VM name :

1. configured as part of the operating system.
2. 15 characters on a Windows virtual machine and
3. 64 characters on a Linux virtual machine.
4. defines a manageable Azure resource, and it's not trivial to change later

VM has price differences between locations.

Compute expenses are priced on a per-hour basis but billed on a per-minute basis.

With the consumption-based option, you pay for compute capacity by the second.

Status of the virtual machine has no relation to the Azure Storage charges that are incurred.

When you stop and deallocate the virtual machine, you can select any size available in your region. Resizing a machine might require a restart that can cause a temporary outage or change configuration settings such as the IP address.

Virtual machines can also have one or more data disks. All disks are stored as virtual hard disks (VHDs).

-- On Windows virtual machines, the temporary disk is labeled as the D: drive by default. This drive is used for storing the **pagefile.sys** file.

-- On Linux virtual machines, the temporary disk is typically /dev/sdb. Azure Linux Agent formats this disk and mounts it to /mnt.

Using multiple data disks gives your applications up to 256 TB of storage per virtual machine.

Azure-managed data disks are stored as page blobs, which are a random IO storage object in Azure. The disk is described as *managed* because it's an abstraction over page blobs, blob containers, and Azure storage accounts. With managed disks, you provision the disk, and Azure takes care of the rest.

The Azure Bastion service (fully platform-managed PaaS service) provides secure and seamless RDP/SSH connectivity to virtual machines directly over SSL. When you connect via Azure Bastion, your virtual machines don't need a public IP address.

Use RDP to connect to Windows VM, SSH for Linux VM and Bastion for any of Windows and Linux machines.

An availability zone in an Azure region is a combination of a fault domain and an update domain.

minimum of three separate zones in all enabled regions.

Category	Description	Examples
Zonal services	Azure *zonal* services pin each resource to a specific zone.	- Azure Virtual Machines - Azure managed disks - Standard IP addresses
Zone-redundant services	For Azure services that are zone-redundant, the platform replicates automatically across all zones.	- Azure Storage that's zone-redundant - Azure SQL Database

Vertical scaling(*scale up and scale down)*, involves increasing or decreasing the virtual machine **size** in response to a workload.

Horizontal scaling(*scale in and scale out)*, involves increasing or decreasing the virtual machine **count** in response to a workload.

Azure Virtual Machine Scale Sets are an Azure Compute resource that you can use to deploy and manage a set of **identical** virtual machines. Helps gain true *autoscaling*

Virtual Machine Scale Sets support the use of Azure Load Balancer for basic layer-4 traffic distribution, and Azure Application Gateway for more advanced layer-7 traffic distribution and SSL termination.

Virtual Machine Scale Sets support up to 1,000 virtual machine instances. If you create and upload your own custom virtual machine images, the limit is 600 virtual machine instances.

In flexible orchestration mode, you manually create and add a virtual machine of any configuration to the scale set. In uniform orchestration mode, you define a virtual machine model and Azure generates identical instances based on that model.

characteristics of deployment slots.

- Deployment slots are live apps that have their own hostnames.
- Deployment slots are available in the Standard, Premium, and Isolated App Service pricing tiers. Your app needs to be running in one of these tiers to use deployment slots.
- The Standard, Premium, and Isolated tiers offer different numbers of deployment slots.
- App content and configuration elements can be swapped between two deployment slots, including the production slot.

validate changes to your app in a staging deployment slot before swapping the app changes with the content in the production slot.

Deployment slots to be used for **Consider validation**. **Consider reductions in downtime**. **Consider restoring to last known good site, Consider Auto swap**.

Connections strings follow the content across the swap.

App Service provide auth mechanism.

- **Allow Anonymous requests (no action)**. Defer authorization of unauthenticated traffic to your application code. For authenticated requests, App Service also passes along authentication information in the HTTP headers. This feature provides more flexibility for handling anonymous requests. With this feature, you can present multiple sign-in providers to your users.
- **Allow only authenticated requests**. Redirect all anonymous requests to /.auth/login/<provider> for the provider you choose. The feature is equivalent to **Log in with <provider>**. If the anonymous request comes from a native mobile app,

the returned response is an HTTP 401 Unauthorized message. With this feature, you don't need to write any authentication code in your app.

Backup and Restore feature need the Standard or Premium tier App Service plan for your app or site.

Azure storage account and container in the same subscription as the app to back up.

Backups can hold up to 10 GB of app and database content.

If your storage account is enabled with a firewall, you can't use the storage account as the destination for your backups.

top-level resource in Azure Container Instances is the **container group**. A container group is a collection of containers that get scheduled on the same host machine

containers in a container group share a lifecycle, resources, local network, and storage volumes.

A container group is similar to a pod in Kubernetes. A pod typically has a 1:1 mapping with a container, but a pod can contain multiple containers. The containers in a multi-container pod can share related resources.

Azure Container Instances allocates resources to a multi-container group by adding together the resource requests of all containers in the group. Resources can include items such as CPUs, memory, and GPUs.

There are two common ways to deploy a multi-container group: Azure Resource Manager (ARM) templates and YAML files.

- **ARM template**. An ARM template is recommended for deploying other Azure service resources when you deploy your container instances, such as an Azure Files file share.
- **YAML file**. YAML file is recommended when your deployment includes only container instances.

Container groups can share an external-facing IP address, one or more ports on the IP address, and a DNS label with an FQDN.

Port mapping isn't supported because containers in a group share a port namespace.

KEDA is a Kubernetes-based Event Driven Autoscaler.

Envoy is an open source edge and service proxy, designed for cloud-native applications

first time you connect to a Windows server VM, it launches Server Manager. This allows you to assign a worker role for common web or data tasks. You can also launch the Server Manager through the **Start** menu.

For inbound traffic, Azure processes the security group associated to the subnet, then the security group applied to the network interface.

Outbound traffic is processed in the opposite order (the network interface first, followed by the subnet).

Deny rules always **stop** the evaluation.

Azure CLI available on macOS, Linux, and Windows, or in the browser using Azure Cloud Shell

Command	Sub-command	Description
vm	create	Create a new virtual machine
	deallocate	Deallocate a virtual machine
	delete	Delete a virtual machine
	list	List the created virtual machines in your subscription
	open-port	Open a specific network port for inbound traffic
	restart	Restart a virtual machine
	show	Get the details for a virtual machine

Command	Sub-command	Description
	start	Start a stopped virtual machine
	stop	Stop a running virtual machine
	update	Update a property of a virtual machine

Type	Sizes	Description
General purpose	Dsv3, Dv3, DSv2, Dv2, DS, D, Av2, A0-7	Balanced CPU-to-memory. Ideal for dev/test and small to medium applications and data solutions.
Compute optimized	Fs, F Remember by F for fast	High CPU-to-memory. Good for medium-traffic applications, network appliances, and batch processes.
Memory optimized	Esv3, Ev3, M, GS, G, DSv2, DS, Dv2, D	High memory-to-core. Great for relational databases, medium to large caches, and in-memory analytics.
Storage optimized	Ls Remember by L for data Lake	High disk throughput and IO. Ideal for big data, SQL, and NoSQL databases.
GPU optimized	NV, NC Remember by Nvidia	Specialized VMs targeted for heavy graphic rendering and video editing.
High performance	H, A8-11 Remember by H for High	Our most powerful CPU VMs with optional high-throughput network interfaces (RDMA).