

DC119

事例から見る、
VMware コンサルタントが明かす
vSAN 活用の最前線と設計ポイント

ヴェイエムウェア株式会社
プロフェッショナルサービス統括本部
コンサルタント 近藤 理恵

#vforumjp

vmware

POSSIBLE
BEGINS
WITH YOU

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

vSAN 活用事例とその効果

vSAN 設計における勘所と設計ポイント

まとめ

Agenda

vSAN 活用事例とその効果

vSAN 設計における勘所と設計ポイント

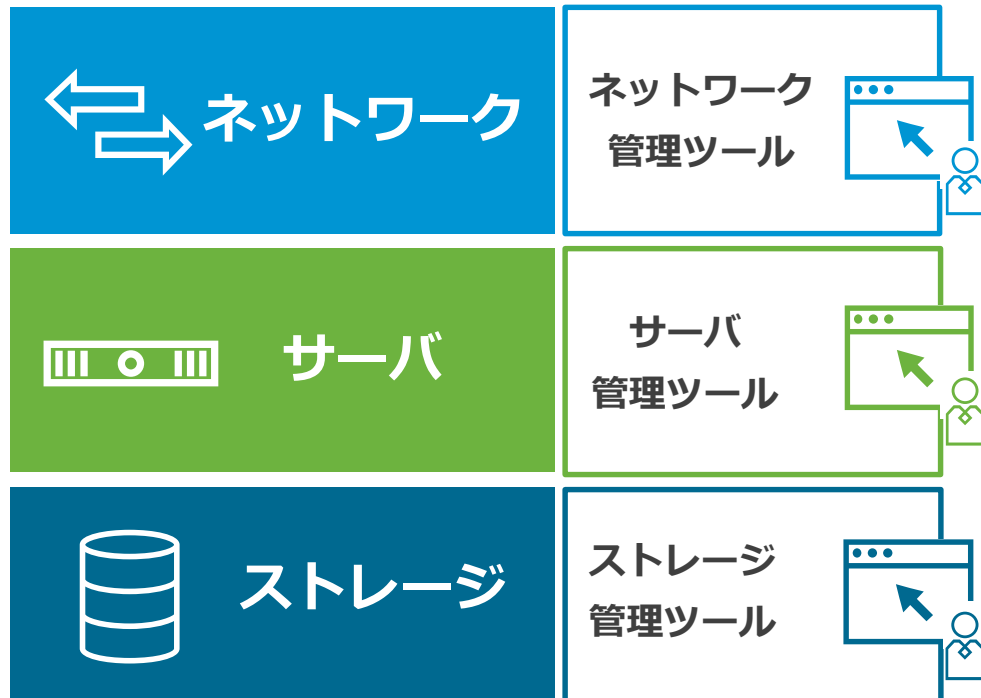
まとめ

ストレージアーキテクチャの潮流

ハードウェアベースからソフトウェアベースのソリューションへの移行

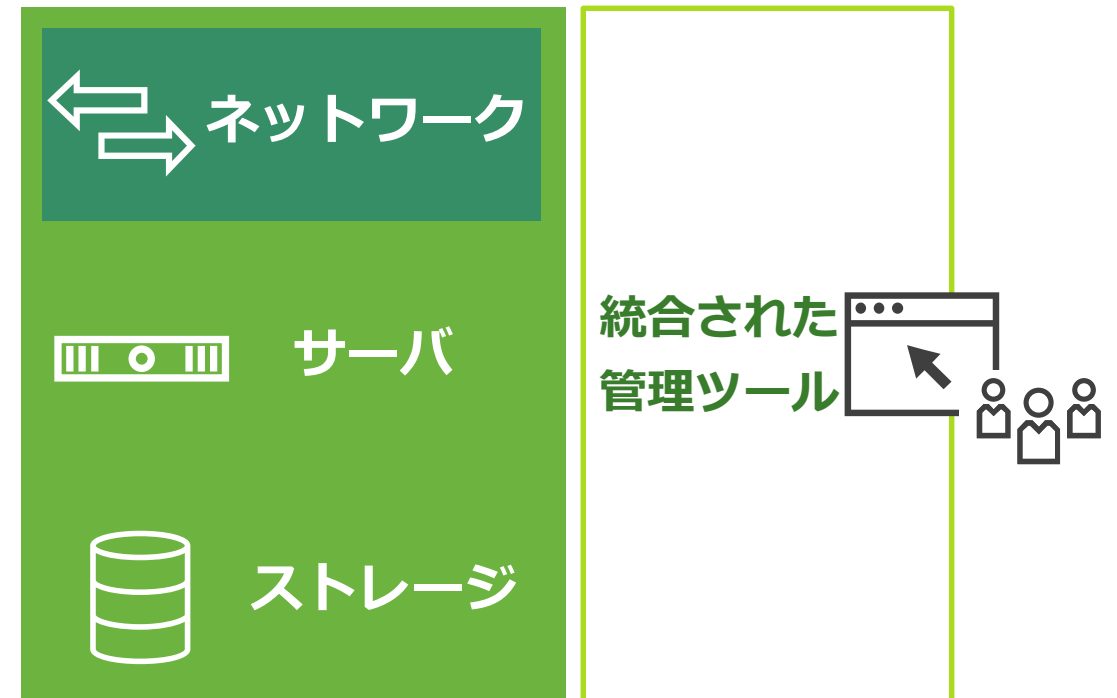
これまでの仮想化環境

各レイヤー毎に機器も管理ツールもサイロ化



Hyper-Converged の環境

Software-Defined Storage (SDS) の技術によりサーバとストレージを統合



※ネットワークは Software-Defined Network (SDN) の技術により統合

ストレージ環境における課題

こんなお悩みありませんか？

- システム更改時の初期投資コストを削減したい

コスト

俊敏性

- 運用や保守のランニングコストを抑えたい

コスト

俊敏性

- 業務停止なしで、システムを簡単に拡張したい

コスト

俊敏性

- 外部ストレージの操作や設定が難しい

俊敏性

信頼性

- サーバ技術者でも簡単に運用できるようにしたい

俊敏性

信頼性

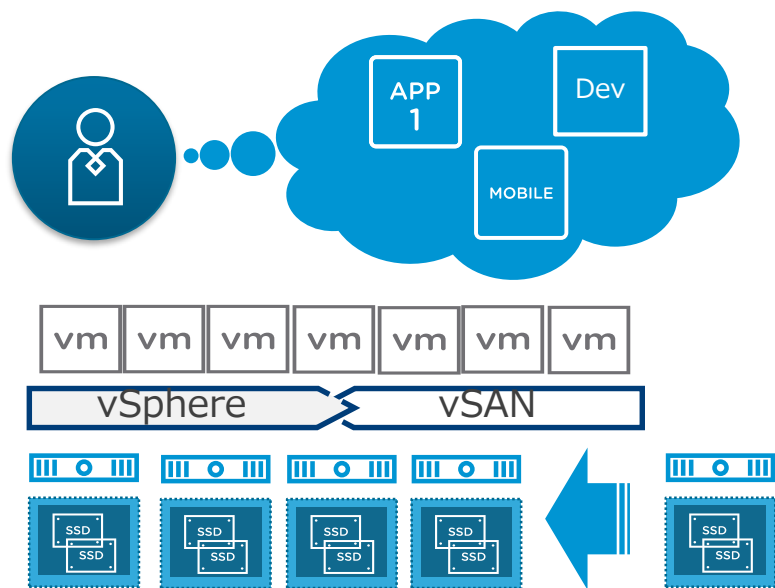
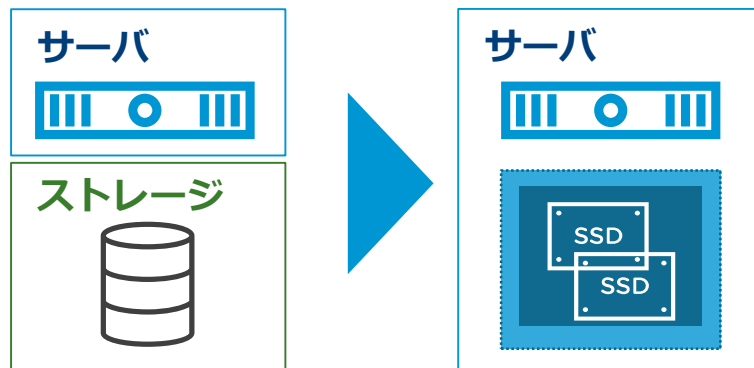
- 災害時のダウンタイムを少なくしたい

信頼性

ビジネスの変化に対応するには、
クラウド環境のように拡張が柔軟に行えて、簡単に運用したい！
それがオンプレミス環境で実現できるといいな



実例1) プライベートクラウド環境



コスト

初期投資・更改コスト削減

ユーザーに提供するクラウド環境は vSAN を利用し **導入費用を削減**
物理ストレージは部分的に利用 (RDM など)

運用コスト削減

物理ストレージ機器の削減
サーバと統合することで保守費用等、 **ランニングコストを削減**

俊敏性

クラウド環境との競合

サーバ単位 (またはディスク単位) のため、 **調達しやすい**
リソース拡張時のダウンタイムなし

サービスの拡充

スモールスタートが可能で、迅速なサービス提供を実現

実例2) 仮想デスクトップ環境

コスト

初期投資・更改コスト削減

デスクトップ環境の高パフォーマンスを
低コストで実現

運用コスト削減

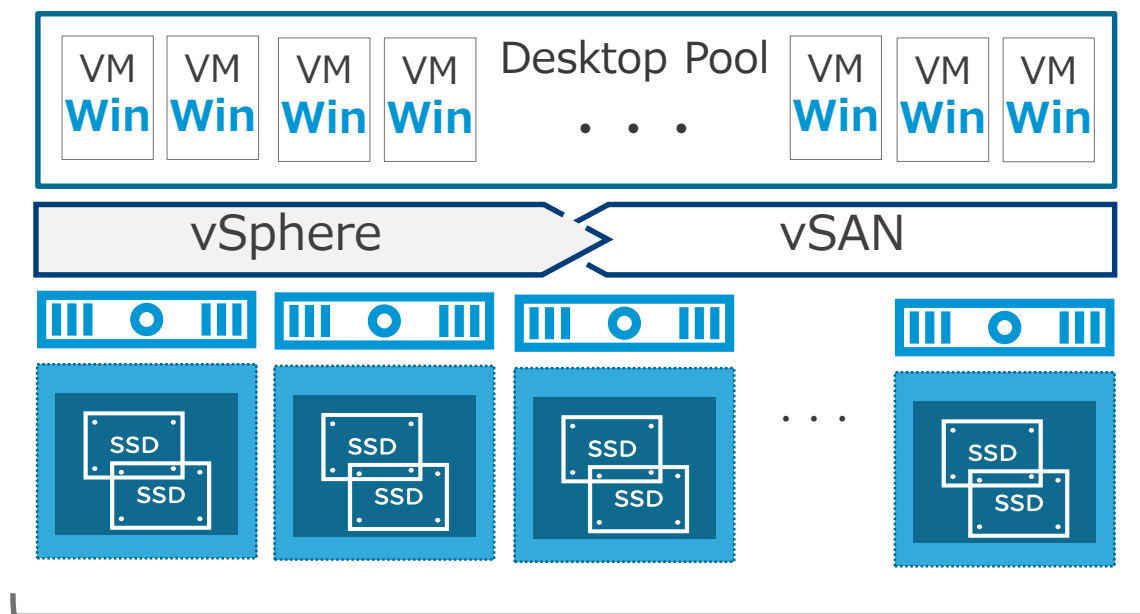
ストレージの**専門知識が不要**
デスクトップ担当者でも運用が容易になる

信頼性

顧客満足度の向上

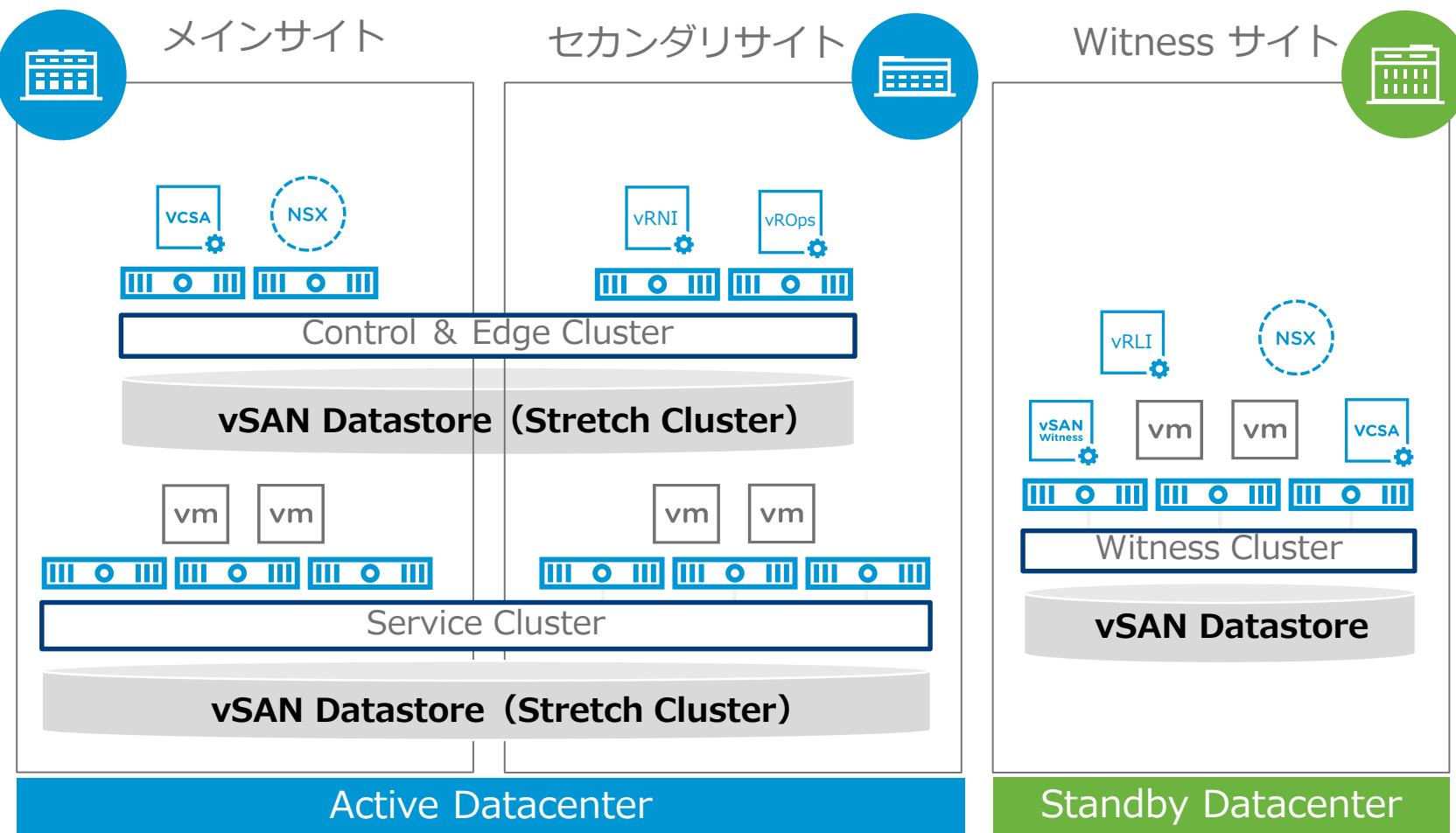
オールフラッシュの採用により
高パフォーマンスを維持

VDI on vSAN



オールフラッシュ採用、重複排除・圧縮の有効化
高パフォーマンスとリソース効率化を実現

実例3) BCP 対応



コスト

初期投資・更改コスト削減

アクティブ・アクティブ DC
構成を**低コスト**で実現

運用コスト削減

DC 間のフェールオーバーを
vSphere HA で実施
サイト切替をシンプル化

信頼性

安全性・信頼性の向上

サイト切替時の **RPO=0**
vSphere HA によるフェールオー
バーによりダウンタイムを軽減

まとめ：vSAN がもたらす、3つの価値

コスト

初期投資・更改コスト削減

- ・ 機器コストの削減

運用コスト削減

- ・ 保守費用の削減
- ・ 運用人員の最適化

俊敏性

クラウド環境との競合

- ・ 運用の標準化
- ・ 柔軟な拡張

サービスの拡充

- ・ 提供スピードの迅速化

信頼性

安全性・信頼性の向上

- ・ 事業継続性の確保

顧客満足度の向上

- ・ 多様化する SLA への柔軟な対応

これらを楽しむには、vSAN の特性を理解した設計が必要です

Agenda

vSAN 活用事例とその効果

vSAN 設計における勘所と設計ポイント

まとめ

vSAN 基本設計の検討ポイント

vSAN 基本設計は、ハード選定にあたる「サイジングフェーズ」と、ハードウェア構成確定後の「vSphere 設計フェーズ」の2つのフェーズに分かれます

サイジングフェーズ

vSAN アーキテクチャ検討

- オールフラッシュ/ハイブリッド

ESXi ブート領域の検討

- ローカルディスク/SD・USB ブート/SATA-DOM

ESXi ログ格納先の検討

- vsantrace
- syslog (Scratch Partition)
- dump file

vSAN ディスク構成

- キャッシュディスク/キャパシティディスク数
- ディスクグループ構成

vSAN ネットワーク構成の検討

- NIC 構成

vSAN クラスタ構成

- ホスト台数
- 管理クラスタ有無
- 許容ホスト障害数と SLA の確認 (RAID 構成)
- 重複排除と圧縮機能の利用有無

vSphere 設計フェーズ

ネットワーク設計

- 仮想スイッチやチーミングポリシーを検討

vSAN クラスタ設計

- HA 設計
- 重複排除と圧縮 (オールフラッシュのみ)
- 保存データの暗号化
- パフォーマンスサービス利用有無

vSAN ストレージポリシー検討

- 許容ホスト障害数、RAID構成を考慮したポリシー設計

ESXi メンテナンスモードオプションの検討

- vSAN データの退避有無

vSAN 障害運用検討

- vSAN コンポーネント再保護の処理開始時間の検討

vSAN 設計はサイジングフェーズに要素が集中
ハードウェアサイジングが重要

vSAN 設計のサイジングフェーズ

ハードウェアサイジングの重要性

なぜ、ハードウェアサイジングが重要なのか？

RFP フェーズで vSAN 設計の大部分が決まる

- ・インフラの RFP 段階で、ハードウェア構成が含まれる
- ・ハードウェア構成変更は費用が発生するため、後から変更しにくい

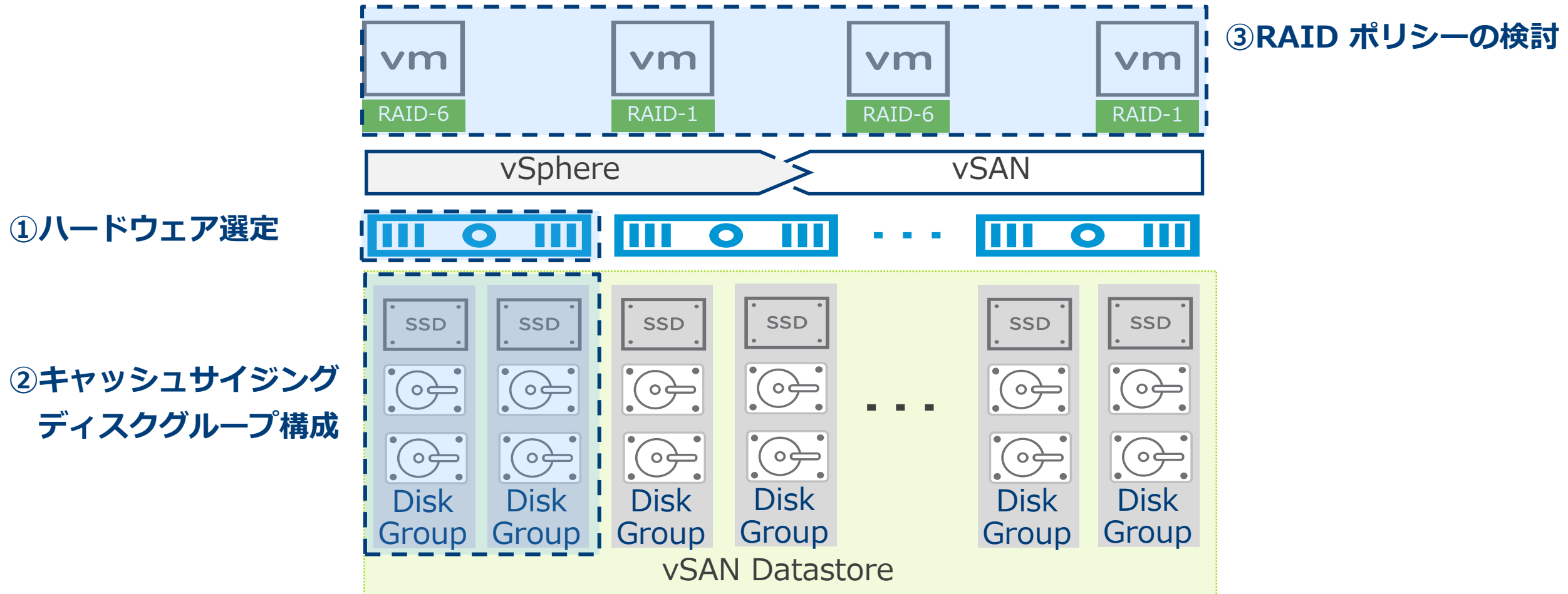
ハードウェア構成で採用できる機能が決まる

- ・重複排除/圧縮機能や RAID-5/6 構成は、オールフラッシュ環境のみ利用可能
- ・RAID 構成により、最小構成台数が異なる（RAID-1：3台、RAID-5：4台、RAID-6：6台）

設計要素を把握せずにハードウェアサイジングを行うと
発注後に有効な機能が十分に使えない等、本来の要件を満たせない可能性があります

vSAN 設計のサイジングフェーズでのポイント

vSAN の検討ポイント



設計ポイント①：ハードウェア選定

VMware vSAN ReadyNode™ 構成の推奨

vSAN ReadyNode で構成することを推奨

vSAN ReadyNode を推奨する理由

ハードウェアの互換性チェックはとても大変です

ハードウェアの各パーツ（サーバ、アレイコントローラー、ディスク、CPU、メモリなど）やファームウェア、ドライバ、すべての互換性を確認する必要がある

vSAN ReadyNode はこれらの互換性をすべてクリアしているため、**事前の互換性チェックが不要**

参考：vSAN ReadyNode とは？

- vSAN ReadyNode とは、サーバベンダー様より提供されている、vSAN の推奨構成
- vSAN と各ベンダー様のハードウェアを組み合わせ、テスト済みの構成のことを示す

ただし、こんな質問もよく耳にします

vSAN ReadyNode のモデルだと、CPU もメモリも足りないし、スペックもあげたい。その場合はどうしたら？？



設計ポイント①：ハードウェア選定

vSAN ReadyNode のカスタマイズ可能なところ、そうでないところ

CPU <ul style="list-style-type: none">同等スペックまたはより良いCPU クロック速度の高いコア数をサポート異なるプラットフォームの世代に関してはサポートされないので注意が必要	キャパシティデバイス <ul style="list-style-type: none">耐久性、性能のともに、同等スペック、またはそれ以上のものをサポート	アレイコントローラー <ul style="list-style-type: none">変更不可（ReadyNode モデルの構成リストに記載されているもの）コントローラーは vSAN 互換性ガイドで指定されたファームウェア、ドライババージョンをサポート
メモリ <ul style="list-style-type: none">メモリ追加可能	NIC <ul style="list-style-type: none">空きスロットがあれば追加可能ESXiとして互換性があること同等スペック、またはそれ以上ものをサポート（25Gや40G）	SAS エクスパンダー <ul style="list-style-type: none">変更不可（ReadyNode モデルの構成リストの注釈に認定有無を記載）
キャッシュデバイス <ul style="list-style-type: none">耐久性、性能のともに、同等スペック、またはそれ以上のものをサポートハイブリット、オールフラッシュのキャッシュ要件（比率）を満たしていること	ブートデバイス <ul style="list-style-type: none">ブートデバイスの変更可能ESXi としてサポートされていること	ストレージプロトコル <ul style="list-style-type: none">プロトコル変更はサポートされない Ex.SAS ドライブ→NVMe など変更不可

最近、日本語KBもできました：<https://kb.vmware.com/s/article/52084?lang=ja>

設計ポイント①：ハードウェア選定

vSANハードウェア選定時のポイント

◎ vSAN ハードウェア選定では、**必ず** vSAN ReadyNode ベースで構成すること

- ・ 一からパーツを組み上げることも可能ですが、ファームウェアやドライバ含む、vSphere、vSAN の互換性を全て確認し、**適切な構成に組み上げるのは非常に大変**です
- ・ 仮に、互換性のない構成で組んだ場合、エラー検知やパフォーマンスなどに影響を与える可能性があり、**基盤として健全性を担保できません**

Tips : vSAN ReadyNode Configuration

1. vSAN のハードウェアを選択するときは、まずは最新の vSAN ReadyNode チェック <http://vsanreadynode.vmware.com/RN/RN>
2. リストより、性能や容量を満たすもの、または近いスペックのものを選択
3. その後、前頁の KB (KB52084) を参考に、不足するリソース等を追加・変更する

設計ポイント②：キャッシュサイジング/ディスクグループ構成

キャッシュデバイスのサイジング

1 つのディスクグループ（以後、DG）に対して、1つの SAS または SATA 半導体ディスク (SSD) または PCIe フラッシュ デバイスが必要です

- ハイブリット構成の場合

- 予想される仮想ディスクサイズの合計の **10 %以上**を推奨

- この値は許容する障害の数を考慮する前の使用量をもとに計算します

- 10 % は一般的に推奨値であり、10 % では多すぎる場合もあれば、十分ではない場合もある

- オールフラッシュ構成の場合

- ユースケースおよびパフォーマンスの要件に基づいて検討が必要

IO 性能/IO 負荷	AF-8 80K IOPS/ホスト	AF-6 50K IOPS/ホスト	AF-4 25K IOPS/ホスト
Rad/Write 70:30 ランダム	800 GB	400 GB	200 GB
Write 30%以上 ランダム	1.2 TB	800 GB	400 GB
Write 100% シーケンシャル	1.6 TB	1.2 TB	600 GB

参照URL : <https://blogs.vmware.com/virtualblocks/2017/01/18/designing-vsan-disk-groups-cache-ratio-revisited/>

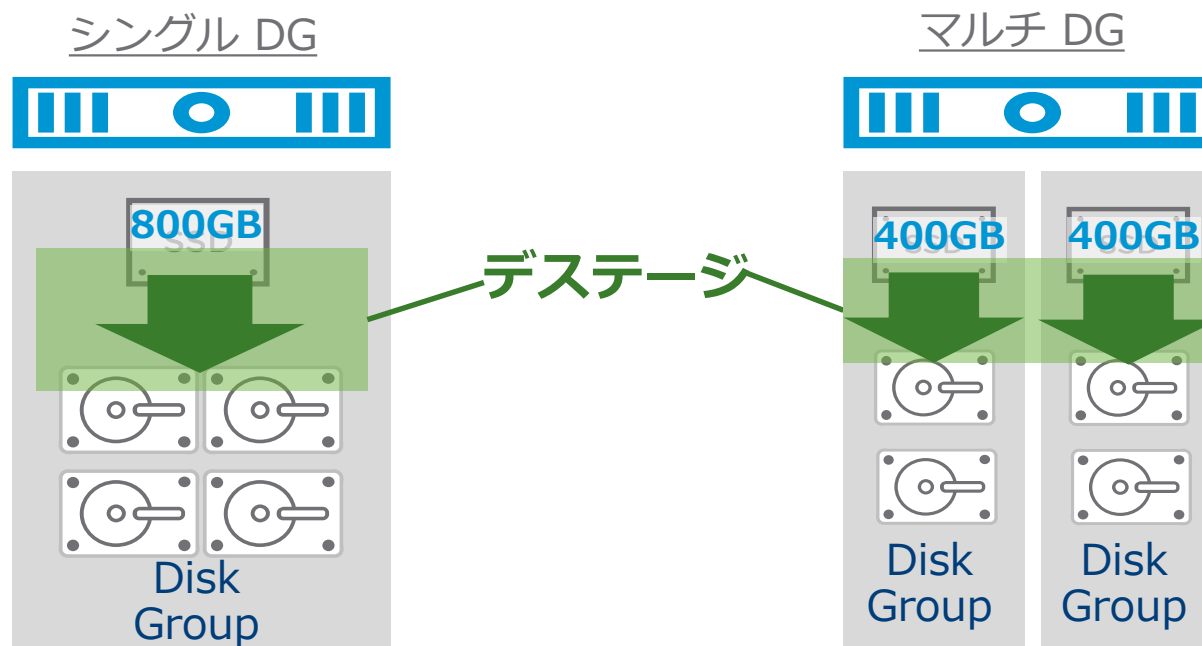
設計ポイント②：キャッシュサイジング/ディスクグループ構成

ディスクグループ構成

◎ ディスクグループは複数分割して構成するのが効果的

理由①：パフォーマンスが向上する

- ・ ディスクグループあたりのディスク容量が大きい場合



デステージ処理が分散される

※デステージ：キャッシュディスク上のデータをキャパシティディスクに書き込む処理

設計ポイント②：キャッシュサイジング/ディスクグループ構成

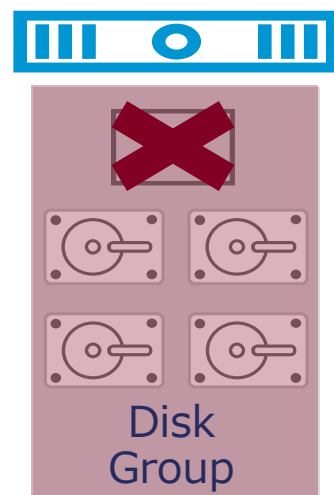
ディスクグループ構成

理由②：障害影響範囲の縮小化

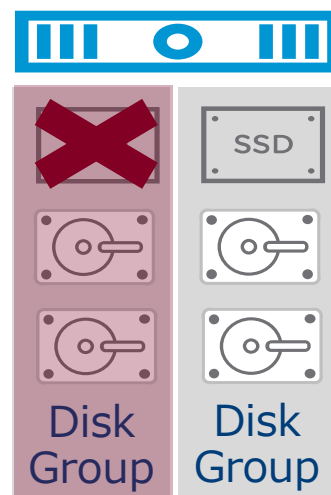
- ディスクグループあたりのディスク容量が大きい場合
- 重複排除/圧縮機能を有効化する場合

キャッシュデバイス障害

シングル DG

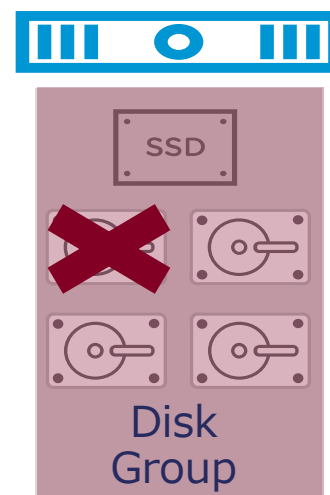


マルチ DG

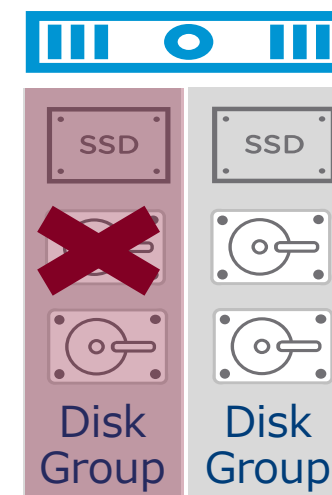


キャパシティデバイス障害（重複排除/圧縮機能の有効化時）

シングル DG



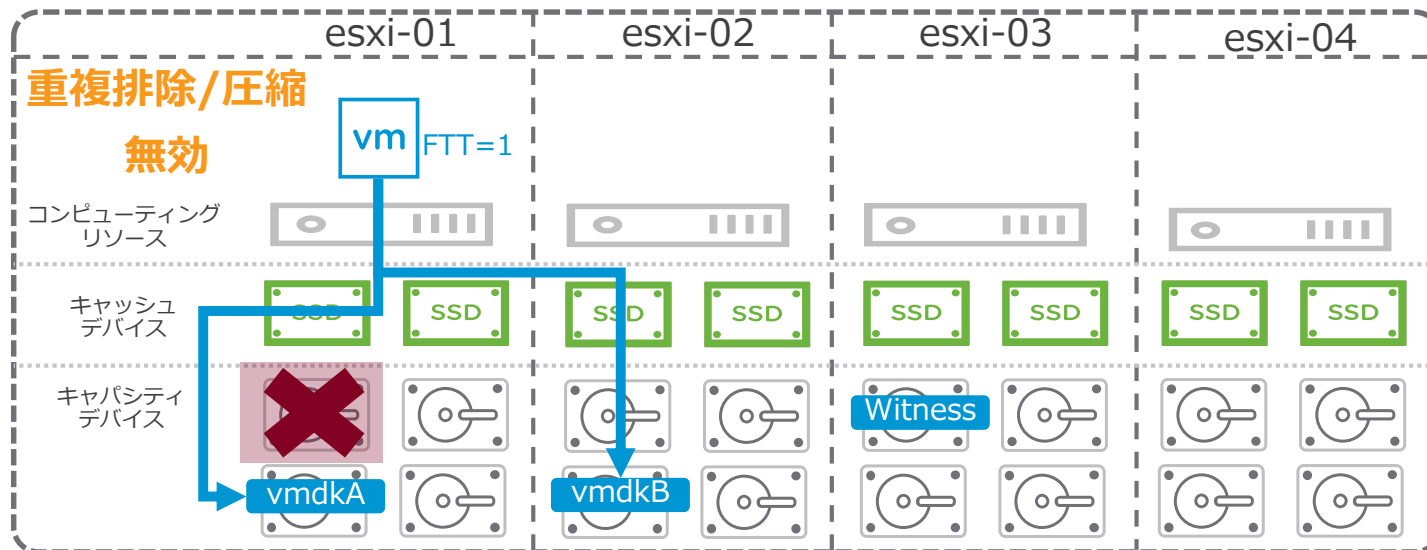
マルチ DG



キャッシュデバイス障害時の影響範囲が少なくなる

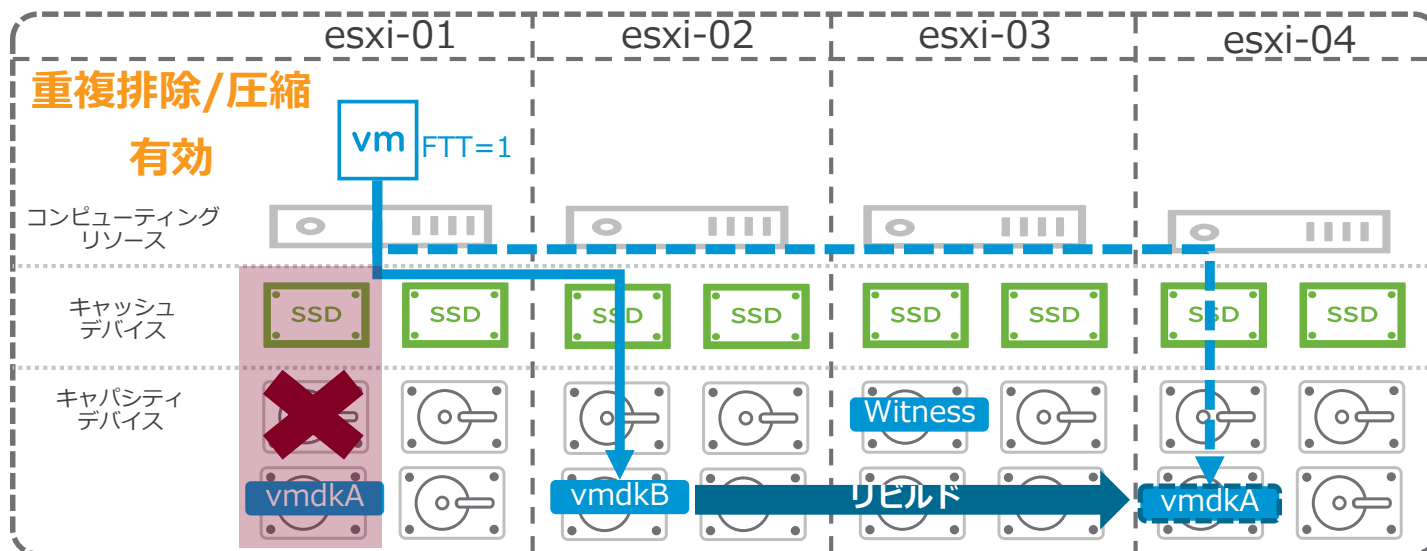
重複排除/圧縮 有効時は、キャパシティデバイス障害時の影響範囲も軽減

(参考) vSAN ディスク障害時の挙動



重複排除/圧縮 無効時

通常（重複排除/圧縮機能 無効時）は、障害が発生したキャパシティデバイスのみがリビルド（再保護）対象となる



重複排除/圧縮 有効時

重複排除および圧縮がディスクグループ単位で実行される

そのため、重複排除/圧縮機能 有効時は、障害が発生したキャパシティデバイスが所属するディスクグループ全体がリビルド（再保護）対象となる

設計ポイント②：キャッシュサイジング/ディスクグループ構成

ディスクグループ構成

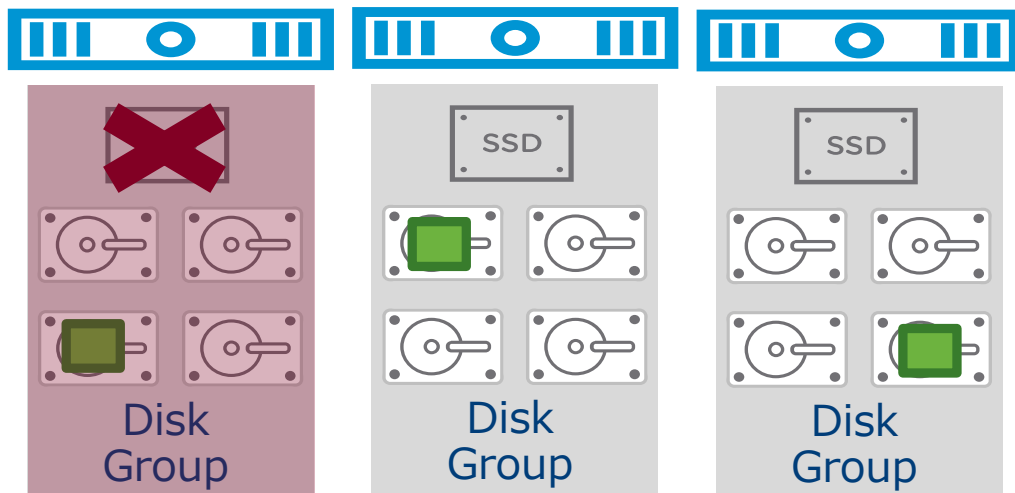
理由③：キャッシュデバイス障害時の保護状態を維持

- ROBO の2ノード構成、3ノード構成など、vSAN ノードが最小構成で構成する場合
 - RAID -1構成の場合、3台のホストにコンポーネントが分散配置されるため、構成可能な最小構成は3ノード

シングル DG



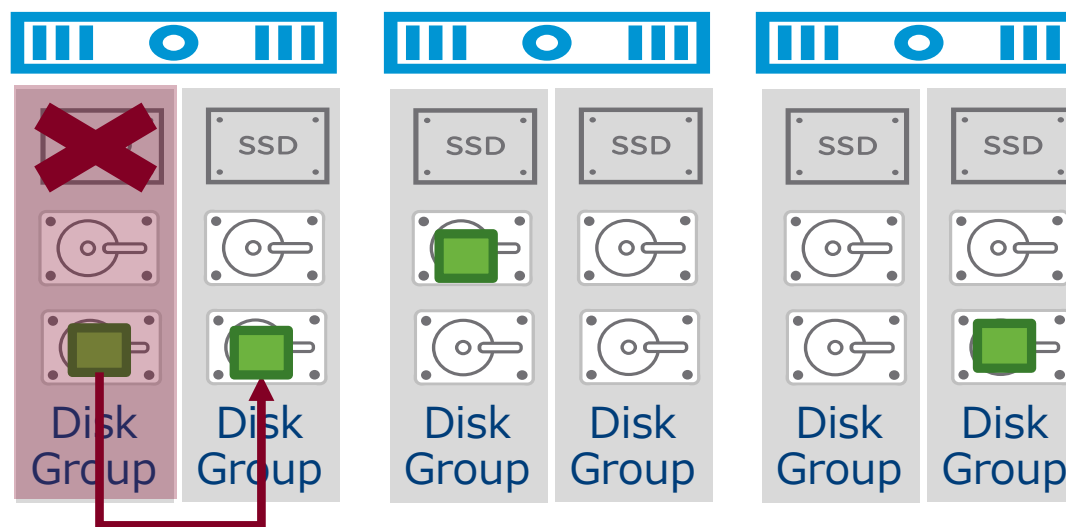
ディスクグループが単一のため、
退避可能なディスクが同一ホストに存在しない
ため、再保護（リビルド）されない



マルチ DG



キャッシュデバイスが障害となっても、
他のディスクグループが同一ホストに存在するため、
再保護（リビルド）が可能



設計ポイント③：RAID ポリシーの検討

RAID-6の使いどころ

◎ サーバ環境における、最近のトレンドはオールフラッシュ+RAID-6 構成

なぜ、RAID-6 の採用が増えているのか？

高 SLA への対応	ホストの2台障害まで許容可能
リソースの効率化	RAID-1の FTT= 2 構成（データを3重で保持する構成）だとディスク容量を消費しすぎてしまう
メンテナンス中の耐障害性を考慮	ESXi のメンテナンスモードの切り替え中は、ホスト1台分縮退状態となるため、メンテナンス中についても耐障害性を担保するため ※詳細次頁参照
性能評価	PoC で性能評価したところ、想定よりも性能劣化しなかった ※一般的なRAID構成と同様に、RAID-1と比較すると、RAID-5/6は性能が落ちる傾向にあるため

オールフラッシュ環境で、vSAN クラスタのホスト数が6台以上の構成が可能な場合は RAID-6の採用をご検討ください (RAID-6 構成は最小構成台数6台となります)

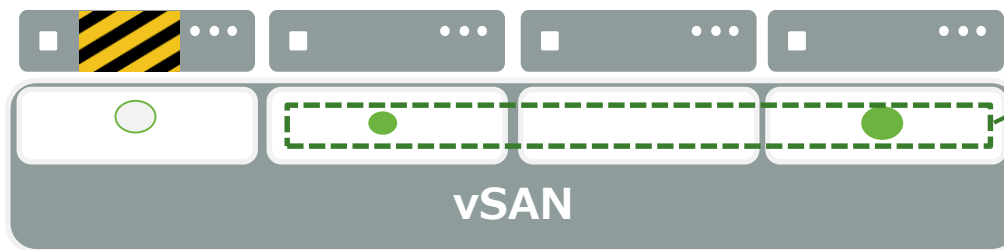
(参考) メンテナンス時の耐障害性の考慮

ホストメンテナンス中、ホスト上にある vSAN データはどうなるのか？

- ・ ホストメンテナンスモードの状態は、一時障害モードとして稼働します（ステータス：Absent）
- ・ RAID1 以上で構成されている場合は、
他のホストに配置されているデータで読み書きを行うため、仮想マシンは停止しません

メンテナンス中のコンポーネント配置

VM object FTT=1 ● Accessible



メンテナンス中のホスト以外に配置されている
Accessibleなコンポーネントのデータにより
読み書きを行う

ただし、ホストのメンテナンス中は**保護状態なし**の縮退稼働となります

メンテナンス中も耐障害性を維持するためには、以下2つのどちらかを選択する必要があります

- ・ FTT（許容する障害の数）を2以上とする（RAID-1でFTT=2以上、または **RAID-6 で構成**する）
- ・ メンテナンスモード切替時にホストのデータを移行させる（メンテナンスモードのオプションで選択）

番外編

VDI 環境における設計ポイント

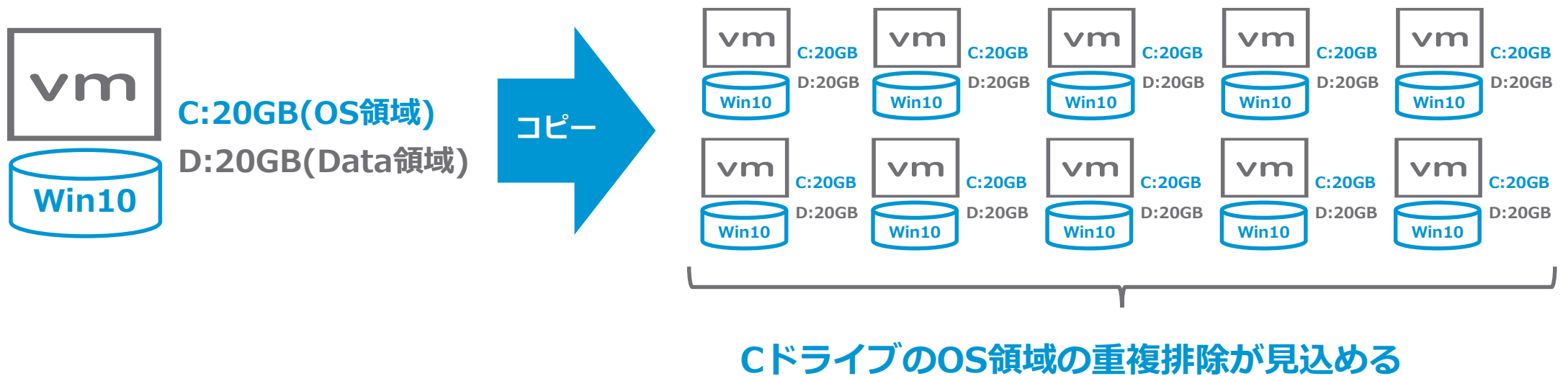
ストレッチクラスタにおける設計ポイント

番外編：VDI 環境における設計ポイント

VDI 環境では、重複排除/圧縮機能の利用を推奨

- ・仮想デスクトップ環境では、各デスクトップはテンプレートとなる仮想マシンからクローンにより生成されるため、重複排除率が高く、効果が期待できる

フルクローンによる展開



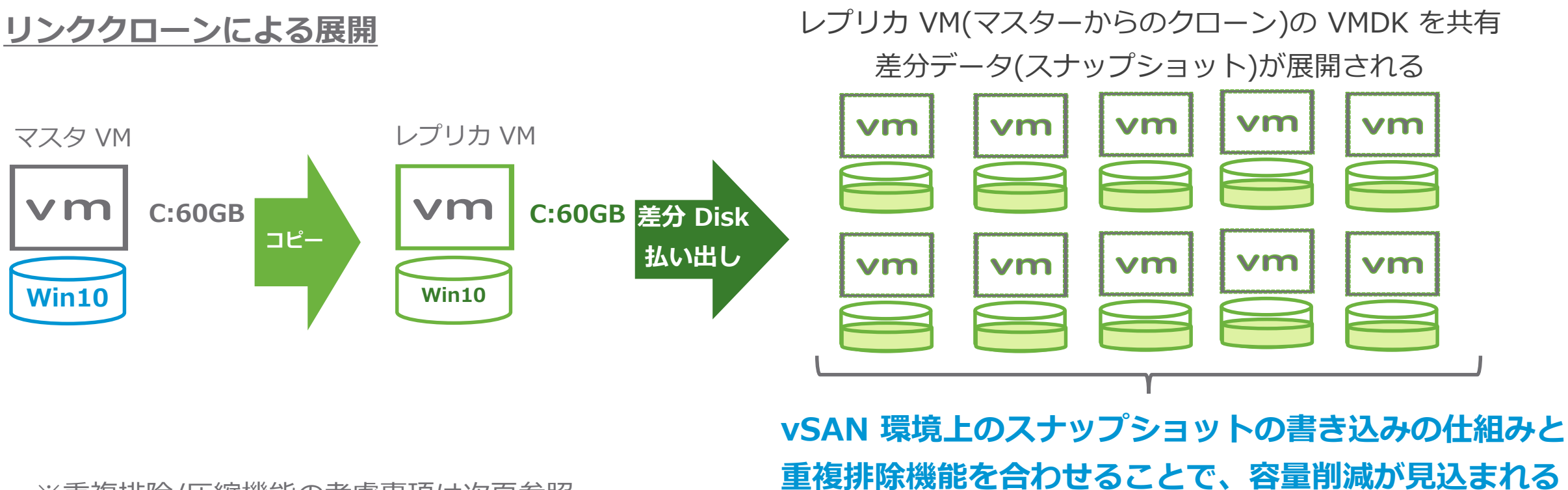
参照URL : VMware Horizon® 7 on VMware vSAN™ Best Practices TECHNICAL WHITE PAPER

<https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/vsan/vmware-horizon-7-on-vmware-vsan-best-practices.pdf>

番外編：VDI 環境における設計ポイント

重複排除/圧縮機能は、一般的に前頁のフルクローン構成での利用が効果的とされているが、リンククローンやインスタントクローン環境においても利用を推奨

リンククローンによる展開



※重複排除/圧縮機能の考慮事項は次頁参照

参照URL：VMware Horizon® 7 on VMware vSAN™ Best Practices TECHNICAL WHITE PAPER

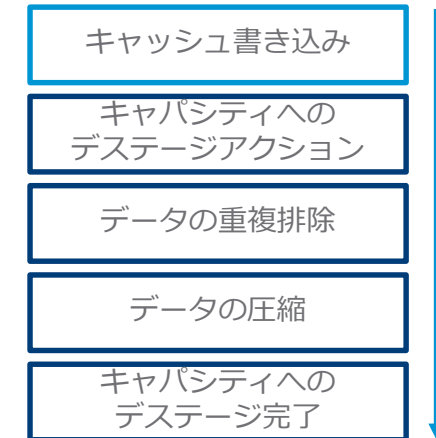
<https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/vsan/vmware-horizon-7-on-vmware-vsan-best-practices.pdf>

(参考) 重複排除/圧縮機能について

重複排除/圧縮 有効時の動作

- ディスクグループレベルで実行
 - クラスタ単位で有効化設定が可能
 - 重複排除とデータ圧縮はディスクグループレベル実行
- キャッシュからキャパシティへデステージング時に重複したデータを除外
- 重複排除後、デステージされる前にデータが圧縮される
 - 既にキャパシティにデステージされているデータは対象外

重複排除/圧縮の流れ



重複排除/圧縮機能の有効時の考慮

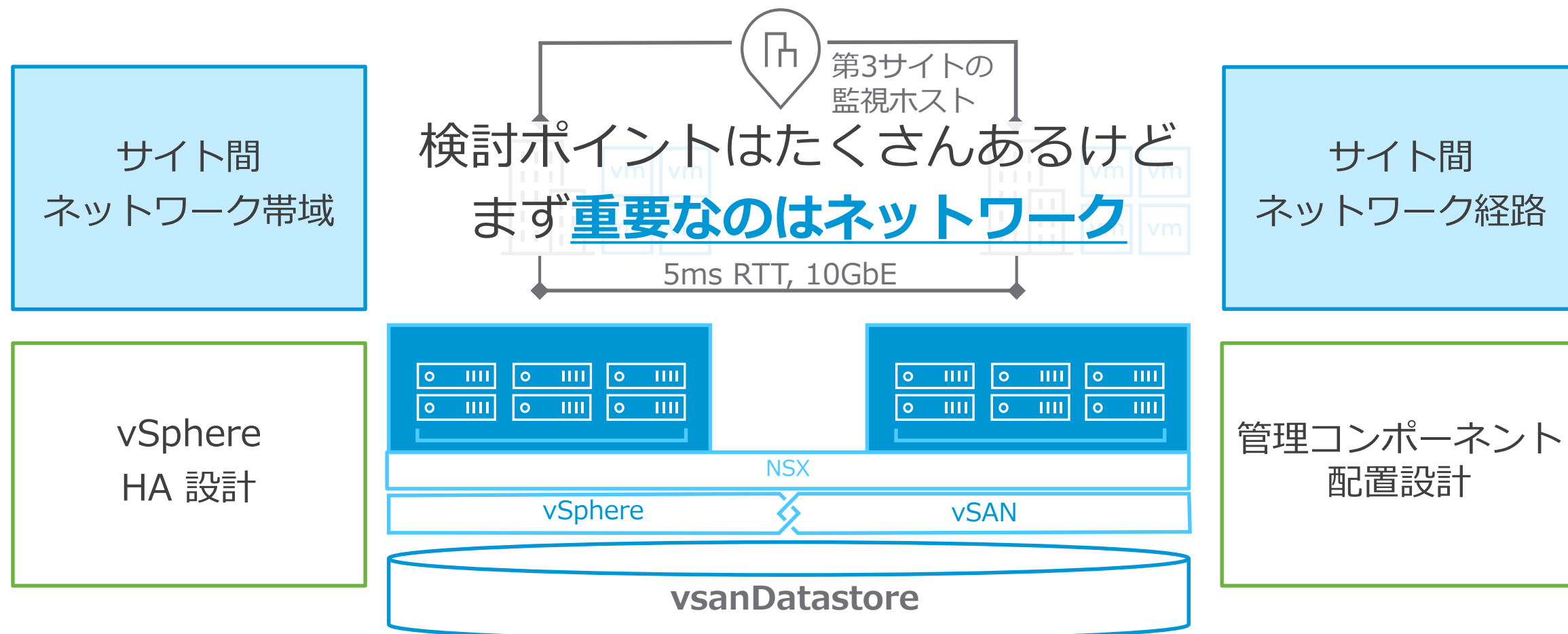
- 一般的なストレージと同様に、無効時と比較するとパフォーマンスの低下が想定されます
重複排除/圧縮を採用する場合には、以下を考慮ください
- パフォーマンスを考慮したサイジング
 - 事前のパフォーマンステストの実施

パフォーマンス要件を満たしているかご確認の上、ご採用ください

番外編：ストレッチクラスタにおける設計ポイント

vSAN ストレッチクラスタ

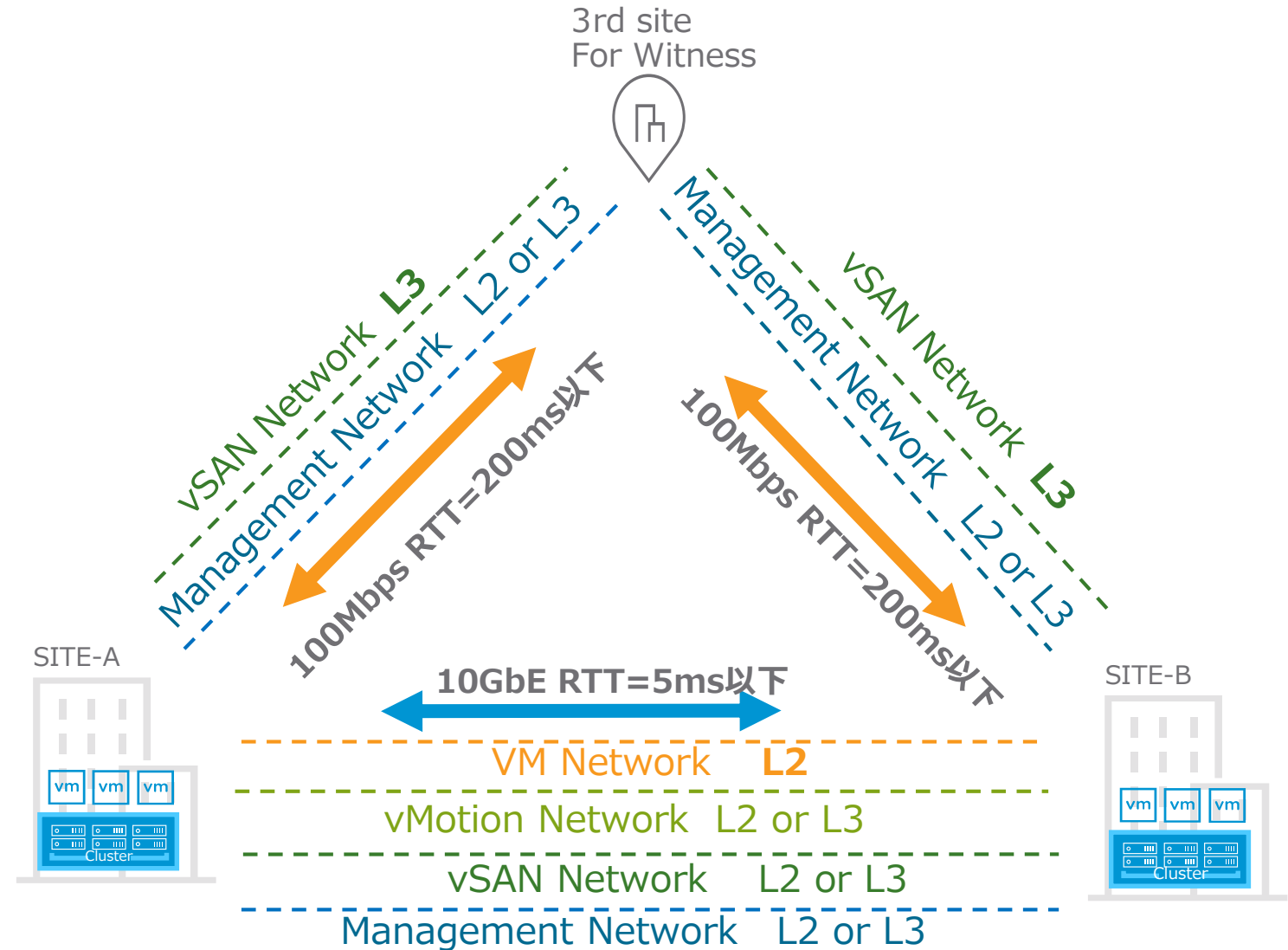
- vSAN データストアをサイト間にて延伸させることで、アクティブーアクティブ DC 構成を可能とする
- サイト障害時も RPO はゼロ、RTO の最小化が可能



番外編：ストレッチクラスタにおける設計ポイント

サイト間ネットワーク要件

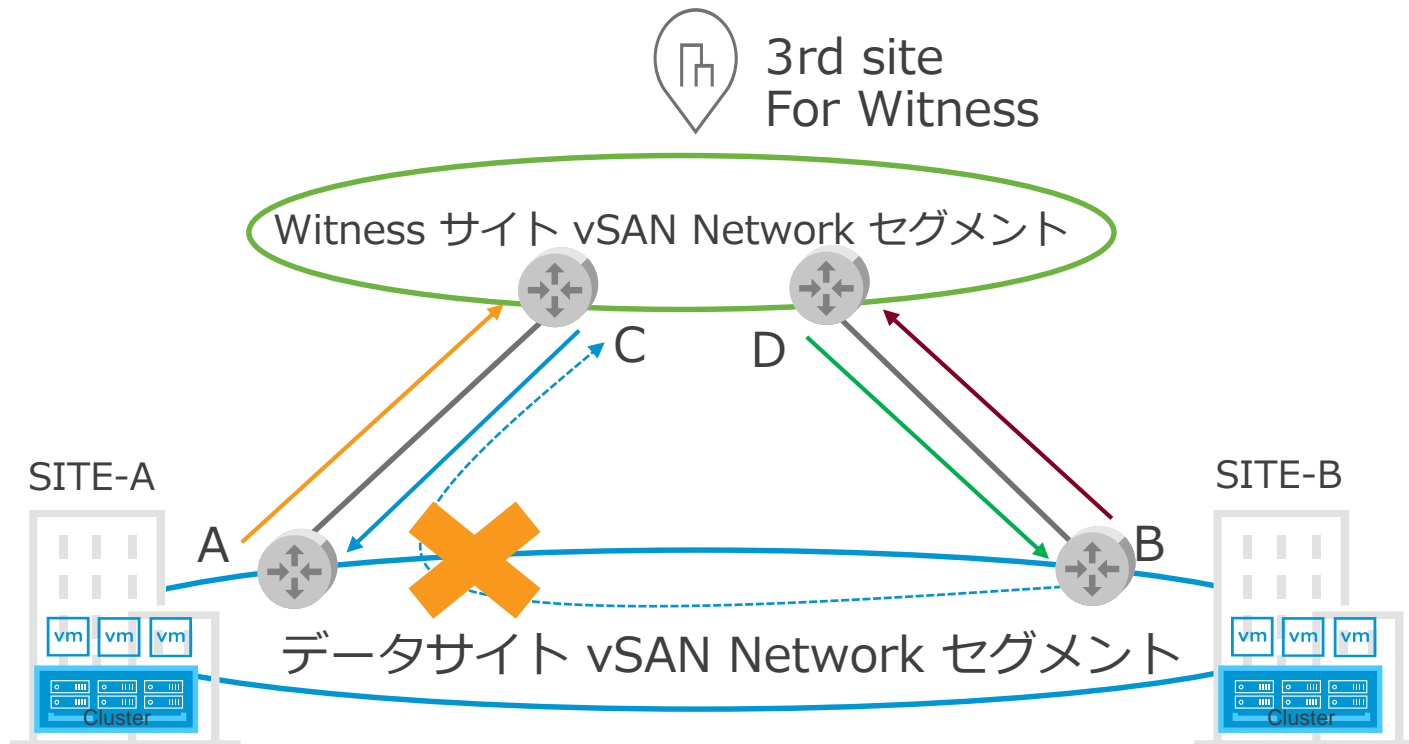
- データサイト間：
5ms RTT 以下の遅延と10GbE の帯域幅で接続(vSAN Networkは性能の向上のため専用10GbE が望ましい)
- 監視ホスト(Witness)：
200ms RTT 以下の遅延と100Mbps の帯域幅をサポートする任意の場所に配置可能
 - 各サイトのホストが11台以上の場合は100ms RTT
- vSAN Network：
専用の Vmkernel と専用の subnet を使用
- 各サイト-監視ホスト間 vSAN Network：L3 で接続することが推奨
- VM Network：
サイト間でVMが移動した場合にも同じ IP アドレスを使用する場合は、L2 で接続が必要



番外編：ストレッチクラスタにおける設計ポイント

サイト間のネットワーク経路

- データサイトと監視サイト間の通信は、必ずスタティックルートを Vmkernel に設定
 - SITE-B の通信がデフォルト GW 設定のままだと SITE-A 側のルーター経由となり、SITE-A の障害時に検知が正常にできなくなる



- SITE-A のホスト
→ Witness サイトへの通信は GW-A を経由
- SITE-B のホスト
→ Witness サイトへの通信は GW-B を経由
- 監視ホスト
→ SITE-A への通信は GW-C を経由
→ SITE-B への通信は GW-D を経由

Agenda

vSAN 活用事例とその効果

vSAN 設計における勘所と設計ポイント

まとめ

まとめ

vSAN の特性を理解した設計により、vSAN の効果を最大化できます

サイジングフェーズ

ハードウェアサイジングが重要

- 必ず **vSAN ReadyNode** ベースで構成すること
- **複数ディスクグループでの構成**が効果的
- オールフラッシュ構成で **RAID-6 採用**がトレンド

vSphere設計フェーズ

vSAN 設計の他、運用管理やテストを含め、**PSO** でご支援可能

番外編

VDI 環境・ストレッチクラスタ構成編

- VDI 環境では重複排除/圧縮の採用がトレンド
- ストレッチクラスタではネットワークが重要

導入および運用に関する支援のご相談は、お近くの VMware コンサルタントまで

本セッション受講の方へのお勧め

DC106

16:25~
Room B

徹底解説!! vSphere 6.7 Deep Dive!!

NS308

14:40~
Room D

**vSANとNSXをフル活用で実現した、
次世代マルチサイトデータセンターの中身とは？**

NS163

17:10~
Room I

**コンサルタントが語る、
マルチデータセンター設計における NSX 活用のツボ**

ご清聴、ありがとうございました。