An MSP's Complete Guide to Azure Resources

Find more: getnerdio.com

The building blocks of an Azure IT environment are called resources. These resources are organized into Resource Groups inside of an Azure subscription. There are billable and non-billable resources; billable resources have a meter attached to them that runs while the resource is provisioned.

In this article, we will explore the three most common types of Azure resources used by MSPs when deploying IT environments: compute (virtual machines), storage, and network.

Every resource used in Azure must be deployed in a geographical location known as a region. An Azure region is a grouping of data centers located in a specific geographic location. Microsoft is constantly growing its global footprint and adding data centers and regions. At this moment, there are 54 regions available in 140 countries, and the list is growing. The most up-to-date map of regions can be viewed here.

Resources deployed in the same region are interconnected with high speed connectivity (think LAN speeds). Resources in different regions can still communicate with each other but are subject to additional WAN latency. The latency depends on how far the regions are from each other.

Compute (Virtual Machines)

Virtual Machines (VMs) in Azure come in predefined sizes that are called families or series. An individual VM is often referred to as an instance. Different VM families are designed for common use-cases and are comprised of certain amounts of CPU cores and GB of RAM. It is not possible to arbitrarily mix and match CPU cores and GB of RAM as can be done with Hyper-V and VMware.

Here, we will focus on the four most commonly used VM families by MSPs: Ds-series, B-series, Esv3-series and NV-series.

Ds-series

These are "general purpose" VMs that can be used for a wide variety of workloads. There are three versions of the DS-series: v1, v2, and v3. Only v2 and v3 should be used.

- **Purpose:** general applications (domain controllers, file servers, application servers, etc.)
- **CPU clock speed:** 2.4Ghz – 3.0GHz (with Intel Turbo Boost)
- **CPU-to-RAM ratio:**
  - V2 – 1:3.5GB (each CPU core gets 3.5GB of RAM)
  - V3 – 1:4.0GB (each CPU core gets 4.0GB of RAM)
- **Storage supported:** Standard and Premium
- **Approximate average list price per CPU:**
  - V2 - $85/month
  - V3 - $77/month
- **Difference between V2 and V3:**
  - V2 VMs use non-hyperthreaded vCPUs (1 vCPU per 1 physical CPU core), which is why they are slightly more expensive. V2 VMs start at a single core size (DS1v2).
  - V3 VMs use hyperthreaded vCPUs (2 vCPUs per 1 physical CPU), which is why they are less expensive. V3 VMs start at a minimum of two vCPUs (D2sv3).

Ds-series VMs are a good fit for workloads that require consistent CPU usage and are not very RAM hungry.

Esv3-series

These are "general purpose, high-memory" VMs that can be used for many workloads that are more RAM hungry rather than CPU hungry.

- **Purpose:** general, RAM bound applications (database servers, application servers, desktops, etc.)
- **CPU clock speed:** 2.3Ghz – 3.5Ghz (with Intel Turbo Boost)
- **vCPU-to-RAM ratio:** 1:8.0GB (each CPU gets 8.0GB of RAM)
- **Storage supported:** Standard and Premium
- **Approximate average list price per CPU:** $88/month

Esv3-series VMs are very similar to Dsv3-series but have double the RAM per CPU and are about 15% more expensive. They are ideal for workloads that consistently utilize the CPU and are memory hungry. Examples are database servers and RDS session hosts.

B-series

These are known as "burstable" VMs. They are very useful but the way they work is a bit complicated. They are used for non-CPU intensive workloads (e.g. domain controllers, file servers) and cost about 50% of an equivalently sized Ds-series VM.

The reason these are cheaper is because Azure imposes a quota on how much of the total CPU cores can be used. This quota is usually a fraction of the total available CPU. For instance, B2m's quota is 60% of a single CPU, which is 30% of the two CPUs visible in the VM. Every second that the VM is using less than its quota (less than 60% of a single CPU), it is "banking credits". These banked credits can be used to burst up to the total available CPUs (100% of two CPUs, in this example) when needed. While bursting, the VM is consuming its banked credits. Once credits run out, the VM's CPU utilization is throttled down to its 60% quota.

Why use B-series VMs? Well, they are cheaper. For approximately the same price that you would pay for a Ds-series VM, you can get a B-series with double the CPUs and double the RAM. However, they should only be used for workloads that are either not CPU intensive or "bursty", meaning they only occasionally need all the CPU but most of the time the CPU is idle.

For instance, an Active Directory domain controller is not utilizing its CPU very heavily on a regular basis. When Windows Updates run, the VM will use all its available CPU horsepower.

B-series are perfect for Domain Controllers since they bank credits while idle and then consume them when needed to update or do some other CPU intensive task.

- **Purpose:** general, non-CPU intensive workloads (e.g. AD domain controllers, file servers)
- **CPU clock speed:** varies
- **vCPU-to-RAM ratio:** varies from 1:1 to 1:4 for VMs larger than B2s
- **Storage supported:** Standard and Premium
- **Approximate average list price per CPU:** ranges from $13/month to $40/month
- **Tips:**
  - Don't use B-series VMs for CPU intensive workloads
  - When a B-series VM is first provisioned, it doesn't have any banked credits and is subject to its quota limit on the CPU -- which means it's slow. Once the VM is running idle for some time, credits get banked and the VM performance improves when it needs to burst.
  - Don't shut down B-series VMs overnight when they are not in use. This will not allow the VMs to bank credits for the following day of usage.

NV-series

These VMs are intended for special use-cases when a dedicated GPU is needed. They include an NVIDIA GRID 2.0 Tesla GPU and are ideal for running graphically intensive workloads like AutoCAD, SolidWorks, and Revit. These are very large and expensive VMs (starting at 6 CPUs and 56GB of RAM) and need to be used with caution and with a specific purpose in mind to not generate unpredictably large Azure compute consumption bills.

- **Purpose:** graphically heavy, visual workloads inside of virtual desktop sessions
- **vCPU-to-RAM ratio:** 6:56GB (each 6 CPUs get 56GB of RAM)
- **vCPU-to-GPU ratio:** 6:1 (each 6 CPUs get 1 M60 GPU)
- **Storage supported:** Standard ONLY (Premium is not supported)
- **Approximate average list price per CPU:** $165/month
- **Tips:**
  - Smallest VM is NV6 (6 CPU / 56GB RAM / 1 GPU)
  - Since only Standard storage is supported, disk performance is not fast
  - Not available in all Azure regions

- New NVv2 VMs are currently in preview and are going to have the following notable improvement once they are generally available:
    - 40% price reduction
    - 2X RAM increase per CPU
    - Support for Premium storage

Now that we understand the different types of VMs, let's talk about how to use them.

The first important thing to understand is that VMs are not stand-alone resources. For example, a VM must have an OS disk (and optionally data disks) attached to it, as well as a virtual network interface (vNIC). A new VM can be created (deployed) using an existing OS disk and vNIC or new disk and vNIC can be created together with the VM. If a VM is deleted, its data (i.e. OS and Data disks) are not deleted. They remain as resource objects in Azure that are not attached to any VM. More on storage resources later.

When deploying a VM, its OS disk must be based on an existing image and cannot be blank. Since you don't have console access to VMs in Azure, the OS cannot be installed on a "blank" OS disk. The OS disk must already have the OS on it. Images could be pulled from the Azure image library or you can create and upload your own custom image as a VHD file to Azure to be used for deploying a VM.

All VMs also come with a temporary D: drive that has locally attached fast storage (SSD). Keep in mind that this disk is temporary, and any data stored on it will likely be erased if the VM is ever shut down or moved to another Azure host in the background. Use this disk for the pagefile and temp data but be sure to never store anything you need to retain on the temporary disk.

After you deploy a VM it becomes provisioned or allocated, meaning it is running on an Azure host, consuming Azure resources and you're consequently being billed for every second that the VM is allocated.

To stop being billed for a running VM, you must stop it. This process causes the VM to become deallocated, which means it is effectively powered off and is not consuming Azure resources. It is possible to shut down a VM and still be paying for it because it stays allocated. When you power off a VM from inside of the OS, it shuts down, but Azure still sees it as allocated and you are being billed. Be sure to stop VMs at the Azure level even if you shut them down at the OS level.

Another important concept to mention when discussing VMs is subscription core quota. To prevent accidental or malicious use of Azure where many VMs are created and a large amount of consumption occurs, Microsoft imposes core quotas on subscriptions by default. The number of CPU cores that can be provisioned in a subscription in total and per VM family are limited.

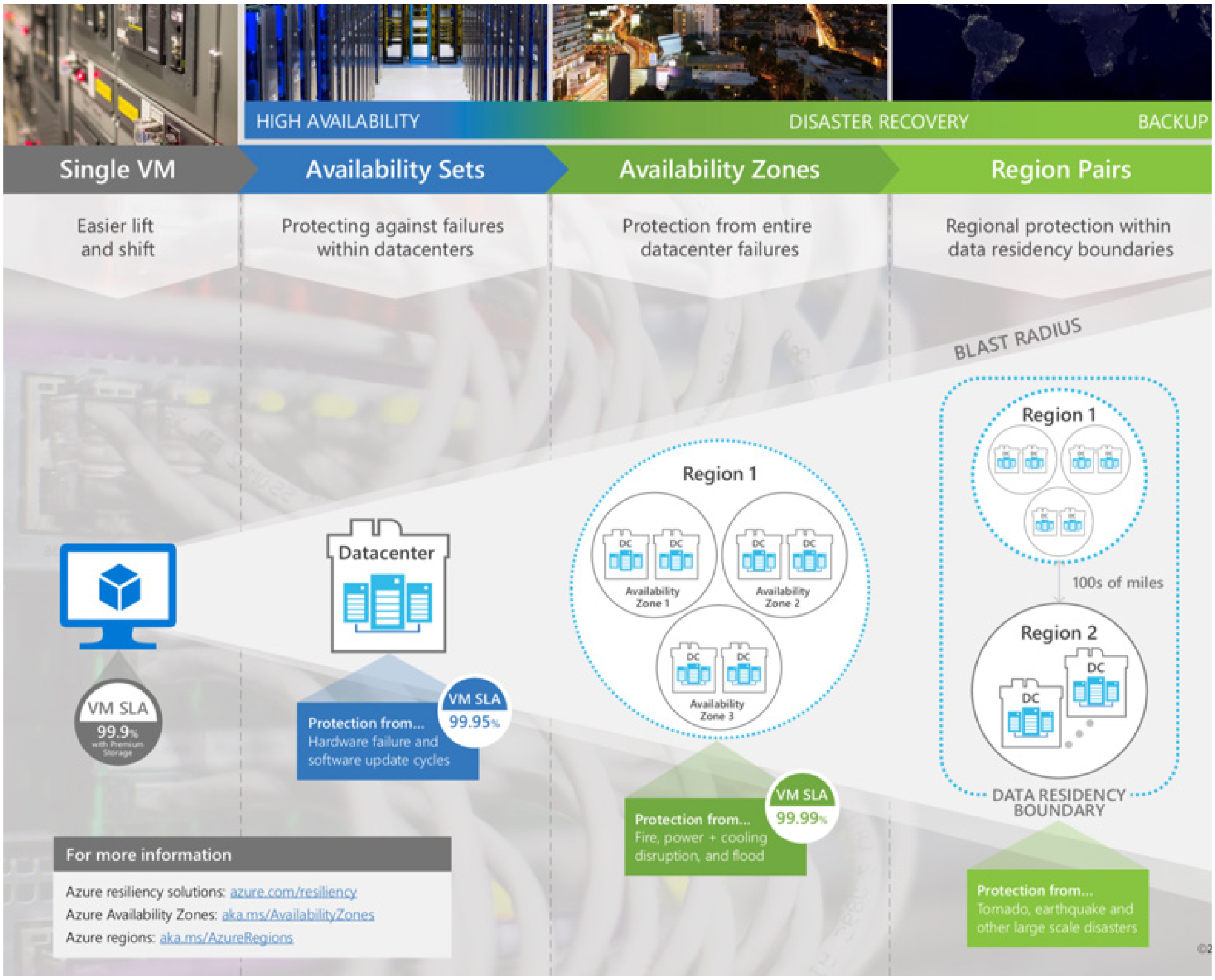
For instance, a free subscription has an overall core quota of 4. Direct Pay-As-You-Go subscriptions have a default core quota of 10, and CSP subscriptions have a core quota of 20. This means that with a CSP subscription, you cannot provision more VMs whose total CPU cores exceed 20. Be mindful of this limit. To increase the core quota limit, you need to submit a request to Microsoft via the Azure portal for a core limit increase.

Finally, it is important to be aware that only some Azure VMs' availability is covered by Microsoft's Service Level Agreement (SLA). VMs not covered by an SLA could be unexpectedly rebooted due to underlying Azure infrastructure upgrades or hardware failure. It has become exceedingly rare to see VMs reboot in Azure, but it was not uncommon in the past.

Presence of an SLA and the availability guarantee (e.g. 99.9% vs. 99.95% vs. 99.99%) is based on several factors that have to do with the type of storage the VM uses for its OS and data disks, as well as if it is deployed in an availability set or an availability zone.

You can learn more about the specifics here.

The diagram below summarizes the available protection options.


HIGH AVAILABILITY
DISASTER RECOVERY
BACKUP
Single VM
Availability Sets
Availability Zones
Region Pairs
Easier lift and shift
Protecting against failures within datacenters
Protection from entire datacenter failures
Regional protection within data residency boundaries
BLAST RADIUS
VM SLA 99.9% with Premium Storage
Datacenter
VM SLA 99.95%
Protection from... Hardware failure and software update cycles
Region 1
DC
Availability Zone 1
Availability Zone 2
Availability Zone 3
VM SLA 99.99%
Protection from... Fire, power + cooling disruption, and flood
Region 1
100s of miles
Region 2
DATA RESIDENCY BOUNDARY
Protection from... Tornado, earthquake and other large scale disasters
For more information
Azure resiliency solutions: azure.com/resiliency
Azure Availability Zones: aka.ms/AvailabilityZones
Azure regions: aka.ms/AzureRegions
©2

For most situations relevant to an MSP, it is important to know that individual VMs ("Single VM" in Microsoft terms) that use any standard storage disks are not covered by any SLA. The chance of outage is very small and even if the VM reboots due to an underlying hardware failure it will restart very quickly elsewhere. However, it is important to remember that no SLA applies.

Critical VMs should use Premium storage only, which will provide them with a 99.9% availability guarantee and improved performance. For additional availability guarantees, distributed workloads that can have multiple VMs participating in the same application, can be placed inside Availability Sets and will then be subject to 99.95% availability guarantee.

An example of such a deployment may be Active Directory. You can have two AD domain controllers in an Availability Set and your AD, as a whole, will have a guarantee of 99.95%. This doesn't mean that each domain controller VM has this guarantee. Rather, the "application" (i.e. AD) is guaranteed to be available 99.95% of the time.

Storage

Azure offers multiple storage options with different performance, redundancy, location and price characteristics. It's easy to get lost in all the available options and to clearly understand what type of storage should be used when. We will focus on three storage resources that are most commonly used by MSPs when deploying IT environments in Azure: **Managed Disks, Backup Vaults, and Files.**

In addition to considering the type of storage resource we need to understand the data redundancy, performance, and cost consideration for each type of storage object.

Data redundancy

- LRS – Locally Redundant Storage
    - Three redundant copies of data stored in one data center
    - 99.99999999% (11 9's) durability
- GRS – Geo-Redundant Storage
    - Six total redundant copies of data. Three copies stored in one region and another three copies are asynchronously replicated to a second region
    - 99.99999999999999% (16 9's) durability
- ZRS – Zone-Redundant Storage
    - Three redundant copies of data stored across two or three data centers within the same Azure region
    - 99.9999999999% (12 9's) durability
- RA-GRS – Read Access GRS
    - This redundancy type is not relevant to the storage objects in this discussion

Performance

There are three Performance tiers: **Standard**, **Premium**, and **Ultra**. Standard storage utilizes inexpensive and slow HDD and recently Microsoft added Standard SSD, which doesn't increase the average performance but makes it more consistent than HDD. Premium storage uses SSD disks and is fast. This type of storage is best for most disk IO intensive applications such as databases and virtual desktops. Ultra SSD is a new type of storage for very high-performance, disk IO intensive applications.

Now that we understand the redundancy and performance characteristics of Azure storage, let's dive into the actual storage resources.

Managed Disks are by far the most commonly used type of storage when deploying an IT environment in Azure using virtual machines. Recall that each VM must have, at a minimum, an OS disk and sometimes one or more additional data disks. These disks that get attached to a VM are known as Managed Disks in Azure. There is an older type of disk called Unmanaged Disk, but for the purposes of our discussion, we will stick to Managed Disks.

If you're interested in learning more about the differences between managed and unmanaged disks. click here.

Managed disks are only available with LRS data redundancy since they are attached directly to VMs and these VMs must be able to communicate with disks in a very high throughput, low latency way. This is why managed disks and VMs that they are attached to must be located in the same region. Disks come in Standard HDD, Standard SSD, Premium SSD, and Ultra SSD performance flavors.

Let's explore each type of managed disk in detail:

Standard HDD (S-type disk – e.g. S4, S10, S20, etc.)

- Available sizes: 32GB – 32TB in discreet increments (e.g. 32GB, 64GB, 128GB, etc.)
- Billed on allocated space, not used space. Creating an S-type disk of a certain size will result in a bill for the entire size, even if it completely unused.
- What you're billed for:
  - Capacity – approximately $0.048/GB/month
  - Operations - $0.0005 per 10,000 transactions
- Performance: up to 500 IOPS and up to 60MB/sec throughput (performance varies significantly and can often be far below this limit)
- When to use?
  - Very low disk IO applications (e.g. ADFS proxy server)
  - Test environments
  - When VM is deallocated but you still want to keep it around, changing it to an S-type disk saves on storage costs

Standard SSD (E-type disk – e.g. E4, E10, E20, etc.)

- Available sizes: 32GB – 32TB in discreet increments (e.g. 32GB, 64GB, 128GB, etc.)
- Billed on allocated space, not used space. Creating an E-type disk of a certain size will result in a bill for the entire size, even if it completely unused.
- What you're billed for:
    - Capacity – approximately $0.075/GB/month
    - Operations - $0.002 per 10,000 transactions
- Performance: up to 500 IOPS and up to 60MB/sec throughput (more consistent performance than S-type disks)
- When to use?
    - Best for most non-disk IO heavy applications because of nice balance between performance consistency and cost (e.g. domain controllers, file servers). Not a good fit for high IO database servers.
    - Production environments, if no SLA is needed
    - Most VDI desktop workloads for typical users

Premium SSD (P-type disk – e.g. P4, P10, P20, etc.)

- Available sizes: 32GB – 32TB in discreet increments (e.g. 32GB, 64GB, 128GB, etc.)
- Billed on allocated space, not used space. Creating a P-type disk of a certain size will result in a bill for the entire size, even if it completely unused.
- What you're billed for:
    - Capacity – approximately $0.15/GB/month
    - Operations – no transaction costs
- Performance: 120 – 7500 IOPS and 25MB/sec – 250MB/sec throughput
- When to use?
    - Best disk performance for any disk IO intensive applications such as databases
    - Great for power user virtual desktops and RDS session hosts with many users
    - Expensive for data storage only when the VM is powered off. Consider converting P to S or E disk if VM is being deallocated and data stored for archival purposes.

Ultra SSD

- High performance and high cost disk option for very disk IO intensive workloads
- Complex billing structure based on provisioned IOPS and throughput in addition to capacity storage
- Not commonly used with typical MSP workloads in Azure

Backup Vaults, as the name implies, are used by the Azure Backup service to store backup snapshots. It is a **Block Blob** storage container and its cost is based on actual consumption. Currently, Azure backup supports only Standard HDD performance tiers and LRS and GRS data redundancy options. The cost of backup vault storage is approximately $0.024/GB/month for LRS and 2X that amount for GRS storage.

Azure Backup is most commonly used by MSPs to protect data on VMs running inside of an Azure IT environment but can also be used to back up data from on-premises systems. To protect Azure VMs, the backup vault must reside in the same region as the VMs that are being backed up to it.

Azure backup can be used to achieve compliance with requirements to save data in multiple geographic locations by selecting the GRS redundancy option when creating the backup vault. This way, there will be multiple copies of the backup data in the same datacenter where the VMs reside as well as multiple copies in another paired region. With GRS, Microsoft has pre-defined region pairs.

Click here for more information on region pairs.

Azure Files is a PaaS offering. The easiest way to think about it is as a Microsoft-managed file server where you can create Windows shares and publish them out to the world. These shares can then be mounted directly on Windows, Linux and macOS devices, either on-premises or in cloud VMs without any special drivers.

Azure Files supports LRS, ZRS and GRS storage and costs range from $0.06/GB/month to $0.10/GB/month plus the cost of operations ($0.015 to $0.03 per 10,000 transactions). Azure Files is currently available with Standard storage only, which significantly limits its performance. However, Premium storage support is in preview and should be available soon.

In summary, Azure offers an almost endless list of storage options with varying redundancy, performance, and cost characteristics. For MSPs, it is important to focus on the storage types that are commonly used for typical IT workloads (managed disks for VMs, Block Blob for Azure

Backup, and Azure Files for creating SMB shares) and avoid confusion around other storage types that are designed for developers creating applications and repositories.

Network

Azure's flexibility when it comes to networking is vast and not without complexity. Many network resources are for advanced use-cases and for developers who are designing new applications.

Here, we will focus on four network resources that are most relevant to an MSP and the way they interrelate with each other: Virtual Networks, Public IP Addresses, Network Security Groups and VPN Gateways.

Before delving into the specifics of these network resources, we need to understand how Azure charges for data transfer (aka bandwidth). The basic rule is that any data coming into an Azure data center is free, while going out of an Azure region will be charged on a per GB basis. It doesn't matter if the data is leaving a region and going into another region or leaving a region and going into some other, non-Azure location. In both cases there is a charge. However, data transfer within the same Azure region (even across different data centers) is free.

How much does outbound data transfer cost? The first 5GB in any given month are free and then it's $0.05 to $0.087 per GB after that. Let's put things in perspective; a 10GB file being downloaded from an Azure hosted VM to your laptop will cost $0.87.

It is important to note that Azure data transfer is not charged per mbps (using 95% percentile or some other method), but rather per transferred GB of data. Let's compare the two methods.

Colocation Provider A charges $50/month for 1mbps of bandwidth using the 95% percentile method. Assuming the line is utilized 95% for the entire month straight, that's equivalent to 60sec/min*60min/hr*24hr/day*30.5days/month * (0.95 * 1mbps) = 2,503,440 megabits per month or 305GB/month. For the same amount of data transfer Azure cost will be $26.48.

Therefore, a useful number for cost comparison between "GB transferred" and "mbps" based pricing is $26 per fully utilized mbps line. Since in a typical hosted IT environment the line is utilized only fractionally, the cost of bandwidth in Azure is relatively low compared to the way other hosting and colocation providers charge for bandwidth.

This data transfer fee applies to all methods of transfer: communicating with a VM in Azure, downloading a file from Azure Files, restoring from a backup to outside of the region where the backup vault resides, using site-to-site VPN, etc. Anytime data leaves the boundaries of an Azure region, there is a charge.

With the cost of data transfer out of the way, let's delve into the way networking is structured in Azure. At the top level there is a Virtual Network (vNet). A vNet has an address space that you as an MSP can define (e.g. 10.1.0.0/16). All objects within a vNet must fall inside of this address space. vNet also contain subnets. These subnets are a way to segment the vNet into smaller sections. For instance, you could have a LAN and DMZ subnets within a vNet.

- vNet – 10.1.0.0/16
    - LAN subnet – 10.1.0.0/17
    - DMZ subnet – 10.1.254.0/24

Subnets that are part of a vNet can have virtual Network Interfaces (vNIC) attached to them. These vNICs are then attached to a VM and this is the way VMs communicate with each other and the rest of the world. VM->vNIC->Subnet->vNet.

Each vNIC has an assigned private IP address (or addresses), DNS settings, an optional public IP address and other network interface properties. In Azure, IP address and DNS settings are not set at the Windows level inside of a VM. Rather, they are set at the vNIC level in Azure. In Windows, the network adapter is set to DHCP and receives its settings from the vNIC that's attached to it. The vNIC itself, could have a statically assigned IP address or a dynamic one given to it by Azure via DHCP.

You can peer (i.e. connect) different vNets together. These vNets can be in the same Azure region or you can use Global vNet Peering to connect vNets in different regions.

Public IP addresses are billable Azure resources that can be assigned to a vNIC. There are dynamic IP addresses and static IP addresses. Dynamic ones have a persistent DNS name that resolves to a dynamic IP, while a static IP address has a fixed IPv4 address and DNS name.

The cost of a public dynamic IP address is $3/month, while the cost of a public static IP address is about $4/month. Assigning a public IP address to a vNIC does not automatically expose the VM to the internet. In order to make it accessible from the internet, a Network Security Group rule must be applied.

Network Security Groups (NSGs) are Azure's basic network firewall. They are non-billable network resources. NSGs are groups of firewall rules that specify what's allowed or denied into and out of a vNet. If an NSG is assigned to a subnet, its rules will apply to all VMs whose vNICs are part of this subnet. Alternatively, NSGs can be assigned directly to a vNIC. In that case, the NSG firewall rules will apply to this single VM only.

VPN Gateway is a service that allows encrypted, site-to-site IPSec VPN connectivity from an on-premises network or another cloud to an Azure vNet. VPN Gateways are Microsoft managed resources that get added to a special subnet in a vNet called the Gateway Subnet. VPN Gateway is a billable network resource and pricing starts at $26/month for a basic gateway with a throughput limit of 100 mbps and support for up to 10 site-to-site VPN tunnels. The largest VPN Gateway is $912/month and supports 1.25 Gbps of throughput with up to 30 tunnels.

For more information on Microsoft Azure, visit our Nerdio Academy, where we regularly publish unique content for Managed Service Provider who are looking to grow their Azure practice.

Contact Us:

Phone: 1-844-463-7346
Email: hello@getnerdio.com
Website: getnerdio.com