to be easily deployed and managed across different cloud environments.

**3. Orchestration:** Cloud native applications are orchestrated using container orchestration platforms like Kubernetes. Orchestration platforms manage the deployment, scaling, and lifecycle management of containers, providing features such as service discovery, load balancing, and automated scaling.

**4. DevOps practices:** Cloud native development embraces DevOps practices, promoting collaboration between development and operations teams. Continuous integration and continuous deployment (CI/CD) pipelines are commonly used to automate the build, testing, and deployment processes.

**5. Elastic scalability:** Cloud native applications are designed to scale horizontally, meaning they can handle increased workloads by adding more instances of the application or its services. This scalability is achieved through the dynamic provisioning and auto-scaling capabilities provided by the cloud platform.

**6. Resilience and fault tolerance:** Cloud native applications are designed to be resilient in the face of failures. They leverage techniques like service redundancy, fault isolation, and automated failure recovery to ensure high availability and fault tolerance.

**7. Cloud-native services:** Cloud native applications take advantage of cloud-native services, such as managed databases, caching, messaging queues, and serverless computing. These services provide additional functionality, reduce the operational burden, and enable developers to focus on core application logic.

**REGIONS AND AVAILABILITY ZONES (AZee):**

In Amazon Web Services (AWS), regions and availability zones are important concepts related to the infrastructure and geographical distribution of AWS services. Here's an explanation of each:

**1. Regions:** AWS divides its global infrastructure into multiple geographic regions. A region is a separate geographical area that consists of multiple data centers (known as Availability Zones). Each region is designed to be isolated from others to provide fault tolerance and resilience. AWS regions are identified by names such as "us-west-2" (Oregon) or "eu-central-1" (Frankfurt). AWS operates regions in various countries around the world.

Regions are used to host AWS services, and customers can select the region where they want to deploy their resources. By choosing a specific region, customers can optimize latency, comply with data residency requirements, and ensure high availability for their applications.

**2. Availability Zones (AZs):** Within each region, AWS further divides its infrastructure into multiple availability zones. An availability zone is essentially a data center or a cluster of data centers located within a region. Each availability zone is isolated from failures in other availability zones and is designed to be independent of one another in terms of power, networking, and cooling.

Availability zones are identified by names like "us-west-2a" or "eu-central-1b". They provide redundancy and fault tolerance, allowing customers to deploy their applications across multiple availability zones to achieve high availability and resilience. By distributing resources across availability zones, customers can protect their applications from single points of failure and mitigate the impact of outages or disruptions.

Availability zones within a region are interconnected through low-latency and high-bandwidth networks, enabling data replication and synchronous communication between resources deployed across different availability zones.

**AWS TOP SERVICES:**

1. Amazon EC2 (Elastic Compute Cloud): Provides resizable compute capacity in the cloud, allowing users to launch and manage virtual servers (EC2 instances) for various applications.

2. Amazon S3 (Simple Storage Service): Offers scalable object storage for storing and retrieving data, such as files, images, videos, and backups, with high durability and availability.

3. AWS Lambda: Allows users to run code without provisioning or managing servers. It enables serverless computing, where functions are executed in response to events or API requests.

4. Amazon RDS (Relational Database Service): Provides managed database services for popular relational database engines like MySQL, PostgreSQL, Oracle, and SQL Server, allowing users to easily set up, operate, and scale databases.

5. Amazon VPC (Virtual Private Cloud): Offers a logically isolated virtual network in the AWS cloud, allowing users to define their network topology, configure IP addressing, and control network access.

6. Amazon SNS (Simple Notification Service): Provides a messaging service that enables the sending and receiving of notifications to a variety of endpoints, such as email, SMS, mobile push, and more.

7. Amazon CloudFront: A global content delivery network (CDN) that securely delivers data, videos, applications, and APIs to users worldwide with low latency and high transfer speeds.

8. Amazon Route 53: A scalable and highly available DNS (Domain Name System) web service that helps users route traffic to various AWS services or external resources.

9. AWS IAM (Identity and Access Management): Offers centralized access management and control for AWS resources. It enables the creation and management of users, groups, roles, and permissions.

10. Amazon SNS (Simple Notification Service): Provides a messaging service that enables the sending and receiving of notifications to a variety of endpoints, such as email, SMS, mobile push, and more.

11. Amazon CloudWatch: A monitoring and observability service that collects and tracks metrics, logs, and events, allowing users to gain insights and take automated actions based on system performance.

12. Amazon DynamoDB: A fully managed NoSQL database service that provides scalable and high-performance storage for structured data. It offers low latency and automatic scaling to handle demanding workloads.

13. AWS CloudFormation: A service that allows users to provision and manage AWS resources using templates, enabling infrastructure as code and simplifying resource deployment and management.

14. AWS Elastic Beanstalk: A platform-as-a-service (PaaS) offering that simplifies the deployment and management of applications, abstracting the underlying infrastructure and automatically handling capacity provisioning, load balancing, and scaling.

15. AWS Glacier: A secure and durable storage service for archiving and long-term backup of data. It offers low-cost storage options for data that is accessed infrequently but needs to be retained.