100 XP

Compile a checklist for creating an Azure Virtual Machine

15 minutes

Performing a migration of on-premises servers to Azure requires planning and care. You can move them all at once, or more likely, in small batches or even individually. Before you create a single VM, you should sit down and sketch out your current infrastructure model and see how it might map to the cloud.

Required resources for IaaS Virtual Machines

Let's walk through a checklist of things to think about.

- Start with the network
- Name the VM
- Decide the location for the VM
- Determine the size of the VM
- Understanding the pricing model
- Storage for the VM
- Select an operating system

Start with the network

The first thing you should think about isn't the virtual machine at all - it's the network.

Virtual networks (VNets) are used in Azure to provide private connectivity between Azure Virtual Machines and other Azure services. VMs and services that are part of the same virtual network can access one another. By default, services outside the virtual network cannot connect to services within the virtual network. You can, however, configure the network to allow access to the external service, including your on-premises servers.

This latter point is why you should spend some time thinking about your network configuration. Network addresses and subnets are not trivial to change once you have them set up, and if you plan to connect your private company network to the Azure services, you will want to make sure you consider the topology before putting any VMs into place.

When you set up a virtual network, you specify the available address spaces, subnets, and security. If the VNet will be connected to other VNets, you must select address ranges that are not overlapping. This is the range of private addresses that the VMs and services in your network can use. You can use unrouteable IP addresses such as 10.0.0.0/8, 172.16.0.0/12, or 192.168.0.0/16, or define your own range. Azure will treat any address range as part of the private VNet IP address space if it is only reachable within the VNet, within interconnected VNets, and from your on-premises location. If someone else is responsible for the internal networks, you should work with that person before selecting your address space to make sure there is no overlap and to let them know what space you want to use, so they don’t try to use the same range of IP addresses.

Segregate your network

After deciding the virtual network address space(s), you can create one or more subnets for your virtual network. You do this to break up your network into more manageable sections. For example, you might assign 10.1.0.0 to VMs, 10.2.0.0 to back-end services, and 10.3.0.0 to SQL Server VMs.

> **Note**
>
> Azure reserves the first four addresses and the last address in each subnet for its use.

Secure the network

By default, there is no security boundary between subnets, so services in each of these subnets can talk to one another. However, you can set up Network Security Groups (NSGs), which allow you to control the traffic flow to and from subnets and to and from VMs. NSGs act as software firewalls, applying custom rules to each inbound or outbound request at the network

interface and subnet level. This allows you to fully control every network request coming in or out of the VM.

Plan each VM deployment

Once you have mapped out your communication and network requirements, you can start thinking about the VMs you want to create. A good plan is to select a server and take an inventory:

- What does the server communicate with?
- Which ports are open?
- Which OS is used?
- How much disk space is in use?
- What kind of data does this use? Are there restrictions (legal or otherwise) with not having it on-premises?
- What sort of CPU, memory, and disk I/O load does the server have? Is there burst traffic to account for?

We can then start to answer some of the questions Azure will have for a new virtual machine.

Name the VM

One piece of information people often don't put much thought into is the **name** of the VM. The VM name is used as the computer name, which is configured as part of the operating system. You can specify a name of up to 15 characters on a Windows VM and 64 characters on a Linux VM.

This name also defines a manageable **Azure resource**, and it's not trivial to change later. That means you should choose names that are meaningful and consistent, so you can easily identify what the VM does. A good convention is to include the following information in the name:

Element	Example	Notes
Environment	dev, prod, QA	Identifies the environment for the resource
Location	uw (US West), ue (US East)	Identifies the region into which the resource is deployed
Instance	01, 02	For resources that have more than one named instance (web servers, etc.)
Product or Service	service	Identifies the product, application, or service that the resource supports

Element	Example	Notes
Role	sql, web, messaging	Identifies the role of the associated resource

For example, `devusc-webvm01` might represent the first development web server hosted in the US South Central location.

What is an Azure resource?

An **Azure resource** is a manageable item in Azure. Just like a physical computer in your datacenter, VMs have several elements that are needed to do their job:

- The VM itself
- Storage account for the disks
- Virtual network (shared with other VMs and services)
- Network interface to communicate on the network
- Network Security Group(s) to secure the network traffic
- Public Internet address (optional)

Azure will create all of these resources if necessary, or you can supply existing ones as part of the deployment process. Each resource needs a name that will be used to identify it. If Azure creates the resource, it will use the VM name to generate a resource name - another reason to be very consistent with your VM names!

Decide the location for the VM

Azure has datacenters all over the world filled with servers and disks. These datacenters are grouped into geographic *regions* ('West US', 'North Europe', 'Southeast Asia', etc.) to provide redundancy and availability.

When you create and deploy a virtual machine, you must select a region where you want the resources (CPU, storage, etc.) to be allocated. This lets you place your VMs as close as possible to your users to improve performance and to meet any legal, compliance, or tax requirements.

Two other things to think about regarding the location choice. First, the location can limit your available options. Each region has different hardware available and some configurations are not available in all regions. Second, there are price differences between locations. If your workload isn't bound to a specific location, it can be very cost effective to check your required configuration in multiple regions to find the lowest price.

Determine the size of the VM

Once you have the name and location set, you need to decide on the size of your VM. Rather than specify processing power, memory, and storage capacity independently, Azure provides different *VM sizes* that offer variations of these elements in different sizes. Azure provides a wide range of VM size options allowing you to select the appropriate mix of compute, memory, and storage for what you want to do.

The best way to determine the appropriate VM size is to consider the type of workload your VM needs to run. Based on the workload, you're able to choose from a subset of available VM sizes. Workload options are classified as follows on Azure:

Option	Description
General purpose	General-purpose VMs are designed to have a balanced CPU-to-memory ratio. Ideal for testing and development, small to medium databases, and low to medium traffic web servers.
Compute optimized	Compute optimized VMs are designed to have a high CPU-to-memory ratio. Suitable for medium traffic web servers, network appliances, batch processes, and application servers.
Memory optimized	Memory optimized VMs are designed to have a high memory-to-CPU ratio. Great for relational database servers, medium to large caches, and in-memory analytics.
Storage optimized	Storage optimized VMs are designed to have high disk throughput and IO. Ideal for VMs running databases.
GPU	GPU VMs are specialized virtual machines targeted for heavy graphics rendering and video editing. These VMs are ideal options for model training and inferencing with deep learning.
High performance computes	High performance compute is the fastest and most powerful CPU virtual machines with optional high-throughput network interfaces.

You're able to filter on the workload type when you configure the VM size in the Azure. The size you choose directly affects the cost of your service. The more CPU, memory, and GPU you need, the higher the price point.

What if my size needs change?

Azure allows you to change the VM size when the existing size no longer meets your needs. You can upgrade or downgrade the VM - as long as your current hardware configuration is allowed in the new size. This provides a fully agile and elastic approach to VM management.

The VM size can be changed while the VM is running, as long as the new size is available in the current hardware cluster the VM is running on. The Azure portal makes this obvious by only showing you available size choices. The command line tools will report an error if you attempt to resize a VM to an unavailable size. Changing a running VM size will automatically reboot the machine to complete the request.

If you stop and deallocate the VM, you can then select any size available in your region since this removes your VM from the cluster it was running on.

> ⚠ **Warning**
>
> Be careful about resizing production VMs - they will be rebooted automatically which can cause a temporary outage and change some configuration settings such as the IP address.

Understanding the pricing model

There are two separate costs the subscription will be charged for every VM: compute and storage. By separating these costs, you scale them independently and only pay for what you need.

Compute costs - Compute expenses are priced on a per-hour basis but billed on a per-minute basis. For example, you are only charged for 55 minutes of usage if the VM is deployed for 55 minutes. You are not charged for compute capacity if you stop and deallocate the VM since this releases the hardware. The hourly price varies based on the VM size and OS you select. The cost for a VM includes the charge for the Windows operating system. Linux-based instances are cheaper because there is no operating system license charge.

> 💡 **Tip**
>
> You might be able to save money by reusing existing licenses for Windows with the **Azure Hybrid benefit**.

Storage costs - You are charged separately for the storage the VM uses. The status of the VM has no relation to the storage charges that will be incurred; even if the VM is stopped/deallocated and you aren't billed for the running VM, you will be charged for the storage used by the disks.

You're able to choose from two payment options for compute costs.

Option	Description

Option	Description
Pay as you go	With the **pay-as-you-go** option, you pay for compute capacity by the second, with no long-term commitment or upfront payments. You're able to increase or decrease compute capacity on demand as well as start or stop at any time. Prefer this option if you run applications with short-term or unpredictable workloads that cannot be interrupted. For example, if you are doing a quick test, or developing an app in a VM, this would be the appropriate option.
Reserved Virtual Machine Instances	The Reserved Virtual Machine Instances (RI) option is an advance purchase of a virtual machine for one or three years in a specified region. The commitment is made up front, and in return, you get up to 72% price savings compared to pay-as-you-go pricing. **RIs** are flexible and can easily be exchanged or returned for an early termination fee. Prefer this option if the VM has to run continuously, or you need budget predictability, **and** you can commit to using the VM for at least a year.

Storage for the VM

All Azure virtual machines will have at least two virtual hard disks (VHDs). The first disk stores the operating system, and the second is used as temporary storage. You can add additional disks to store application data; the maximum number is determined by the VM size selection (typically two per CPU). It's common to create one or more data disks, particularly since the OS disk tends to be quite small. Also, separating out the data to different VHDs allows you to manage the security, reliability, and performance of the disk independently.

The data for each VHD is held in **Azure Storage** as page blobs, which allows Azure to allocate space only for the storage you use. It's also how your storage cost is measured; you pay for the storage you are consuming.

What is Azure Storage?

Azure Storage is Microsoft's cloud-based data storage solution. It supports almost any type of data and provides security, redundancy, and scalable access to the stored data. A storage account provides access to objects in Azure Storage for a specific subscription. VMs always have one or more storage accounts to hold each attached virtual disk.

Virtual disks can be backed by either **Standard** or **Premium** Storage accounts. Azure Premium Storage leverages solid-state drives (SSDs) to enable high performance and low latency for VMs running I/O-intensive workloads. Use Azure Premium Storage for production workloads, especially those that are sensitive to performance variations or are I/O intensive. For development or testing, Standard storage is fine.

When you create disks, you will have two options for managing the relationship between the storage account and each VHD. You can choose either **unmanaged disks** or **managed disks**.

Option	Description
Unmanaged disks	With unmanaged disks, you are responsible for the storage accounts that are used to hold the VHDs that correspond to your VM disks. You pay the storage account rates for the amount of space you use. A single storage account has a fixed-rate limit of 20,000 I/O operations/sec. This means that a storage account is capable of supporting 40 standard virtual hard disks at full utilization. If you need to scale out with more disks, then you'll need more storage accounts, which can get complicated.
Managed disks	Managed disks are the **newer and recommended disk storage model**. They elegantly solve this complexity by putting the burden of managing the storage accounts onto Azure. You specify the size of the disk, up to 4 TB, and Azure creates and manages both the disk *and* the storage. You don't have to worry about storage account limits, which makes managed disks easier to scale out.

Select an operating system

Azure provides a variety of OS images that you can install into the VM, including several versions of Windows and flavors of Linux. As mentioned earlier, the choice of OS will influence your hourly compute pricing as Azure bundles the cost of the OS license into the price.

If you are looking for more than just base OS images, you can search the Azure Marketplace for more sophisticated install images that include the OS and popular software tools installed for specific scenarios. For example, if you needed a new WordPress site, the standard technology stack would consist of a Linux server, Apache web server, a MySQL database, and PHP. Instead of setting up and configuring each component, you can leverage a Marketplace image and install the entire stack all at once.

Finally, if you can't find a suitable OS image, you can create your disk image with what you need, upload it to Azure storage, and use it to create an Azure VM. Keep in mind that Azure only supports 64-bit operating systems.

Next unit: Exercise - Create a VM using the Azure portal

Continue >