Computer Services

We can view compute services as one of the technical pillars of Azure.

Technical pillars:

- Compute services
- Networking services
- Storage services
- Databased services

Note: These are infrastructure/foundational level services.

Compute Services

Basically, this is refereeing to executing code in the cloud. It could be a website, app, banking functionality for a bank (logging/check bank balance).

MS Azure dozens of Compute Types. Here are a few:

- Virtual Machines (VM)
- VM Scale Sets (VMSS)
- App services (Web apps)
- Azure Container Instances (ACI)
- Azure Container Apps
- Azure Kubernetes Service (AKS)
- Azure Virtual Desktop

Virtual Machines:

This is the closest analogue to a “server” (computer) in cloud computing. But it is virtual. That is, a single physical machine has been subdivide into slices and you get to rent a single slice of it.

Standalone Server Analogy:

Think of this as a single-detached house. You can do whatever you want with it (generally). It is very difficult to do anything that affects your neighbors (soundproof walls). You do not share any services with your neighbors (besides garbage, sewer, water, electricity – provided by the city).

Virtualization Analogy:

Think of a "host" as an apartment building on the same land. A "virtual machine" is an apartment in that building. You are using common services (garbage, sewer, water, electricity) and other services available (shared gym, heating/cooling). It is generally cheaper to rent (but you would have same the feeling as you in a house / get the benefits of a house).

Thus, virtual machines:

- Infrastructure as a service – IaaS
- Take an existing machine (desktop / server) from your environment into the cloud (a copy)
- Windows / Linus operating systems – several of each
- A "slice" of a physical machine shared with other customers
- Full control over it, as if it was your machine

Note: In AWS, a Virtual Machine is called Elastic Compute Cloud (EC2).

Virtual Machine Types:

You cannot arbitrary enter how many CPG, RAM, disk size, … you need. Microsoft provide predetermined options (more than 700).

Scaling Azure VMs

You can increase the size of a VM easily (scale up) OR yu can add more VMs and have them work together to handle the work (scale out).

Note: Scale up has some limitations (e.g. maximum CPU). Thus, preferred method is scale out.

One approach to do this is “Virtual Machine Scale Sets”. Here a group of virtual machines that can grow and shrink in quantity based one a predefined rule (based one monitoring demand / time (schedule) / other factors).

VM Scale Sets:

- Elasticity
- Two or more VM running the exact same code
- With a “load balancer” (traffic cop) in front to direct traffic randomly to one of the machines
- Able to add more machines as demand grows (autoscaling)
- Able to add more machines as demand slows
- Can handle up to 100 VMs in a single scale set
- If need, you can create more scale sets

Two main concepts when running multiple VMs are: Availability Sets and Proximity Groups

Availability Sets:

You have multiple VM that have an identical function.

You want to signify to Azure that these machines are critical and they should be kept separated each other. The reason for this is “fault isolation”. In this way if something goes wrong, it may affect one of the VM.

Fault isolation is done using:

- Fault domain: This is some type of unplanned outages (power / network) which would affect entire groups / rack of computers and then bring down all machine run on that rack. So, we do not want all the VMs run on that rack.
- Update domain: These are planned outages. Example: MS roll out new version of planform, and they do not do it to all the servers at once (different update domains). Thus, we can put our VMs in different update domains.
- Separated power sources and network switches
- Updated one at a time, not all together

The opposite of availability set is a Proximity Group. Here,

- Multiple VMs that have identical function
- Arranged in a VM Scale Set
- You want them placed together
- Less availability for more performance
- Fastest inter-server communications