

vFORUM **2019**

HC122

あなたが知らない
VMware Cloud on AWS
(Technical Deep Dive)

ヴェイエムウェア株式会社
ソリューションビジネス本部
クラウドソリューション技術統括部
シニアクラウドスペシャリスト 薛 昊

vmware®

©2019 VMware, Inc.

Make
Your
Mark

免責事項

- このセッションには、現在開発中の製品/サービスの機能が含まれている場合があります。
- 新しいテクノロジーに関するこのセッションおよび概要は、VMware が市販の製品/サービスにこれらの機能を搭載することを約束するものではありません。
- 機能は変更される場合があるため、いかなる種類の契約書、受注書、または販売契約書に記述してはなりません。
- 技術的な問題および市場の需要により、最終的に出荷される製品/サービスでは機能が変わる場合があります。
- ここで検討されているまたは提示されている新しいテクノロジーまたは機能の価格およびパッケージは、決定されたものではありません。

Agenda

概要

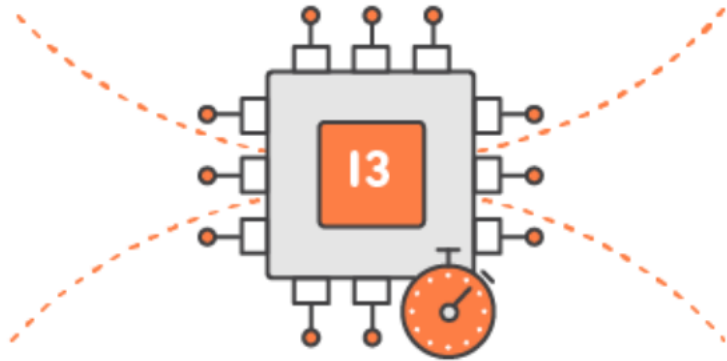
Compute/Availability

Storage

Network

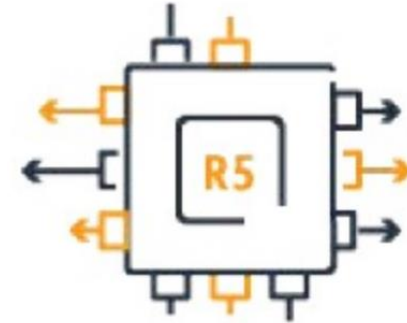
概要

ベアメタルスペック概要



AWS i3.metal

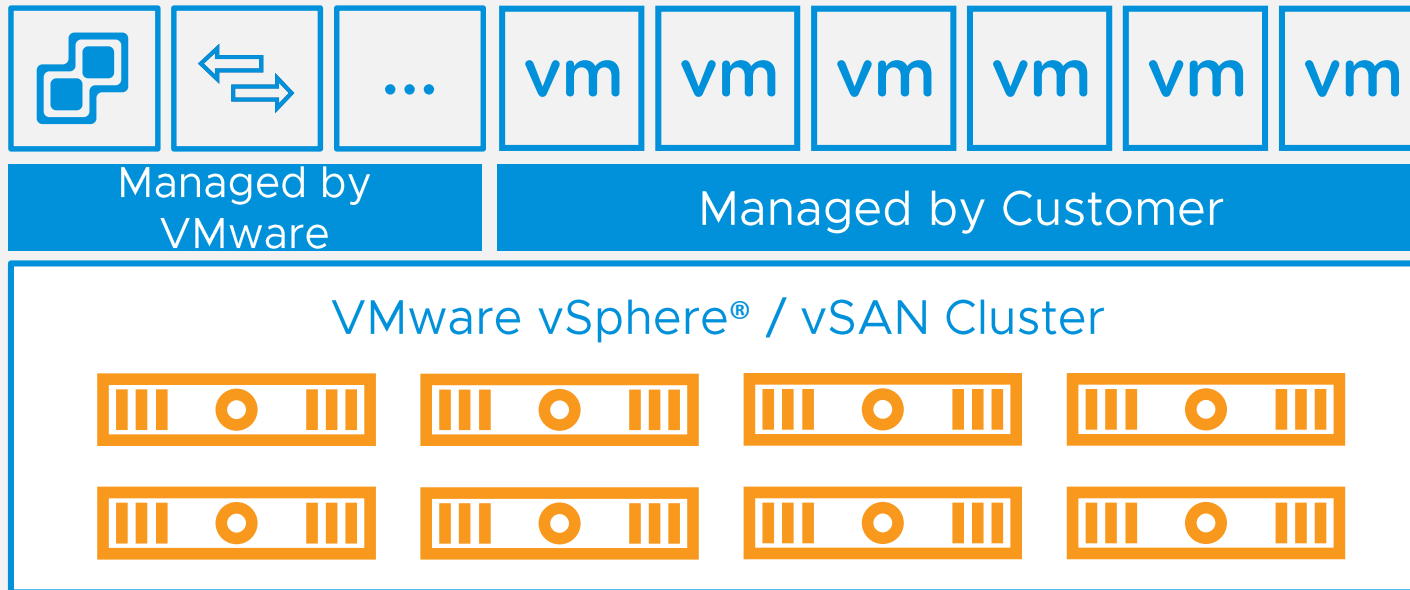
- Intel Xeon E5-2686 v4 processors
- 36 cores
- 2.3 GHz
- 512 GiB RAM
- 10TB NVMe flash
- 25 Gbps Networking



AWS r5.metal

- Intel® Xeon® Platinum 8000 Series (Skylake-SP)
- 48 cores
- 2.5 GHz
- 768 GiB RAM
- EBS Storage only (15-35 TiB)
- 25 Gbps Networking

クラスタコンフィグ



VMware vSAN™ をベースとしたクラスタ基盤、管理系およびワークロード系 VM はデプロイされます

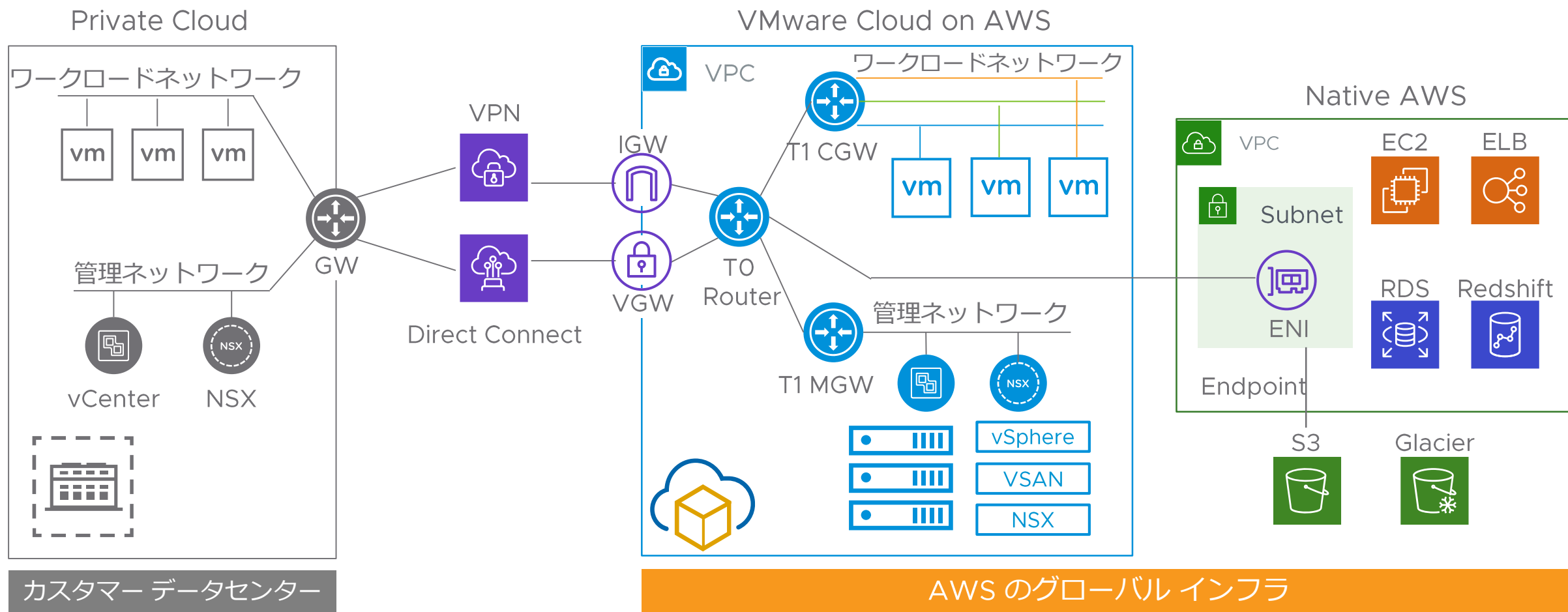
- VMware vCenter®
- VMware NSX®
- HCX

クラスタサイズは 3-16 ホストとなるホストの追加削除は自動的に行う

vSphere High Availability 機能

vSphere DRS 機能

VMware Cloud on AWS ネットワーク構成概要



Compute / Availability

vSphere HA および DRS

ストレッチクラスタ

Compute / Availability

vSphere HA および DRS

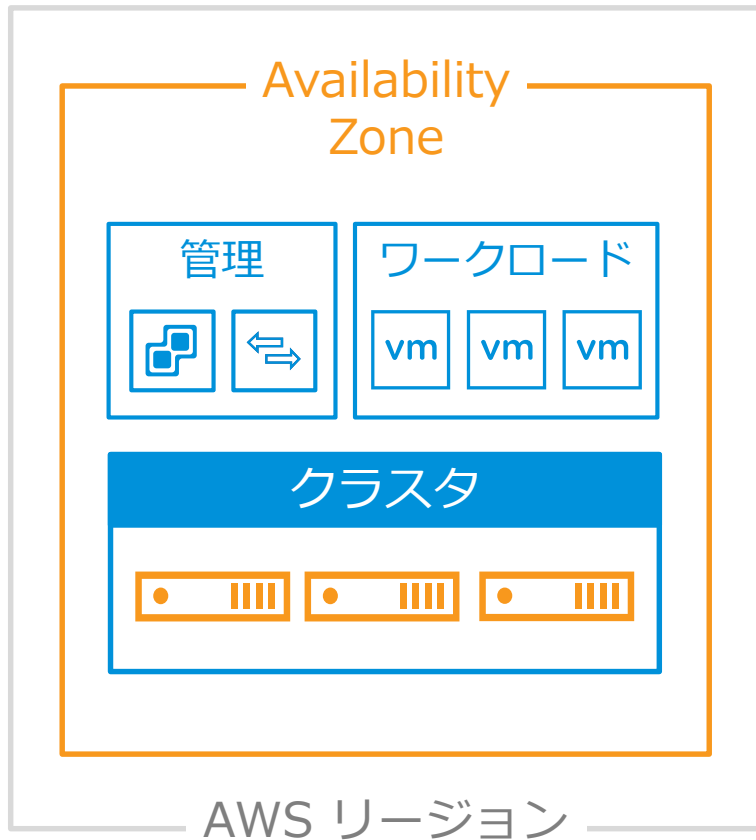
vSphere HA DRS		
	設定項目	設定値
	ホスト監視の有効化	チェック
	ホストのフェイルオーバー キャパシティの 定義基準	クラスタ リソースの割合 (%)
	クラスタで許容するホスト障害	1
	仮想マシンの監視	仮想マシンとアプリケーションの監視
	ホスト隔離への対応	仮想マシンをパワーオフして再起動
	vSphere DRS	オン
	移行の閾値	優先順位 3
	電源管理 (DPM)	オフ

最適化されたサービスを提供できる設定、ユーザーによる変更は不可

Compute / Availability

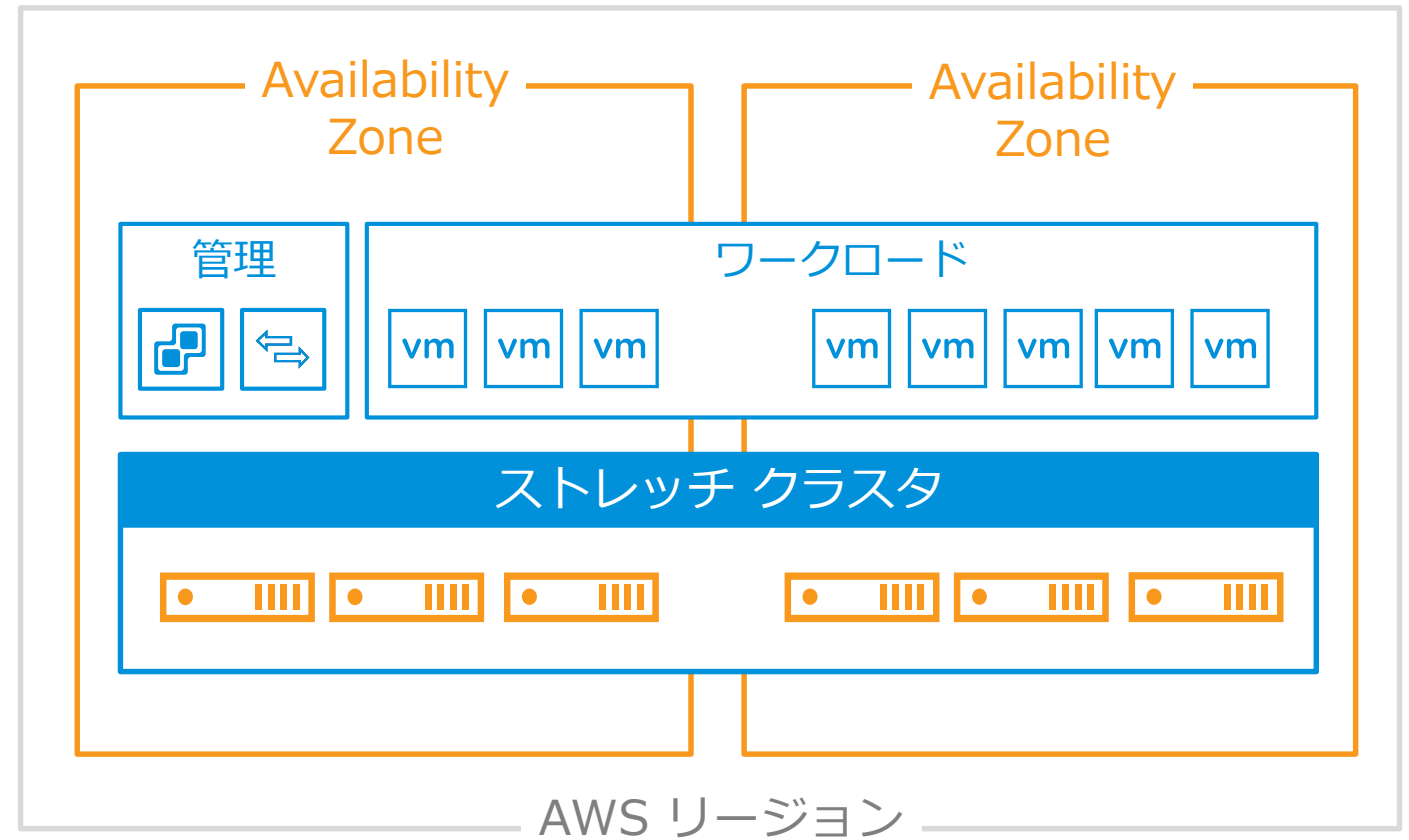
ストレッチ クラスターの構成概念

標準クラスター構成
(Non-Stretched Cluster)



AWS のグローバル インフラ

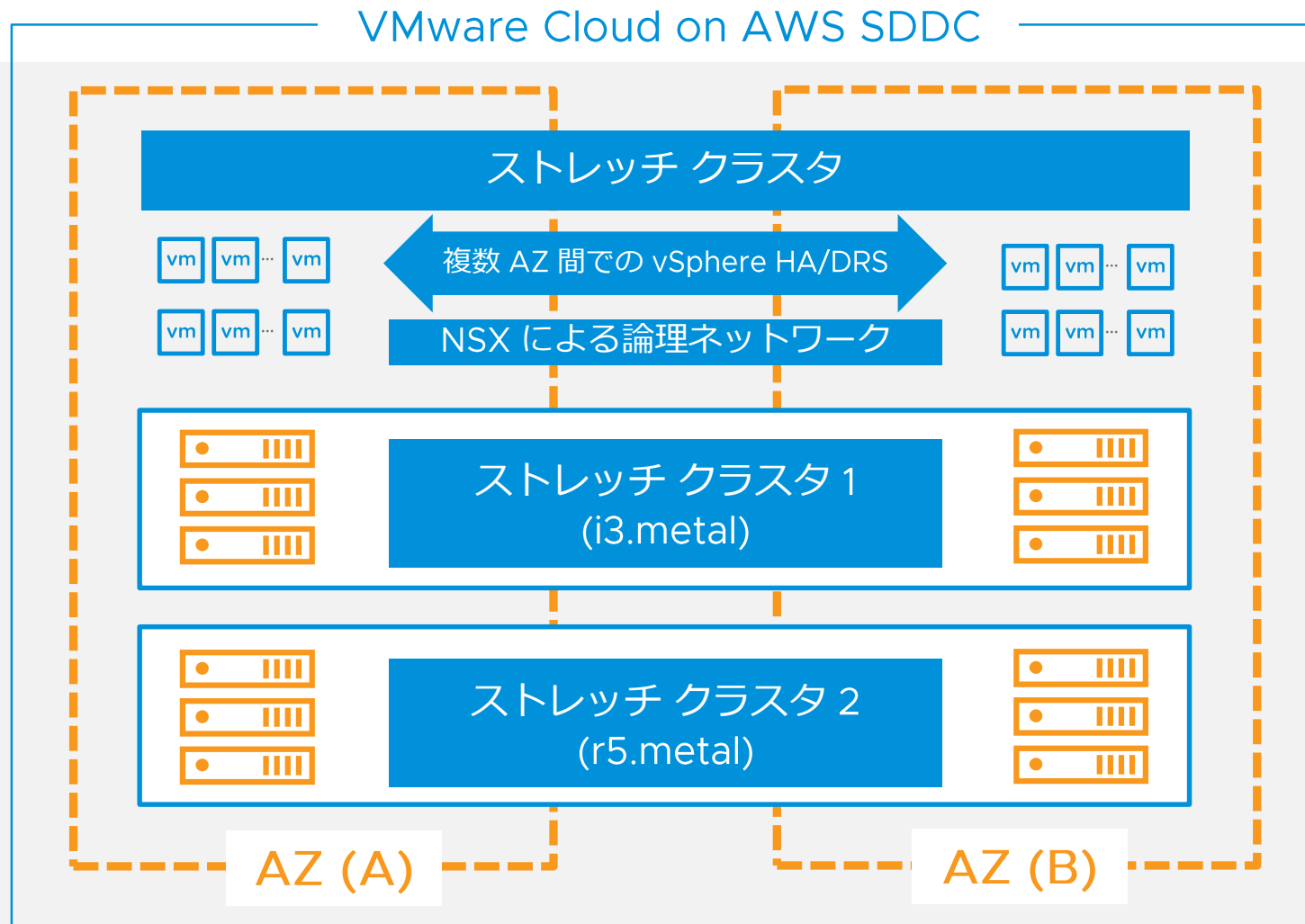
ストレッチ クラスター構成
(Stretched Cluster)



AWS のグローバル インフラ

Compute / Availability

高い可用性を提供するストレッチ クラスタ構成



AZ を跨いでストレッチ クラスタを構成可能 (SLA 99.99%)

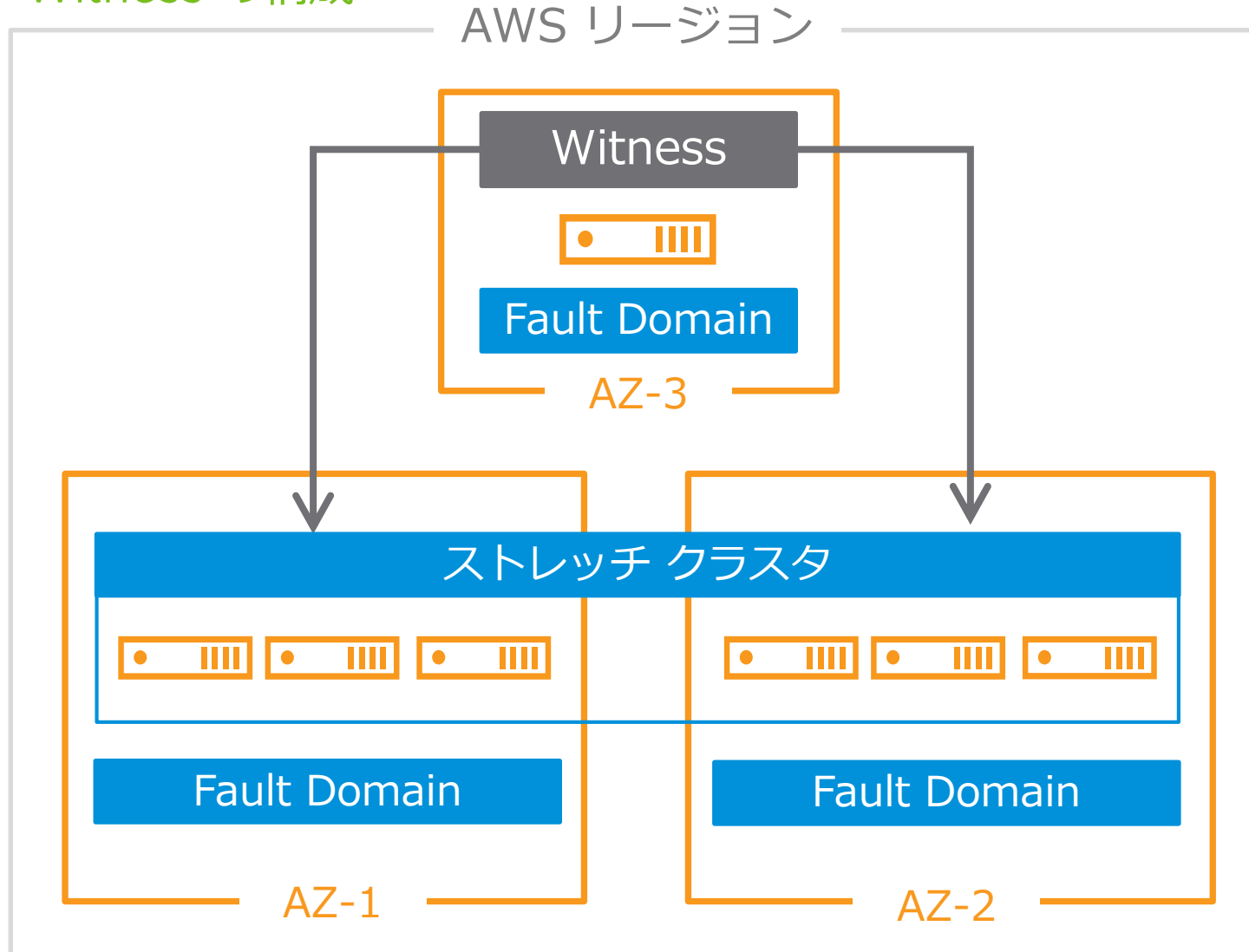
AZ 障害時は、もう一方の AZ で仮想マシンを再起動

1 つの SDDC 内に複数の "タイプの異なるストレッチ クラスタ" をサポート

i3.metal ストレッチクラスタと r5.metal ストレッチ クラスタを混在可能

Compute / Availability

Witness の構成



ストレッチ クラスタは、一つのリージョンで三つの AZ を使用する（三つの Fault Domain 構成）

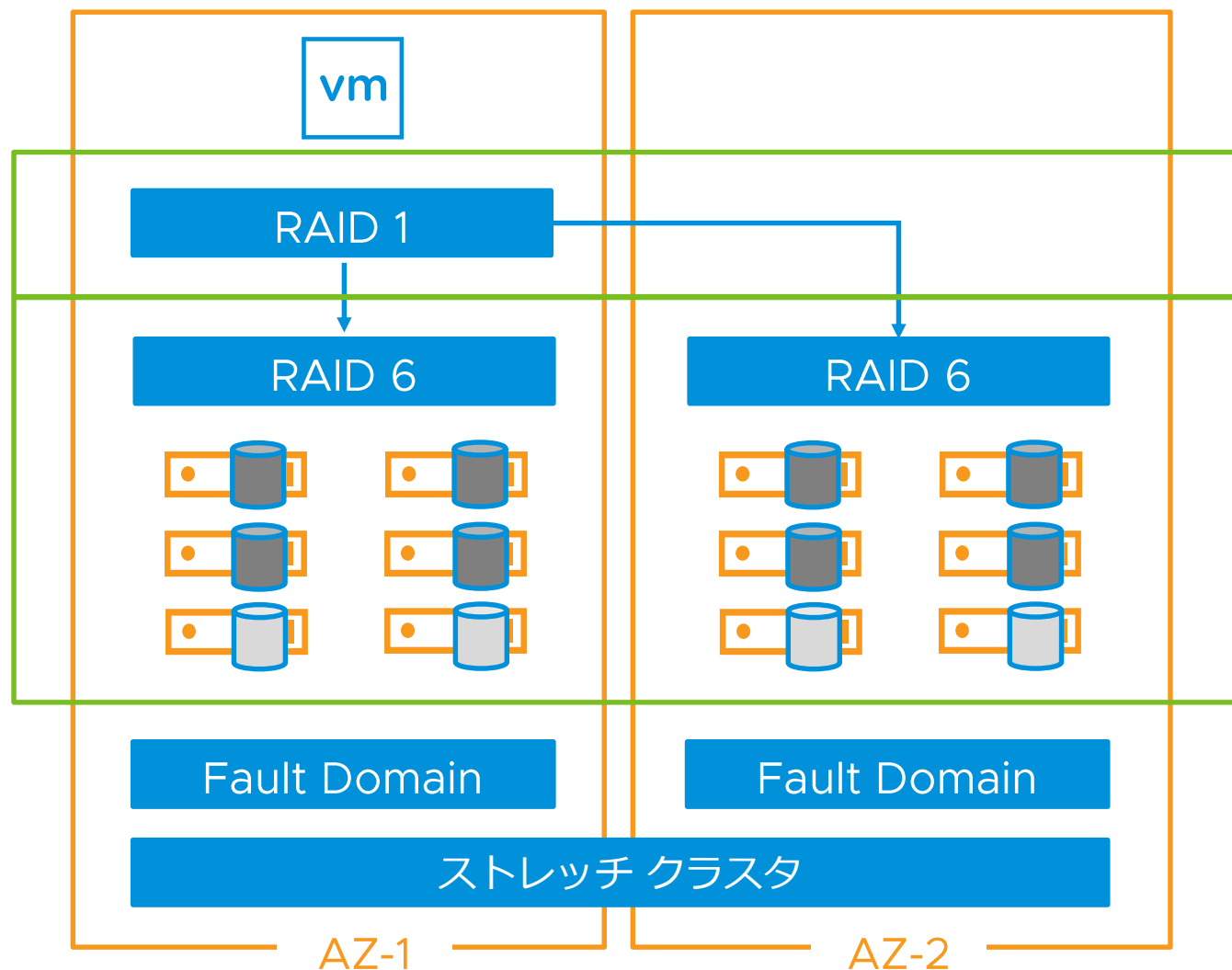
- そのうち一つを Witness で使用する

Witness はスプリットブレインを防ぐための Tie-Breaker として機能する

Witness はストレッチ クラスタ作成時に Amazon EC2 インスタンスとして自動的に展開される

Compute / Availability

ストレッチクラスタによるサイト内とサイト間マルチレイヤで保護



サイト内とサイト間の冗長化

- サイト間 Dual Site Mirroring サイト内最大 RAID6、FFT=2 で高可用性を確保

サイト障害とホスト障害それぞれ独立

サイト間トラフィックを減らしたサイトローカルリディロジックを最適化

- ローカルリード
- リモート同期ライト

Compute / Availability

二レイヤでサイトを保護

vSAN

可用性

詳細なポリシー

ルール

タグ

サイトの耐障害性 ⓘ	デュアル サイト ミラーリング (ストレッチ クラスタ) ✓
許容される障害の数 ⓘ	2 件の障害 - RAID-6 (イレージャ コーディング) ▼

100 GB の仮想マシン ディスクで使われるストレージ容量は 133.33 GB

キャンセル

戻る

次へ

サイトアフィニティ

どのサイトにデータをキープする
のを決める

ローカル保護

vSAN ホスト上にどのようにデー
タをキープするのを決める

Compute / Availability

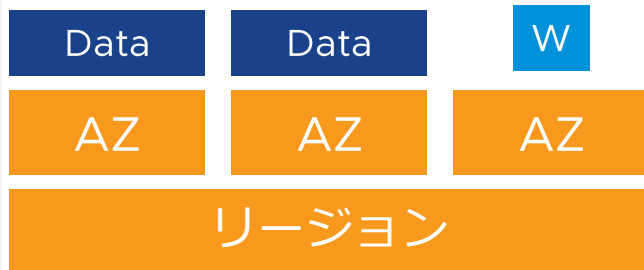
ポリシーベースのデータプレースメント

Dual Site Mirroring

全てのデータを両サイトにミラーリングされる



ストレッチクラスタ



None – Keep data on Preferred

全てのデータをプリファードサイトにキープする



ストレッチクラスタ

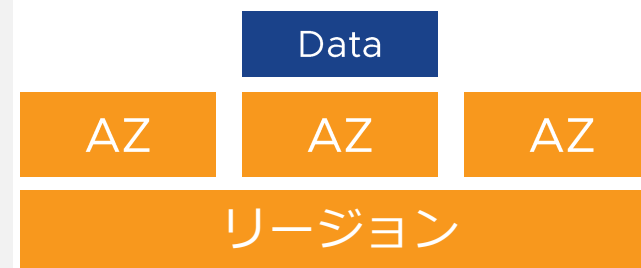


None – Keep data on non-preferred

全てのデータをノンプリファードサイトにキープする

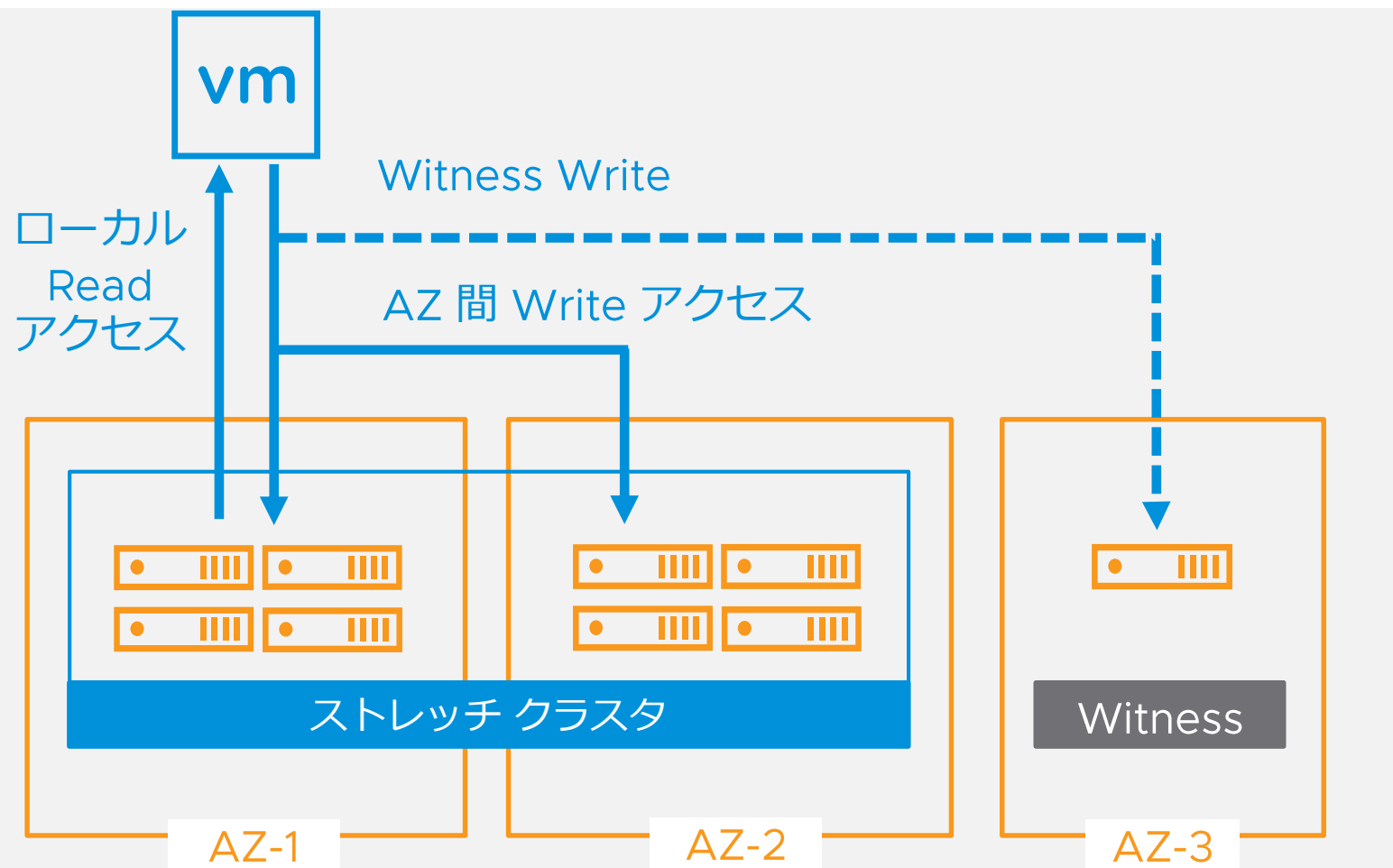


ストレッチクラスタ



Compute / Availability

サイト間トラフィック

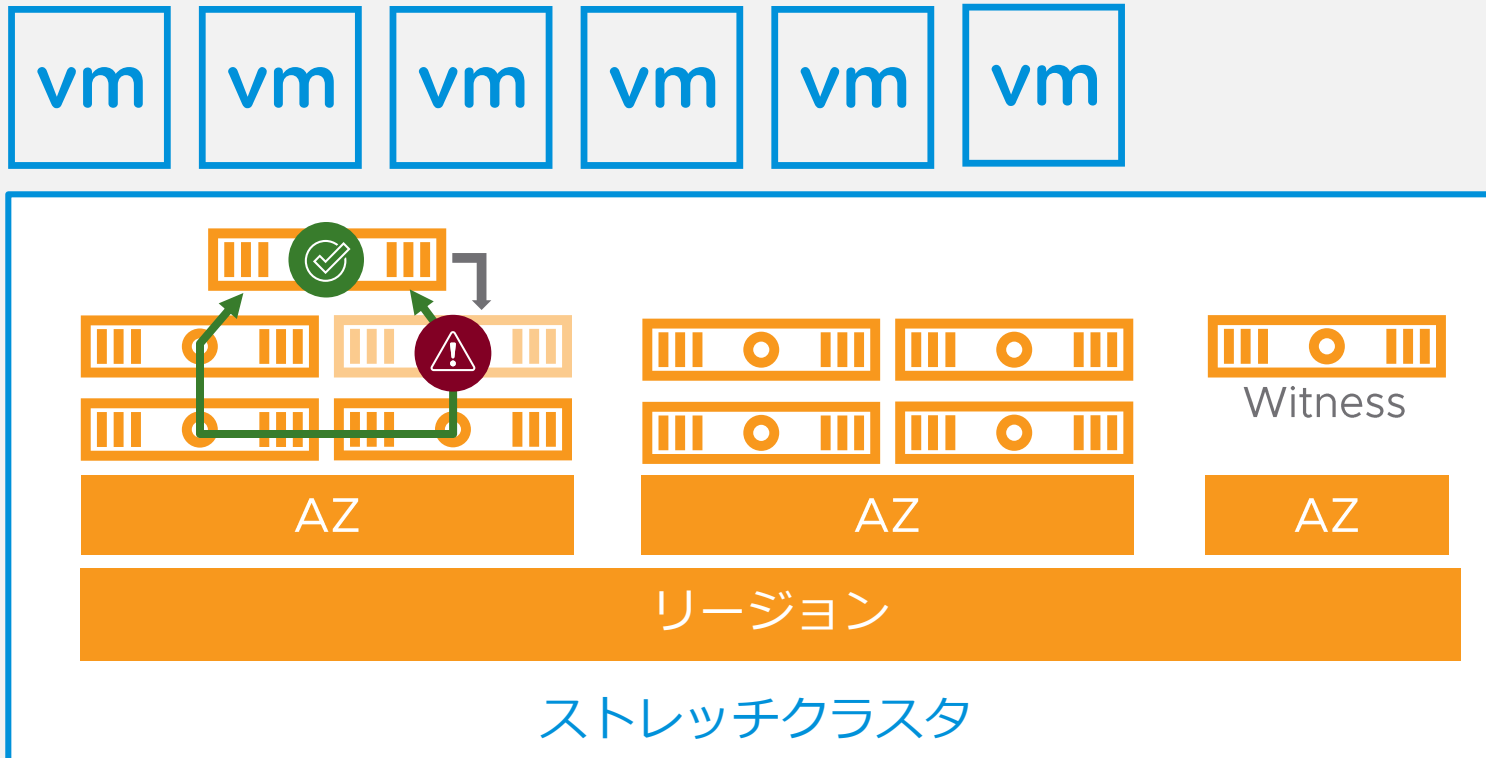


サイト間トラフィックを減らしたサイトローカリティロジックを最適化

- ローカルリード
- AZ 間リモート同期ライト

Compute / Availability

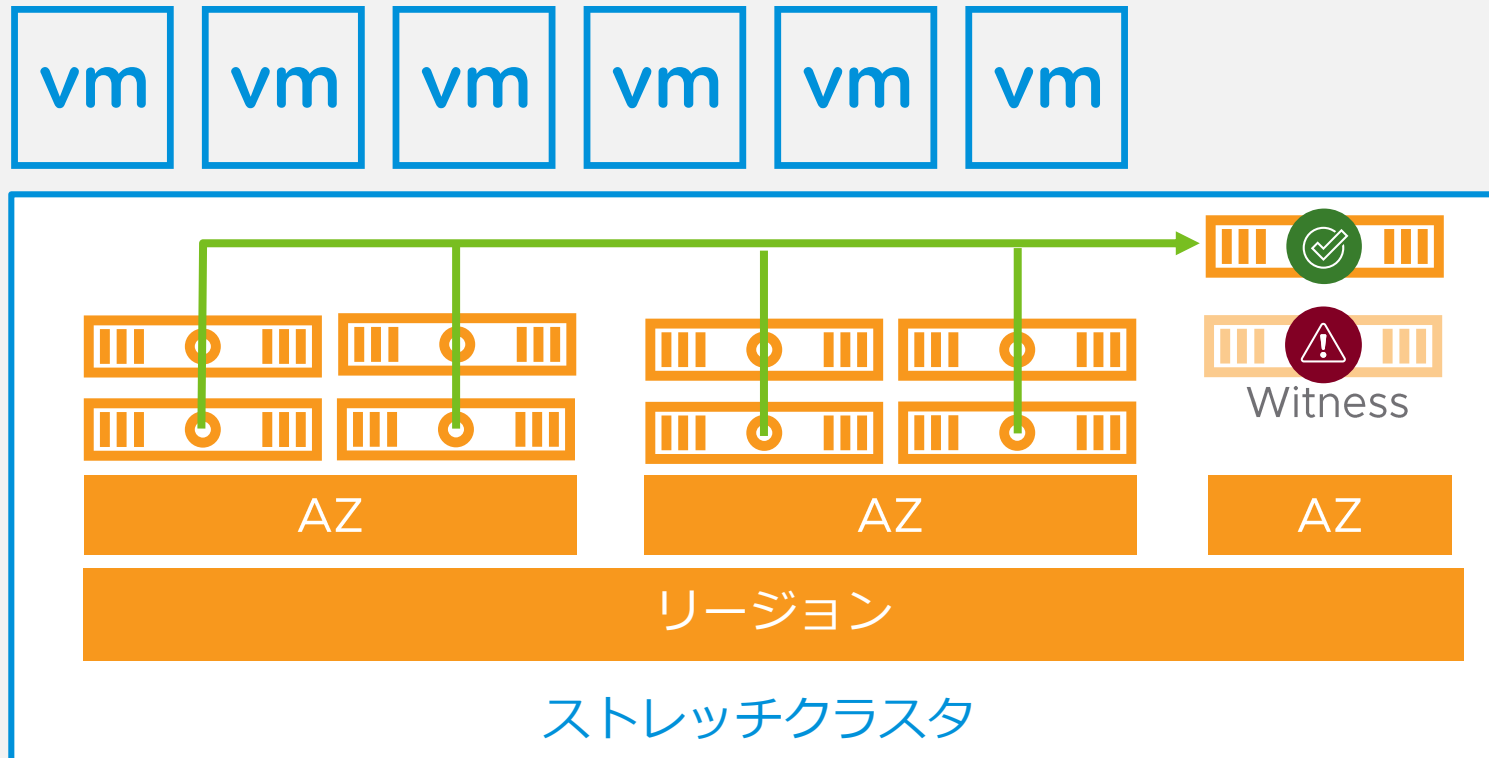
ホスト障害の動き – ストレッチクラスタ



1. ホスト障害もしくは、問題検出
2. 新しいホストをクラスタに追加される
3. 問題あったホストのデータを新ホストでリビルドもしくはマイグレートされる
4. 障害ホストをクラスタから退避、新ホストにリプレイスされる

Compute / Availability

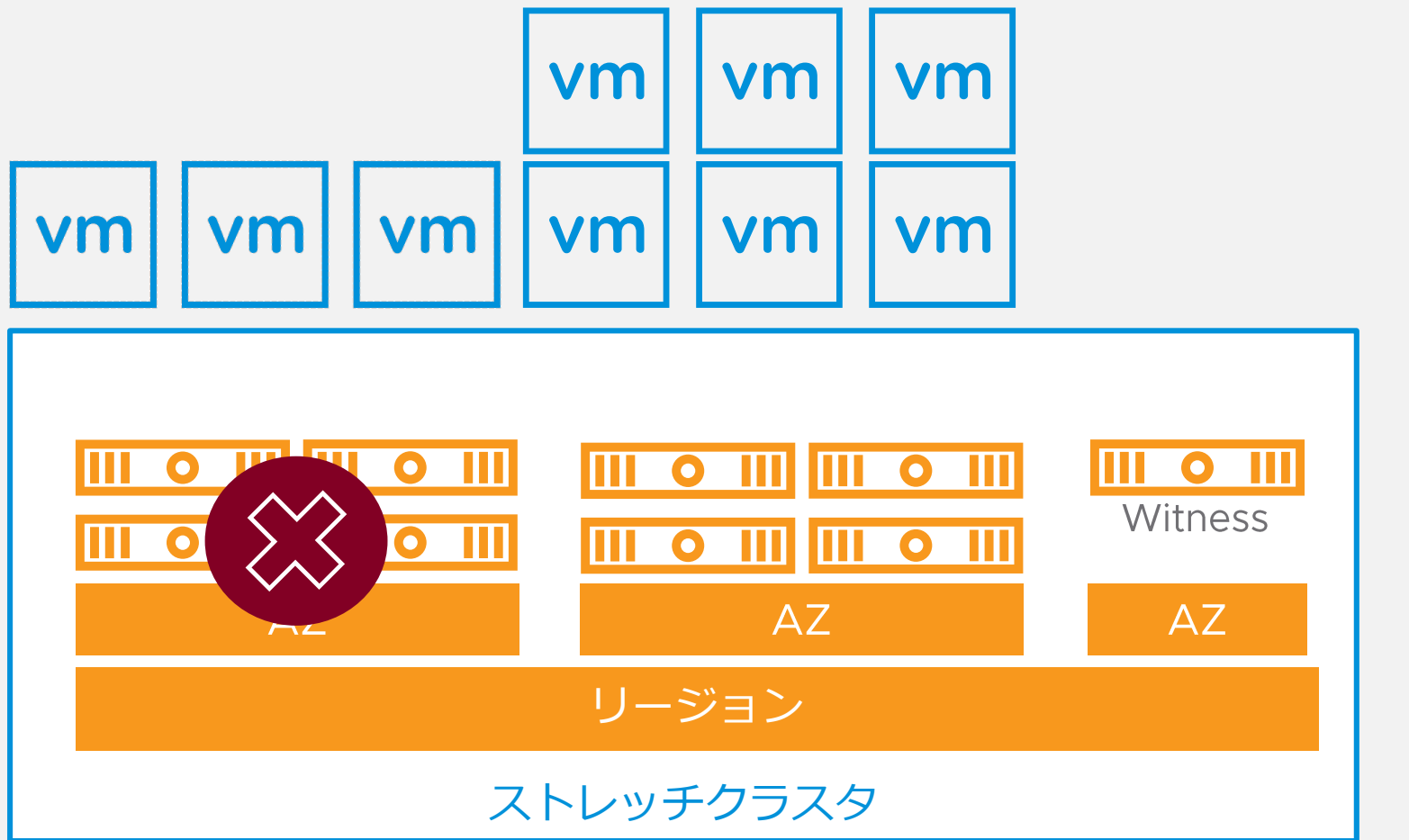
Witness 障害時の動き – ストレッチクラスタ



1. Witness 障害もしくは、問題検出
2. 新しい Witness をクラスタにプロビジョンされる
3. サイト間 Witness コンポーネントを新しい Witness にリビルドされる
4. 障害した Witness を新 Witness にリプレイスされる

Compute / Availability

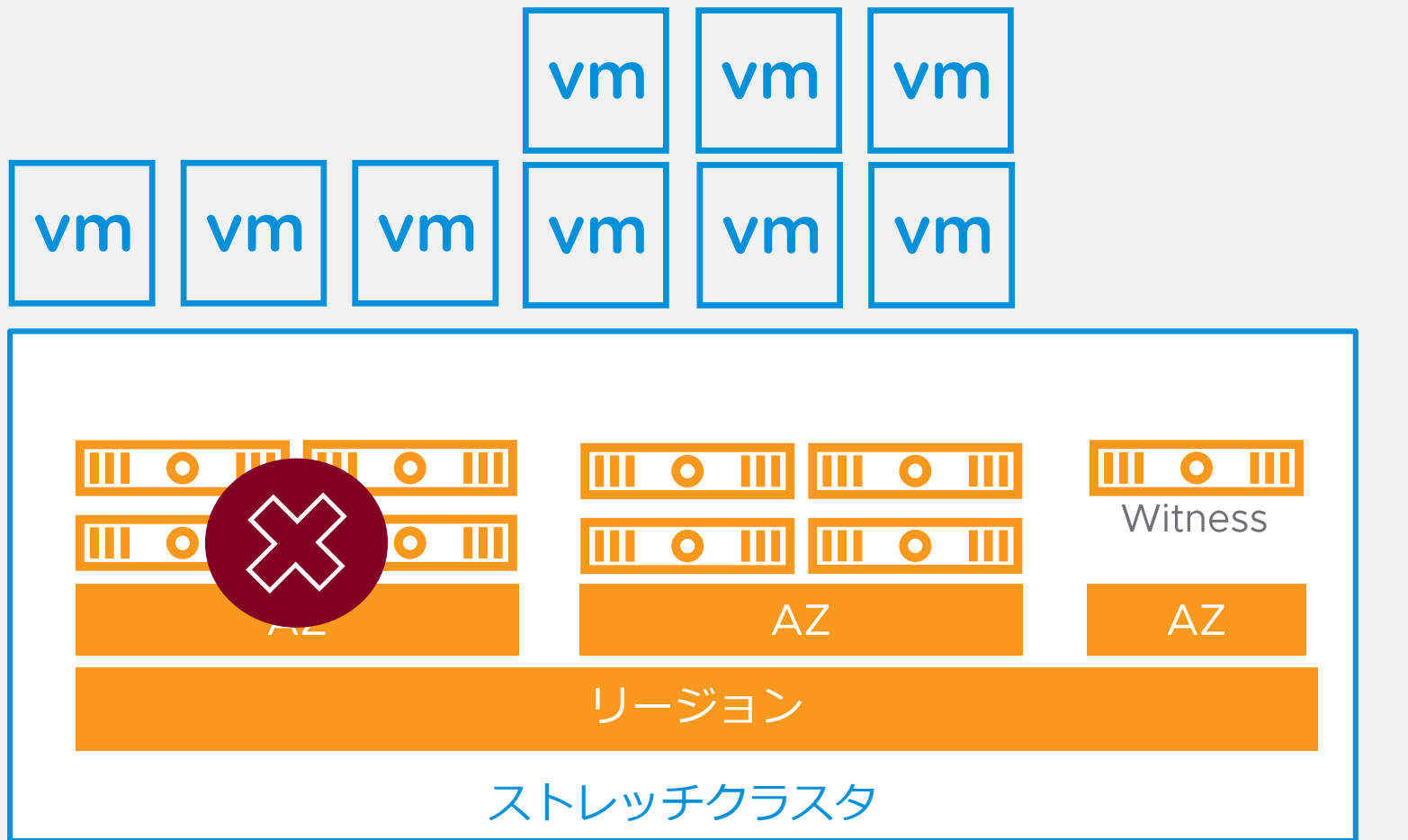
AZ 障害時の動き – ストレッチクラスタ



1. 片方サイトに障害発生
2. vSphere HA VM を正常サイトに再起動させる

Compute / Availability

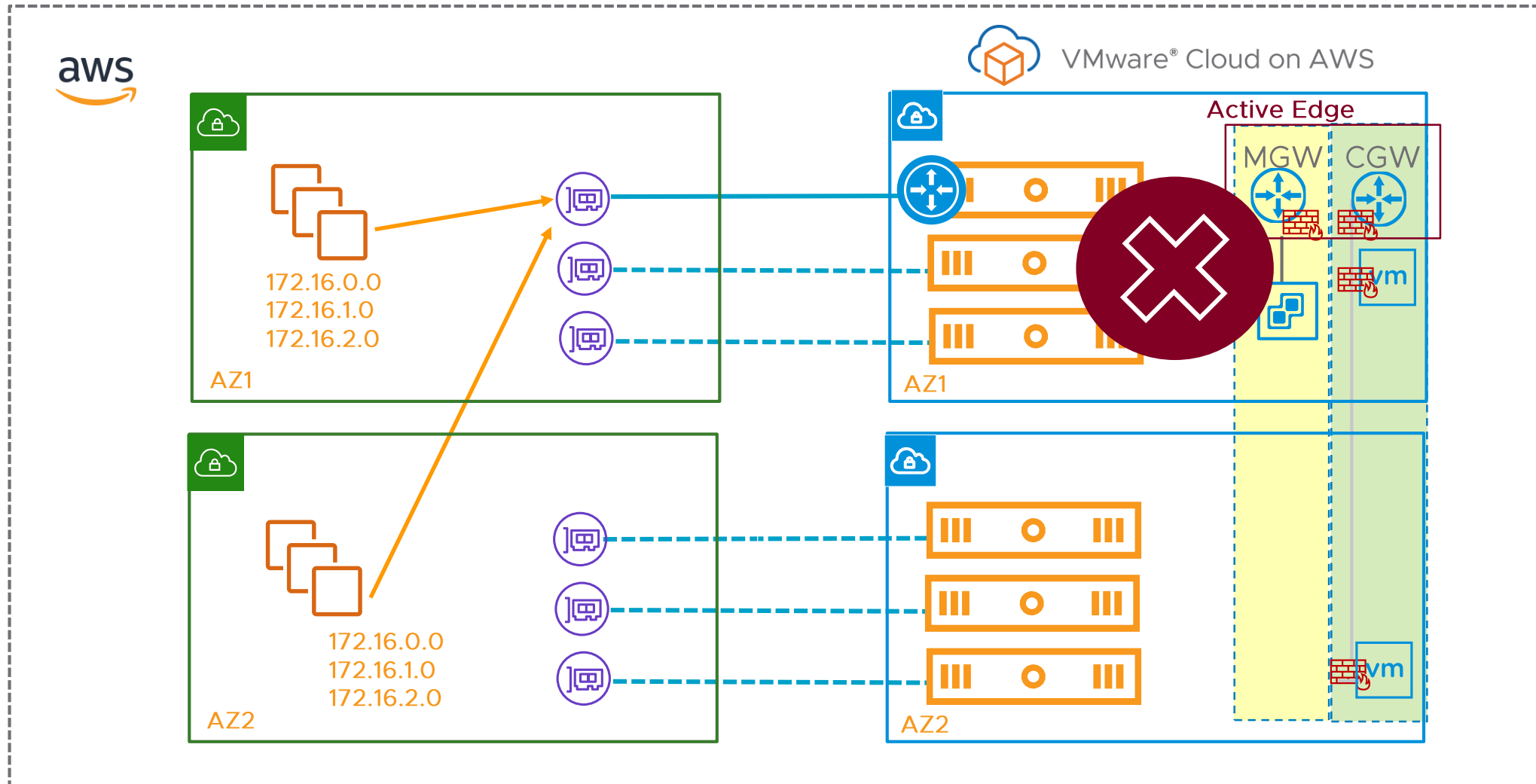
AZ 障害時の動き – ストレッチクラスタ



1. 片方サイトに障害発生
2. vSphere HA VM を正常サイトに再起動させる
3. AZ 復旧、データを再同期
4. データをリストアされる

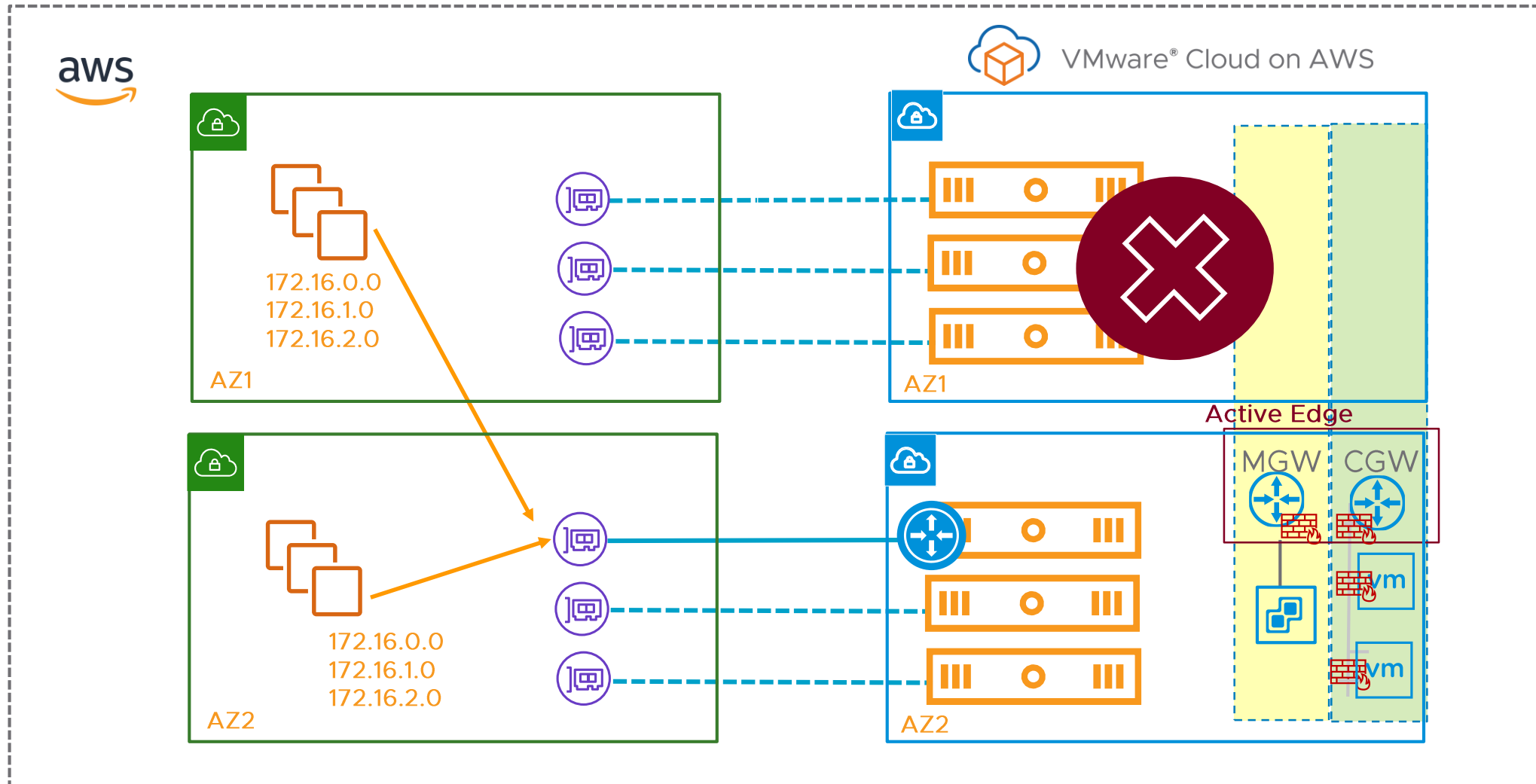
Compute / Availability

障害時ネットワークの切り替え – ストレッチクラスタ



Compute / Availability

障害時ネットワークの切り替え – ストレッチクラスタ



Storage

i3.metal と r5.metal の比較まとめ

パフォーマンス比較

i3.metal と r5.metal 障害時の自動復旧

Elastic DRS

Storage

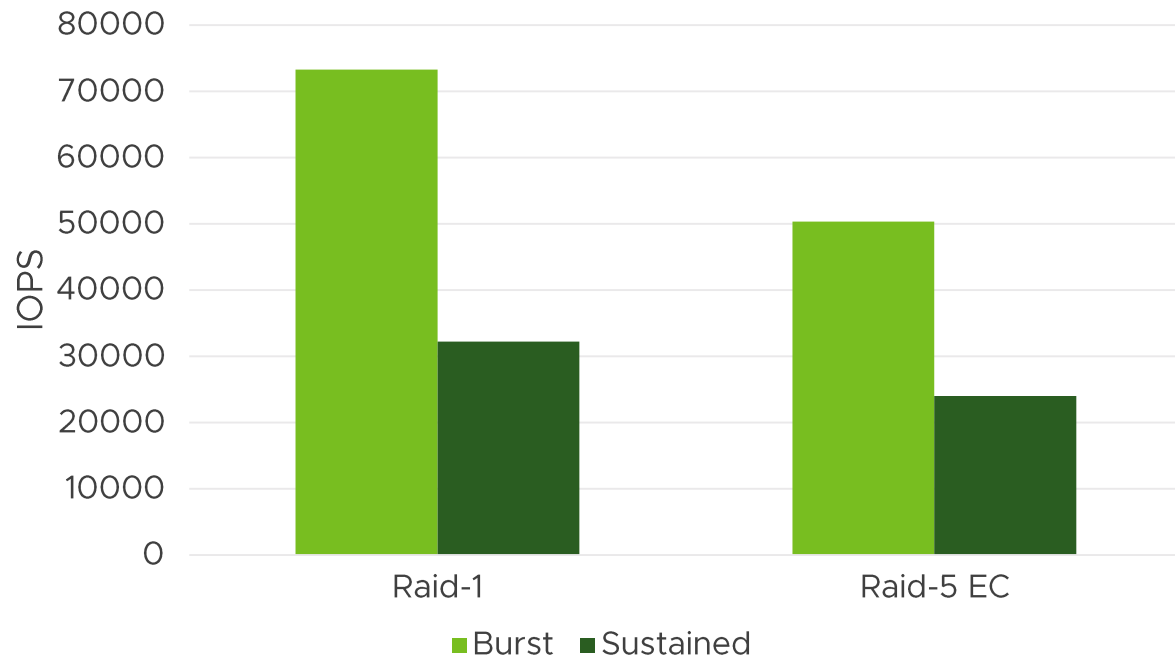
i3.metal と r5.metal の比較

項目	i3.metal	r5.metal (EBS バックエンド)
ディスクグループ	2	3
キャパシティ	容量固定	単独にストレージを追加できる
性能	高 read / write 性能	高キャパシティ
障害時リビルド	ホスト障害にフルリビルドが必要	フルリビルド不要
暗号化	Nitro ファームウェアレイア暗号化+KMS により vSAN ストレージ暗号化	
重複排除	○	ゼロブロックのみ
圧縮	○	○

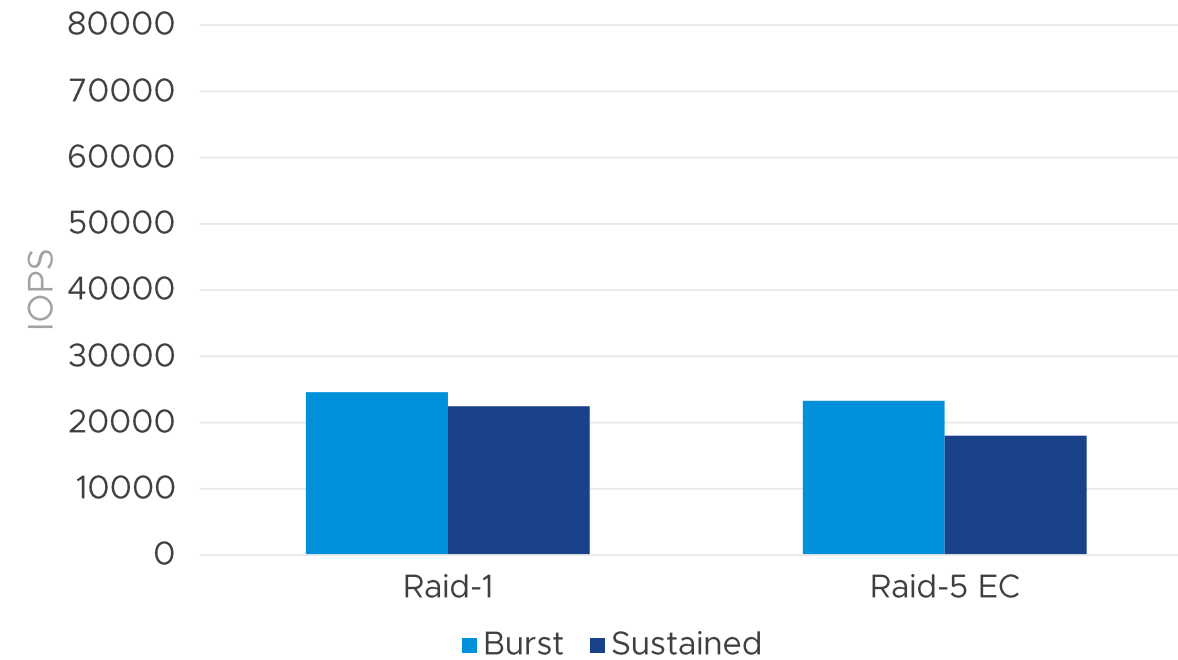
Storage

ホスト単位パフォーマンス

i3.metal



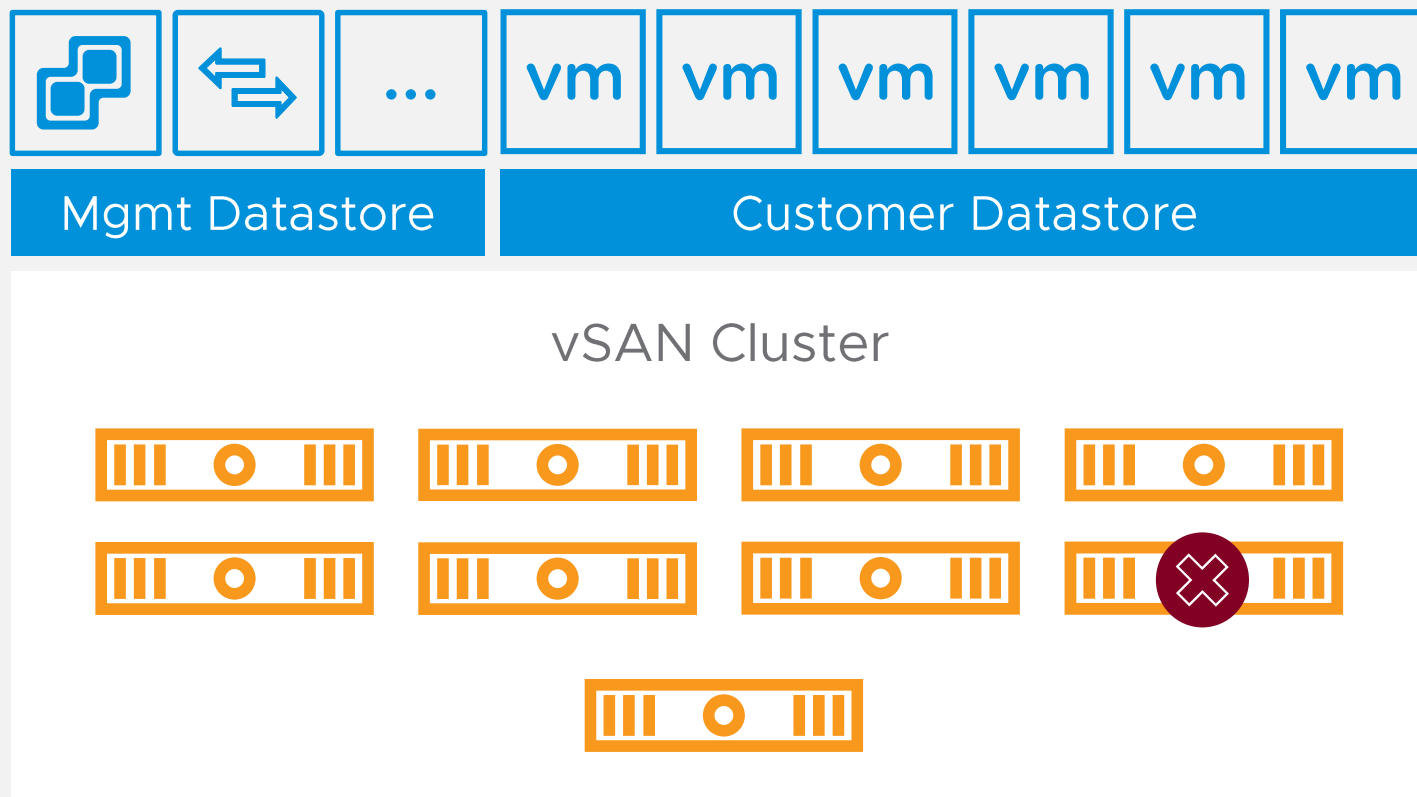
r5.metal



16KB, 70% Read, 800GB wss/host

Storage

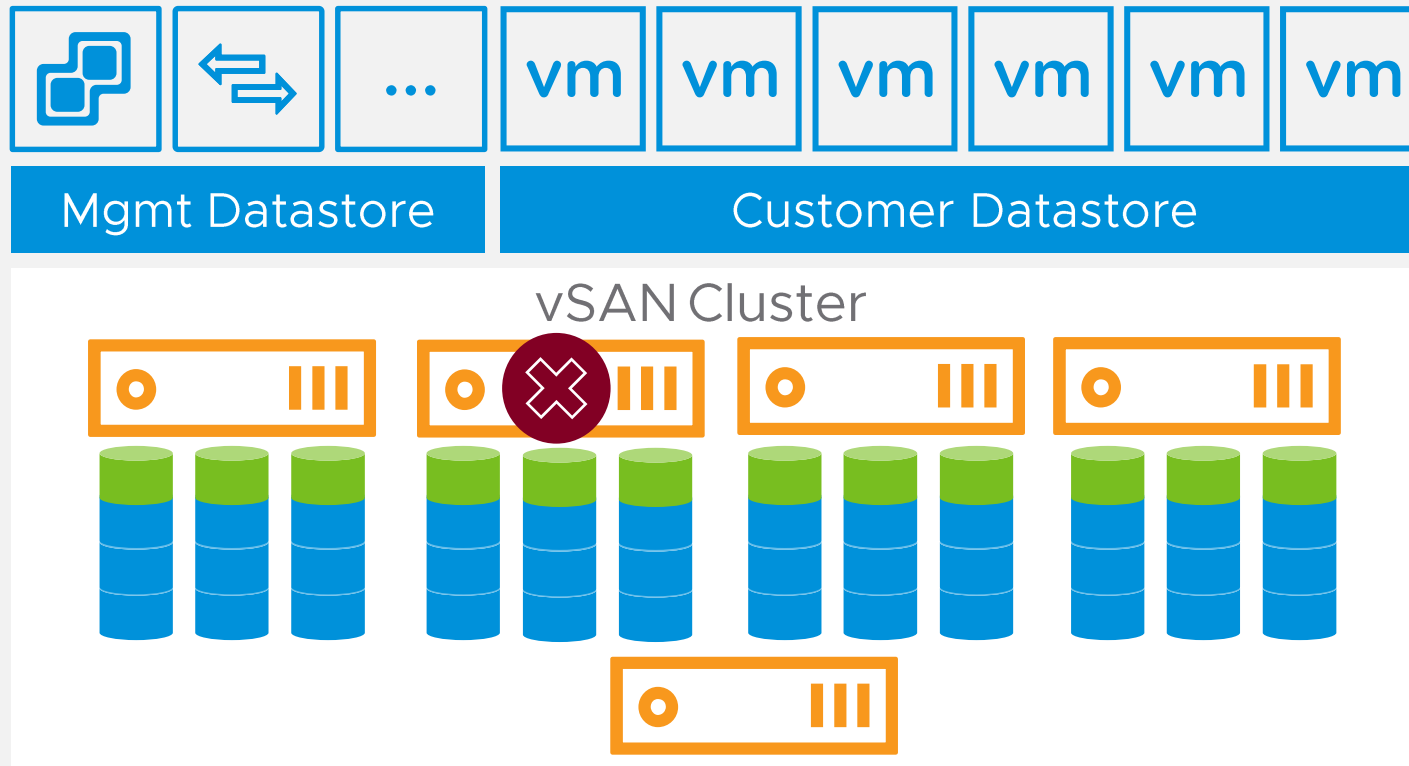
i3.metal ホスト障害におけるクラスタ自動復旧



1. 障害検出
2. ホスト追加される
3. データリビルド
4. 障害ホストの削除

Storage

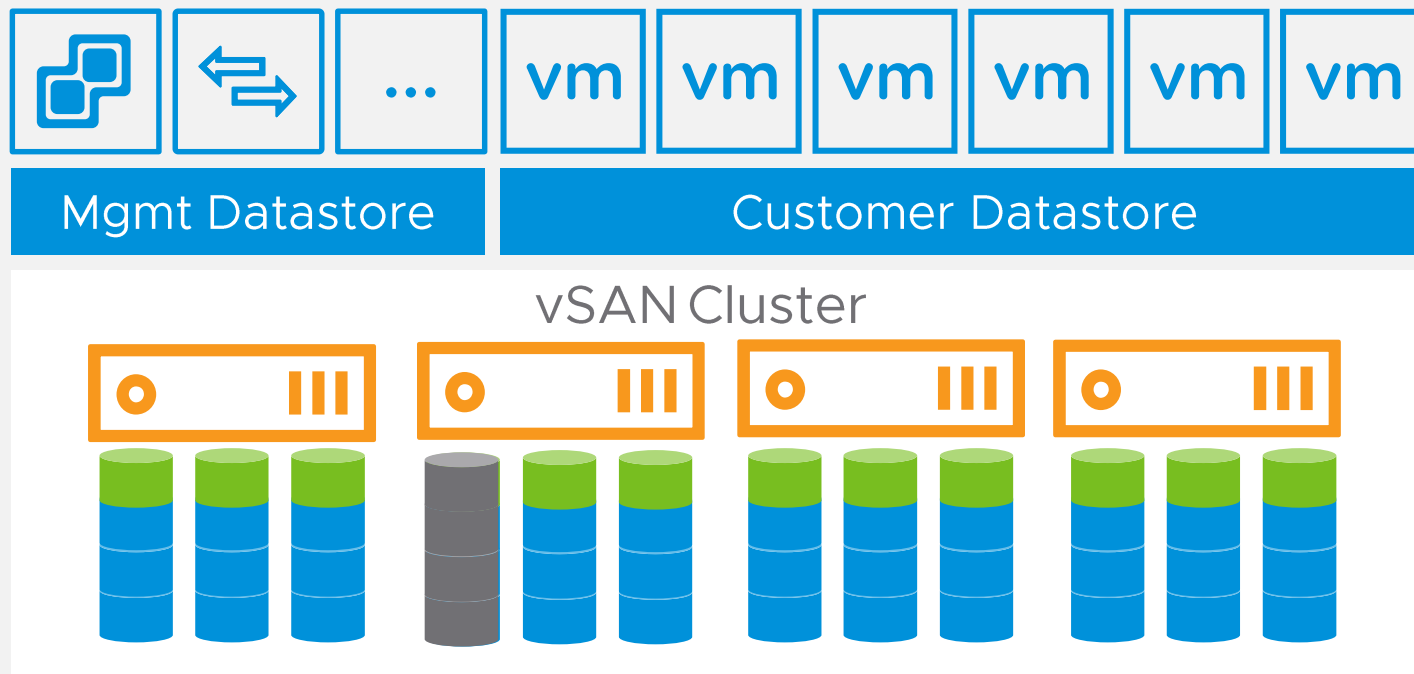
r5.metal ホスト障害におけるクラスタ自動復旧



1. 障害検出
2. テンポラリーホスト追加
3. 障害VMを再起動
4. 障害ホストを削除
5. 新しいホストを追加し、EBSをアタッチ
6. 変更点を再同期
7. テンポラリーホストを削除

Storage

r5.metal EBS障害におけるクラスタ自動復旧



1. 問題検出
2. ディスクグループ削除
3. 新しいディスクグループが作成される
4. データリビルド

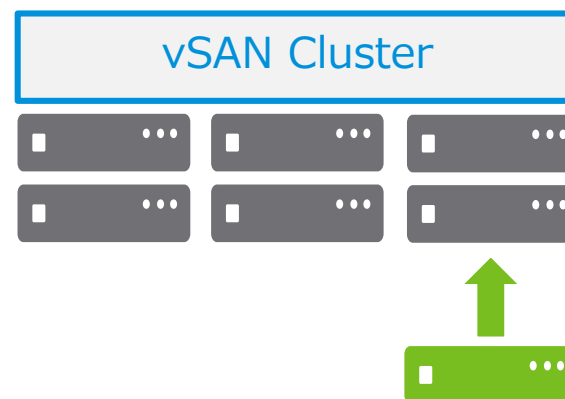
Storage

Elastic DRS 動作概要

1. 指定のしきい値内で クラスターが稼働



2. しきい値を超過 追加のホストをプロビジョニング



3. 指定のしきい値内での負荷に クラスターが復帰



Storage

Elastic DRS 設定



< Elastic DRS の設定： VMC-SET-JP-TOKYO-16： キャパシティ： Cluster-1 の編集

Elastic DRS を使用すると、最適な数のホストをパワーオン状態で維持することができます。これにより、クラスタの使用率が高いまま、必要な CPU、メモリ、およびストレージのパフォーマンスを確保できます。

Scale Up for Storage Only

We'll add a host when cluster storage utilization becomes critical.

パフォーマンスを最大にするための最適化

クラスタの使用状況に基づいてホストをすばやく追加し、また時間をかけてホストを削除することでパフォーマンスを最大にします。

最小クラスタ サイズ: 3

最大クラスタ サイズ: 4

コストを最小にするための最適化

クラスタの使用状況に基づいて、すばやくホストを削除し、また時間をかけてホストを追加することで継続的にコストを削減します。

最小クラスタ サイズ: 3

最大クラスタ サイズ: 4

保存

キャンセル

Storage

Elastic DRS 閾値

DRS 設定	コンポーネント	ホスト追加の閾値	ホスト削除の閾値
パフォーマンス	CPU 使用率	90%	50%
	メモリ使用率	80%	50%
	ストレージ容量	70%	20%
コスト	CPU 使用率	90%	60%
	メモリ使用率	80%	60%
	ストレージ容量	70%	20%

全てが
下回ったら
ホスト削除

Elastic DRS はクラスタ毎の設定

最小ホスト 3、最大ホスト 16

5 分間隔で使用率を取得し閾値を計算

スケールアウト後の 30 分はスケールアウトなし、かつ 3 時間はスケールインなし

Network

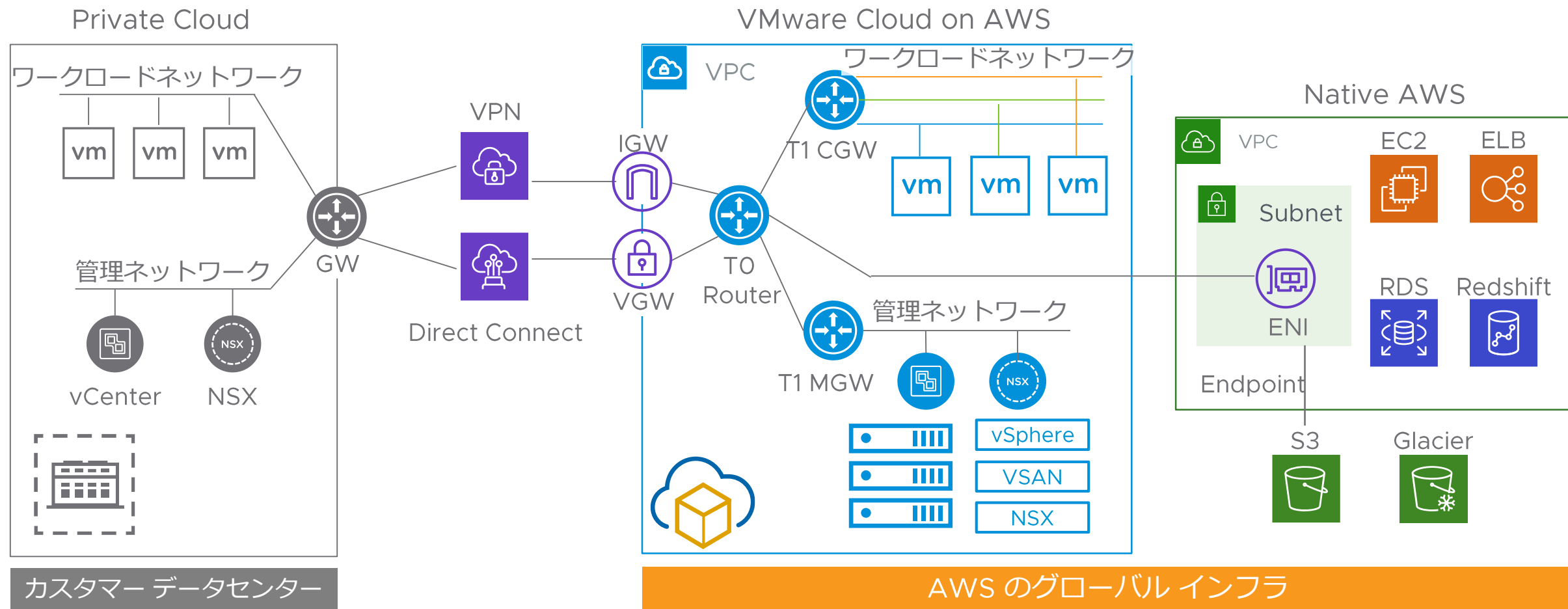
構成概要

North – South / East - West ルーティング

ファイアウォール

Networking

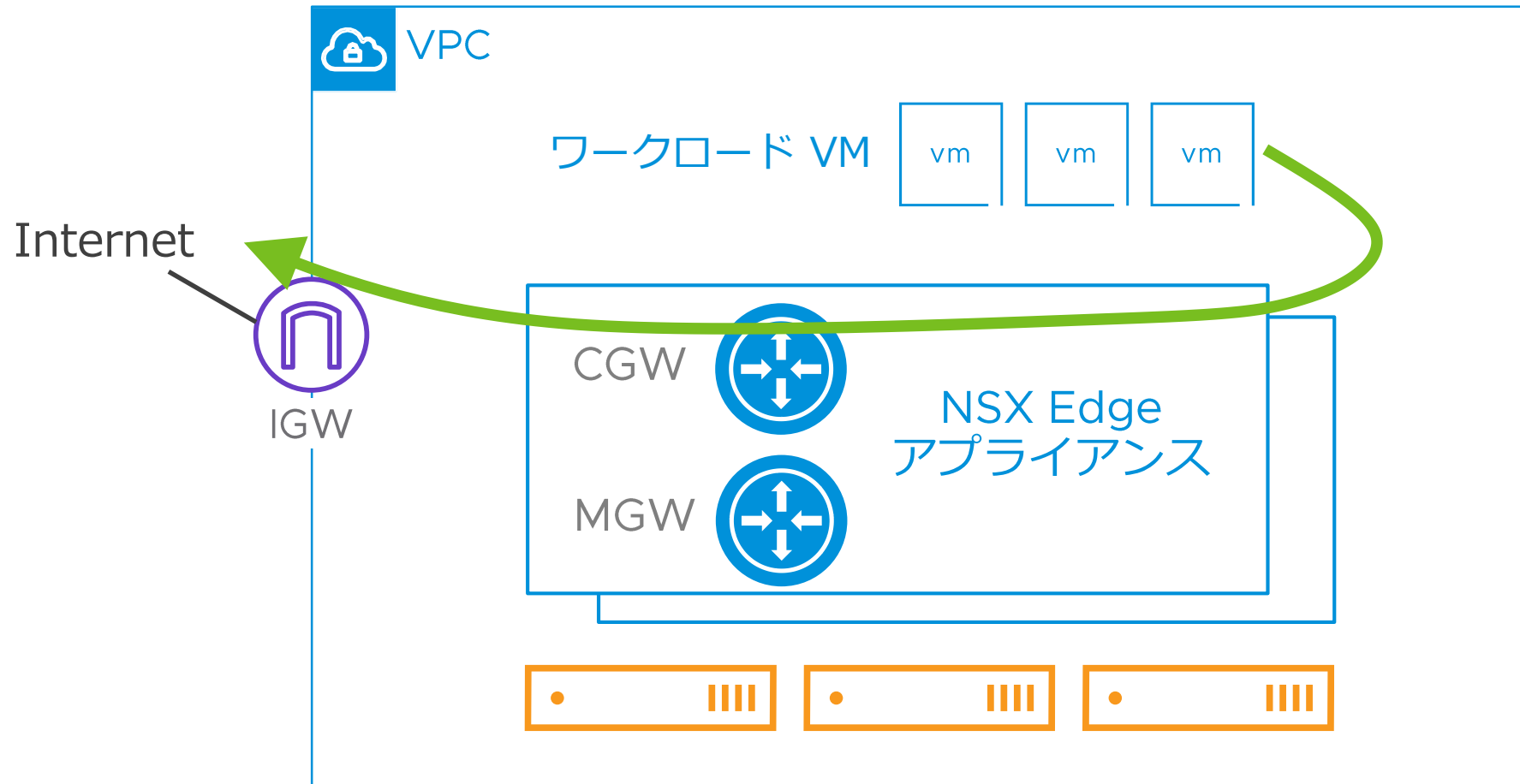
VMware Cloud on AWS ネットワーク構成概要



Networking

North-South ルーティング

VMware Cloud on AWS

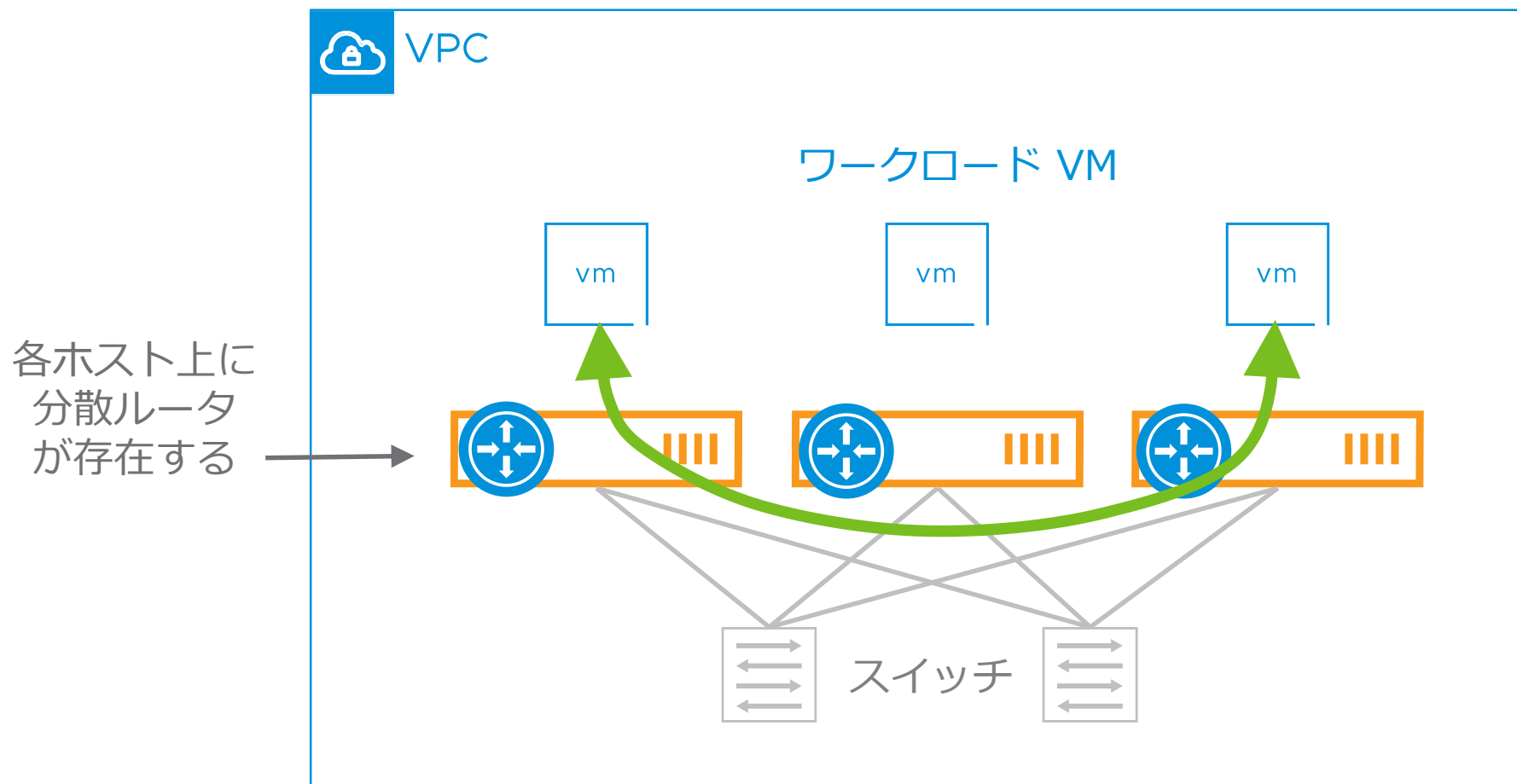


VMware NSX®
Edge™ ゲート
ウェイにより
North-South
ルーティングを
提供する

Networking

East - West ルーティング

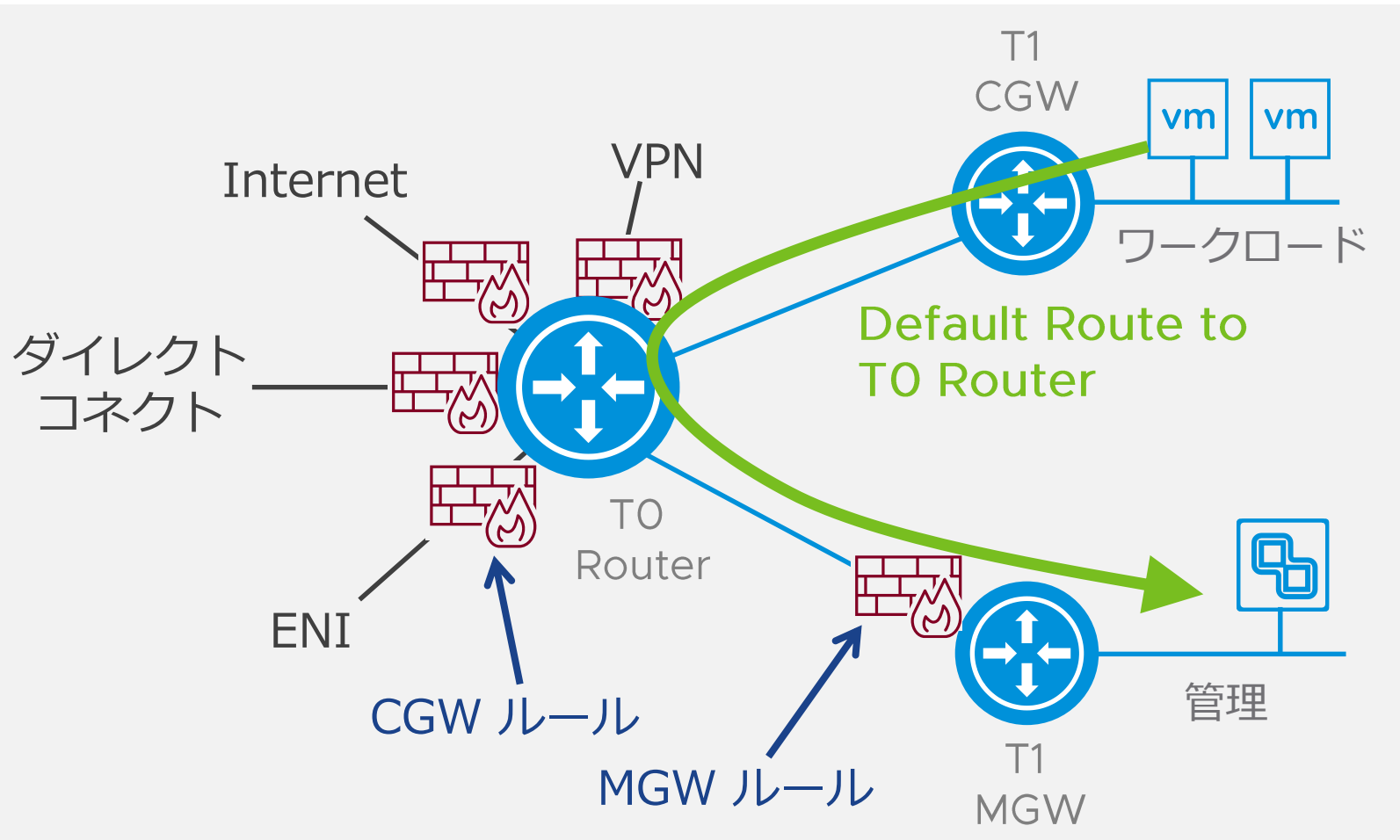
VMware Cloud on AWS



NSX 分散ルーティング
により、East-West
ルーティングを提供する

Networking

Edge ファイアウォール



CGW のファイアウォール
ルールは T0 Edge 側のイン
ターフェースに適用される

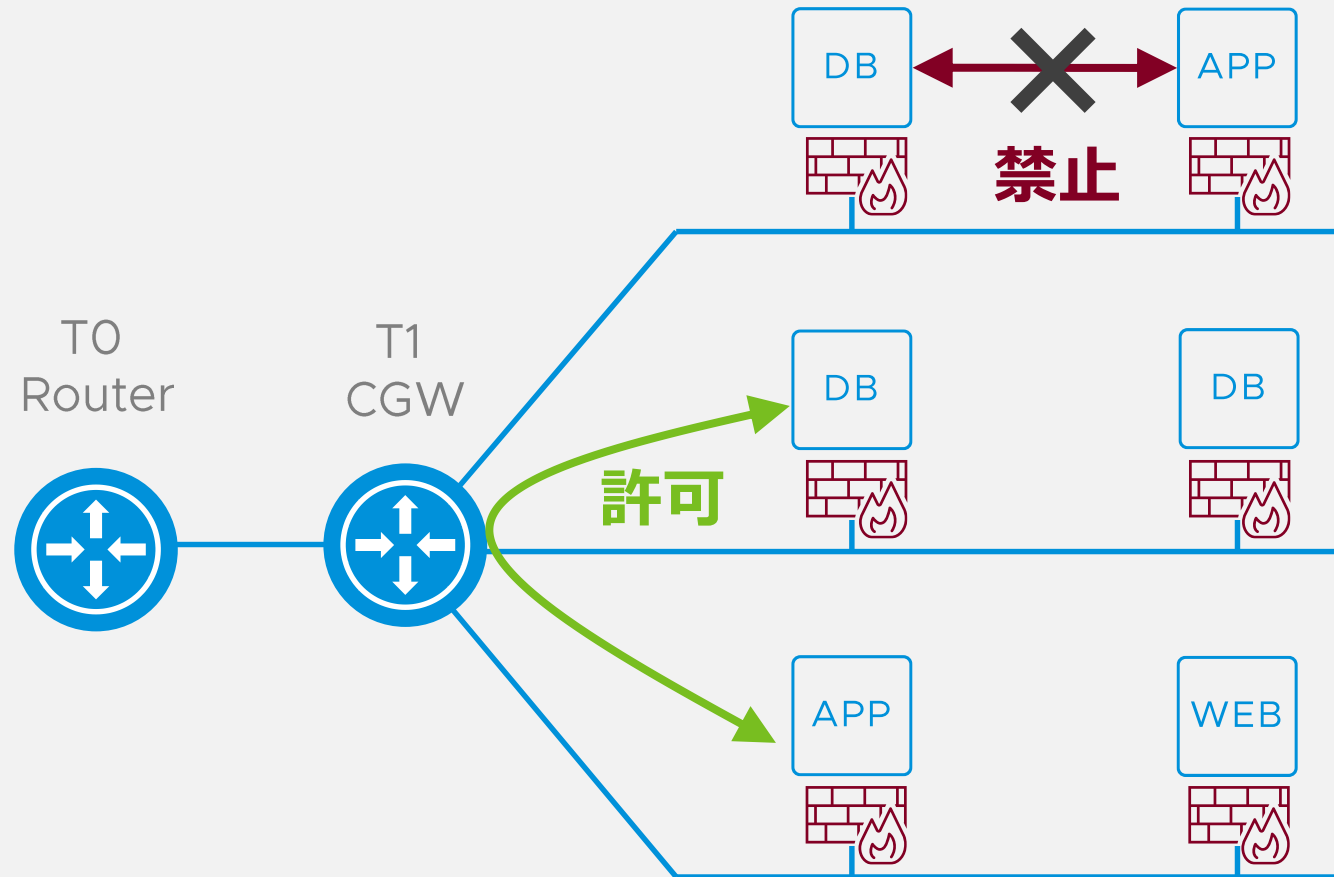
各インタフェースごとに細か
くファイアウォール ルールを
定義できる

ファイアウォールの設定は、
全て管理コンソールから一元
的に定義する

インターネット向けトラ
フィックはデフォルトでSNAT
されます

Networking

分散ファイアウォール



VM の vNIC レベルでポリシーを定義し、East-West トラフィックを細かくコントロールできる

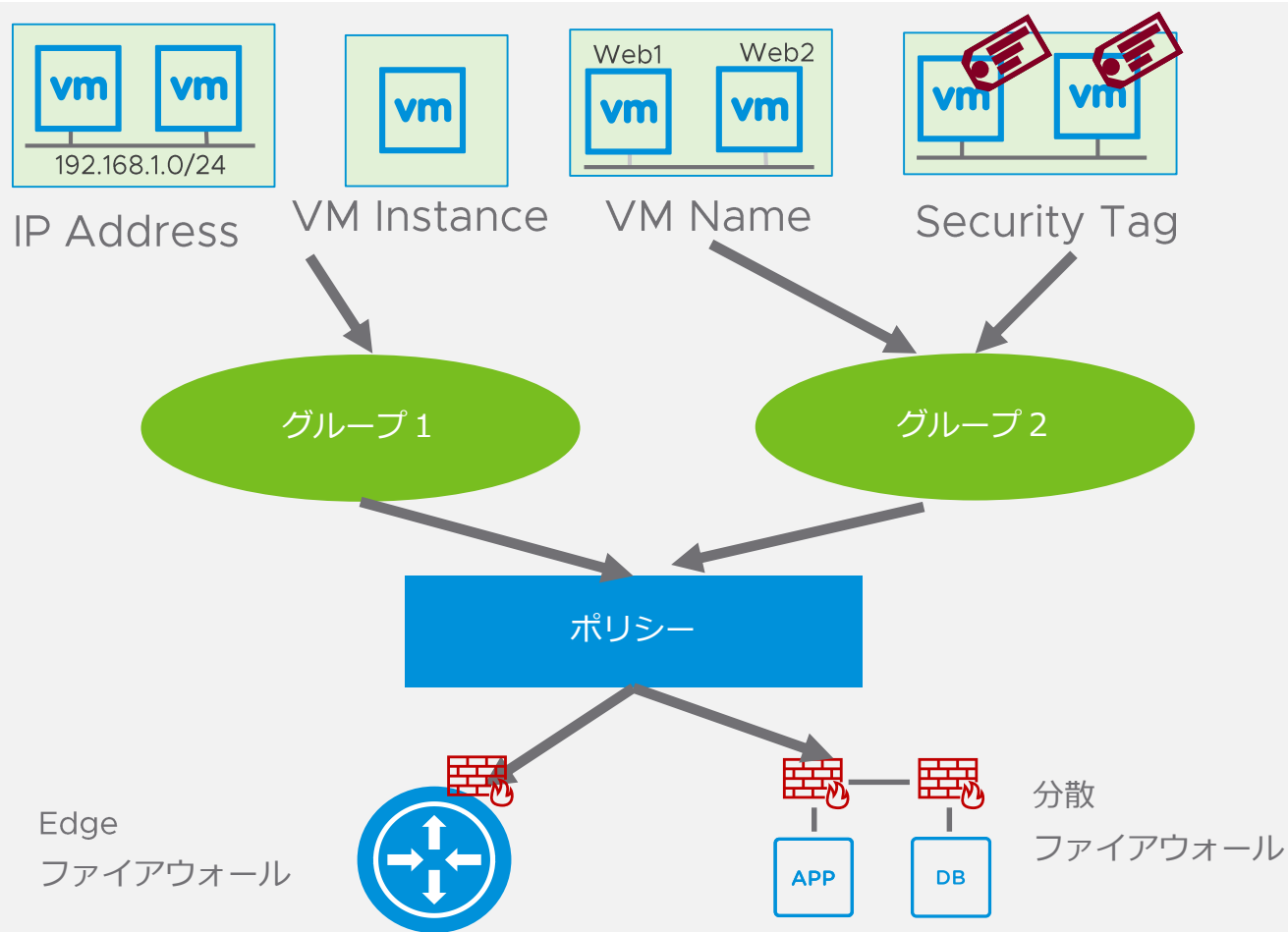
アプリケーションの特性による
シンプルなポリシーベースの定義

- 例) VM 名やユーザ定義のタグなど

簡単にネットワークを隔離

Networking

セキュリティグループベースのファイアウォールポリシー



全てのオブジェクト（マシン名、OS名、IP アドレスなど）をグループ化できる

簡単かつわかりやすくポリシー定義を作成できる

Edge/分散ファイアウォール両方とも定義したグループからポリシーベースのファイアウォールルールを設定できる



Thank You