**Azure Availability Set**

**What is it?** An **Availability Set** is a logical grouping of Azure virtual machines (VMs) that helps ensure the availability of your application during planned or unplanned maintenance. It’s used to distribute VMs across multiple physical servers within an **Azure datacenter** to avoid a single point of failure.

**Key Concepts:**

* **Fault Domains (FDs)**: These represent **physical racks** in the data center. Each VM is placed in a different fault domain to ensure that if one rack goes down, the others continue running.
* **Update Domains (UDs)**: These represent different **update or maintenance groups**. VMs are placed in different update domains to ensure that not all VMs are updated at the same time, which can minimize downtime.

**Why use it?**

* **High availability**: If one VM or server fails, others are still running.
* **Distribute resources**: It ensures that your VMs are placed across multiple fault domains and update domains.

**Azure Load Balancer**

**What is it?** Azure Load Balancer is a **network load balancing service** that distributes incoming traffic to multiple VMs or instances of an application to ensure that no single VM gets overloaded. It helps ensure high availability and reliability by balancing traffic across resources.

**Key Concepts:**

* **Internal Load Balancer (ILB)**: Distributes traffic within a **virtual network** (VNet).
* **Public Load Balancer**: Distributes incoming internet traffic to VMs in your Azure environment.
* **Health Probes**: Azure Load Balancer continuously checks the health of backend VMs, ensuring traffic is only routed to healthy VMs.

**Why use it?**

* **Scalability**: Balances traffic among multiple instances to handle more requests.
* **Resilience**: Increases fault tolerance by redirecting traffic to healthy instances.

**2. VMSS (Virtual Machine Scale Sets)**

**What is it?** Azure **VMSS** is a service that allows you to automatically deploy and manage **a group of identical VMs**. It helps to **scale applications** horizontally (i.e., adding or removing VMs based on traffic demand).

**Key Concepts:**

* **Autoscaling**: VMSS can automatically increase or decrease the number of VMs based on metrics like CPU usage, memory, or incoming traffic.
* **Load Balancer Integration**: VMSS works seamlessly with Azure Load Balancer to distribute traffic across the VMs in the scale set.
* **Consistent Configuration**: All VMs in a scale set are identical in configuration, which ensures consistency.

**Why use it?**

* **Scalability**: Automatically adjusts the number of VMs based on demand.
* **High availability**: Distributes VMs across availability zones, ensuring uptime during failures.
* **Simplified management**: Easily manage and configure the entire set of VMs at once.

**3. Azure Capture > creating AMI**

**What is it?** **Azure Capture** is a feature that allows you to **capture** and **save the image** of a running virtual machine (VM) for later use. This image can be used to **recreate the VM** or to **create a similar VM** with the same configuration.

**Key Concepts:**

* **Capture Image**: A snapshot of the OS and data disks of a VM.
* **VM Template**: The captured image can be used as a template to deploy identical VMs.
* **Automation**: It helps with automation and consistency when deploying multiple VMs with the same configuration.

**Why use it?**

* **Consistency**: Ensure that multiple VMs have identical configurations.
* **Backup**: Create backups of VMs for disaster recovery or redeployment.
* **Speed**: Quickly deploy similar VMs using captured images.

**4. Azure Site Recovery and Backup**

**Azure Site Recovery (ASR)**

**What is it?** Azure **Site Recovery** (ASR) is a disaster recovery solution that **replicates your applications and workloads** from on-premises or from one Azure region to another. It helps protect against major outages by enabling you to failover to a secondary region in case of disaster.

**Key Concepts:**

* **Replication**: Continuous replication of your VMs, servers, and data to a secondary location.
* **Failover**: In the event of a failure in your primary location, you can **failover to the secondary region** with minimal downtime.
* **Recovery Plans**: You can define automated recovery plans, which include steps to recover resources and applications.

**Why use it?**

* **Business continuity**: Ensure your business keeps running even during major outages.
* **Disaster recovery**: Quickly recover applications and workloads after a disaster.
* **Cost-effective**: No need to maintain expensive secondary data centers.