



Azure Margin Optimization for MSPs

Find more:
getnerdio.com



For an MSP building a cloud practice in Azure one of the most important considerations is to understand how to transact in Azure. This includes becoming familiar with the various types of Azure resources and how they work, how each type of resource is billed, how to architect an Azure IT environment, and most importantly **how to make money.**

Azure is a public cloud with publicly listed prices. Any customer can go to azure.com and look up the price of VMs, storage and networking. It will not be as easy or straightforward as you may think because of the complexity and vast number of options--but the prices are public. It is important to keep this in mind when creating an Azure offer, customer proposals, and quotes.

Two common approaches to creating an Azure offer are **bundling** and **resale**. Some MSPs prefer to provide their customers with an all-in IT solution that includes infrastructure running in Azure, software licensing, and managed services – all for a predictable monthly price on a per-user, per-device basis--or some other unit of value. Other MSPs bundle their managed services and resell Azure infrastructure separately from the managed services agreement.

Why would a customer choose to buy Azure through an MSP rather than directly from Microsoft if the price is the same? First, the MSP is an Azure expert and can guide the customer towards the most efficient use of the cloud. By purchasing Azure through the CSP program, the MSP maintains management capabilities on the Azure subscription to quickly resolve issues for the customer. Most importantly, MSPs have access to Microsoft support

resources, which can be instrumental in quickly addressing issues; a benefit the customer cannot receive without paying extra. For these reasons, most customers will choose to purchase Azure through their MSP.

In terms of the MSP's offer of cloud services to their end customer, optimizing the cost of Azure infrastructure will have significant impact on the MSP's gross margins and it is therefore important to spend some time understanding the levers that are available to optimize Azure costs and increase your margin.

There are many levers available to reduce Azure costs drastically relative to list prices. We're not talking here about 5%-10% points but high double-digit numbers, in some cases. Here we will review the top 5 levers that can make the most difference in driving MSP profitability. These five will take you most of the way towards really making money offering and reselling Azure as a core part of your cloud practice.

In this article, we will focus on levers that reduce Azure infrastructure costs—but there is another important contributor to optimizing cost and MSP needs to keep in mind—the human capital costs of quoting, deploying, managing and optimizing each customer's Azure environment which can be quite significant and, in some cases, exceed the costs of the infrastructure. We will focus on strategies to drastically reduce these labor costs in future articles.

**...optimizing the cost of
Azure infrastructure will
have significant impact on
the MSP's gross margins...**

Lever #1: Become a CSP

Margin impact: Increase of 10%–20% margin of total Azure consumption

This one is somewhat of a no-brainer. Not only will it be easier to transact Azure (i.e. create customer subscriptions), it will also allow you to increase your margin by receiving a discount off Azure list prices via your CSP Provider. You will also be eligible for various incentives that Microsoft makes available to their CSPs based on their growth objectives. These incentives are incremental to the discount you'll receive on Azure consumption and can be in the 10% range, or more, when added up.

There are two types of CSPs: Direct CSPs and CSP Resellers. Direct CSPs purchase Azure directly from Microsoft but are required to maintain high consumption volume and need to overcome many qualification hurdles. CSP Resellers purchase Azure through an Indirect CSP Provider. These providers are large distributors that help CSP Resellers understand Azure, its pricing, and billing. Any MSP can become a CSP Reseller without any volume commitments or hurdles.

Direct CSPs and CSP Resellers are eligible for the following benefits:

- Azure consumption discounts: 7%–15%
- Azure Reserved Instances (more on this below) discounts: 1%–5%
- CSP Software Subscriptions (more on this below) discounts: 5%–15%
- Office 365, Dynamics 365, and Microsoft 365 discounts: 10% – 20%
- Local, regional, and other accelerators (i.e. rebates): these vary by year and Microsoft priorities. Some are available to Direct CSPs only.

There are many good reasons to take advantage of the CSP program and increase your margins by 10% to 20% on Azure consumption.

This Nerdio webinar contains a section that delves into details of Microsoft's incentives, rebates and discounts:



Everything Azure for MSP's

Reach out to your favorite IT distributor and ask them to become a CSP Reseller or visit <https://partner.microsoft.com/en-US/cloud-solution-provider/indirect> for more detailed information.

Lever #2: Azure Reservations (aka Reserved Instances)

Margin impact: Increase of 20%–57% margin on Azure Compute consumption

The cost of Virtual Machines in Azure is the single biggest component of a typical MSP's IT environment. Therefore, focusing on reducing this large consumption component is where we'll start. The savings are significant but carry a bit of complexity and need for up-front planning to take advantage of them.

Microsoft's hyper-scale data center strategy has allowed the company to deploy many global regions (54 as of this writing) at a great cost to the company. As Azure adoption continues to rapidly grow, Microsoft needs to forecast demand in the various regions, which is far from trivial since public cloud is primarily advertised as a pay-as-you-go utility where you could run a VM one day and turn it off the next. To help with this forecasting challenge and to reward customers who are willing to commit to a certain amount of compute capacity in a specific region for an extended period, Microsoft introduced **Reserved Instances (RIs)**.

RIs are reservations of a specific type of compute capacity (i.e. VM family/series) in a specific geographic location (i.e. Azure region) for a predefined period (12 or 36 months). Depending on the VM family, duration of the reservation, and region, these RIs can save you from 20% to 57% relative to the list Pay-As-You-Go (PAYG) price. The trade-off is that you have to pre-pay for the reservation in advance. This is great news for MSPs because typical IT workloads they deploy in Azure on behalf of their customers are persistent and customers are generally open to making 1 or 3-year commitments. No customer wants to move an IT environment from one cloud to another every month!

Let's explore how these RIs work.

**These RIs can
save you from
20% to 57%
relative to the list Pay-As-
You-Go (PAYG) price**

When you provision a VM in Azure, two billing meters start running: **base compute** and **Windows Server license**. The published PAYG rate includes both of these components and every plain, vanilla VM you power on will bill you for both. RIs stop the base compute meter.

RIs are purchased with a lump sum payment via the CSP program and are applied to your tenant or subscription. Any VM running inside of that subscription that “matches” the RI will have its base compute rate zeroed out on the next invoice. Remember that RIs are purchased on a per-VM-family, per-region basis. This means that it will only match to a VM or set of VMs if they are of the same family and in the same region as the RI.

With Instance Size Flexibility, Microsoft will automatically apply any reservations in the most advantageous way to reduce the bill – if the VMs are running in the same region and have the same family as the RI. For instance, a reservation for a D4sv3 VM will be applied to a D4sv3 VM or two D2sv3 VMs. Examples of VM families include Ds, B, Es, NV and others. Keep in mind that RIs are a billing concept. There is nothing that needs to be done on the VM itself to stop its base compute meter and utilize the reservation. Azure does that automatically upon issuance of the invoice.

With Instance Size Flexibility, Microsoft will automatically apply any reservations in the most advantageous way to reduce the bill

What if you need to change your reservation from one VM family to another or move your VMs to another Azure region?

No problem! RIs can be exchanged without any fees or penalties. Any unused portion of an RI will be applied as a credit towards the purchase of a new RI for a different family, in a different region, or both. For instance, imagine you bought a reservation for D4sv3 for 12 months and used it for 3 months and now need to upgrade this VM to E4sv3 to get more RAM. You'll get a credit for 75% of your purchase price of the D3sv3 VM towards the purchase of a new E4sv3 VM. Keep in mind that you can only buy RIs in 12- or 36-month increments so you're effectively extending your reservation by 3 months (15 months instead of the original 12).

What if you need to cancel a reservation?

This is also possible, but there is a cancellation fee. The cancellation fee is 12% of your purchase price. The unused portion of your reservation will be refunded to you minus the 12% cancellation fee. There are some limitations to this on an annual basis. For example, your cancellations cannot exceed \$50,000 in a year.

Let's stop and think about this for a minute from the perspective of an MSP. The worst-case scenario is a 12% cancellation fee on a reserved instance of a VM they may no longer need in the future. However, the savings is anywhere from 20% to 57%. Therefore, cash flow considerations aside (more on that later), it makes sense to reserve all VMs even if they may need to be exchanged (no fee) or cancelled later (12% fee).

What about an MSP who wants to re-assign a reservation from one customer to another?

Imagine you sign up a new customer who wants to be migrated to Azure. They sign up for a 3-year managed service agreement and you purchase 3-year RIs for their VMs to really maximize your margin. After 6 months, the customer gets acquired and the new parent company has a large IT department with their own Azure deployment, and they decide to terminate your managed services agreement. You're now stuck holding a bunch of 3-year (30 months now) reservations that you already paid for. What should you do?

Good news—you have multiple options. But first let's understand the concept of reservation scope. When you purchase and apply a reservation to your Azure tenant you have the option of setting its scope to either a **specific Azure subscription** or to **shared scope**. Subscription scope will limit the reservation to stop base compute meters for matching VMs inside a single subscription – the one you assigned it to. Shared scope, on the other hand, allows reservations to stop the base compute meters for any matching VM in any subscription within the same Azure account (aka tenant).

Therefore, when setting up your Azure practice you should consider having a single tenant for all of your hosted environments with a dedicated subscription for each of your customers. This way you'll have the maximum flexibility and isolation of each customer's billing (per subscription) but also be able to "pool" your reservations together at the tenant level (i.e. shared scope) so they can be maximally utilized.

If your customer cancels your managed service agreement, you have three options:

1.

If you have no other customers who can make use of Azure VMs you will be forced to cancel the reservations and pay the 12% early termination fee. However, remember that your savings should more than pay for the early termination fee even in this worst-case scenario.

2.

If you have other customers or are bringing in new customers who can make use of reserved Azure VMs, but who need different types of VMs or need to be in a different region, then you would exchange your reservations – at no charge – and extend them to 12 or 36 months.

3.

If you have other customers or are bringing in new customers who can make use of the reserved instances in the same region and same VM family, then there is nothing else for you to do. By setting the RIs to shared scope and having all your customers under one tenant with individual subscriptions, the RIs will just keep working for you and stopping the base compute meter on VMs.

What about cashflow?

You may be concerned that having to come up with 36 or even 12 months worth of Azure VM fees is a burden on your company. There are financing companies specializing in working with MSPs who will finance the purchase of RIs. This way you'll get the benefit of the RI discount but keep the cash outlay monthly. There is obviously going to be a financing fee associated with this, but with savings of up to 57% it's still worthwhile. Reach out to us if you'd like to be put in touch with such financing companies.

In summary, RIs or reservations are a significant lever to save up to 57% of compute costs, which is the single biggest cost component of an Azure IT environment, and dramatically increase your margins. They do require some advance planning, budgeting, and structuring of your Azure account the right way, but can significantly increase the profitability of your Azure practice. On top of the inherent savings you get with reservations, you may also get anywhere from a 2%-5% discount as a CSP Direct or CSP Reseller. As you can see, discounts start to stack up and free up margin to be used in better ways.

This Nerdio webinar on the topic delves into details of RIs:



Demystifying Azure costs...

Lever #3: Azure Hybrid Usage and CSP Software Subscriptions

Margin impact: Increase of 10% to 49% margin on Azure Compute consumption

People often look at the monthly cost of an Azure VM, multiply it by 36 to get its 3-year cost and conclude that it is much more expensive than buying a server. What most don't realize is that the Pay As You Go list price of a VM includes in it the "premium of optionality" (i.e. you can turn it off at any time and stop paying) and a subscription to the Windows Server operating system. By reserving the base compute and bringing your own license for the operating system, the cost of the VM drops by 80%!

Microsoft has created a special entitlement called **Azure Hybrid Usage (AHU)** that allows MSPs to pay for Windows Server via another licensing program and not through Azure. AHU is a benefit unique to Azure. You can't bring your own Windows server license to AWS or GCP, for example. The cost of a VM in other public cloud will include the cost of the license, even if you already own Windows.

AHU can be enabled on a per-VM basis (or on all VMs using an automation platform like Nerdio) in the Azure portal. Turning it on stops the Windows OS billing meter and the cost of the VM becomes lower. How much lower? That depends on the VM family but can be up to half of the VM's cost. Let's explore this in detail.

In reality, by reserving the base compute and bringing your own license for the operating system, the cost of the VM drops by 80%!

Remember the two billing meters that start running whenever a VM is powered on are base compute and Windows OS. Reservations stop the base compute meter and AHU stops the Windows OS meter. So, if you have both an RI for a VM and AHU turned on, the monthly cost of this VM is \$0. This doesn't mean that it's "free" since you paid for the reservation and Windows

license up front. But with the monthly cost being \$0, the upfront cost of the reservation and Windows license can be as little as 20% of the total PAYG cost of that VM when spread out over 36 months.

To be entitled to turn AHU on for a VM you need to have the right kind of Windows Server license that has this benefit built-in. The first option is Windows Server licensed via EA or another volume license agreement with active Software Assurance. This is an uncommon scenario for MSPs so we won't discuss it here.

The second, and more relevant option, is purchasing the Windows Server license via the CSP program. This is known as CSP Software Subscriptions. You can also buy SQL Server and RDS licenses through this program. Let's see how it works.

A CSP Software Subscription can be purchased via the Microsoft Partner Center (for Direct CSPs) or via a CSP Provider.

Since each license "pack" is for 8 CPU cores, you need a single license to cover any VM with up to 8 CPU cores (e.g. DS1v2, D2sv3, D8sv3). If a VM has more than 8 CPU cores, then you need 2 or more license packs. So, a 16 core VM will need 2 license packs and a 20 core VM will need 3 license packs. Always round up to the closest multiple of 8. Pricing automation tools, like Nerdio's Cost Estimator, do the math for you so you know exactly how many cores you need to license.

When should you use AHU licensed with CSP Software Subscriptions? This depends on the

Here are some important characteristics of CSP Software Subscriptions:

- **License:** Windows Server Standard Core (this covers every VM in Azure even if the VM is running Windows Server Datacenter)
- **Term:** 1 year or 3 years (no month-to-month option)
- **# of CPU cores per license:** 8 (this is important, we'll see how it applies shortly)
- **List price (USD)**
- **1-year subscription:** \$291 (\$24.25/month)
- **3-year subscription:** \$582 (\$16.17/month)
- **CSP Discounts:** up to 15% depending on your CSP Provider and volume
- **Cancellation:** full refund within 60 days—no refund after that

VM family and what percentage of the time the VM is running. As an example, if you run a D8sv3 VM for 790 hours (just over a month) then it's cheaper to buy a 1-year Subscription and use AHU rather than paying for Windows Server via Azure billing meter. If you run the VM for 12 months on a 1-year subscription, you'll save \$2,937 as compared to not using Software Subscriptions. Here are some more specific examples:

VM Instance	PAYG Cost (monthly)	Base Compute only (AHU)	Cost of Windows OS (per VM per month)	3-year Software Subscription Cost (amortized monthly)	Monthly Savings (with 3 year sub)	1-year Software Subscription Cost (amortized monthly)	Monthly Savings (with 1 year sub)
Dsv3-Series							
D2sv3 (2C/8GB)	\$152.57	\$85.41	\$67.16	\$16.17	\$50.99	\$24.25	\$42.91
D4sv3 (4C/16GB)	\$305.14	\$170.82	\$134.32	\$16.17	\$118.15	\$24.25	\$110.07
D8sv3 (8C/32GB)	\$610.28	\$341.64	\$268.64	\$16.17	\$252.47	\$24.25	\$244.39
D16sv3 (16C/64GB)	\$1,220.56	\$683.28	\$537.28	\$32.33	\$504.95	\$48.50	\$488.78
D32sv3 (32C/128GB)	\$2,441.12	\$1,366.56	\$1,074.56	\$64.67	\$1,009.89	\$97.00	\$977.56
D64sv3 (64C/256GB)	\$4,882.24	\$2,733.12	\$2,149.12	\$129.33	\$2,019.79	\$194.00	\$1,955.12
B-Series							
B1s (1C/1GB)	\$11.97	\$9.05	\$2.92	\$16.17	(\$13.25)	\$24.25	(\$21.33)
B1ms (1C/2GB)	\$21.02	\$18.10	\$2.92	\$16.17	(\$13.25)	\$24.25	(\$21.33)
B2s (2C/4GB)	\$42.05	\$36.21	\$5.84	\$16.17	(\$10.33)	\$24.25	(\$18.41)
B2ms (2C/8GB)	\$78.11	\$72.42	\$5.69	\$16.17	(\$10.48)	\$24.25	(\$18.56)
B4ms (4C/16GB)	\$156.22	\$144.54	\$11.68	\$16.17	(\$4.49)	\$24.25	(\$12.57)
B8ms (8C/32GB)	\$313.17	\$289.81	\$23.36	\$16.17	\$7.19	\$24.25	(\$0.89)
Esv3-Series							
E2sv3 (2C/16GB)	\$175.20	\$108.04	\$67.16	\$16.17	\$50.99	\$24.25	\$42.91
E4sv3 (4C/32GB)	\$350.40	\$216.08	\$134.32	\$16.17	\$118.15	\$24.25	\$110.07
E8sv3 (8C/64GB)	\$701.53	\$432.89	\$268.64	\$16.17	\$252.47	\$24.25	\$244.39
E16sv3 (16C/128GB)	\$1,403.06	\$865.78	\$537.28	\$32.33	\$504.95	\$48.50	\$488.78
E32sv3 (32C/256GB)	\$2,805.39	\$1,730.83	\$1,074.56	\$64.67	\$1,009.89	\$97.00	\$977.56
E64sv3 (64C/432GB)	\$5,092.48	\$2,943.36	\$2,149.12	\$129.33	\$2,019.79	\$194.00	\$1,955.12

As you can see, it doesn't always save you money to use AHU. Specifically, in the case of B-series VMs. However, the savings are quite significant with other VMs types like Dsv3 and Esv3.

In summary, combining RIs with AHU and CSP Software subscriptions can reduce the cost of your VMs up to 5X (80%). It goes without saying that the margin impact to an MSP from such significant cost reductions cannot be overlooked. As with RIs, some planning, budgeting, and cash flow considerations must be evaluated. However, if building a scalable Azure practice is your goal then it is worthwhile to check out AHU and Software Subscriptions.

This Nerdio partner webinar for MSPs delves into details of AHU and Software Subscriptions:



Discussing Azure Hybrid Usage...

Lever #4: Auto-scaling for cost optimization

Margin impact: Increase of up to 72% margin on Azure Compute and OS storage consumption

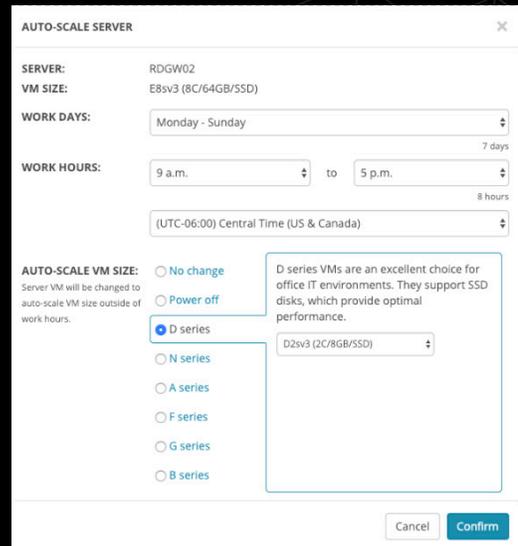
The value proposition of the public cloud is its utility-like consumption billing model. “Pay only for what you use”. This is great in concept but taking advantage of this aspect of Azure is not as easy as it sounds. To pay for only what you use, you need a mechanism to know what you need when and a system that automatically resizes your Azure workloads to fit the demand at any given time. This means that if a VM doesn't need to be on, a system needs to be in place to know it and act on it by shutting down the VM at the appropriate time and then turn it back on when it's needed again.

Newly designed cloud-native applications build in special logic to optimize the consumption of Azure and leverage things like Azure VM Scale Sets that can dynamically grow (scale out) and shrink (scale in) based on demand. Unfortunately, most applications that MSPs deal with for their customers (e.g. Active Directory, database servers, Remote Desktops, etc.) are not natively equipped to be automatically scaled.

Imagine you have an application server that requires 8 CPU cores to support the user load. You select an E8sv3 VM to run the application since it has 8 CPU cores and 64GB of RAM. The cost of this VM is \$0.872/hour or \$636.56/month. However, the application is only used during regular business hours of Monday-Friday 9am to 5pm. That's 40 hours out of a total of 168 hours in a week - only 23.8% of the time. **If you had a mechanism to automatically shut the VM down outside of business hours and turn it back on before business hours, you'd save up to 76.2% of the VM cost! Not at all insignificant.**

Now what if you can't have the application be completely unavailable outside of business hours and on the weekend but need it to be available to a smaller number of users who may need to perform a task or have access to it at night or on the weekend? Wouldn't it be great if the VM could automatically be resized from its baseline size of E8sv3 at 5pm to D2sv3, which is 21.6% of the cost? Unfortunately, there is no native Azure functionality that can do this for you.

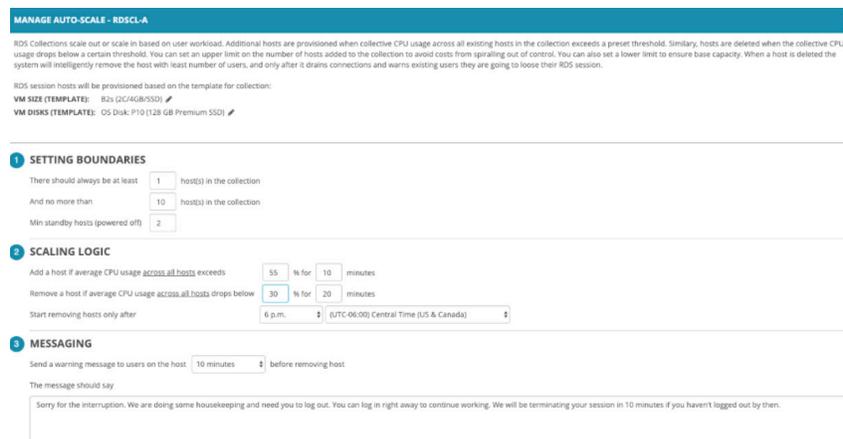
Azure automation platforms – like [Nerdio for Azure](#) – can help. You can set business hours for each VM and tell the system what to do with the VM outside of those hours: leave it alone, shut it down, or change it to something smaller. The system will then automatically execute these instructions—resizing the VM after the end of business hours and prior to the start of business hours. This leads to significant savings that may even exceed the savings provided by RIs and AHU.



What about Remote Desktop Services (RDS)?

Imagine you have an RDS Collection that is supporting 1,000 users. During peak hours, when everyone is on their desktop, there are 10 E32sv3 VMs being utilized. Outside of peak hours there may be up to 100 users accessing their remote desktop at any one time. Wouldn't it be great if there could be just one VM running all the time and up to 10 VMs would get automatically built and powered up when they are needed?

Nerdio can help here as well. Instead of resizing or shutting down individual VMs, Nerdio's RDS Collection auto-scaling is integrated with Azure VM Scale Sets. This allows the RDS collection to grow as CPU utilization grows and shrink after business hours as CPU utilization declines as users log off. Here is an example:



The advantage of Auto-Scaling as compared to Reserved Instances and AHU lies in the fact that there are no commitments and upfront purchases to be made. You use Azure in a pay-as-you-go manner and the auto-scaling keeps things efficient.

Lever #5: Burstable VM instances

Margin impact: Increase of up to 50% margin on Azure Compute

B-series VMs 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 they 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, the B2ms quota is 60% of a single CPU, which is 30% of the 2 CPUs visible in the VM. Every second that the VM is using less than its quota (less than 60% of a single CPU, in our example) it is “banking credits”. These banked credits can be used to burst up to the total available CPUs (100% of 2 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 should an MSP use B-series VMs?

They are cheaper. For approximately the same price 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. However, when Windows Updates run, the VM will use all its available CPU horsepower. B-series VMs 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.

Here are some important characteristics of Burstable or B-Series VMs:

- **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

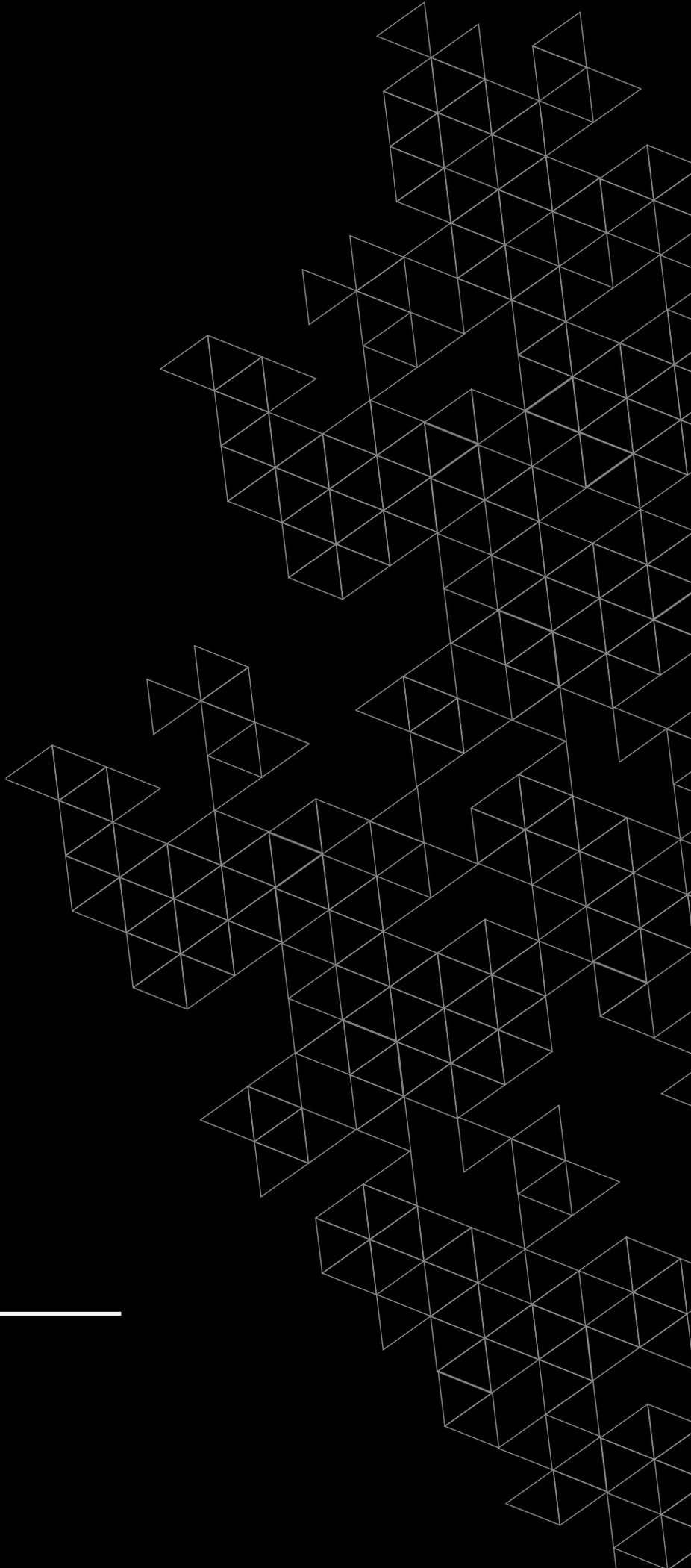
Tips

- 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.

As you can see there are multiple levers that can help MSPs optimize their Azure consumption and make IT environments running in Azure be more affordable than any other alternative. Not to mention the advantage of hyper-scalability, high security, flexibility, and Microsoft's solid reputation. Understanding these levers will help MSPs start to model what their Azure offers could look like, how much margin they can achieve, and how to build a successful and profitable cloud practice in Azure.

Tools like Nerdio for Azure can help automate much of this complex manual work with its pricing and packaging automation, automated deployment engine, single-pane-of-glass management portal, and cost optimizing auto-scale technology.

Check us out at www.getnerdio.com



Contact Us:

Phone: 1-844-463-7346

Email: sayhello@getnerdio.com

Website: getnerdio.com