10分で理解するMicrosoft Ignite 2024 Azure Virtual Machine、Virtual Machine Scale Sets アップデート情報

2024/12/06 第5回 JAZUG for Women

Kazuki Yamabe

アジェンダ

- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

アジェンダ

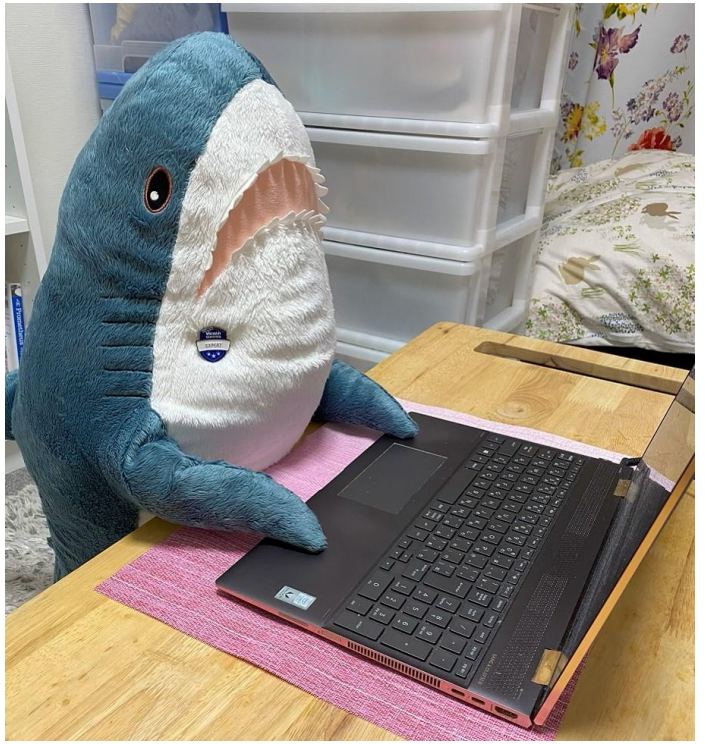
- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

自己紹介

名前：Kazuki Yamabe

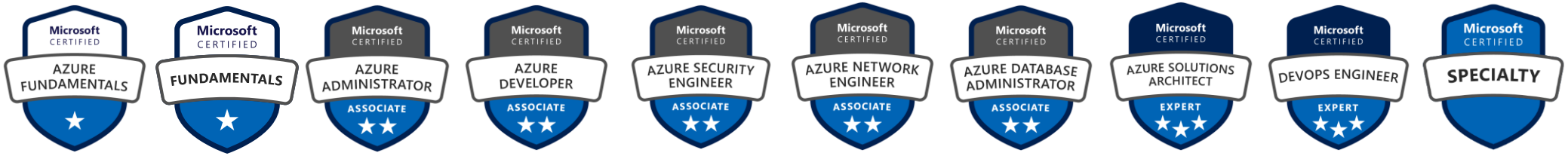
職種：インフラエンジニア

得意分野：Azure、Microsoft 製品、サーバーを含めたインフラ全般、
IaC、CI/CD、など



■ ブログ・SNS

- ブログ：https://www.kdkwakaba.com/
- X：@kdk_wakaba
- Linkedin：kdk-wakaba



注意事項

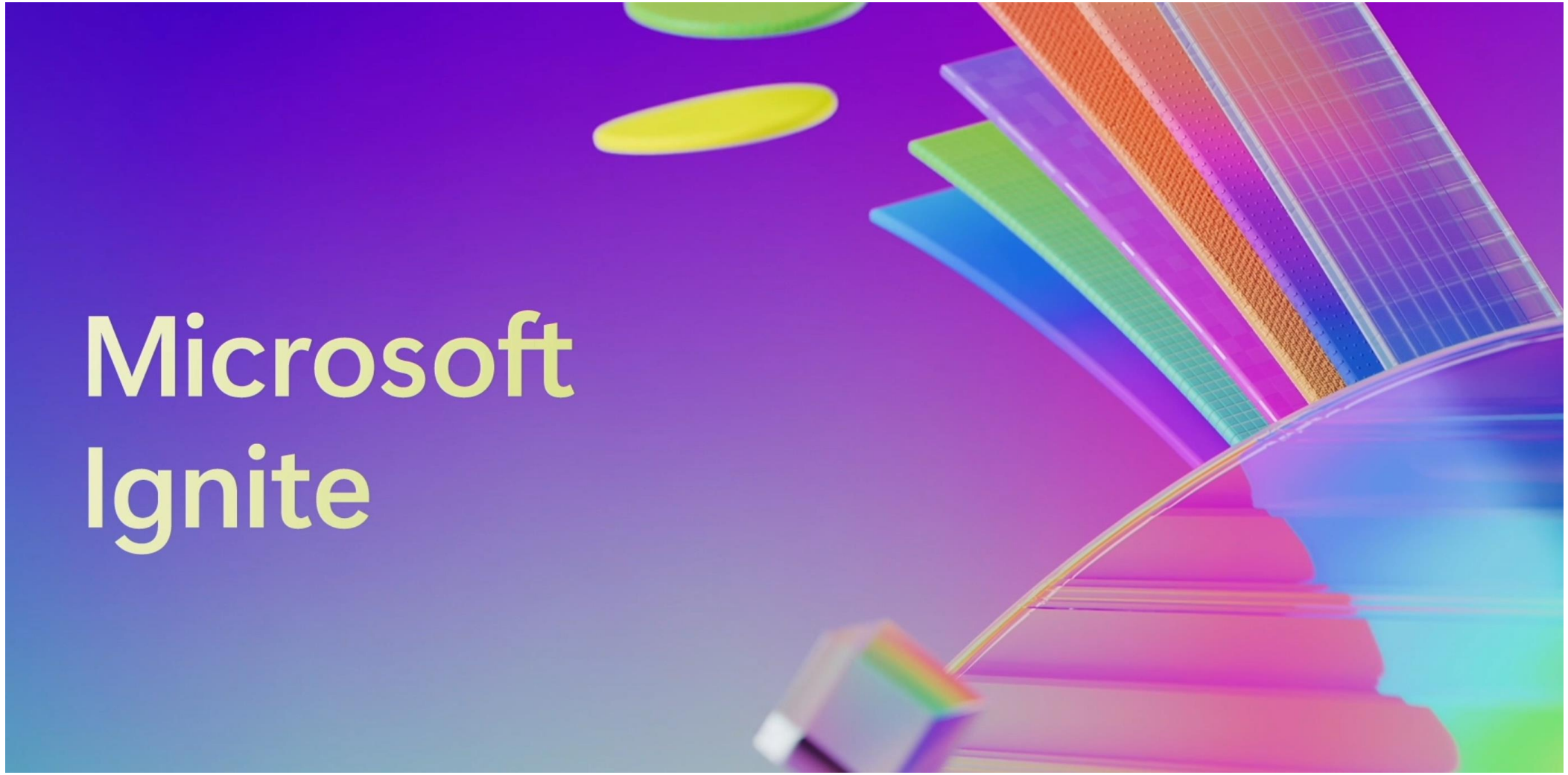
- 本内容はパブリックプレビュー、プライベートプレビューの内容が含まれています。今後のアップデートにより仕様変更となる可能性があるためご留意ください
- 発表者が見逃しているアップデート情報がありましたら、やんわりと指摘いただければ幸いです

アジェンダ

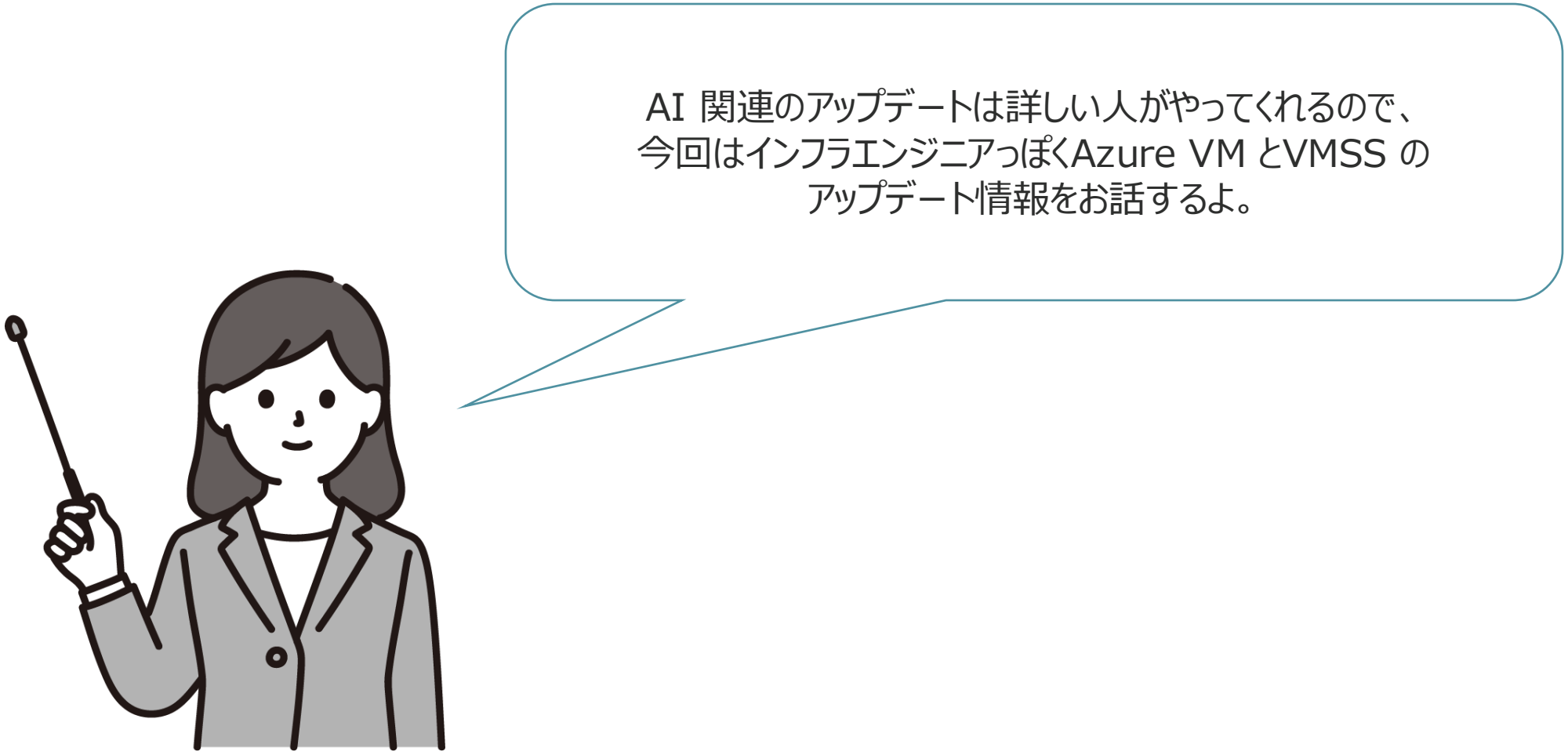
- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

Microsoft Ignite 2024

みなさん、Microsoft Ignite 2024 に参加しましたか？様々なアップデート情報、セッションがありましたね。



Microsoft Ignite 2024



アジェンダ

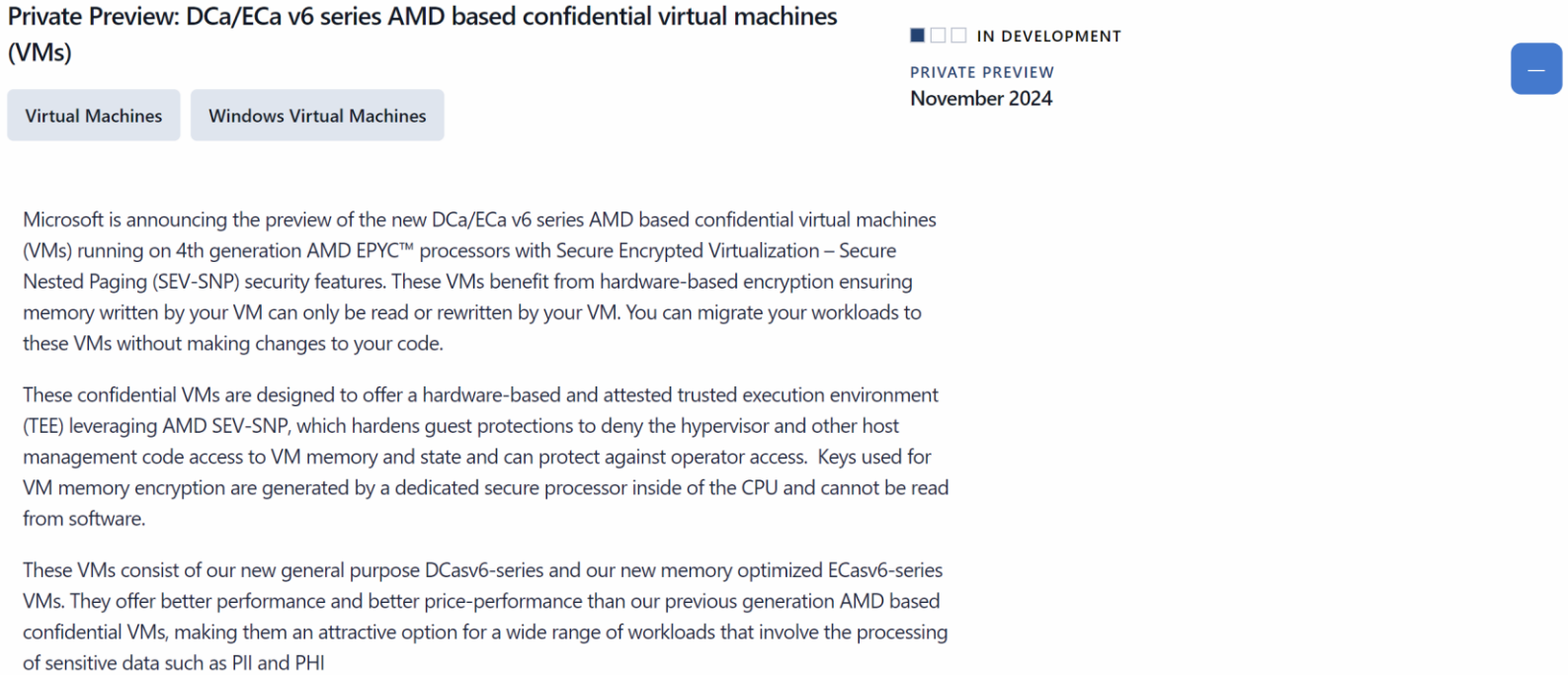
- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

Azure VM Update

- AMD EPYC CPU 搭載のConfidential Virtual Machine (CVM) (Private Preview)
- Compute Fleet のマルチリージョンデプロイ (Public Preview)
- Compute Fleet の属性ベースのVM (Public Preview)
- (番外編) Windows Server 2025 (Generally Available)

Azure VM Update - AMD EPYC CPU 搭載のConfidential Virtual Machine (CVM)

- 機密データやセキュリティ要件の厳しいワークロードでARM ベースのAzure VM が利用可能
  - AES-256によるメモリ暗号化がデフォルト有効化
  - Virtualization-based Security (VBS)を利用したキー保護機能の追加
- 汎用のDCasv6/DCadsv6、メモリ最適化のECasv6/ECadsv6 が追加



Private Preview: DCa/ECa v6 series AMD based confidential virtual machines (VMs)

IN DEVELOPMENT

PRIVATE PREVIEW

November 2024

Virtual Machines | Windows Virtual Machines

Microsoft is announcing the preview of the new DCa/ECa v6 series AMD based confidential virtual machines (VMs) running on 4th generation AMD EPYC™ processors with Secure Encrypted Virtualization – Secure Nested Paging (SEV-SNP) security features. These VMs benefit from hardware-based encryption ensuring memory written by your VM can only be read or rewritten by your VM. You can migrate your workloads to these VMs without making changes to your code.

These confidential VMs are designed to offer a hardware-based and attested trusted execution environment (TEE) leveraging AMD SEV-SNP, which hardens guest protections to deny the hypervisor and other host management code access to VM memory and state and can protect against operator access. Keys used for VM memory encryption are generated by a dedicated secure processor inside of the CPU and cannot be read from software.

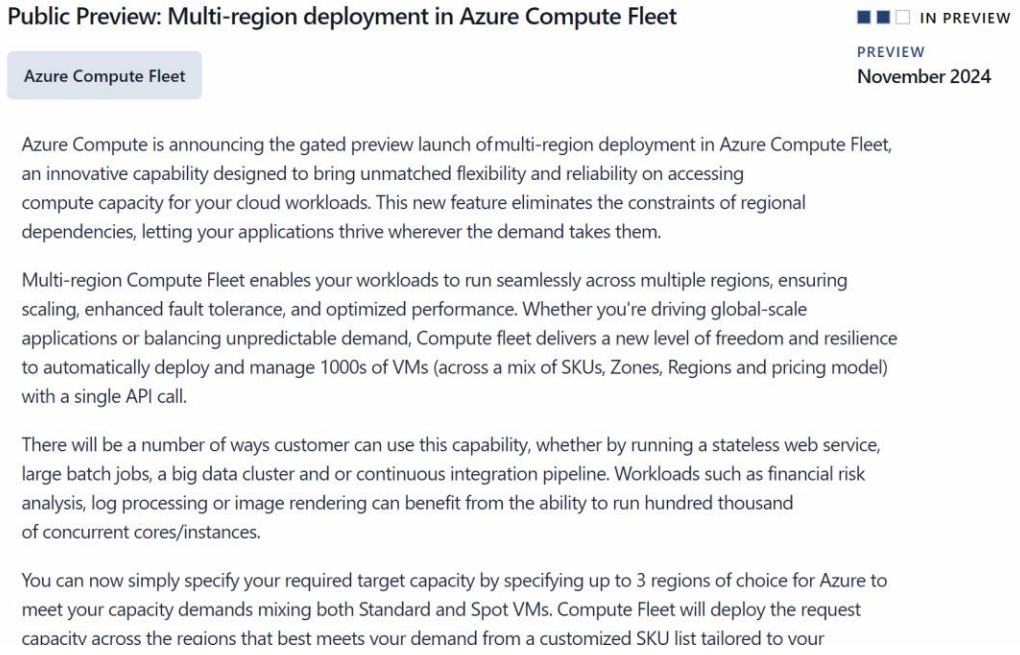
These VMs consist of our new general purpose DCasv6-series and our new memory optimized ECasv6-series VMs. They offer better performance and better price-performance than our previous generation AMD based confidential VMs, making them an attractive option for a wide range of workloads that involve the processing of sensitive data such as PII and PHI

Private Preview

Update page： DCa/ECa v6 シリーズ AMD ベースの機密仮想マシン (VM)

Azure VM Update - Compute Fleet のマルチリージョンデプロイ

- Spot VMs やStandard VMs で最大3つのデプロイリージョンを指定できる機能
  - このアップデートまで複数リージョン構成には複数のCompute Fleet が必要だった
  - リージョン障害が発生したら完全停止となるが複数Compute Fleet を作るのもコストがかさむ…
- 複数のリージョン指定により可用性向上、費用対効果の高いスケーリングが可能



Public Preview: Multi-region deployment in Azure Compute Fleet

IN PREVIEW

Azure Compute Fleet

PREVIEW
November 2024

Azure Compute is announcing the gated preview launch of multi-region deployment in Azure Compute Fleet, an innovative capability designed to bring unmatched flexibility and reliability on accessing compute capacity for your cloud workloads. This new feature eliminates the constraints of regional dependencies, letting your applications thrive wherever the demand takes them.

Multi-region Compute Fleet enables your workloads to run seamlessly across multiple regions, ensuring scaling, enhanced fault tolerance, and optimized performance. Whether you're driving global-scale applications or balancing unpredictable demand, Compute fleet delivers a new level of freedom and resilience to automatically deploy and manage 1000s of VMs (across a mix of SKUs, Zones, Regions and pricing model) with a single API call.

There will be a number of ways customer can use this capability, whether by running a stateless web service, large batch jobs, a big data cluster and or continuous integration pipeline. Workloads such as financial risk analysis, log processing or image rendering can benefit from the ability to run hundred thousand of concurrent cores/instances.

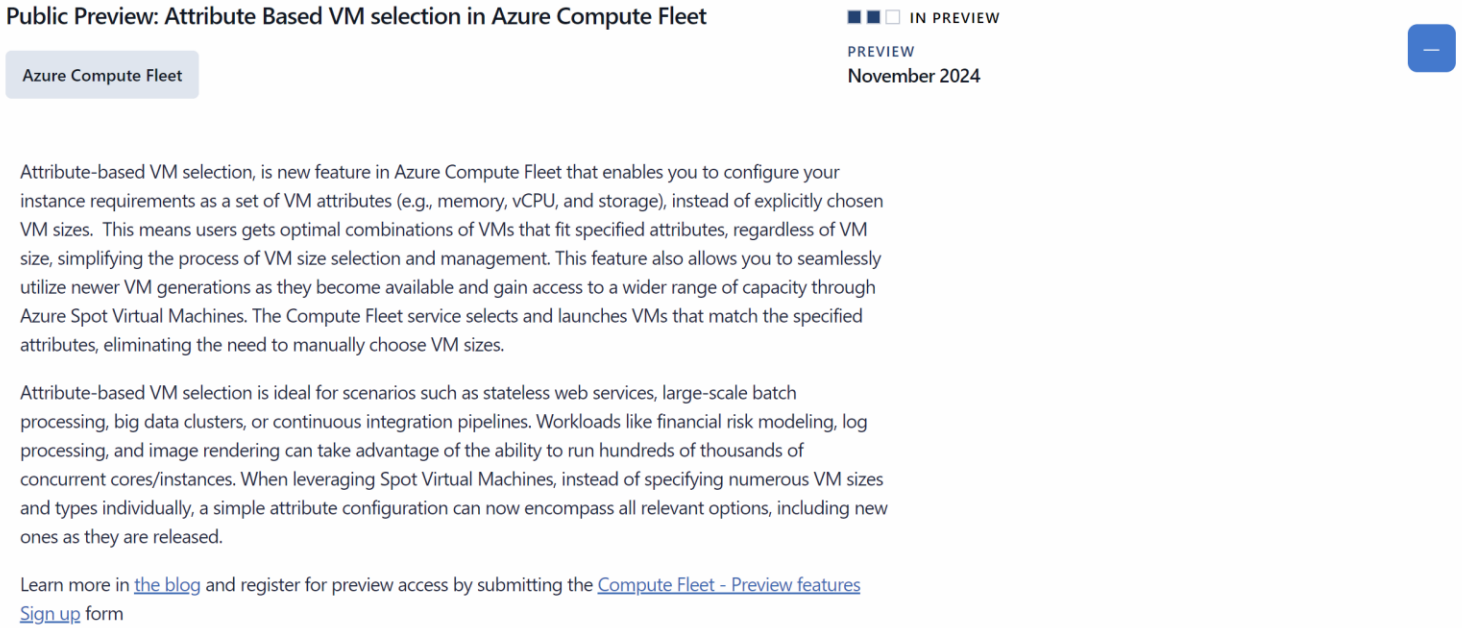
You can now simply specify your required target capacity by specifying up to 3 regions of choice for Azure to meet your capacity demands mixing both Standard and Spot VMs. Compute Fleet will deploy the request capacity across the regions that best meets your demand from a customized SKU list tailored to your

Public Preview

Update page： [Azure Compute Fleet でのマルチリージョンデプロイ]

Azure VM Update - Compute Fleet の属性ベースのVM

- メモリ、vCPU、ストレージなど設定した属性のVM サイズを自動的に選択し起動する機能
  - このアップデートまではVM サイズを1つずつ確認し選択する必要があった
  - ワークロードに合わせたVM サイズを判断できる人の確保、構築時の工数が課題だった
- Spot VMs は容量最適化と最低価格、Standard VMs は最低価格の割り当て戦略を設定可能



Public Preview: Attribute Based VM selection in Azure Compute Fleet

Azure Compute Fleet

IN PREVIEW

PREVIEW

November 2024

Attribute-based VM selection, is new feature in Azure Compute Fleet that enables you to configure your instance requirements as a set of VM attributes (e.g., memory, vCPU, and storage), instead of explicitly chosen VM sizes. This means users gets optimal combinations of VMs that fit specified attributes, regardless of VM size, simplifying the process of VM size selection and management. This feature also allows you to seamlessly utilize newer VM generations as they become available and gain access to a wider range of capacity through Azure Spot Virtual Machines. The Compute Fleet service selects and launches VMs that match the specified attributes, eliminating the need to manually choose VM sizes.

Attribute-based VM selection is ideal for scenarios such as stateless web services, large-scale batch processing, big data clusters, or continuous integration pipelines. Workloads like financial risk modeling, log processing, and image rendering can take advantage of the ability to run hundreds of thousands of concurrent cores/instances. When leveraging Spot Virtual Machines, instead of specifying numerous VM sizes and types individually, a simple attribute configuration can now encompass all relevant options, including new ones as they are released.

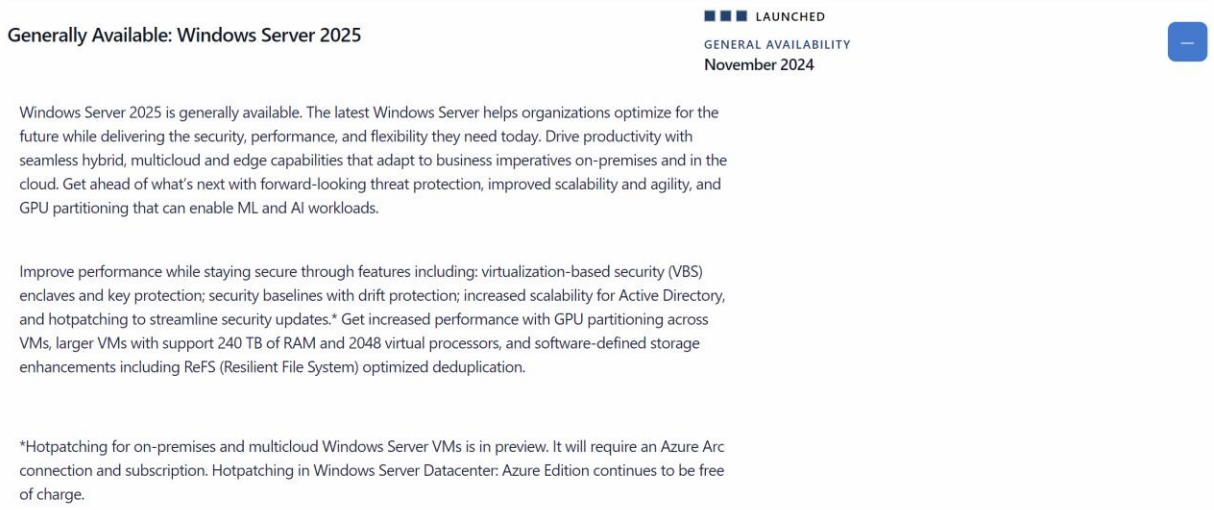
Learn more in the blog and register for preview access by submitting the Compute Fleet - Preview features Sign up form

Public Preview

Update page： Compute Fleet の属性ベースのVM の選択

Azure VM Update – Windows Server 2025 (番外編)

- 2024年11月1日にWindows Server 2025 がGA
- Active Directory、SBM などインフラ領域におけるセキュリティの強化
- AI や機械学習に関するGPU パーティショニングやHyper-V のパフォーマンス向上
- Azure Arc を利用したオンプレミス、マルチクラウドのHotpatch 機能のサポート
  - 再起動の数を減らせるため、メンテナンスコストを削減できる



Generally Available: Windows Server 2025

LAUNCHED

GENERAL AVAILABILITY
November 2024

Windows Server 2025 is generally available. The latest Windows Server helps organizations optimize for the future while delivering the security, performance, and flexibility they need today. Drive productivity with seamless hybrid, multicloud and edge capabilities that adapt to business imperatives on-premises and in the cloud. Get ahead of what's next with forward-looking threat protection, improved scalability and agility, and GPU partitioning that can enable ML and AI workloads.

Improve performance while staying secure through features including: virtualization-based security (VBS) enclaves and key protection; security baselines with drift protection; increased scalability for Active Directory, and hotpatching to streamline security updates.* Get increased performance with GPU partitioning across VMs, larger VMs with support 240 TB of RAM and 2048 virtual processors, and software-defined storage enhancements including ReFS (Resilient File System) optimized deduplication.

*Hotpatching for on-premises and multicloud Windows Server VMs is in preview. It will require an Azure Arc connection and subscription. Hotpatching in Windows Server Datacenter: Azure Edition continues to be free of charge.

個人的な注目機能

Generally Available

Update page： Windows Server 2025

アジェンダ

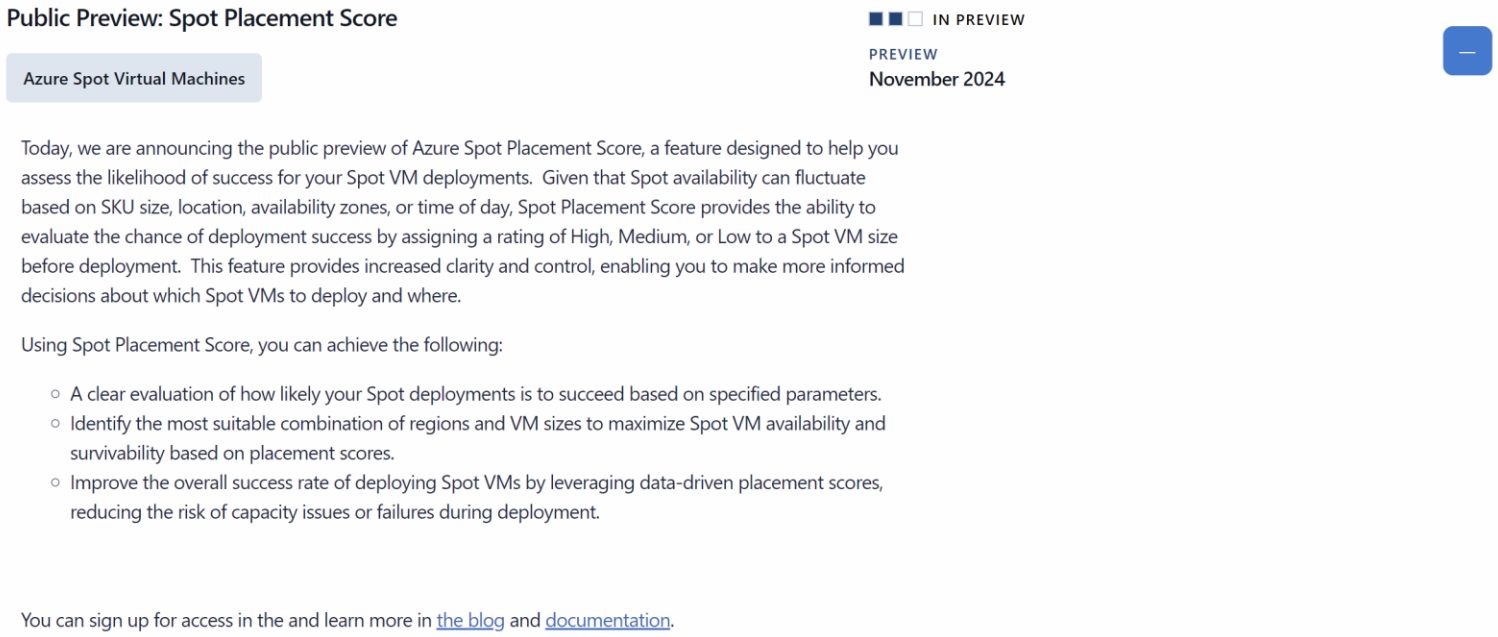
- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

Azure VMSS Update

- スポット配置スコア (Public Preview)
- VMSS アップグレードポリシー (Generally Available)
- VMSS MaxSurge アップグレード (Generally Available)
- VMSS ローリングアップグレードのカスタムメトリック (Public Preview)
- VMSS 回復力のある作成と削除 (Public Preview)
- VMSS ゾーン拡張 (Generally Available)

Azure VMSS Update – スポット配置スコア

- VMSS のスポットインスタンスにおけるデプロイ成功率を高、中、低でスコアリングする機能
  - どのSKU を選べばよいか、安定しているリージョンはどこか、など参考にできるためデプロイ成功率を向上できる
  - スポットインスタンス作成したのにすぐ終了した、みたいな事象を減らせる
- 表示されるのは現時点のスコアのみのため、定期的なスコア収集や変動時のアクション実行はひと手間必要



Public Preview: Spot Placement Score

Azure Spot Virtual Machines

IN PREVIEW

PREVIEW

November 2024

Today, we are announcing the public preview of Azure Spot Placement Score, a feature designed to help you assess the likelihood of success for your Spot VM deployments. Given that Spot availability can fluctuate based on SKU size, location, availability zones, or time of day, Spot Placement Score provides the ability to evaluate the chance of deployment success by assigning a rating of High, Medium, or Low to a Spot VM size before deployment. This feature provides increased clarity and control, enabling you to make more informed decisions about which Spot VMs to deploy and where.

Using Spot Placement Score, you can achieve the following:

- A clear evaluation of how likely your Spot deployments is to succeed based on specified parameters.
- Identify the most suitable combination of regions and VM sizes to maximize Spot VM availability and survivability based on placement scores.
- Improve the overall success rate of deploying Spot VMs by leveraging data-driven placement scores, reducing the risk of capacity issues or failures during deployment.

You can sign up for access in the and learn more in the blog and documentation.

Public Preview

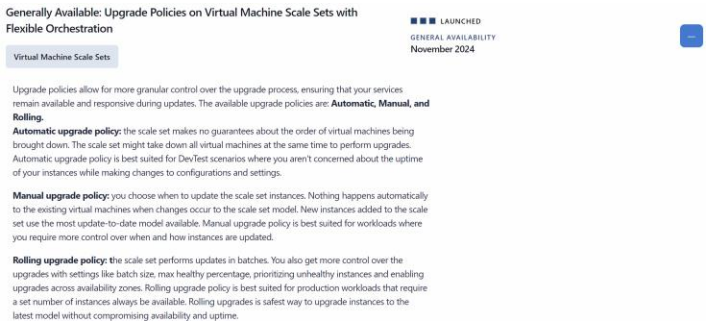
Update page： スポット配置スコア

Azure VMSS Update – VMSS アップグレードポリシー

- VMSS 上のインスタンスのアップグレード方法を設定する機能

項目	説明
手動	ユーザー側でインスタンスの手動アップグレードを行う。更新タイミングを詳細に制御したい場合に適している。デフォルトの設定。
自動	Azure 側でインスタンス停止タイミングを判断し自動アップグレードを行う。開発/検証、可用性を気にしない環境に適している。
ローリング	アップグレードの割合、間隔などの設定に合わせて自動バッチ処理のアップグレードを行う。インプレースとMaxSurge を設定できる。一定数のインスタンスを起動させておきたいケースに適している。

- 特定のインスタンスのみ更新タイミングを分けたい場合は、ローリングアップデートのカスタムメトリックが必要


Generally Available: Upgrade Policies on Virtual Machine Scale Sets with Flexible Orchestration

Virtual Machine Scale Sets

LAUNCHED
GENERAL AVAILABILITY
November 2024

Upgrade policies allow for more granular control over the upgrade process, ensuring that your services remain available and responsive during updates. The available upgrade policies are: **Automatic, Manual, and Rolling.**

Automatic upgrade policy: the scale set makes no guarantees about the order of virtual machines being brought down. The scale set might take down all virtual machines at the same time to perform upgrades. Automatic upgrade policy is best suited for DevTest scenarios where you aren't concerned about the uptime of your instances while making changes to configurations and settings.

Manual upgrade policy: you choose when to update the scale set instances. Nothing happens automatically to the existing virtual machines when changes occur to the scale set model. New instances added to the scale set use the most update-to-date model available. Manual upgrade policy is best suited for workloads where you require more control over when and how instances are updated.

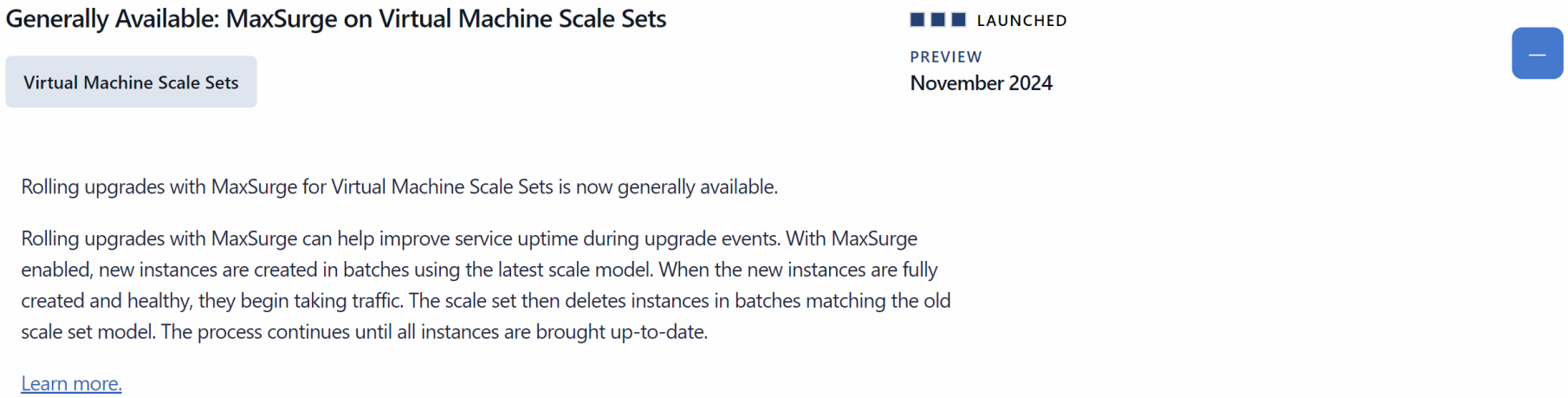
Rolling upgrade policy: the scale set performs updates in batches. You also get more control over the upgrades with settings like batch size, max healthy percentage, prioritizing unhealthy instances and enabling upgrades across availability zones. Rolling upgrade policy is best suited for production workloads that require a set number of instances always be available. Rolling upgrades is safest way to upgrade instances to the latest model without compromising availability and uptime.

Generally Available

Update page : 柔軟なオーケストレーションを備えたVMSS のアップグレードポリシー

Azure VMSS Update – VMSS MaxSurge アップグレード

- アップグレードポリシーのローリングで更新後のVM 追加後に古いVM を削除するアップグレード方法
  - このアップデートまでローリングのポリシーはインプレースだけだった
  - アップグレード時にVM 数減少、一斉停止など可用性に課題があった
- アップグレード時に可能な限り可用性に影響を与えず更新可能
- MaxSurge アップグレードで追加されるVM もクォーターの対象のため注意が必要

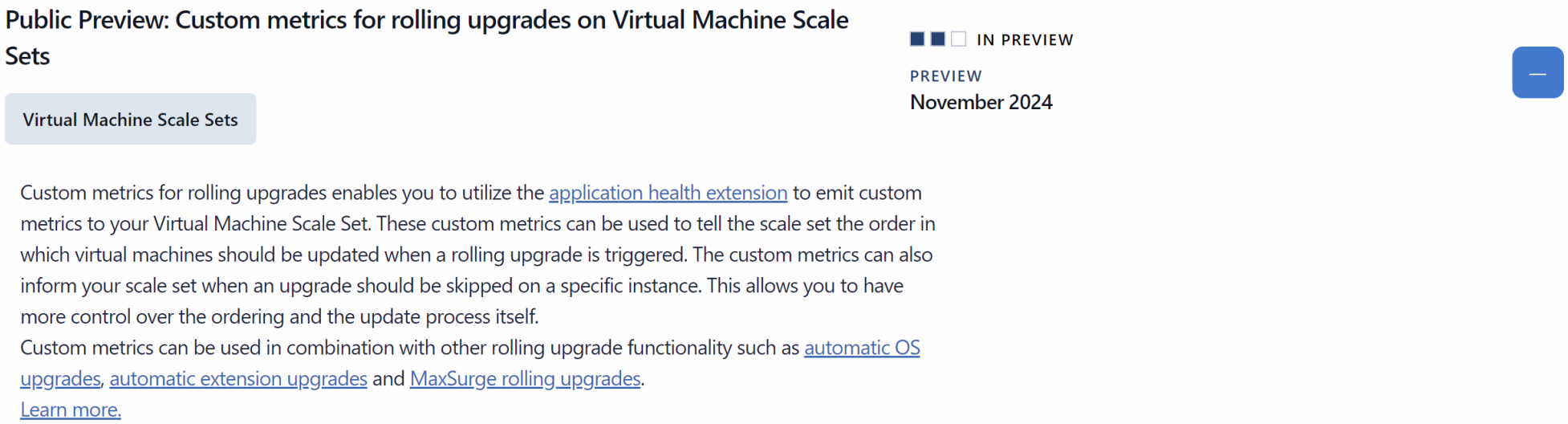


Generally Available

Update page： VMSS のMaxSurge

Azure VMSS Update – VMSS ローリングアップグレードのカスタムメトリック

- アプリケーションの正常性拡張機能を使い、VMSS のインスタンスにカスタムメトリックを出力させる機能
  - インスタンスのアップグレード順序を指定したい、 特定のインスタンスをスキップさせたい、などに使う
- 拡張機能のインストール後、VM 側でBash またはPowerShell で応答用スクリプトが必要
  - 公式サンプル： GitHub - Azure-Samples/application-health-samples

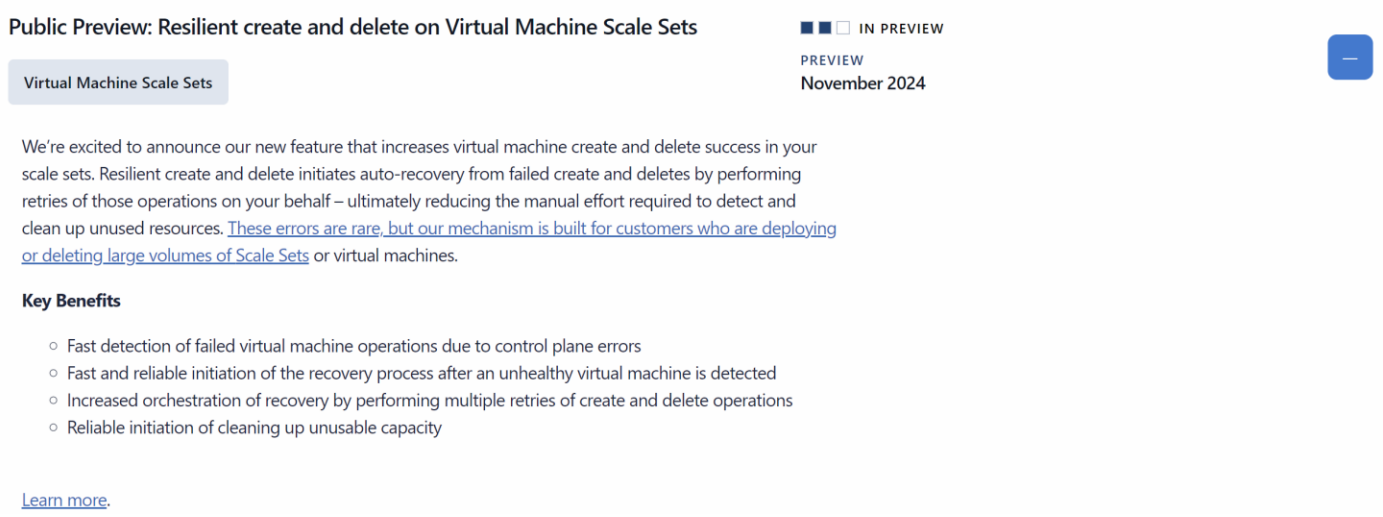


Public Preview

Update page： Virtual Machine Scale Sets のローリングアップグレードのカスタムメトリック

Azure VMSS Update – VMSS 回復力のある作成と削除

- VMSS の新規作成、スケールアウト時、作成や削除に失敗したVM を自動的に再試行する機能
  - 作成失敗のVM 手動削除、削除失敗のVM クリーンアップを自動化できる
- 回復力のある作成、削除で設定項目が分かれているため、片方のみ設定なども可能
- 失敗時の再試行回数は5回、作成失敗時のタイムアウトはOS によって変わる
  - Windows は20分、Linux は8分



Public Preview: Resilient create and delete on Virtual Machine Scale Sets

IN PREVIEW

Virtual Machine Scale Sets

PREVIEW
November 2024

We're excited to announce our new feature that increases virtual machine create and delete success in your scale sets. Resilient create and delete initiates auto-recovery from failed create and deletes by performing retries of those operations on your behalf – ultimately reducing the manual effort required to detect and clean up unused resources. These errors are rare, but our mechanism is built for customers who are deploying or deleting large volumes of Scale Sets or virtual machines.

Key Benefits

- Fast detection of failed virtual machine operations due to control plane errors
- Fast and reliable initiation of the recovery process after an unhealthy virtual machine is detected
- Increased orchestration of recovery by performing multiple retries of create and delete operations
- Reliable initiation of cleaning up unusable capacity

Learn more.

Public Preview

Update page： VMSS での回復力のある作成と削除

Azure VMSS Update – VMSS ゾーン拡張

- ゾーン拡張は作成済みのVMSS で可用性ゾーンの拡張、新規可用性ゾーンを追加できる機能
  - 従来は可用性ゾーン追加にVMSS の再作成が必要だった
- 拡張後のゾーン削除はサポートされていないため、拡張時の判断は慎重に行う必要がある

Generally Available: VMSS Zonal Expansion

Virtual Machine Scale Sets

LAUNCHED

GENERAL AVAILABILITY
November 2024

We are pleased to announce that the Azure VMSS Zonal Expansion feature is now generally available. This significant enhancement allows customers to transition their virtual machine scale sets from a regional to zonal resiliency strategy or add additional zones to a scale set. By distributing VMs across multiple zones, customers can significantly enhance their business continuity and resilience, achieving a higher availability Service Level Agreement (SLA) of 99.99% compared to the previous 99.95%.

The Azure VMSS Zonal Expansion feature simplifies the migration process, enabling customers to update their scale set environment to the Azure recommended zone-redundant availability strategy without the need to delete and recreate instances or experience application downtime. This means that businesses can now take full advantage of the improved availability and reliability offered by Azure availability zones, ensuring their applications remain highly available and resilient to potential disruptions.

For more information, visit the docs.

Generally Available

Update page： VMSS ゾーン拡張

アジェンダ

- 自己紹介
- Microsoft Ignite 2024
- Azure VM Update
- Azure VMSS Update
- まとめ
- 参考資料

まとめ

- 既存のAzure VM、VMSS も現在に合わせた機能追加や改善のアップデートが入っている
- Ignite のアップデートではIaaS の運用コスト、可用性向上に関連する内容が多い印象
- ARM ベースのVM 周辺のアップデートとAzure VM もAI に合わせてアップデートされている

参考資料

- Azure の機密 VMについて
- マルチリージョン Compute Fleet (プレビュー)
- Azure Compute Fleet の属性ベースの VM 選択 (プレビュー)
- Windows Server 2025 の新機能
- Spot Placement Score (Preview)
- Virtual Machine Scale Sets のアップグレード ポリシー モード
- Virtual Machine Scale Sets の MaxSurge でのローリング アップグレード
- Virtual Machine Scale Sets 上でローリング アップグレードのカスタム メトリックを構成する (プレビュー)
- Resilient create and delete for Virtual Machine Scale Sets (Preview)
- 可用性ゾーンを使用する仮想マシン スケール セットを作成する

ご清聴ありがとうございました。