

DC110

# vSphere Persistent Memory

～次世代メモリ技術 NVDIMM と vSphere 6.7 ～  
「バイト・アドレスابل」に進化する不揮発性メモリ技術の活用

グイエムウェア株式会社

ストラテジック・アライアンス  
プリンシパル テクニカル・アライアンス・マネージャ

齋藤 康成

#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。

# Brand Name / Approved Short Name

BRAND NAME	APPROVED SHORT NAME
VMware vSphere®	vSphere
VMware ESXi™	ESXi
VMware vSphere® vMotion®	vSphere vMotion
VMware vSphere® Storage vMotion®	vSphere Storage vMotion
VMware vSphere® Distributed Resource Scheduler™	vSphere DRS
VMware vSphere® Fault Tolerance	vSphere FT

## Agenda

記憶デバイスの進化

vSphere Persistent Memory

PMEM の動作モードとパフォーマンス

業界における VMware の取り組み

まとめ

# Agenda

## 記憶デバイスの進化

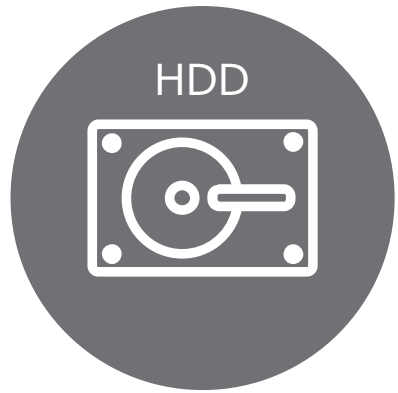
vSphere Persistent Memory

PMEM の動作モードとパフォーマンス

業界における VMware の取り組み

まとめ

# 記憶デバイスの変遷



永続的

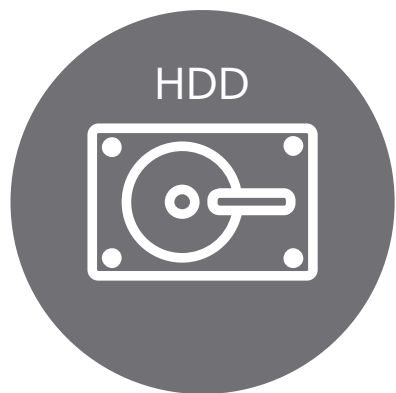
低速



一時的

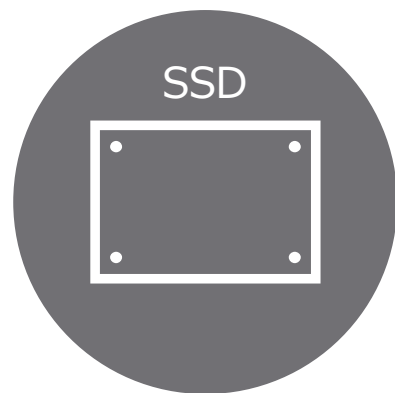
高速

# 記憶デバイスの変遷



永続的

低速



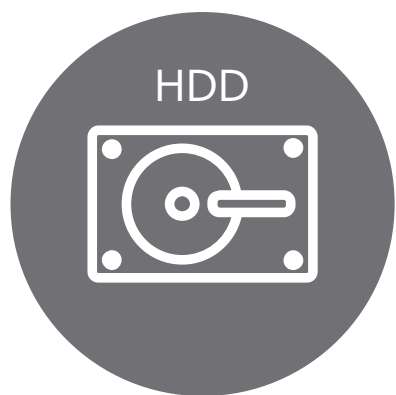
永続的



一時的

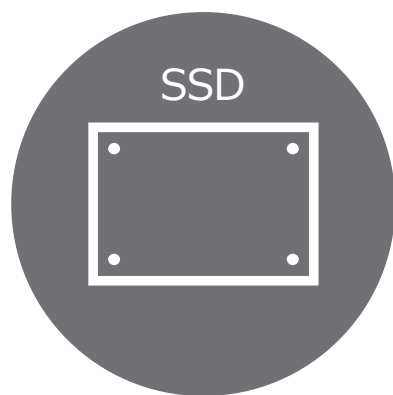
高速

# 記憶デバイスの変遷



永続的

低速



永続的



永続的



一時的

高速

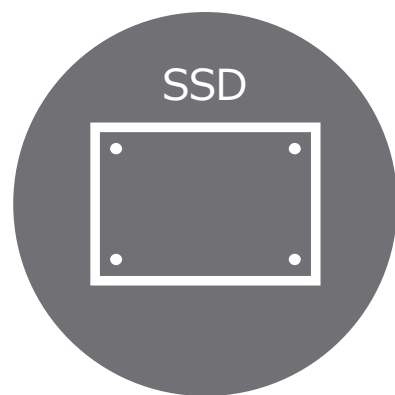


# 記憶デバイスの変遷



永続的

低速



永続的



永続的



永続的



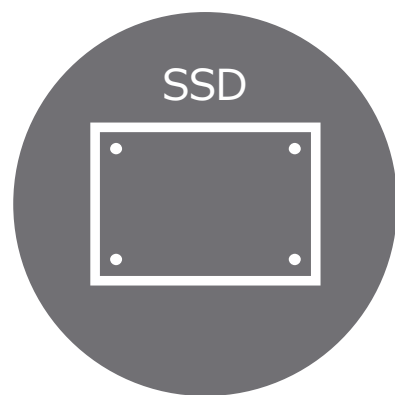
一時的

高速

# 記憶デバイスの変遷



永続的



永続的



永続的



永続的

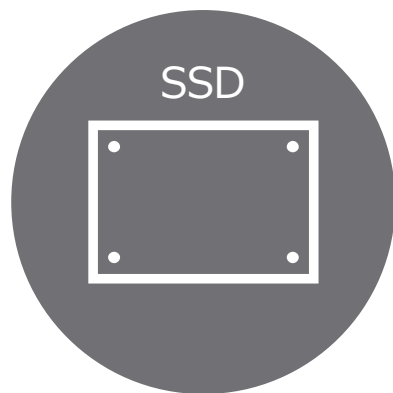
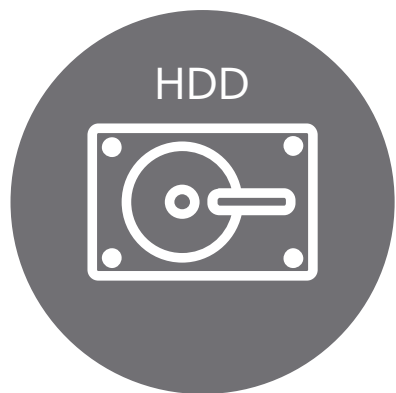


一時的

CPU から遠い

CPU に近い

# 記憶デバイスの変遷

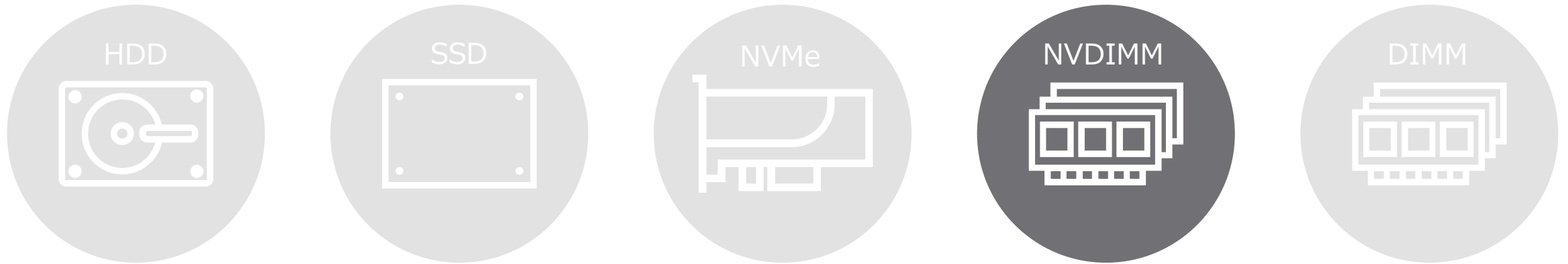


ブロックモードアクセス

ランダムアクセス

バイト・アドレスابل

# 記憶デバイスの変遷



DIMM スロットに装着される不揮発性メモリ  
高速・低遅延であるにも関わらず永続的  
バイト・アドレスサブル  
vSphere 6.7 よりサポート

# Agenda

記憶デバイスの進化

## **vSphere Persistent Memory**

PMEM の動作モードとパフォーマンス

業界における VMware の取り組み

まとめ

# vSphere Persistent Memory

## 永続的メモリデバイスを vSphere 環境で活用

- メモリデバイスとしての特性
  - DIMM スロットに装着
  - 高速、低遅延
  - バイト・アドレスサブル
- 不揮発性メモリとしての特性
  - 永続的記憶領域として利用可能
- vSphere 6.7 の新機能の 1 つ



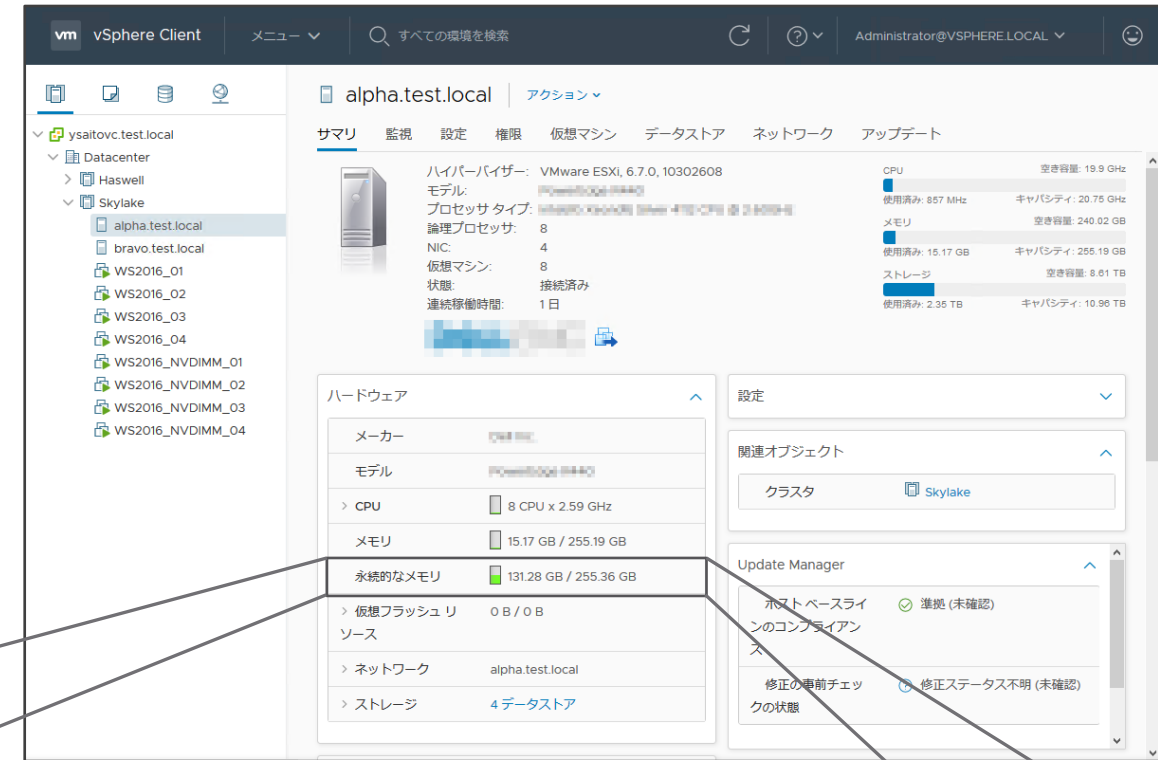
## NVDIMM (≡ PMEM)

正確には NVDIMM は PMEM の実装形態の 1 つであるが、本セッションでは原則として NVDIMM と PMEM は同一の意味で用いる。

# Persistent Memory (PMEM) の認識

ESXi は PMEM を自動的に  
認識・構成

- Free PMEM Area を有する  
ホスト上で ESXi 6.7 を起動
- ESXi は PMEM Datastore  
を自動構成



永続的なメモリ



131.28 GB / 255.36 GB

# 自動構成された PMEM Datastore

```
[root@localhost:~] ls -l /vmfs/volumes
total 3080
drwxr-xr-x    1 root    root           8 Jan  1  1970 31b7b67c-bc1ac09a-7fea-4b76c50b2d3d
drwxr-xr-x    1 root    root           8 Jan  1  1970 49f7e377-a1eb795c-5068-70ac7ab52839
drwxr-xr-x    1 root    root           8 Jan  1  1970 5a376463-655aa289-11a6-e41f13bc7b60
drwxr-xr-t    1 root    root       73728 Dec 18  06:47 5a37646d-1b00f39d-48ed-e41f13bc7b60
drwxr-xr-x    1 root    root           8 Jan  1  1970 5a376470-ae69b1cf-ca29-e41f13bc7b60
drwxrwxrwx   32 root     1         4096 Nov 16  00:10 5f5d7d04-559cb55f
lrwxr-xr-x    1 root    root          40 Dec 19  02:37 PMemDS-3ae1736f-dcb7-0d4f-beac-bf9a0
dc281d3 -> pmem:5a376e5b-11b3c415-b0f0-e41f13bc7b60
drwxrwxrwx   33 root    root         4096 Dec 19  01:56 d1f39063-ad1ac6be
lrwxr-xr-x    1 root    root          17 Dec 19  02:37 img3 -> 5f5d7d04-559cb55f
lrwxr-xr-x    1 root    root          35 Dec 19  02:37 local001 -> 5a37646d-1b00f39d-48ed-
e41f13bc7b60
lrwxr-xr-x    1 root    root          17 Dec 19  02:37 nfs2 -> d1f39063-ad1ac6be
drwxr-xr-t    1 root    root       1680 Dec 19  02:00 pmem:5a376e5b-11b3c415-b0f0-e41f13bc
7b60
[root@localhost:~]
```



# 自動構成された PMEM Datastore

## ホストの PMEM Datastore

- ESXi Embedded Host Client から確認できる

vmware ESXi™ root@alpha.test.local | ヘルプ | 検索

ナビゲータ

- ホスト  
管理  
監視
- 仮想マシン 8
- ストレージ 4
  - PMemDS-d455f9fb-9bc...
  - 監視
  - その他のストレージ...
- ネットワーク 4

alpha.test.local - ストレージ

データストア アダプタ デバイス 永続的なメモリ

新しいデータストア 容量の増加 | 仮想マシンを登録 データストアブラウザ | 更新 | アクション 検索

名前	ドライブ...	キャパシ...	プロビジ...	空き	タイプ	シンプロ...	アクセス
alpha_local01	非 SSD	924 GB	7.06 GB	916.94 GB	VMFS6	サポート対象	単一
img3	不明	4.81 TB	2.12 TB	2.69 TB	NFS	サポート対象	単一
PMemDS-d455f9fb-9bc9-0b47-9aca-d5...	不明	255.36 GB	131.28 GB	124.08 GB	PMEM	未サポート	単一
Shared_ysaito	非 SSD	5 TB	87.98 GB	4.91 TB	VMFS6	サポート対象	単一

4 アイテム

# 自動構成された PMEM Datastore

## ホストの PMEM Datastore

- ESXi Embedded Host Client から確認できる



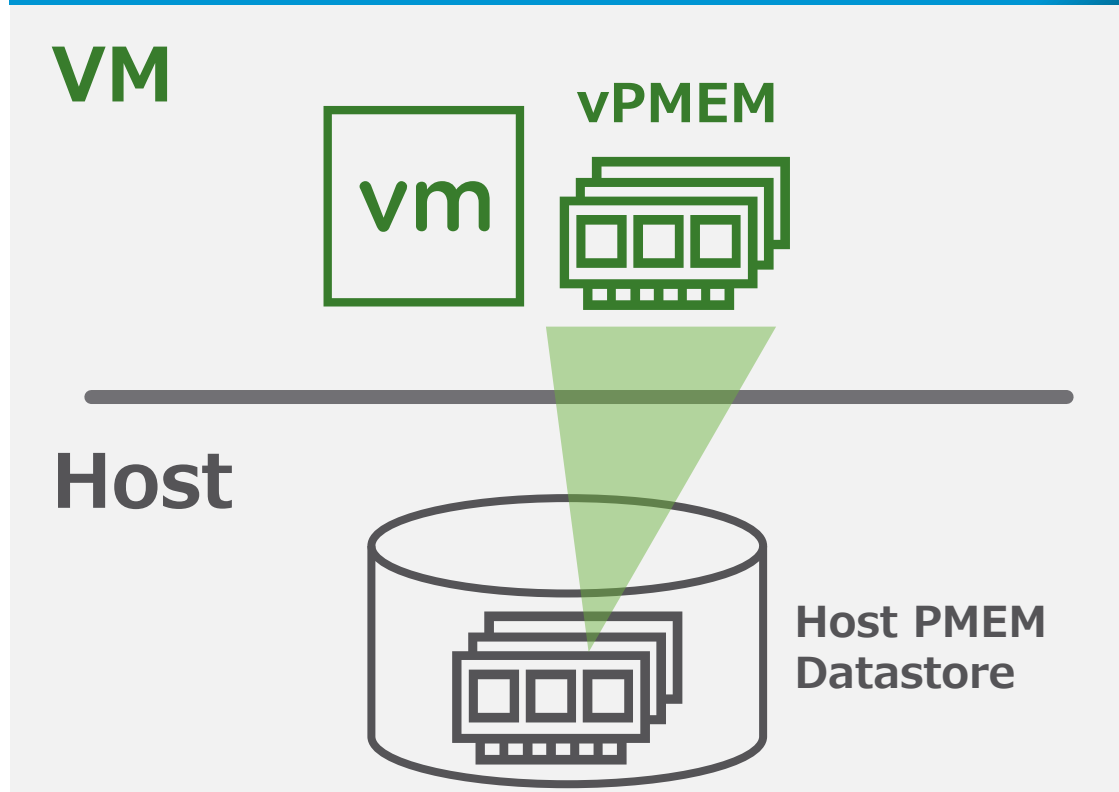
# PMEM Datastore が構成されない場合

## ESXi による PMEM Datastore の構成

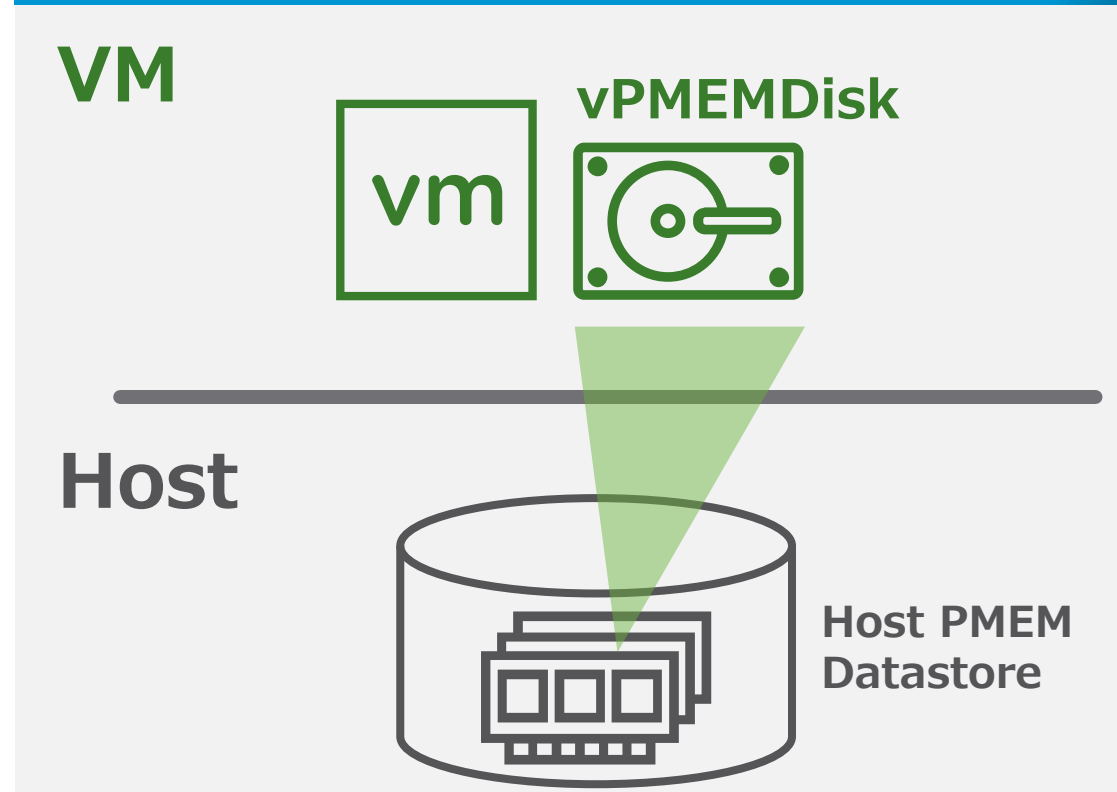
- ESXi は、他の OS (e.g., Windows, Linux …) により構成された PMEM 領域の再構成は行わない
- 他の OS で利用した PMEM 領域を ESXi で利用する場合は、サーバ BIOS から PMEM 上のパーティション、ネームスペースを削除後、ESXi を起動

# 仮想マシンから PMEM 資源を利用する方法

## 1. 仮想 PMEM として構成

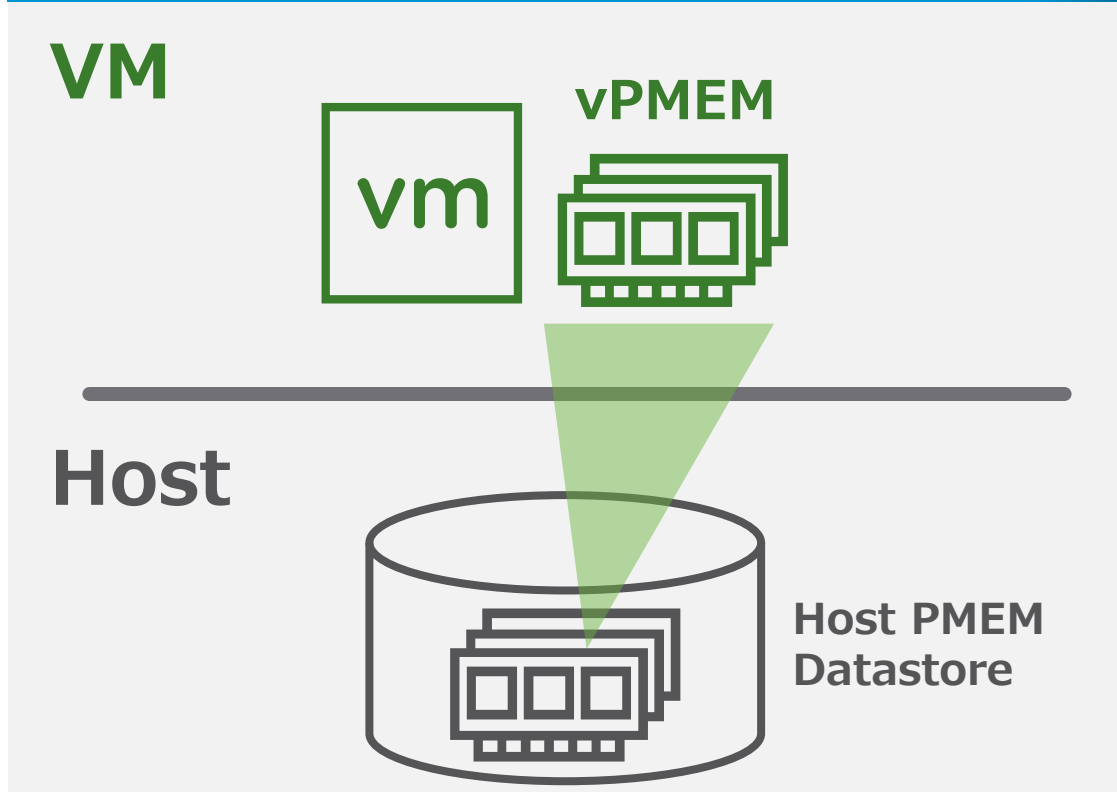


## 2. 仮想ディスクとして構成



# Option 1 : 仮想 PMEM

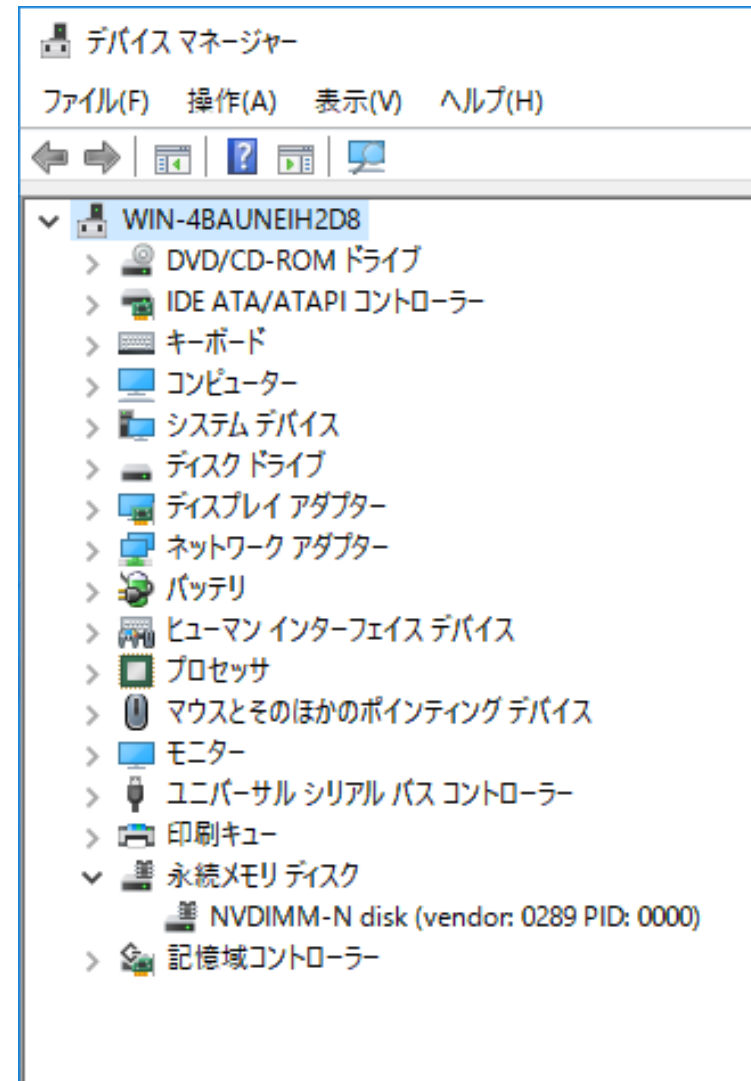
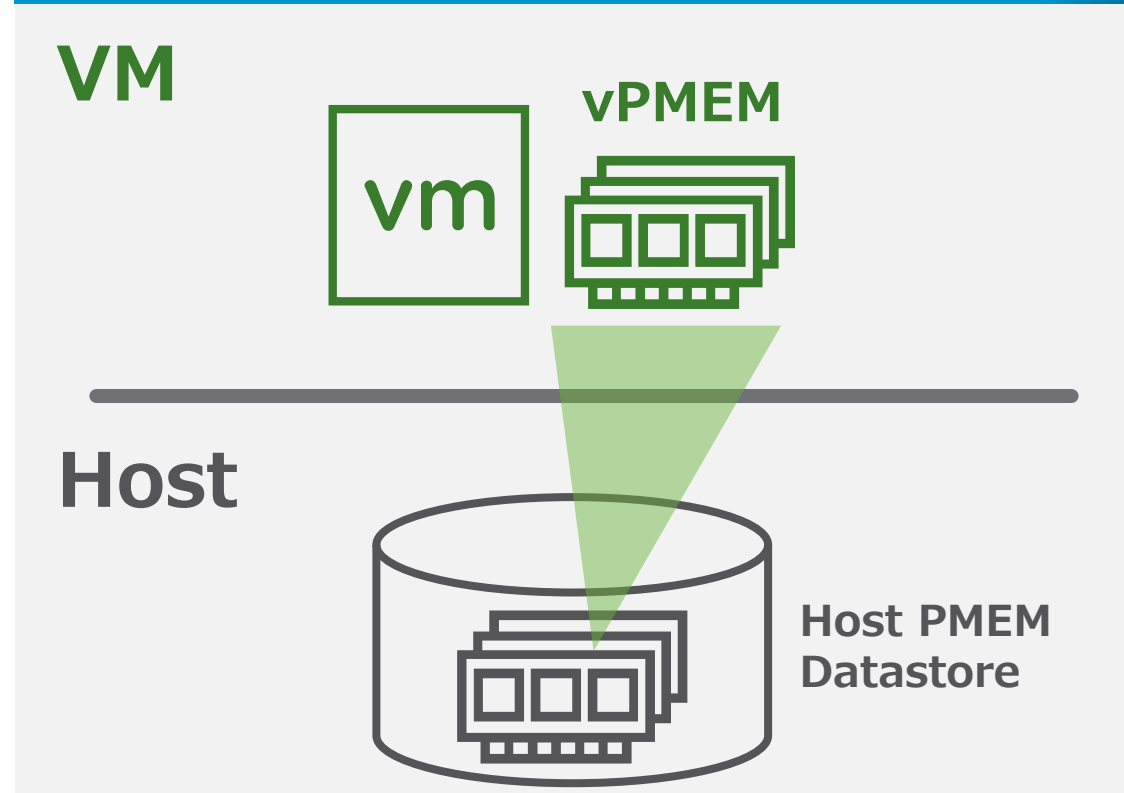
## 1. 仮想 PMEM として構成



- vPMEM デバイスを構成
- ゲスト OS からは NVDIMM として認識される
  - ゲスト OS 側が NVDIMM をサポートしている必要がある
- バイト・アドレスサブルという特性を享受できる
  - ゲスト OS 並びにアプリケーションが Direct Access Mode (DAX) に対応している場合

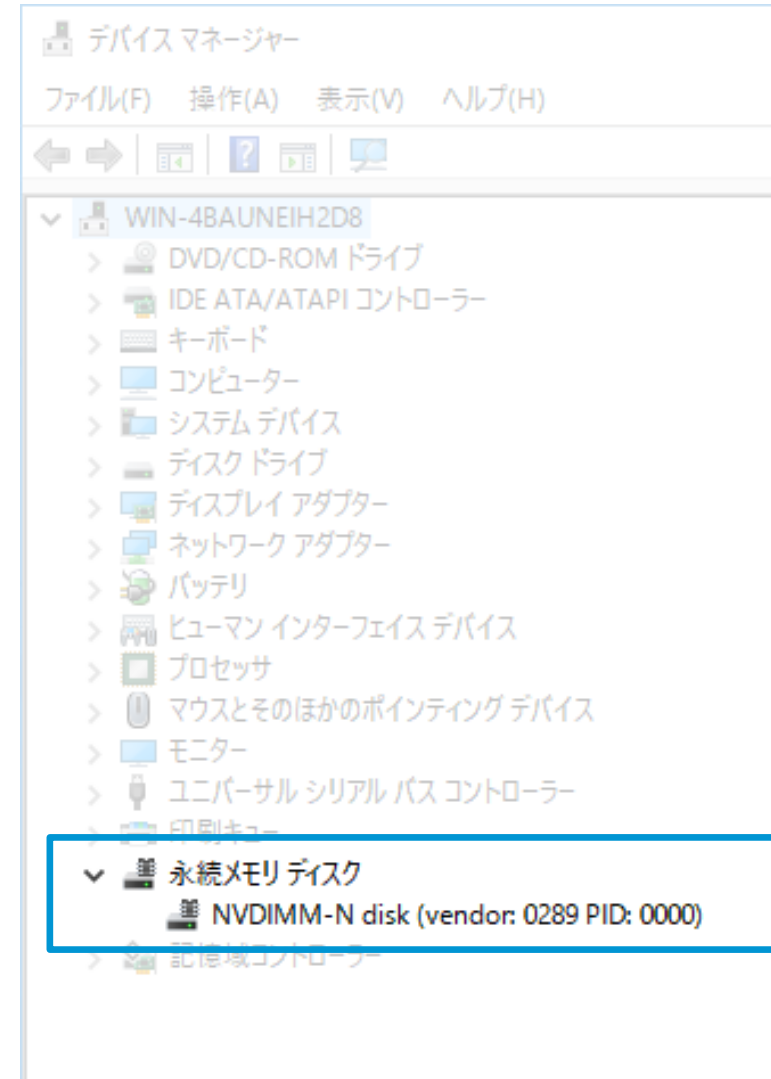
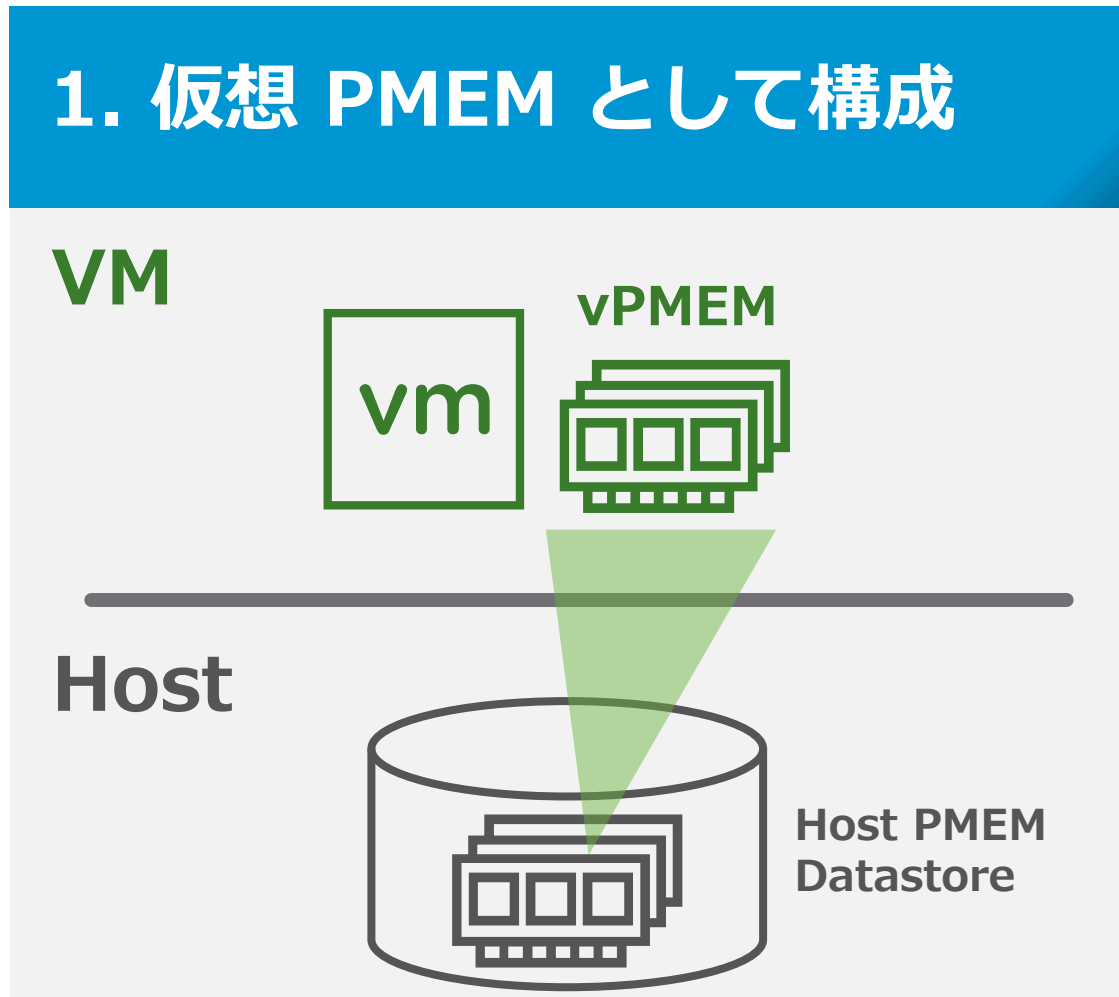
# Option 1 : 仮想 PMEM

## 1. 仮想 PMEM として構成



# Option 1 : 仮想 PMEM

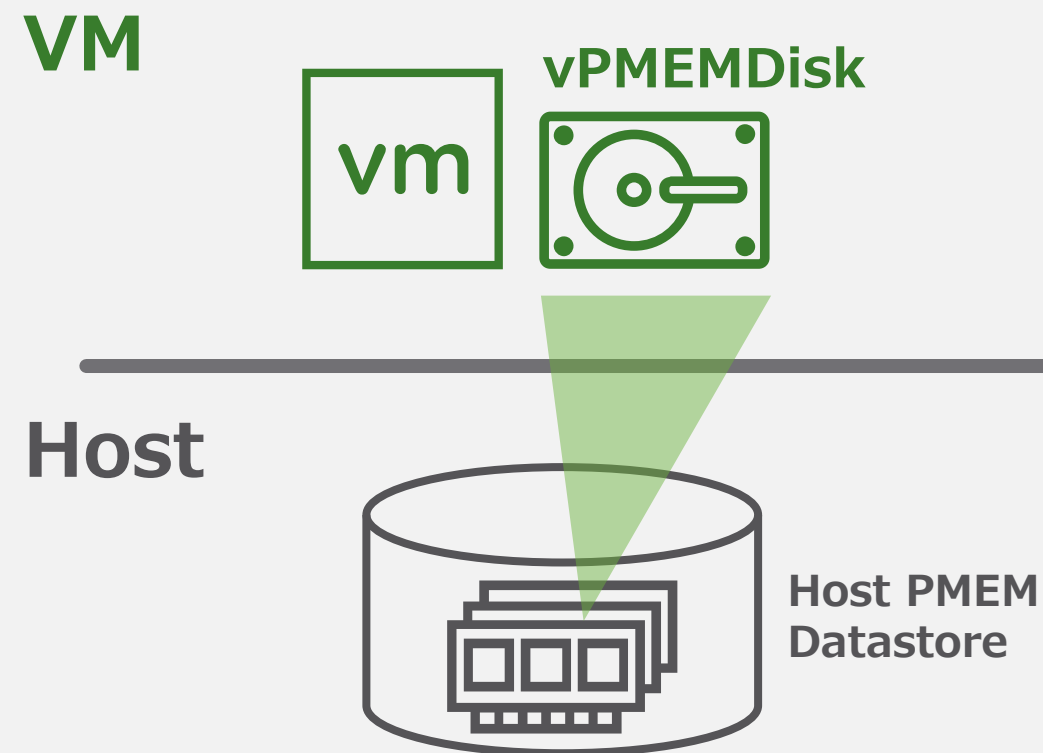
## 1. 仮想 PMEM として構成



# Option 2 : 仮想ディスクとして構成

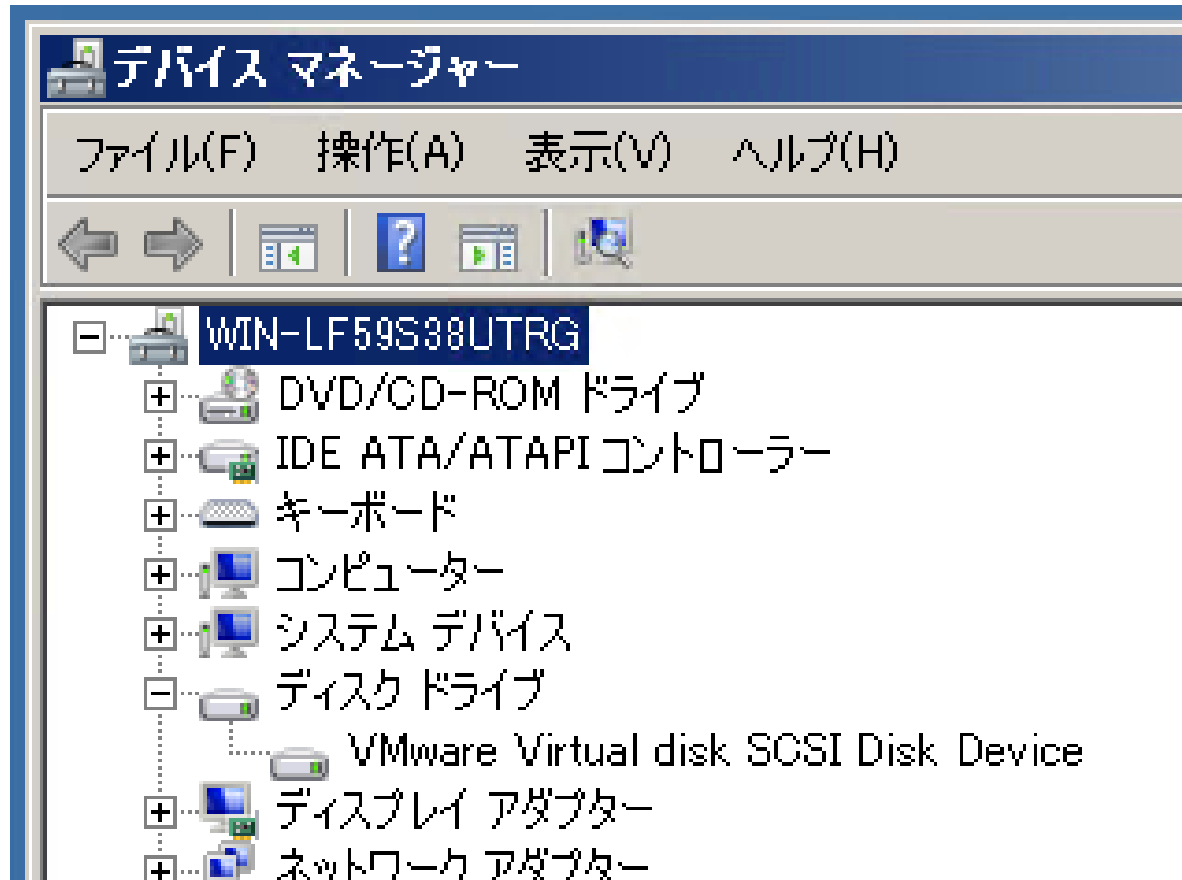
- 仮想ディスクとして構成
  - ゲスト OS が PMEM / NVDIMM に未対応であっても利用可能
- ホスト側バックエンドには PMEM Datastore が利用される
  - 高速・低遅延のデータストア
- ブロックモードアクセス
  - ゲスト OS は仮想 SCSI コントローラ経由でアクセス

## 2. 仮想ディスクとして構成

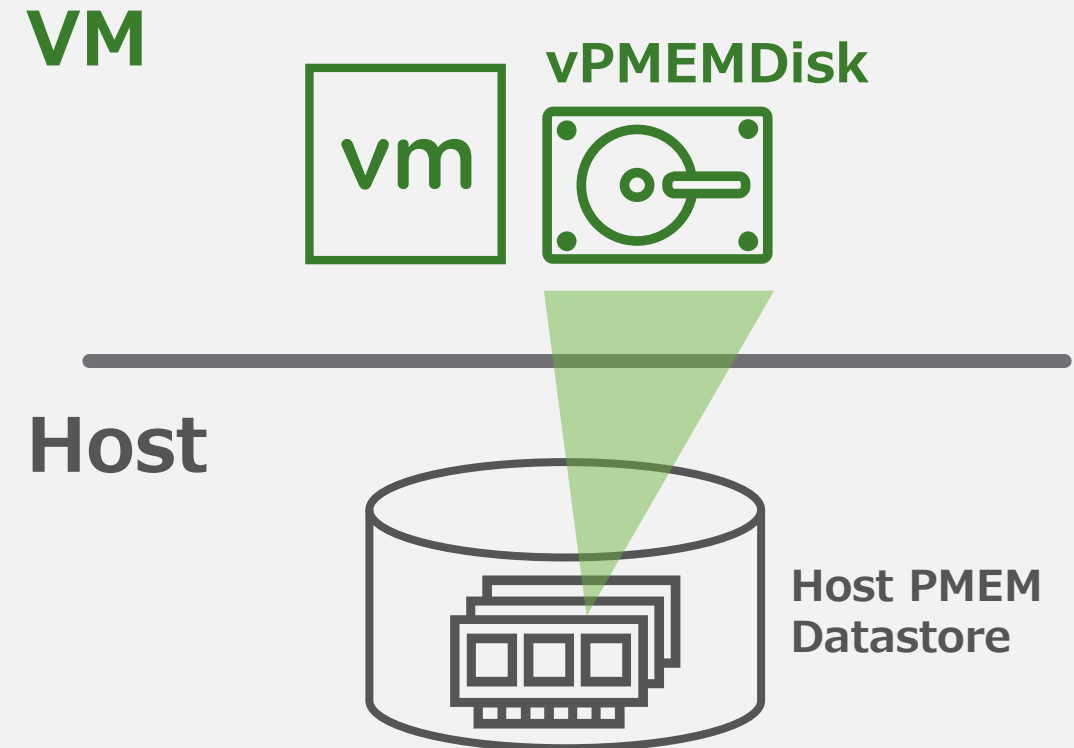




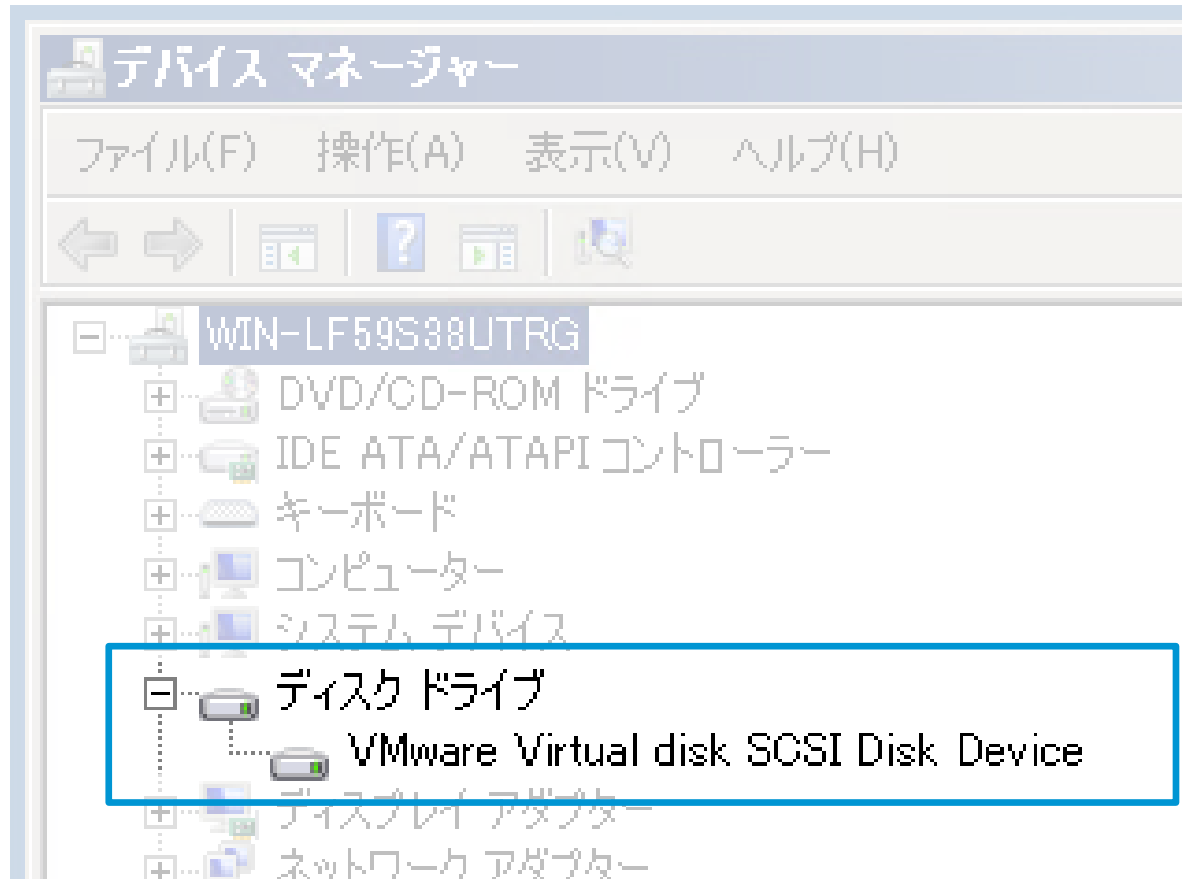
# Option 2 : 仮想ディスクとして構成



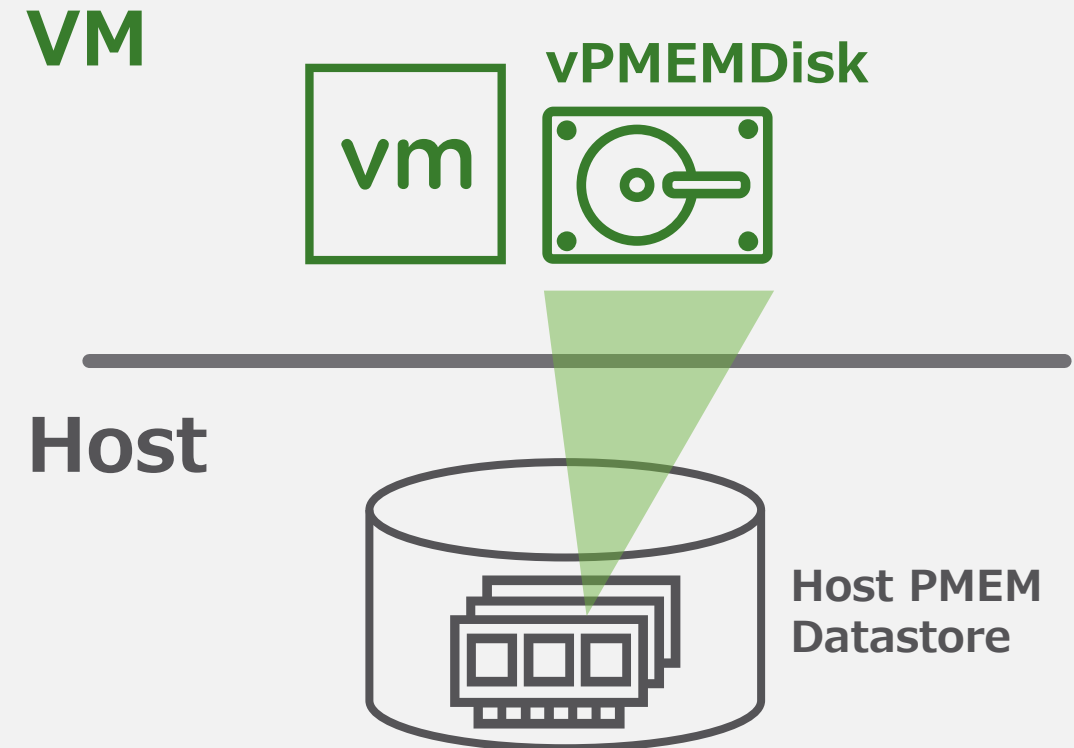
## 2. 仮想ディスクとして構成



# Option 2 : 仮想ディスクとして構成



## 2. 仮想ディスクとして構成



# vPMEM の利用

## 要件

- ホストに PMEM データストアが構成されている
- 仮想 PMEM をサポートする Guest OS Type が選択されている
  - [VMware Knowledge Base 51903](#) を参照
    - Windows Server 2016 Build 14393 以降
    - RedHat Enterprise Linux 7.4 以降など
- 仮想ハードウェアバージョン 14 以降

# vPMEM の構成

設定の編集 | W2K16\_with\_NVDIMM-N

仮想ハードウェア | 仮想マシン オプション

新規デバイスを追加

> CPU	2	▼
> メモリ	8	GB ▼
NVDIMM 1	32	GB ▼
▼ NVDIMM コントローラ		
使用可能な永続的メモリ (PMEM)	11.93 GB	
ゲストでサポートされています	4 TB	
この仮想マシンの容量:	1 NVDIMM 32768 MB	
> ハードディスク 1	100	GB ▼
> SCSI コントローラ 0	LSI Logic SAS	

- CD/DVD ドライブ
- ホストの USB デバイス
- ハードディスク
- RDM ディスク
- 既存のハードディスク
- ネットワーク アダプタ
- SCSI コントローラ
- USB コントローラ
- SATA コントローラ
- NVDIMM
- NVMe コントローラ
- Trusted Platform Module

キャンセル OK

# vPMEM - .vmx 構成ファイル

```
      :  
nvdimm0.present = "TRUE"  
nvdimm0:0.present = "TRUE"  
nvdimm0:0.fileName = "/vmfs/volumes/vmfs/volumes/pmem:5a376  
e5b-11b3c415-b0f0e41f13bc7b60/W2K116_with_NVDIMM-N/W2K16_wi  
th_NVDIMM-N_1.vmdk"  
nvdimm0:0.size = "32768"  
nvdimm0:0.node = "0"  
      :
```

# vPMEMDisk の利用

## vPMEMDisk

- バックエンド側に ESXi Host PMEM Datastore を用いる仮想ディスク
- ESXi Host PMEM Datastore が構成されていれば利用可能
  - ゲスト OS の種別、仮想ハードウェアバージョンに非依存
- ゲスト OS は仮想 SCSI コントローラ経由でアクセス
  - ブロックモードアクセス

# vPMEMDisk の構成

## 新規仮想マシン

- ✓ 1 作成タイプの選択
- ✓ 2 名前とフォルダの選択
- ✓ 3 コンピューティング リソー...
- 4 ストレージの選択**
- 5 互換性の選択
- 6 ゲスト OS を選択
- 7 ハードウェアのカスタマイズ
- 8 設定の確認

### ストレージの選択

設定とディスク ファイルを保存するデータストアを選択します

☐ 標準 ☒ PMEM ⓘ

ⓘ 仮想マシンのホームを除くすべてのディスク ファイルは、デフォルトで PMEM データストアに保存されます  
以下の仮想マシンのホームに使用するストレージ ポリシーおよびデータストアを選択してください

☐ この仮想マシンを暗号化 ⓘ

新規に仮想マシンを作成する際に指定する場合は、  
データストア種別として PMEM を選択

# vPMEMDisk の構成

▼ 新規ハード ディスク *	2	GB ▼
最大サイズ	1 TB	
仮想マシン ストレージ ポリシー	データストアのデフォルト ポリシー ▼	
場所	データストアのデフォルト ポリシー	
ディスク プロビジョニング	Host-local PMem Default Storage Policy	
共有	Management Storage Policy - Large	
	Management Storage Policy - Regular	
	VM Encryption Policy	
	vSAN Default Storage Policy	
	VVOL No Requirements Policy	

仮想ディスクを個別に追加する場合は、仮想マシンストレージポリシーとして Host-local PMem Default Storage Policy を選択



# vPMEMDisk - .vmx 構成ファイル

```
    :  
    scsi0:0.present = "TRUE"  
    scsi0:0.deviceType = "scsi-hardDisk"  
    scsi0:0.fileName = "/vmfs/volumes/pmem:5a376e5b-11b3c415-b0  
f0-e41f13bc7b60/W2K8R2_on_PMem/W2K8R2_on_PMem.vmdk"  
    sched.scsi0:0.shares = "normal"  
    sched.scsi0:0.throughputCap = "off"  
    scsi0:0.redo = ""  
    :
```

# PMEM を構成した VM の相互運用性と制限事項

## サポートされる機能

- ライブ・マイグレーション機能
  - vSphere vMotion、vSphere Storage vMotion など
- VM プロビジョニング機能
  - クローン操作、コールドマイグレーション、コンテンツライブラリ、OVF 操作など
- vSphere DRS
- vSphere FT
  - プライマリ VM が fail したときの動作はフェイルオーバーのみ

# PMEM を構成した VM の相互運用性と制限事項

vSphere 6.7 ではサポートされていない機能

- vSphere HA
  - vSphere HA は共有データストアの利用が前提
- 仮想マシンスナップショットの取得
  - VADP など、スナップショット機能に依存する機能も利用不可

vPMEM、vPMEMDisk 双方において上記制限事項が発生する

# Agenda

記憶デバイスの進化

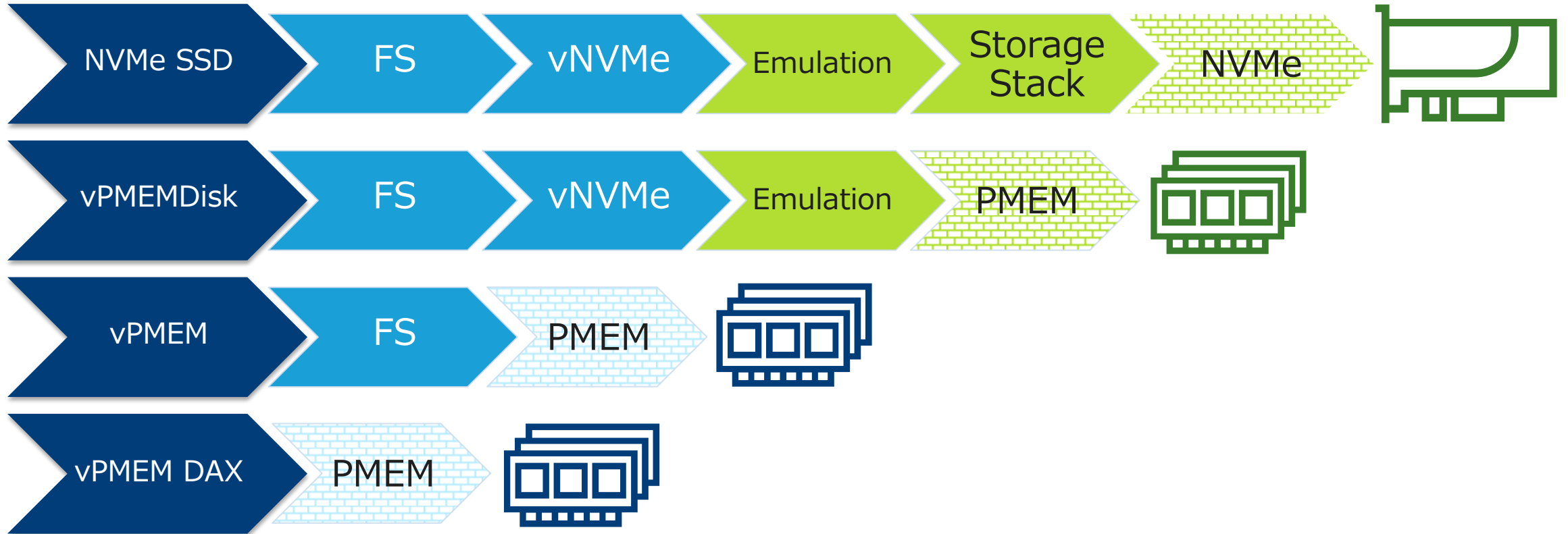
vSphere Persistent Memory

**PMEM の動作モードとパフォーマンス**

業界における VMware の取り組み

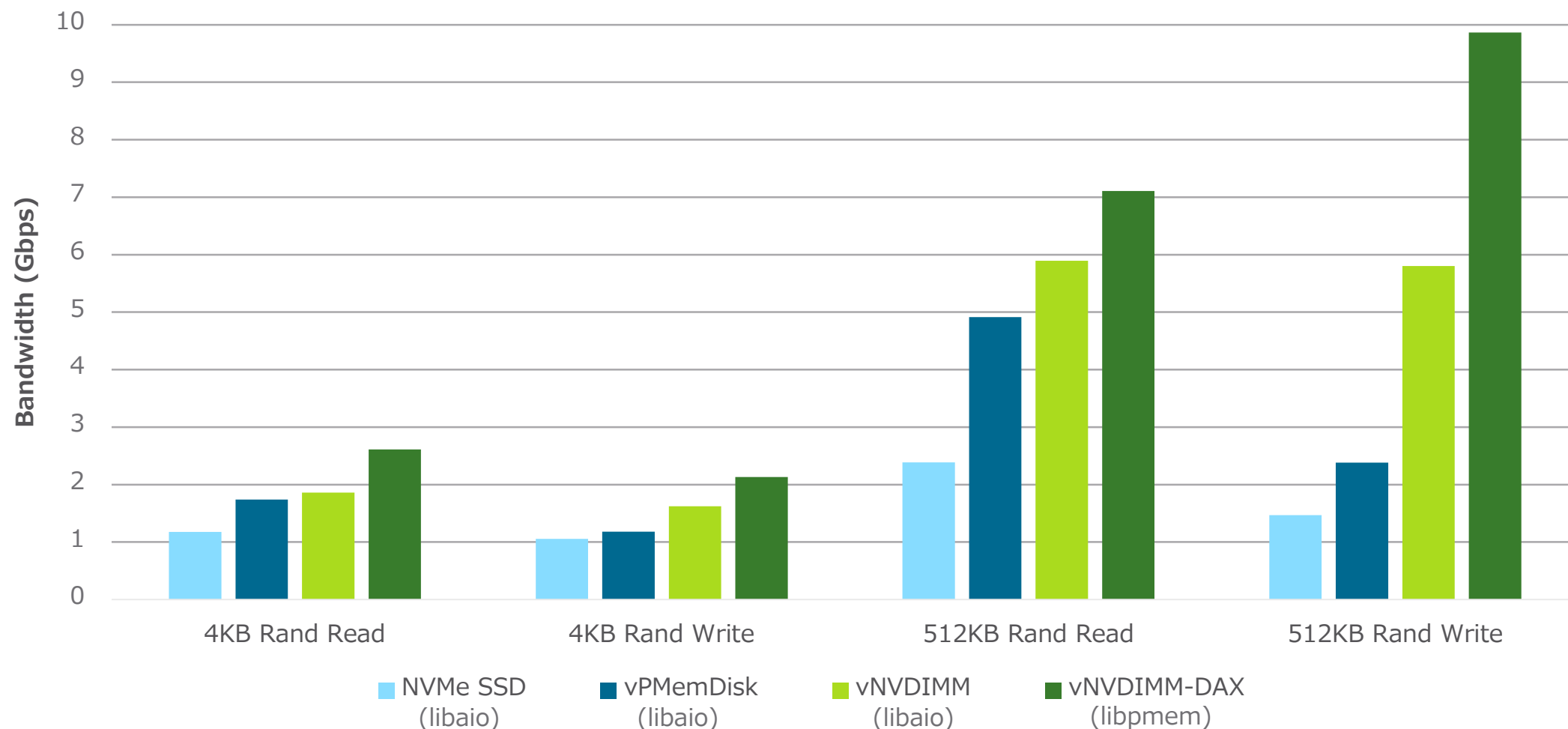
まとめ

# アクセス方式の違い

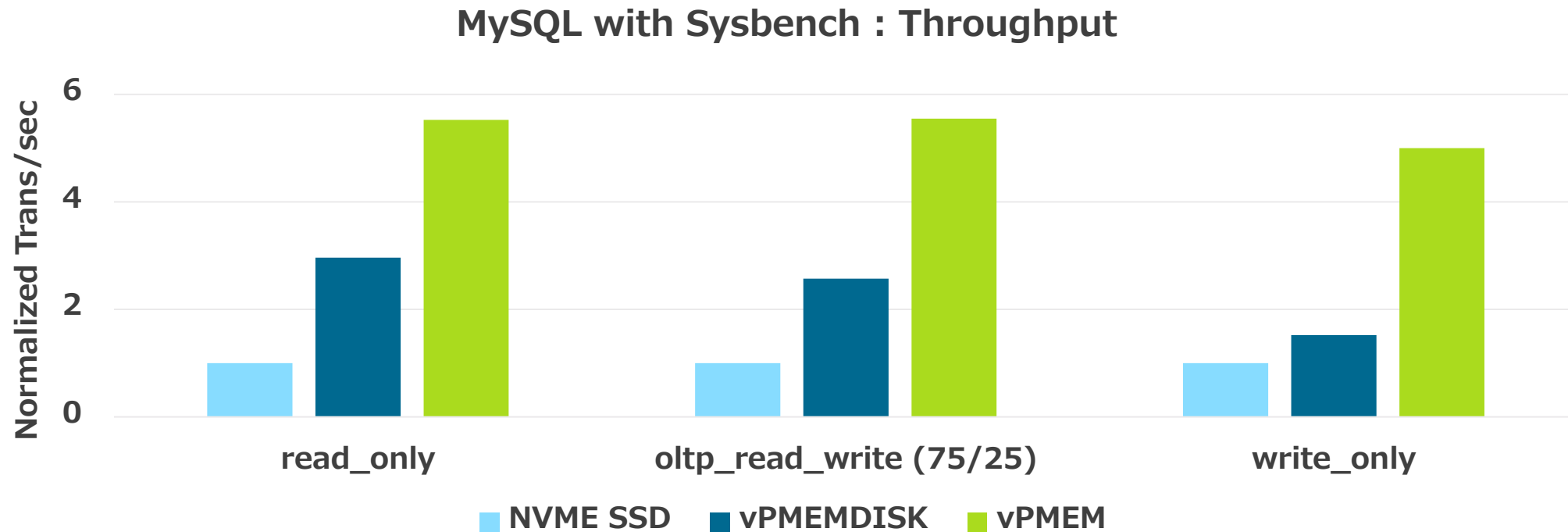


# Flexible I/O Test による Raw I/O 性能測定結果

Fedora VM – FIO Single Thread Test Results



# Sysbench を用いた DB 性能の測定結果

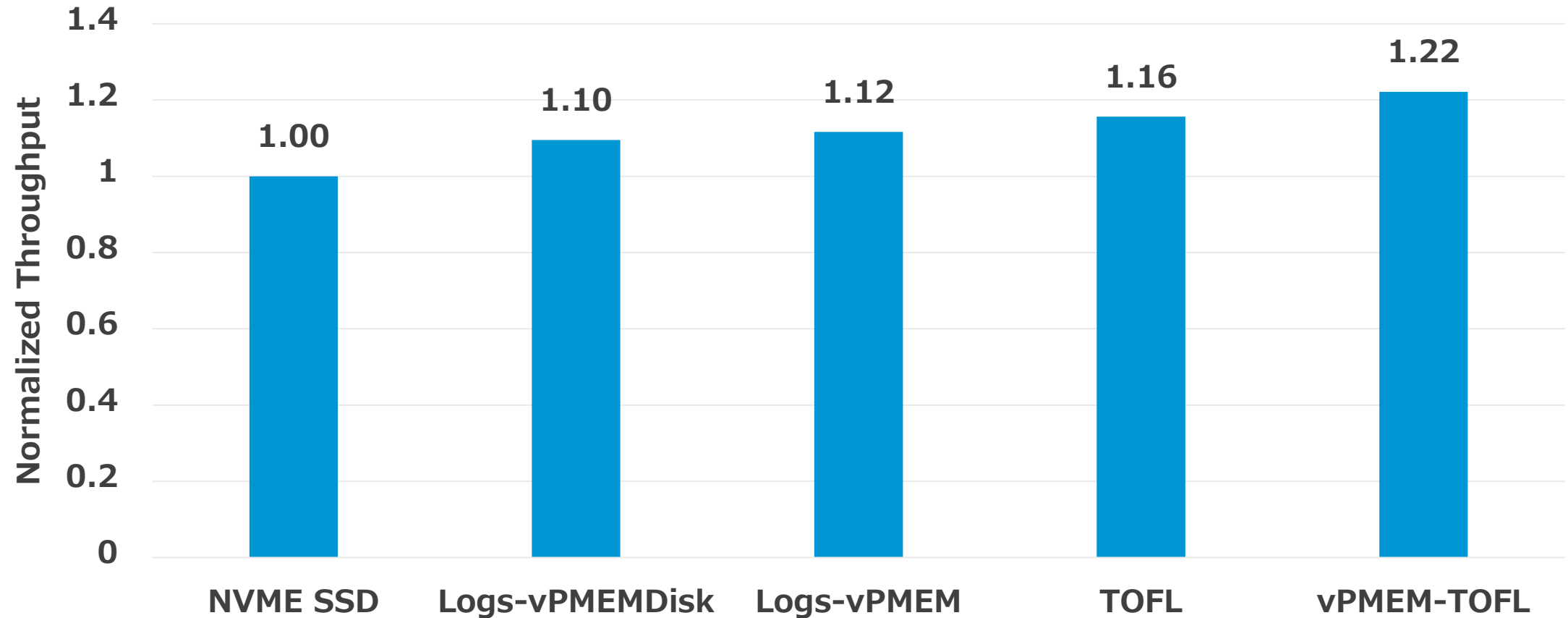


**vPMEM では 5 倍 ~ 5.5 倍程度のスループットを記録**

出典: VMworld 2018 US, Breakout Session, VIN2183BU

# SQL Server 2016 SP1, HammerDB

Windows Server 2016 + SQL Server 2016 SP1, HammerDB Test Results



出典: VMworld 2018 US, Breakout Session, VIN2183BU



# SQL Server 2016 SP1, HammbereDB (補足)

## SSD

- DB と Logs の双方を NVMe SSD 上に配置

## Logs-vPMEMDisk

- Logs を vPMEMDisk 上に配置

## Logs-vPMEM

- Logs を vPMEM 上に配置

## Tail-of-the-log (TOFL)

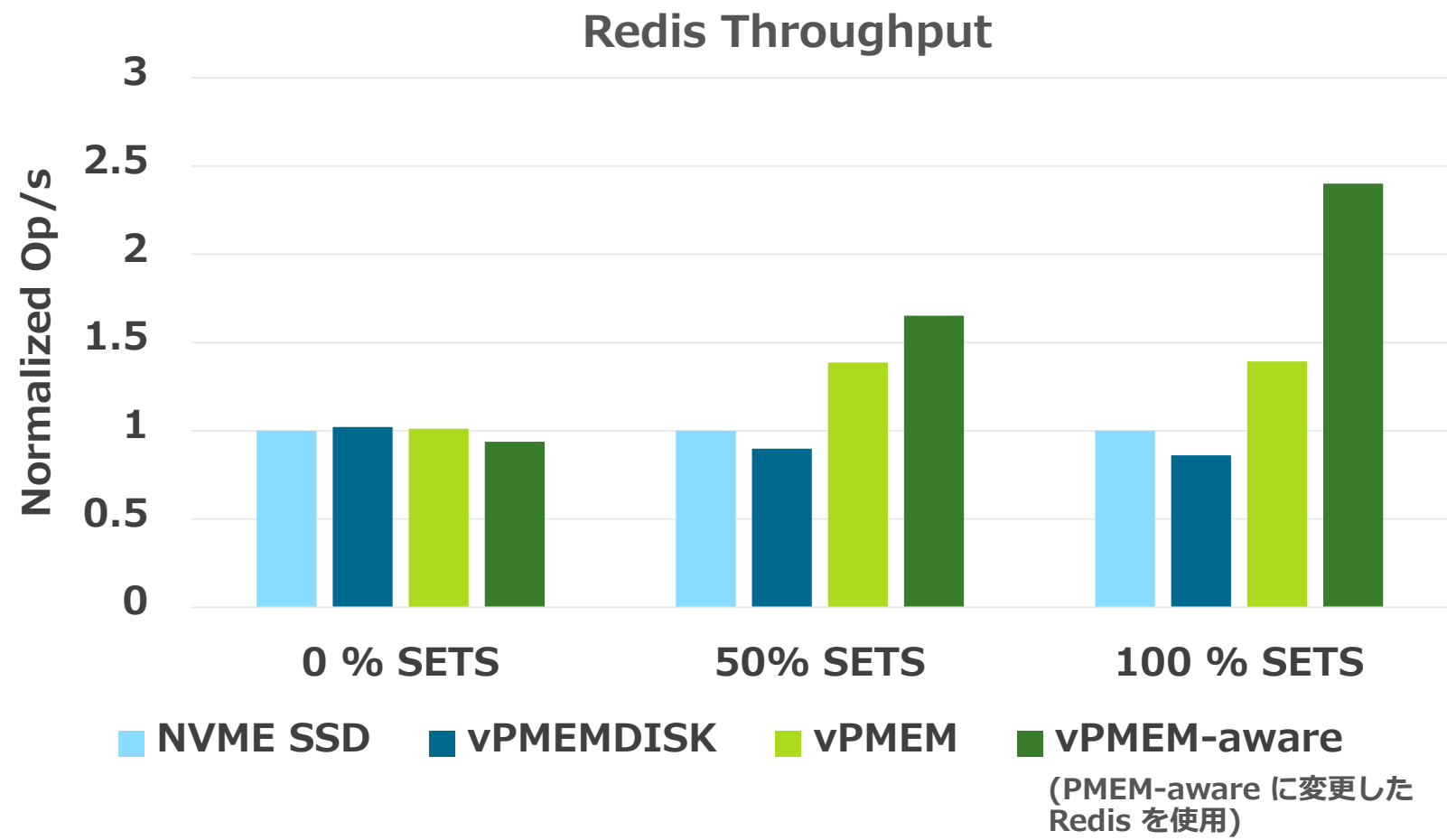
- Tail-of-the-Log (TOFL) を vPMEM DAX ボリューム上に配置

## vPMEM-TOFL

- DB と Logs を vPMEM 上に配置、TOFL を vPMEM DAX ボリューム上に配置

出典: VMworld 2018 US, Breakout Session, VIN2183BU

# Redis (key-value store) の性能測定結果



# Agenda

記憶デバイスの進化

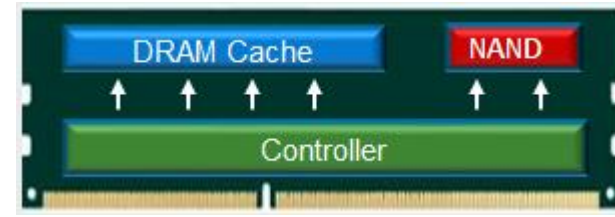
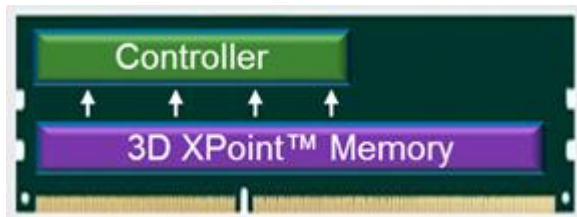
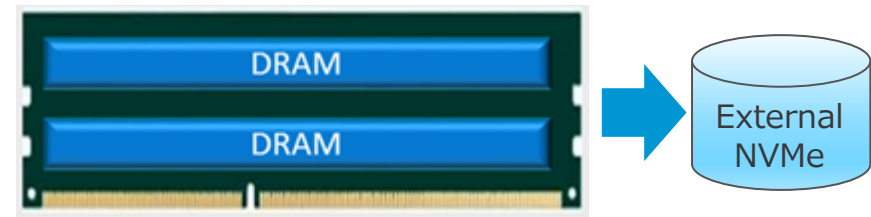
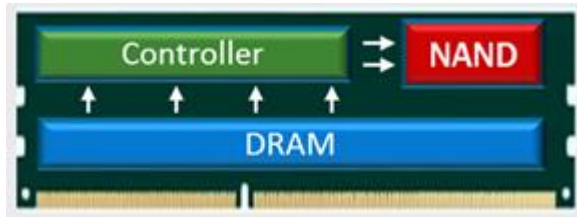
vSphere Persistent Memory

PMEM の動作モードとパフォーマンス

**業界における VMware の取り組み**

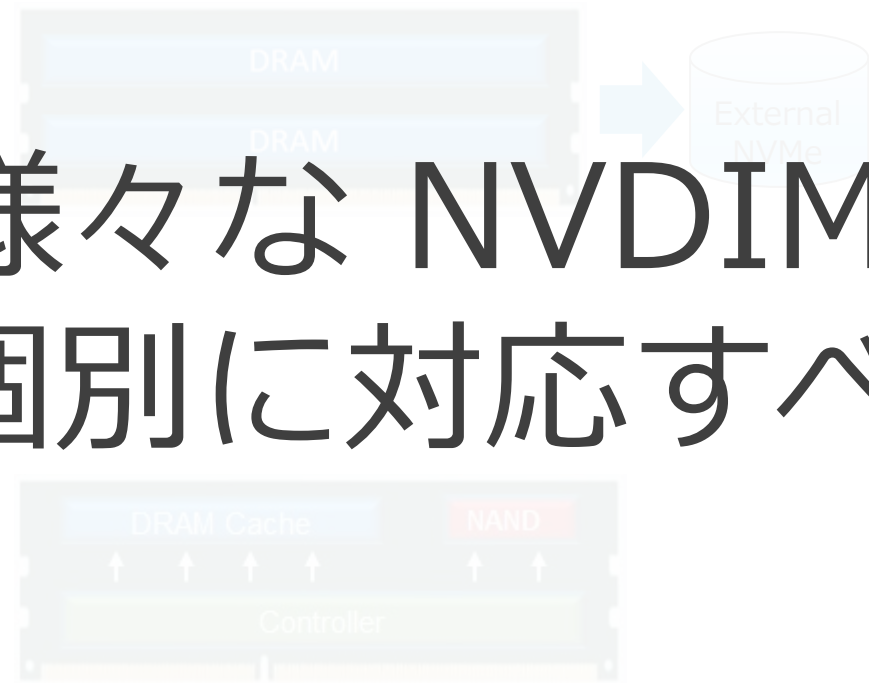
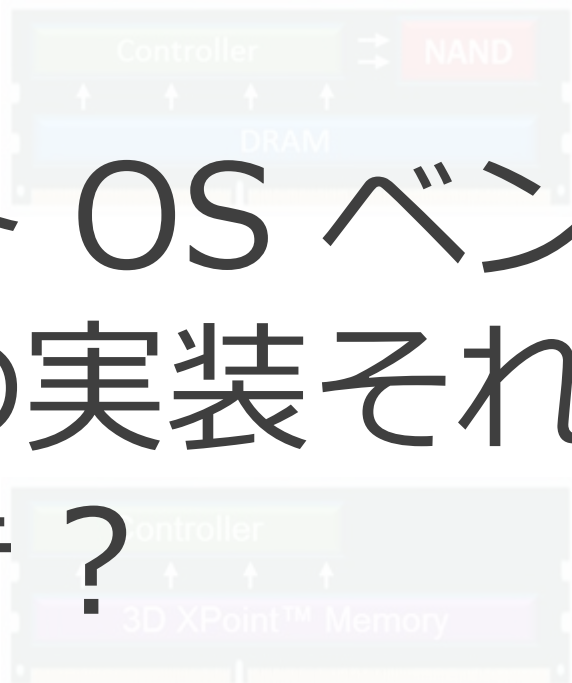
まとめ

# NVDIMM - 複数のベンダ、様々なテクノロジー



# NVDIMM - 複数のベンダ、様々なテクノロジー

各 OS ベンダは、様々な NVDIMM  
の実装それぞれに個別に対応すべ  
き？



# NVDIMM に固有の考慮事項

## Enumeration

- OS は NVDIMM をどのように検出し、メモリ領域と関連付けすれば良いのか

## 名前空間

- NVDIMM は論理的に分割され名前空間を構成する

## ARS (Address Range Scrubbing)

- NVDIMM に固有の RAS 機能

## Health

- SMART 情報、ヘルスステータスの取得

# 標準化に向けた取り組み

VMware は様々なハードウェア・パートナー様と協業し、標準化仕様の確立に貢献

- ACPI と UEFI の標準仕様に多数のプロポーザルを提出
  - NVDIMM Firmware Interface Table (NFIT)
  - NVDIMM Namespace Standardization
  - NVDIMM Platform Capabilities
  - NVDIMM Health Methods
  - Address Range Scrub (ARS) Interface Updates

# Evolution of Industry Wide Common Standards

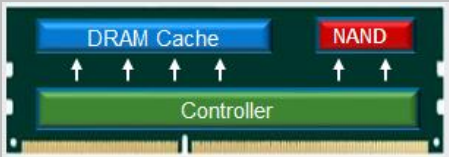
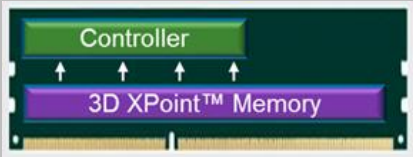
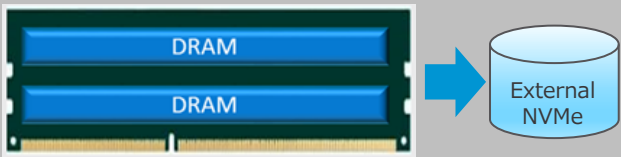
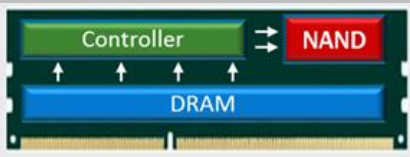
Operating Systems

Common NVDIMM Module

Firmware

Common NVDIMM Interface

Hardware





# Agenda

記憶デバイスの進化

vSphere Persistent Memory

PMEM の動作モードとパフォーマンス

業界における VMware の取り組み

**まとめ**

# まとめ

- PMEM (NVDIMM) と呼ばれるデバイスの出現
- vSphere 6.7 より Persistent Memory 機能を提供
- PMEM の高速、低遅延、バイト・アドレスサブルという特性の活用が今後期待できる
- VMware は、NVDIMM インターフェ이스の標準化に大きく貢献

# 本セッション受講の方へのお勧め

DC 113

17:10~  
Room F

**サポート流 vSphere トラブルシューティングアプローチ**

DC 134

17:20~  
Room A

**Deep Dive into "VMware Cloud on AWS" '18**

# 本セッションに関連する展示・ハンズオンラボのご紹介

お客様のペースで  
利用いただける  
セルフペースラボです。

以下の製品については  
日本語マニュアルも  
提供されています。

- What's new in vSphere 6.7
- vSphere 6.7 Challenge Lab
- VMware vSAN 6.7 Getting Started
- VMware Cloud on AWS
- What's New in vRealize Operations and vRealize Log insight
- VMware vSphere Automation PowerCLI
- VMware vRealize Automation 7.0

VMware  
ハンズオン  
ラボ

VMware  
ソリューション展示

データセンターからクラウドまで一貫したインフラストラクチャの構築と運用・管理、そして、これらのクラウドへの高いアクセス性と高度な分散化に対応するセキュリティについてご紹介します。

ご清聴、ありがとうございました。