

# DTW2024を振り返る

～脚光を浴びたAIと日陰に押し出された??  
ハイブリッドクラウド?その真実に迫る!～

デル・テクノロジーズ株式会社  
パートナー事業本部  
中野 修 & 石塚 智規



このセッションでお話しすること

# Dell Technologies World 2024 サマリー

## ハイブリッドクラウド戦略よ、いずこへ？



# Dell Technologies World 2024サマリー

# Dell Technologies Experience Overview / World



■ 全体登録者：  
約10,000人以上

■ セッション数：  
Keynote : 2  
Trailblazer : 2  
Luminary : 3  
Breakout Session : 150

■ Private Session：  
Japan NDA Session : 5  
個別EBC, 個別ミーティング

覚えていますか？ AIの4つの方針を・・・

AIを製品に組み込む

**AI IN**

AIをデルのプラットフォーム  
で稼働させる

**AI ON**

今年はこちらの具現化された製品やソリューション、サービスが  
ドーンと発表されました

社内でAIを活用する

**AI FOR**

エコパートナーと  
AIビジネスを革新する

**AI WITH**

# 1日目のキーノート：マイケル・デル

- デル創立40周年
  - AI時代は既に始まっている
  - 「何を得るのか」がリアルな課題へ
  - 帰ってきた「Data is Fuel」
- 登場したソリューション
    - AI Factory
    - PowerStore Prime
    - PowerScale F910
    - PowerEdge XE9680L

# 1日目のキーノート：ユーザー登壇



- ✓ 「AI時代」には全ての企業はインテリジェントカンパニーになる
- ✓ 競争が激化するなかで勝つためにはAIを活用したスピードアップは不可欠

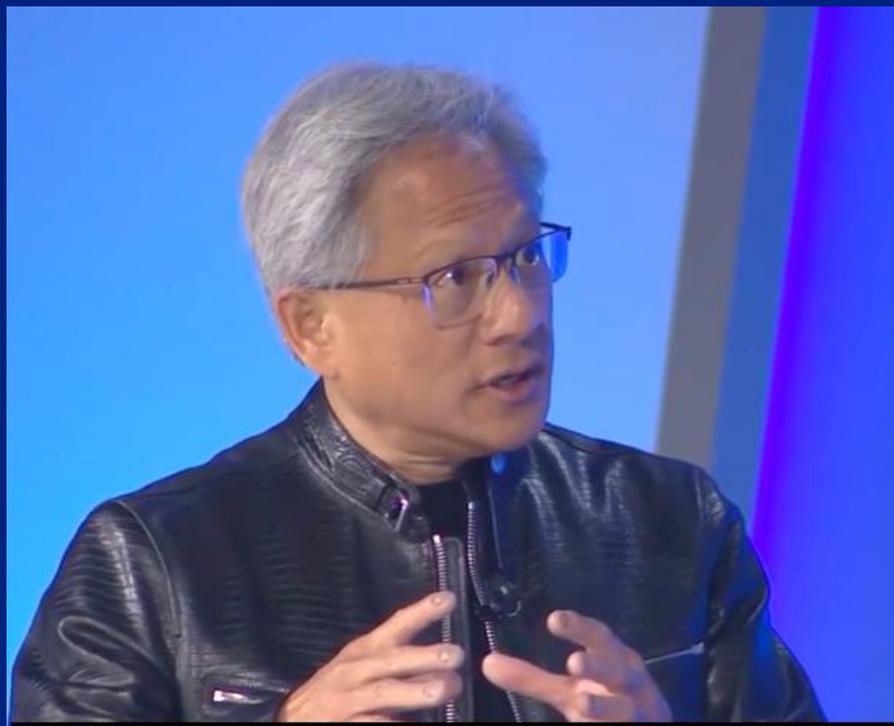
ServiceNow 会長兼CEO Bill McDermott



- ✓ 提供する業務SaaSに独自モデルのAIを組み込み、業務を自動化する支援を提供し始めた

Samsung SDS CEO Seung Woo Hwang

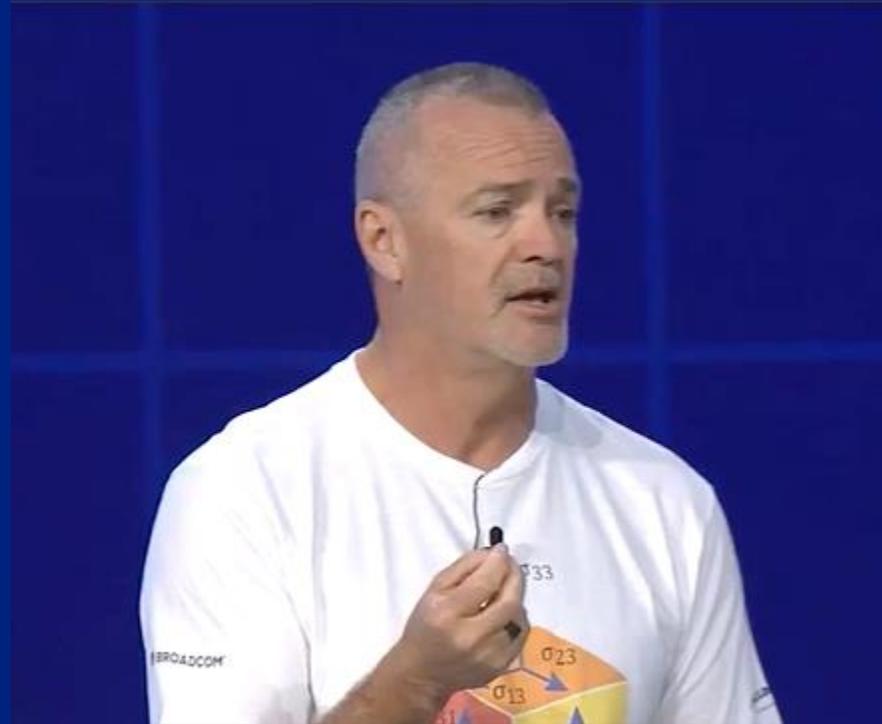
# 1日目のキーノート：パートナー登壇



- ✓ 様々なデータ種からその情報の意味を学習できるようになった
- ✓ さらに、あるモダリティを別のモダリティに変換したり生成できるようになった
- ✓ AIは新しい産業革命

NVIDIA 創業者兼CEO Jensen Huang

## 2日目のキーノート：ジェフ・クラーク



- ✓ 2日目は実際のソリューションの話しが中心
  - ✓ AI Revolution → Feed the Beast
  - ✓ Accelerate the adoption of AI
    - データは差別化要素
    - データをAIに如何に取り込むか
    - 適切なITの活用
    - 革新速度に対応できるスケール性と柔軟性
    - 広範囲でオープンなエコシステム
- プロダクト&オペレーション担当副会長 Jeff Clark

## 2日目のキーノート：ユーザー登壇



- ✓ アプリケーション開発からテストをAI PCを活用して迅速化
  - プロセスと品質スピードの改善、生産性の向上、エラーの削減、プライバシーとセキュリティの向上
  - 1時間かかっていたテストが30分へ

Deloitte CCO Dounia Senawi



- ✓ AI Factoryの活用状況を紹介
  - 工場での組み立て、レース中のタイヤ選択の決断に活用
  - 経営側でも営業から人事、財務までAIを活用
  - 限られた予算を効率的に使う（支払う）観点でもAIがかなり重要な役割を担っている

McLaren Racing CEO Zak Brown

## 2日目のキーノート：パートナー登壇



- ✓ Llama 3をAI Factoryのエコシステムとして利用できる
- ✓ Hugging Faceとして推奨オンプレインフラ

meta社 Sy Chaudhary  
Hugging Face製品責任者Jeff Baudier



- ✓ データを生成し、利用するEdgeと、処理を行うxPUコンピューティングリソースを繋ぐネットワークの要素は大きくなる
- ✓ 求められるのはハイスピードは言わずもがな、スケール性と低消費電力

Broadcom President Charlie Kawwas

振り返ってみると・・・

ホントにAIばかりでしたね・・・

そして大きな疑問が残りました・・・

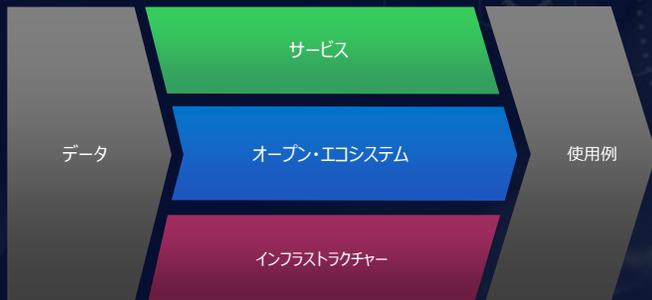
結局 **AI Factory** ってなんぞ？

その謎を解明すべく、我々はAIの奥地へ向かった・・・

—  
AI Factoryってなんで？

# Dell Technologies World 2024 Key Announcement

## Dell AI Factory



## Dell AI Factory with NVIDIA



## Dell Technologiesイノベーション

Edge

Introducing  
**PowerStore Prime**  
The smart choice for all-flash storage

**DELL**Technologies

**A P E X**

# Dell AI Factory概要

あらゆる規模の組織におけるAIイノベーションの加速を支援するデルのアプローチ

## データ：AI Factoryへの燃料供給

データはAIファクトリーを動かす原料である

最も貴重なデータはオンプレミスとエッジにある

デルは、お客様の最も貴重なデータにAIを導入し、そのデータを保存、保護、管理するリーダーです。

## エキスパートAIサービス

効果的なAIファクトリーの成功には熟練したチームが必要だが、AIに対応できるスキルは不足しており、エコシステムも多様化

デルは、お客様のAIジャーニーをガイドし、ビジネス目標に沿ったAIの成果を加速させながら、適切な技術ソリューションを大規模に活用してきた豊富な経験を有しています。



## ユースケースによる成果

AIファクトリーは、最優先のユースケースに基づくビジネス成果を生み出します。

デルは、検証済みのソリューションとカスタマイズされたサービスにより、最も重要なAIユースケースの導入を簡素化します。

## AIに最適化されたインフラ

インフラはAIファクトリーのコア

AIのユースケースは多様で、AI技術は急速に進化している

デルは、クライアントからクラウドまで、最も幅広いAIポートフォリオを提供しているため、AIへの投資を適切な規模に設定でき、AIをあらゆる場所で実行できる柔軟性を備えています。

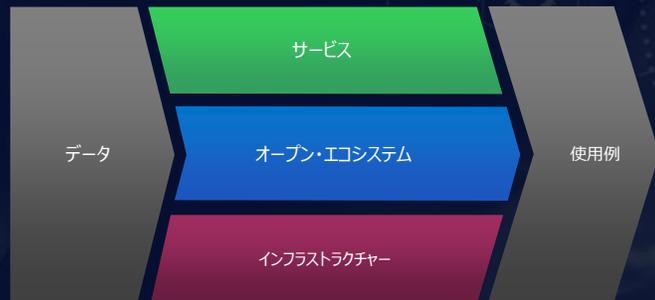
## AIエコシステム

単一のベンダーですべてのAIニーズを満たすことはできない

デルは、オープンなエコシステムを育成してきた実績があり、幅広いAIエコシステムの構築に取り組んでいる。

# Dell Technologies World 2024 Key Announcement

## Dell AI Factory



## Dell AI Factory with NVIDIA



## Dell Technologiesイノベーション

Edge

Introducing  
**PowerStore Prime**

The smart choice for all-flash storage



**DELL**Technologies

**A P E X**

# Dell AI Factory

データ

クラウド

データセンター

エッジ

## サービス

Microsoft Copilotサービス

Dell Enterprise Hub向けサービス

## オープン・エコシステム

Qualcomm

Hugging Face

nVIDIA

Meta

Microsoft Azure

## インフラストラクチャー

Dell AI PC

Dell PowerSwitch

Dell PowerScale

Dell Data Protection

使用例

コンテンツとCODE生成

デジタルアシスタント

データ作成

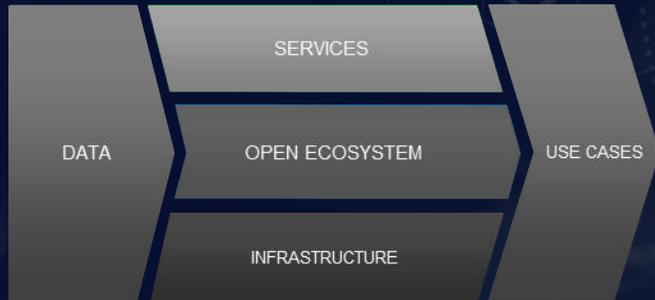
コンピュータビジョン

デジタル・ツインズ

その他...

# Dell Technologies World 2024 Key Announcement

## Dell AI Factory



## Dell AI Factory with NVIDIA



## Dell Technologiesイノベーション

Edge

Introducing  
**PowerStore Prime**

The smart choice for all-flash storage

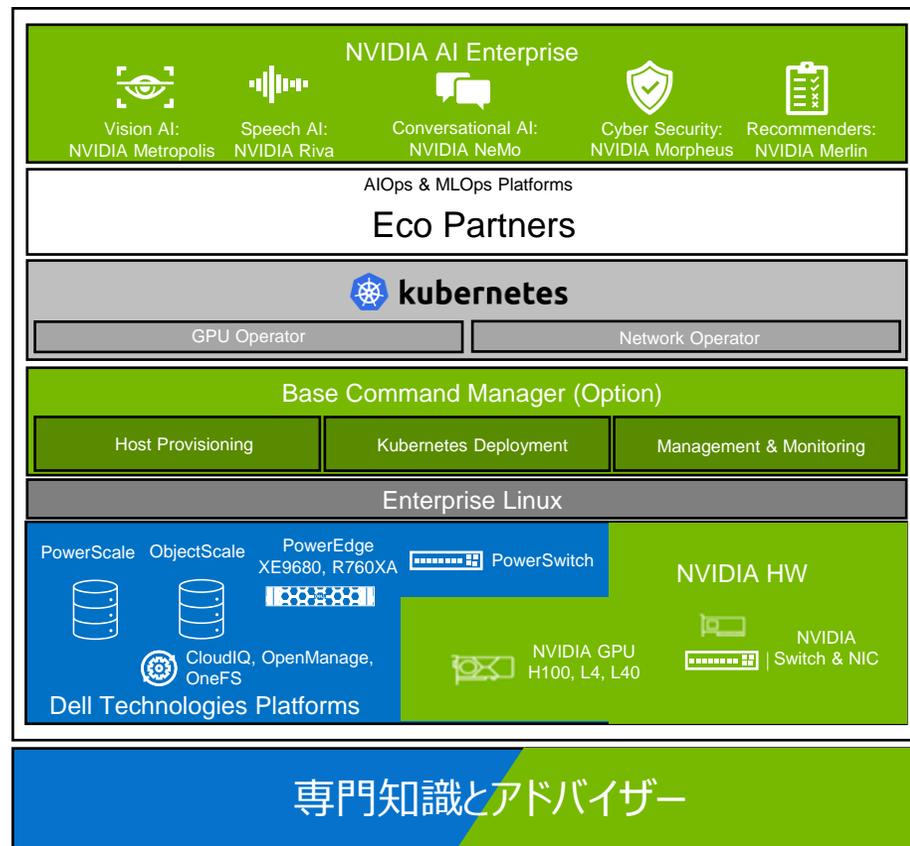


**DELL**Technologies

**A P E X**

# Dell Validated Design for Generative AI with NVIDIA

「Dell Validated Design for Generative AI with NVIDIA」は、NVIDIAと共同で構築した推論のための設計図で、エンタープライズ生成AI向けに、安全で拡張性に優れたモジュール式のプラットフォームの迅速な展開を実現できるように最適化されています



- アクセラレータ技術
- AI ソフトウェア
- 学習済みモデルと専門知識

**DELL**Technologies

- 基盤技術
- ソフトウェア
- サービス

**DELL**Technologies  
PARTNER PROGRAM

# Dell Validated Design for Generative AI with NVIDIA

## Validated Designs for AI Portfolio Overview

Solution	Purpose	Benefits	Use Case
<b>Generative AI: Inferencing (with NVIDIA)</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	スケーラブルなモジュール式アーキテクチャにより大規模言語モデル (LLM) 推論を簡素化し、さまざまなユースケースをサポートします。	<ul style="list-style-type: none"> <li>事前トレーニングされた GenAI モデルを迅速に実装します。</li> <li>デルと NVIDIA の共同アーキテクチャで出力と価値を生成します。</li> </ul>	コンテンツ作成、サポートアシスタント、自然言語検索、プロセス自動化など。
<b>Generative AI: Model Customization and Tuning (with NVIDIA)</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	特定のユースケースに合わせて既存の (事前トレーニング済み) GenAI モデルを再トレーニングする方法を学びます。	<ul style="list-style-type: none"> <li>データとユースケースに応じてカスタマイズおよび微調整することで、GenAI モデルのパフォーマンスを向上させます。</li> <li>デルと NVIDIA の共同アーキテクチャを実装します。</li> </ul>	コンテンツ作成、サポートアシスタント、自然言語検索、プロセス自動化など。
<b>AI for Virtualized Environments</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	マルチインスタンス GPU を使用した NVIDIA と VMware の共同ソリューションにより、従来のワークロードと並行して AI を実行することが容易になります。	<ul style="list-style-type: none"> <li>高速化された AI ワークロードを既存のエンタープライズ アプリケーションとシームレスに統合します。</li> <li>AI 固有の個別のシステムを作成することは避けてください。</li> </ul>	ソフトウェア定義のワークロード。高密度仮想化と AI ワークロード、HPC とパフォーマンス ワークロード、プライベート クラウド環境。
<b>AI MLOps with cnvrg.io</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	cnvrg.io と共同開発した Dell Validated Design for AI を使用して、研究から実稼働へのシームレスな移行のために ML パイプラインを標準化します。	<ul style="list-style-type: none"> <li>最適化されたハードウェアとソフトウェア スタックを迅速に導入します。</li> <li>ML パイプラインを合理化して、より迅速な開発とスムーズな運用移行を実現します。</li> <li>エンジニアリングでテストされた機械学習オペレーション (MLOps) ソリューションを自信を持って導入します。</li> </ul>	AI ワークロードは、機械学習、AI オペレーション、データ分析に重点を置いています。
<b>Automatic Machine Learning</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	Dell Validated Design for AI — 自動機械学習 (AutoML) を使用して AI モデルのトレーニングを自動化し、アルゴリズムの選択、機能生成、ハイパーパラメータ調整、反復モデリング、モデル評価を合理化します。	<ul style="list-style-type: none"> <li>データサイエンティストの AI モデルトレーニングを簡素化します。</li> <li>合理化された MLOps により、AI の本番環境への移行が迅速化されます。</li> <li>エンジニアリングでテストされた MLOps ソリューションを自信を持って導入します。</li> </ul>	AI ワークロードは、機械学習、AI オペレーション、データ分析に重点を置いています。
<b>Conversational AI</b> <a href="#">Solution Brief</a> <a href="#">Validated Design</a>	Dell Validated Design for AI を使用して、高度な仮想アシスタントの導入を簡素化し、加速します。会話を通じてコンピュータと人間の間のシームレスな対話を促進するように最適化された会話型 AI です。	<ul style="list-style-type: none"> <li>会話型ユーザー インターフェイスと仮想アシスタントを簡単に作成します。</li> <li>会話型 AI のユースケースには、事前トレーニングされた仮想アシスタントを利用します。</li> <li>ビジネス ユーザーや非開発者がノーコード機能を使用して AI アシスタントを構築できるようにします。</li> </ul>	会話およびサポートのユースケース: (人事、従業員のワークスペース、銀行業務、ヘルスケア、検索支援など)



# AIと言えはNVIDIA (Why the focus on PowerSwitch?)

# AI ワークロードの種類と設計上の考慮事項

## AI モデルの構築

### ワークロード

- **トレーニング**
  - 一般知識の基礎を構築するためにデータを大量消費する
  - 処理には数日から数週間の時間かかる
  - インフラ要件の規模と規模を考えると、「AI工場」やサービス・プロバイダーの領域となる
- **ファイン チューニング**
  - 特定の用途に特化した知識でベースモデルを改良する
  - データ・セキュリティと内部IPの保護を確保するため、オンプレミスで運用される
  - 小規模クラスターのため、**企業顧客の範囲**

## AI モデルの使用

- **推論**
  - ビジネス・アプリケーションへのモデルの活用
  - ビジネスにとって情報や結果が重要なので、オンプレミスで運用される
  - 会話形式のフローで結果を出すために、素早いシステム応答性が必要

### 要件

#### 大規模、数千のGPU

- 何千サイクルも繰り返し実行されるバースト的なエレファントフロー（大量の通信）
- 混雑や遅延が発生すると、テールレイテンシーが長くなり、ジョブの完了時間が大幅に延びる
- ファブリックには**高度に最適化されたトポロジーと負荷分散機能が必要**

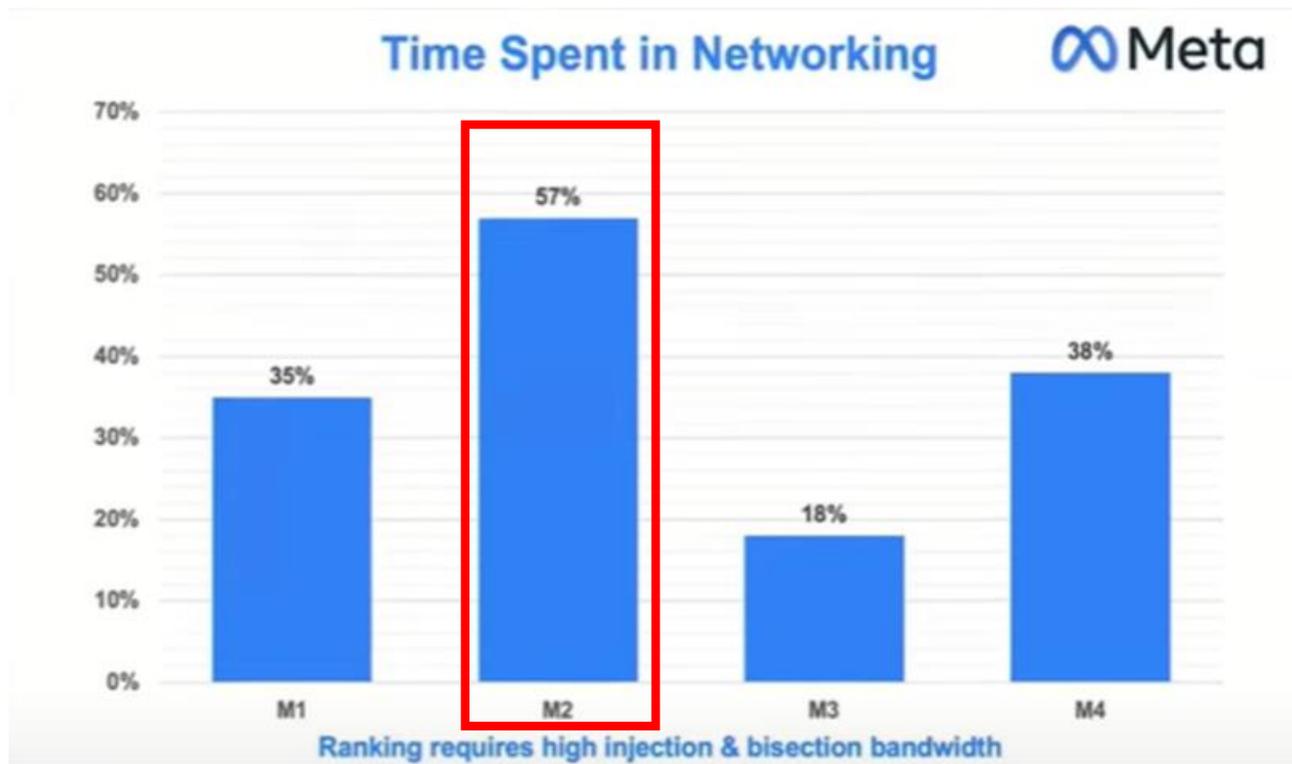
#### 小規模、数十から数百のGPU

- 一般的なベースモデルを出発点として導き出される
- トレーニング時と同様の作業負荷パターンだが、データセットが小さく、反復回数は少ない
- ジョブ完了時間全体に対する**帯域レイテンシーの影響は少ない**

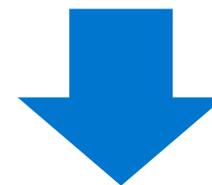
#### 最低スケールは、CPUまたはGPU

- ファブリックへの負荷が最も少ない
- リクエストに応え、迅速に対応する
- ほとんどの推論環境は**1台のホスト**で実行できる

# 生成AI特有の負荷でNetworkの存在感が増してきた



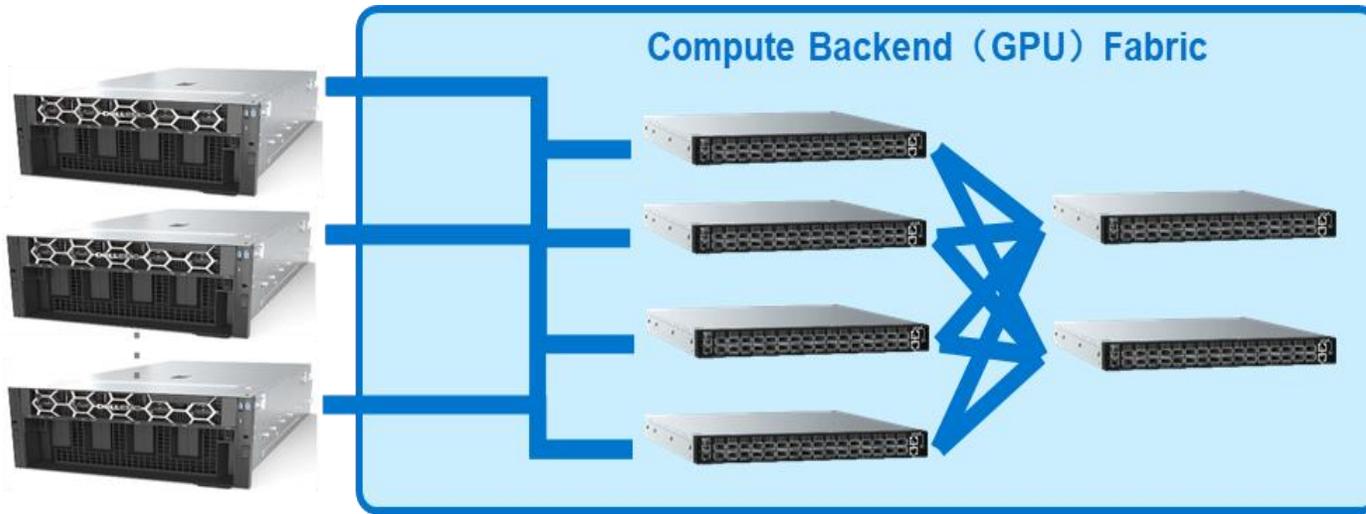
生成AIの推論・学習による  
Networkの滞留時間は50%を  
超えるケースも出てきた



Networkの性能が生成AIの  
パフォーマンスに影響する

2022 OCP Global Summit. 「Infrastructure for Large Scale AI: “Empowering Open” by Alexis Bjorlin (Meta). <https://www.opencompute.org/events/past-events/2022-ocp-global-summit>(参照2024.3/4)

# 生成AIにおけるNetworkへの要求



1. GPUクラスター間で膨大なワークロードが発生

→より低遅延・広帯域へ

2. 巨大なGPUクラスター構築

→拡張性

→コスト・入手性

3. サーバーストレージ含む  
大規模インフラを制御

→自動化・オーケストレーション

→ITインフラ包括的なサポート

# Dellで提供できる AI ネットワークの選択肢

	Dell イーサネット	NVIDIA InfiniBand	NVIDIA イーサネット
スイッチ	Dell PowerSwitch (Broadcom Tomahawk 搭載)	Nvidia Quantum-2	Nvidia Spectrum-4
サーバーNIC	Broadcom PCIe Broadcom Thor 2 Nvidia Connect-X 7 VPI Intel Gaudi Integrated NIC	Broadcom PCIe Nvidia Connect-X 7 NDR Nvidia BlueField-3 IB	Broadcom PCIe Nvidia Connect-X 7 VPI Nvidia BlueField-3
ケーブルと トランシーバー	DellのSKU	NvidiaのSKU	NvidiaのSKU
サービス	エンドツーエンドで Dellがサービス提供	サーバーのサービスはDell ファブリックのサービスはNvidia	サーバーのサービスはDell ファブリックのサービスは Nvidia
位置づけ	AMDおよびIntel GPUソリューションの推奨構成 Nvidia GPUソリューションでも選 択できる	最高のパフォーマンスと大規模な トレーニングワークロードのため には、Nvidia GPUソリューションが 推奨されます。	Nvidia独自のAI最適化イーサ ネット・ソリューションをお求めのお 客様向け

# AI用イーサネット ソリューションのための機能

継続的な高トラフィックに対応するため

## 帯域の有効活用 (アダプティブルーティング)

- Dell : IBヘッダでのハッシング (TH4 & TH5)  
コグニティブルーティング (TH5, 2H 2024)  
対応機種に注意！
- NVIDIA : パケットスプレー & NIC (BF3)での並び替え

分散計算の効率化 (テールレイテンシを重視) のため

## 効率的なロスレス転送

- Dell : DCQCN、スイッチでパケットに  
マークし、サーバーでフロー単位で出力抑制
- NVIDIA : DPUとスイッチで輻輳検知し、  
送信側NIC (BF3)でフロー単位にバッファ

## 広帯域・低遅延なハードウェア

- Dell : 400G - Z9664 (TH4)  
800G - Z9864 (TH5, Aug 2024)
- NVIDIA : 400G - SN5400 (Spectrum-4)  
800G - SN5600 (Spectrum-4)

## ネットワークの管理 (運用とモニタリング)

- Dell : SmartFabric Manager (Aug 2024)  
BeyondEdge、Augtera (3<sup>rd</sup> Party)
- NVIDIA : Unified Fabric Manager  
NetQ

イーサネットを利用するための、ROCEv2対応  
低遅延のための、カットスルー対応

# AI用イーサネット ソリューション ハードウェア

## NVIDIA Spectrum-X

### Components:

- NVIDIA SN5600 Switch
  - 64 x 800GE / 128 x 400GE ports (OSFP)
  - 51.2 Tb/s Switching capacity
  - NVIDIA **Spectrum-4** ASIC



- NVIDIA **BlueField-3** Super-NIC
  - 1 x 400GE / 2 x 200GE / 4 x 100GE



## Dell AI イーサネット

### Components:

- Dell Z9864F Switch
  - 64 x 800GE / 128 x 400GE ports (OSFP)
  - 51.2 Tb/s Switching capacity
  - Broadcom **Tomahawk-5** ASIC

2024年8月予定

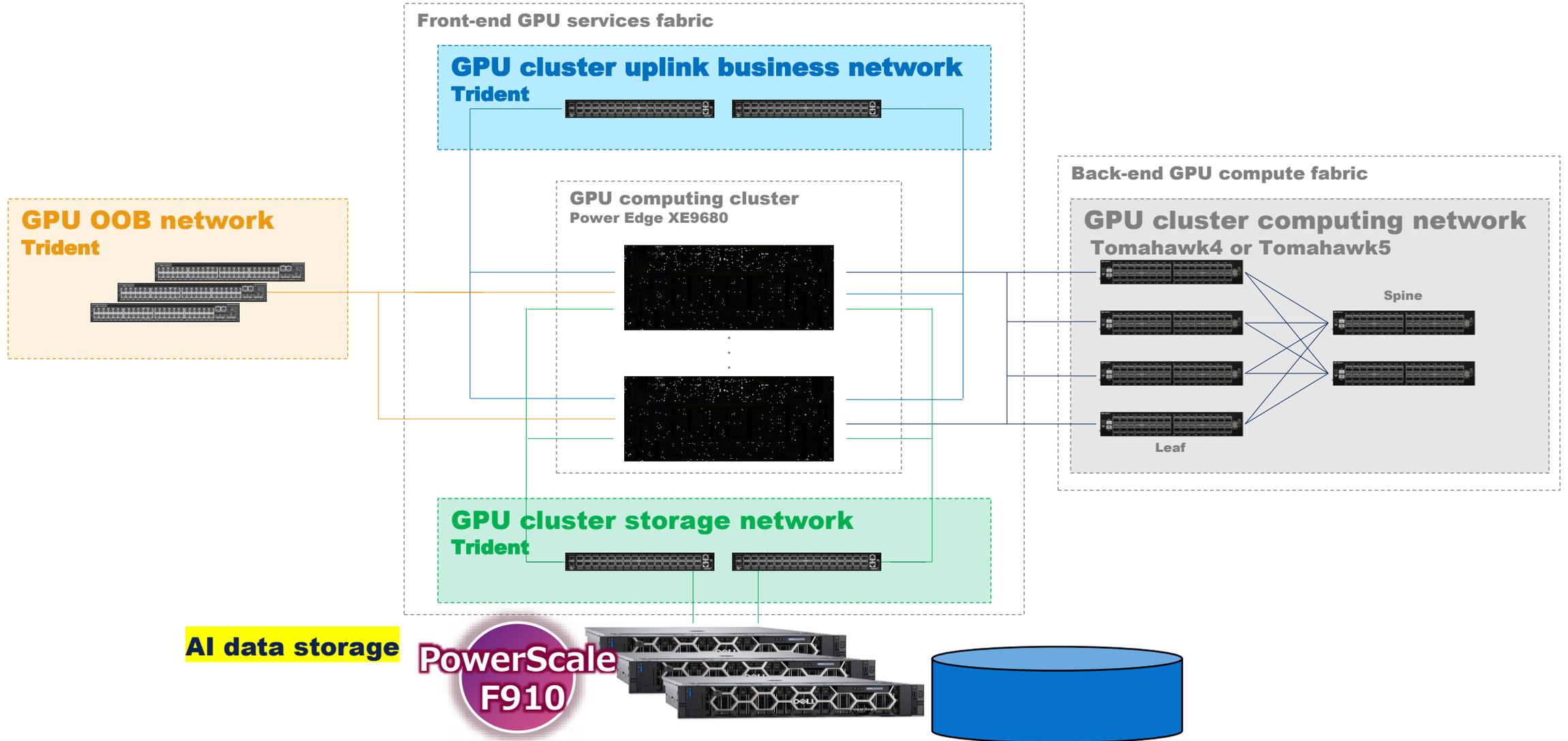


- Dell Westwood (Broadcom **Thor-2**) NIC
  - 1 x 400GE / 2 x 200GE / 4 x 100GE

2024年7月予定



# AI用イーサネット ソリューション 構成例



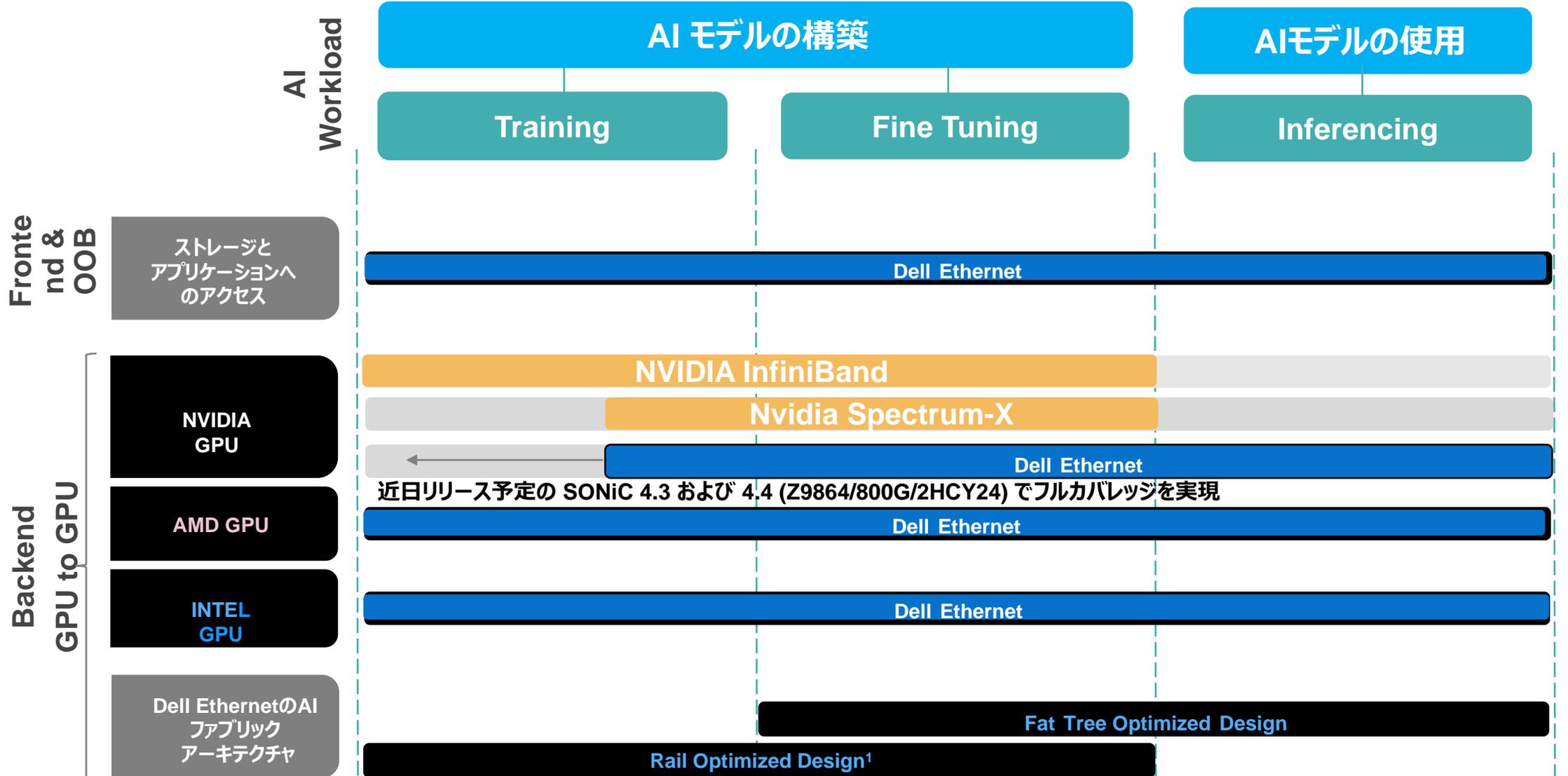
・ GPU cluster uplink business network : NPU = Trident 系機器での設計

・ GPU OOB network : NPU = Trident 系機器での設計

・ GPU cluster storage network : NPU = Trident 系機器での設計

・ GPU cluster computing network : NPU = Tomahawk 系機器での設計

# AI ワークロードからファブリックへのマッピング オプション



<sup>1</sup> Roadmap for late Q2/early Q3FY25

—  
ハイブリッドクラウド戦略よ、いずこへ？

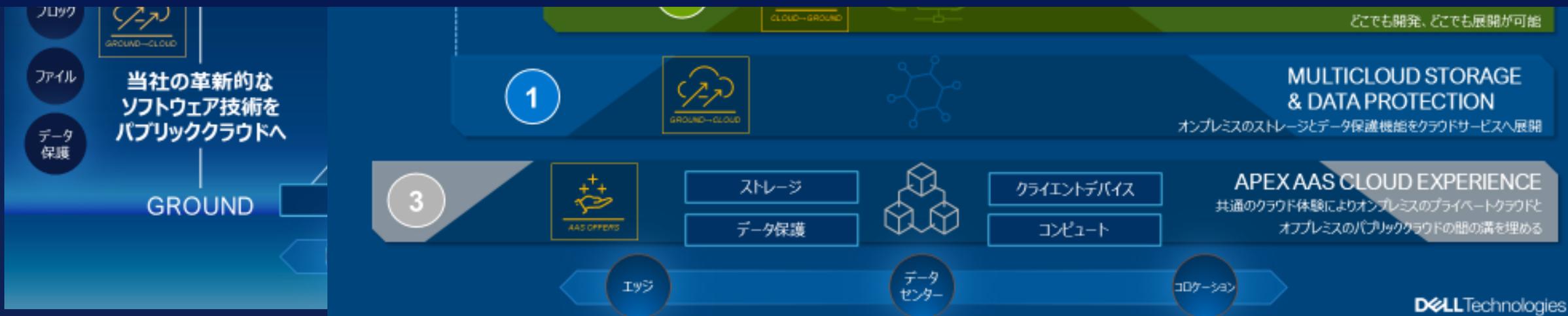
# そもそもDell Technologiesのハイブリッド戦略とは？

## クラウド最適化を実現するマルチクラウド戦略

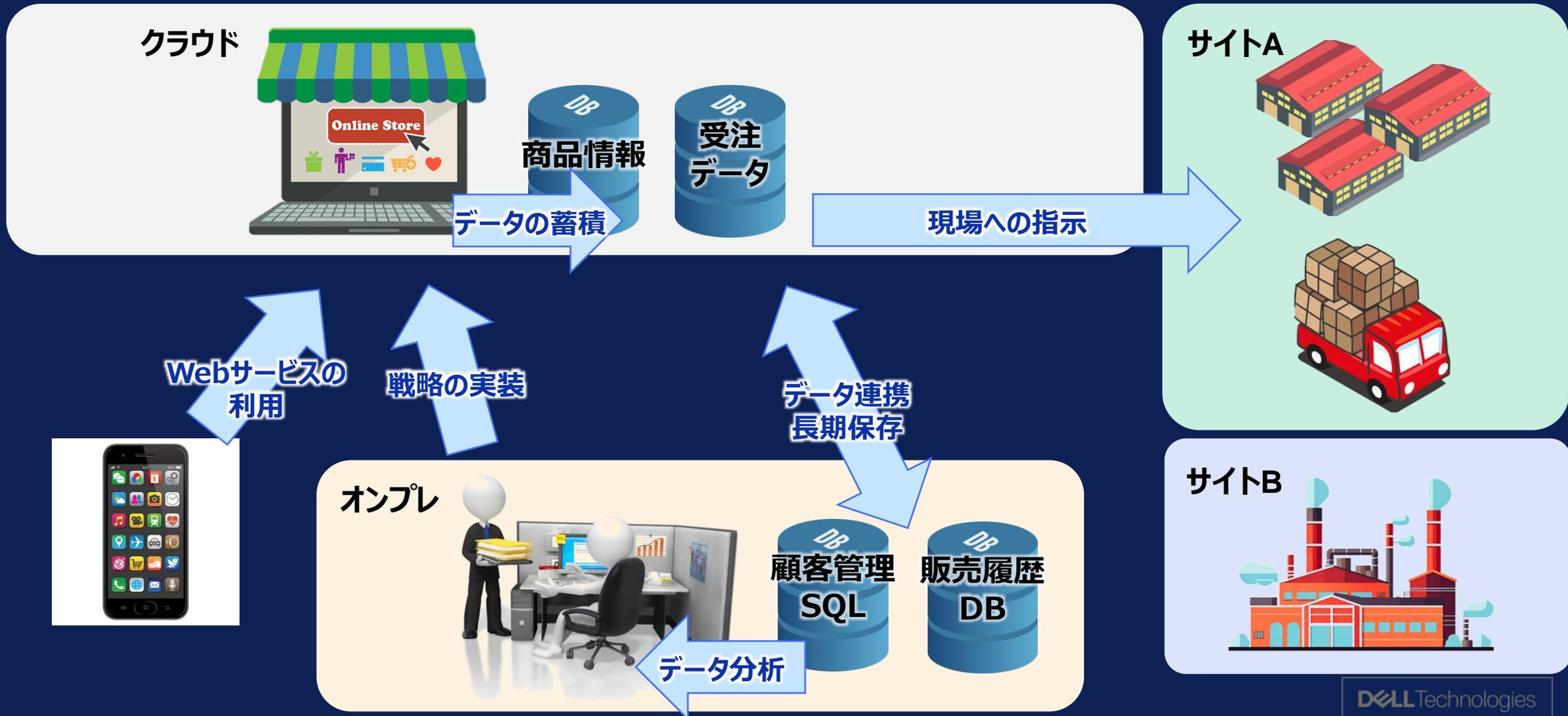
信頼できる



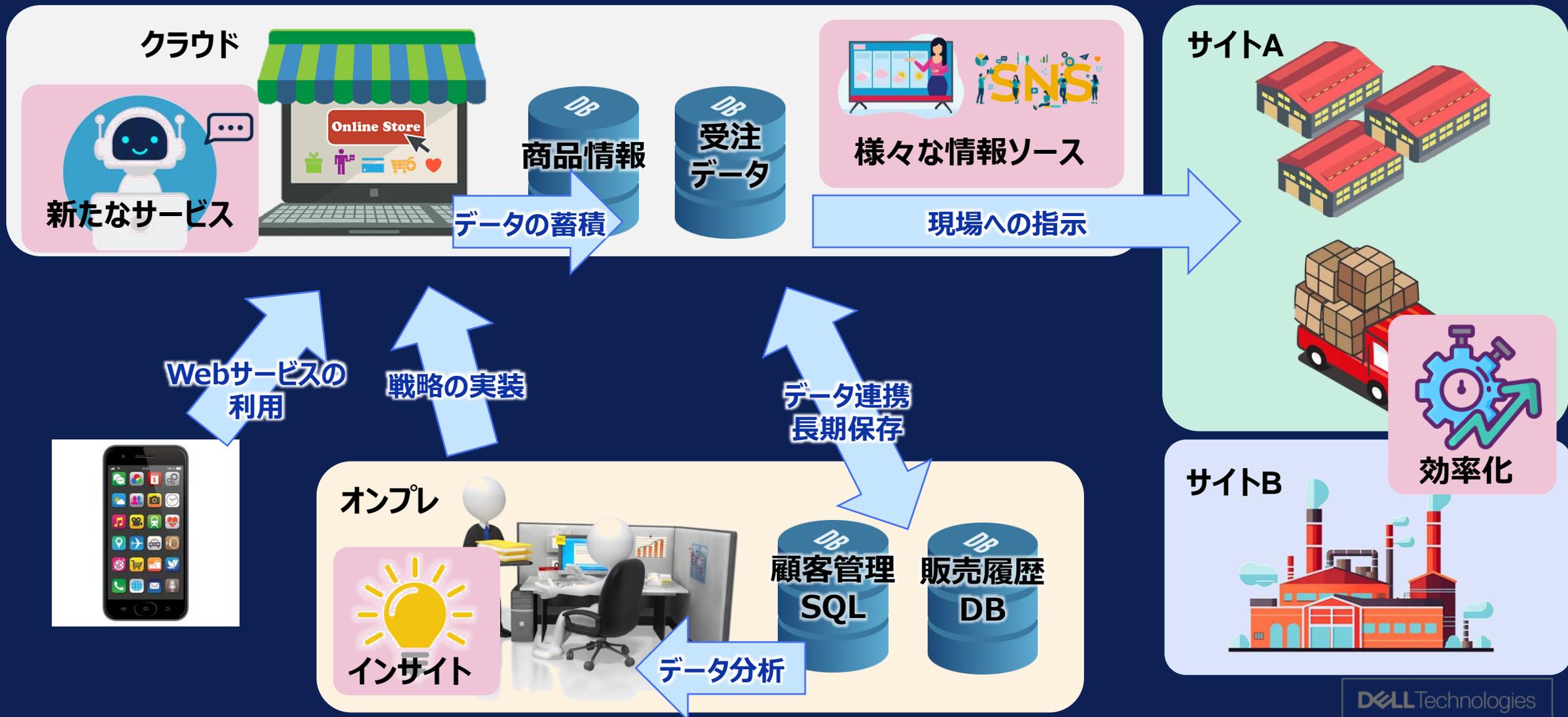
これはインフラ側からの視点なので  
ユーザは「どう使うことになるのか」がイマイチ分からない・・・



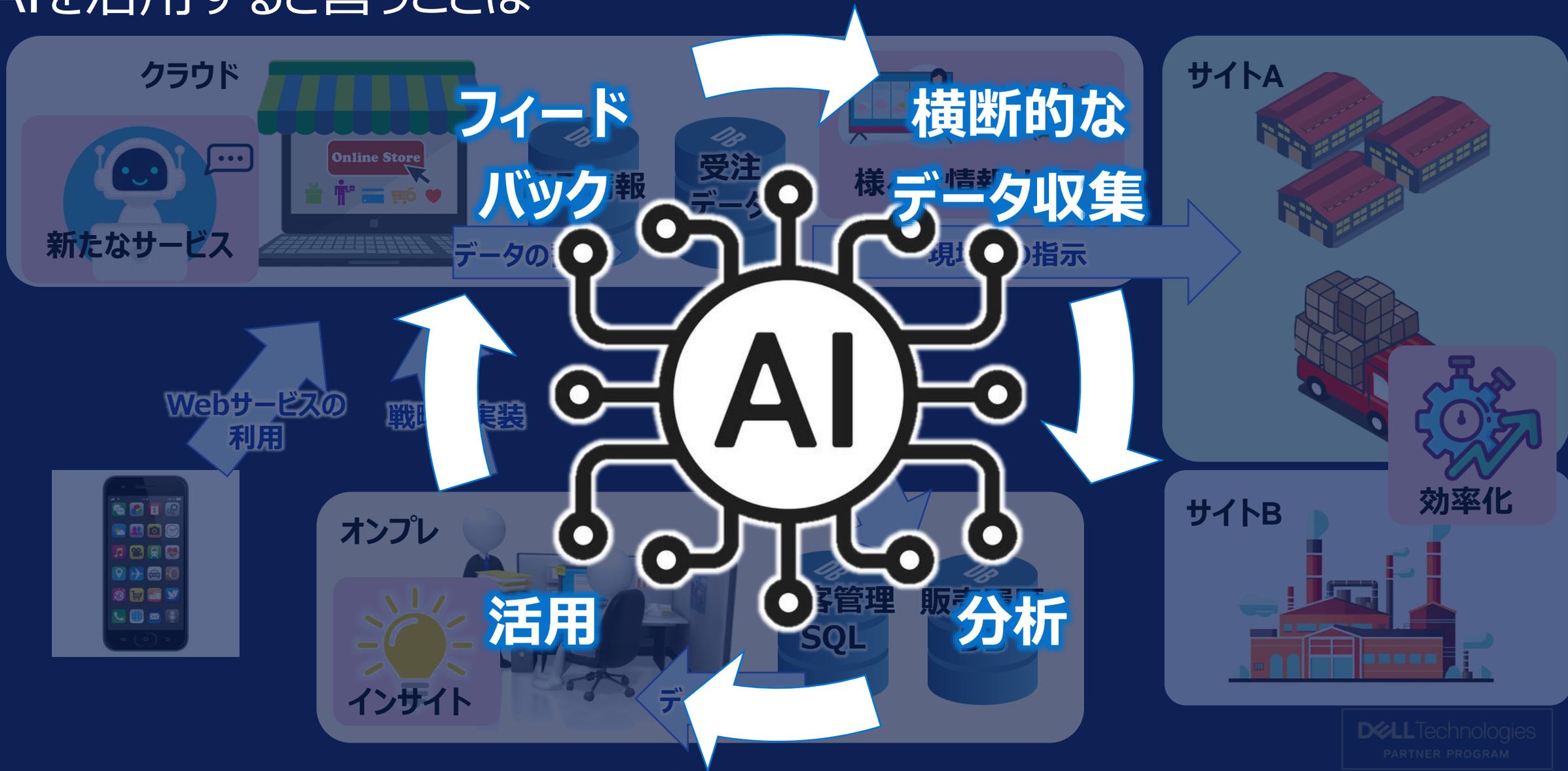
# 今まではアプリ視点で場所を選んでましたよね？



# AIを活用するということは・・・



# AIを活用するということは・・・



実際にAI活用を検討し始めると直面する問題・・・

スキル

育成 or アウトソース

コスト

CAPEX or OPEX

運用

利用時間 & 管理

環境

電源 & 空調

# スキルと運用の解決 = AI Factory



# コストの解決 = APEX Subscriptions

"OUTCOME" OFFERS

APEX Data  
Storage Services

APEX Compute

APEX Private Cloud

APEX Custom\*  
Flex on Demand (FOD)



## A P E X Subscriptions

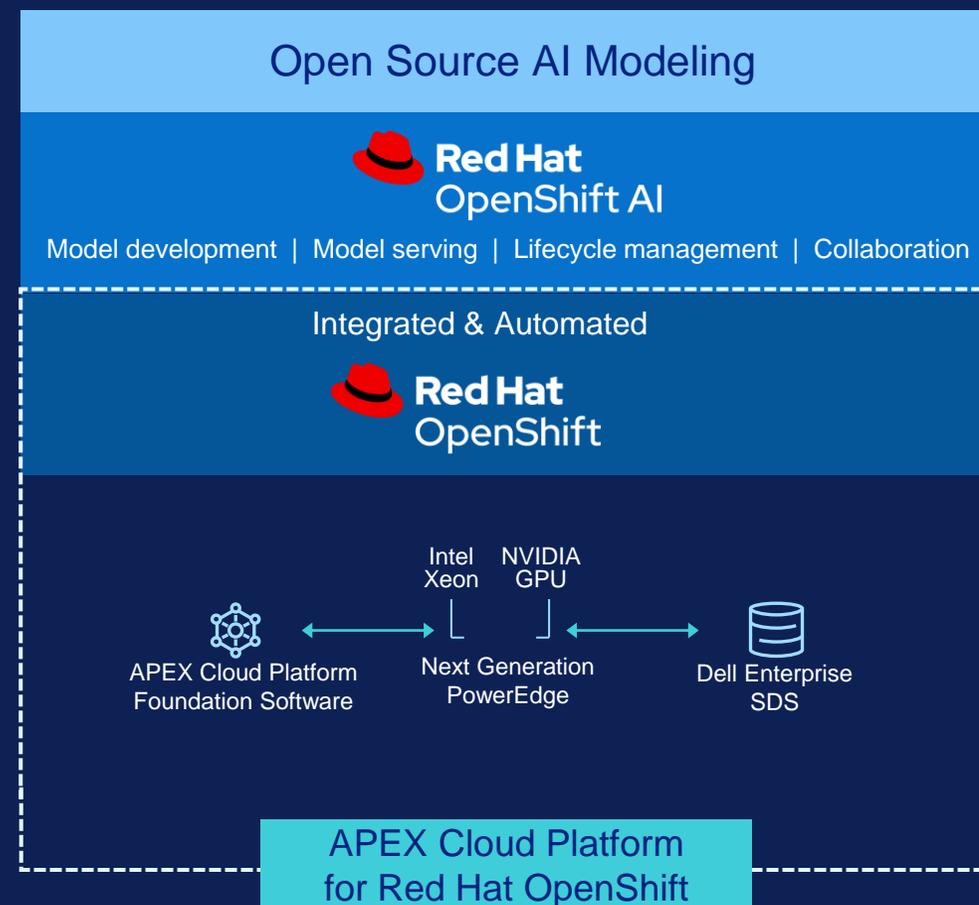
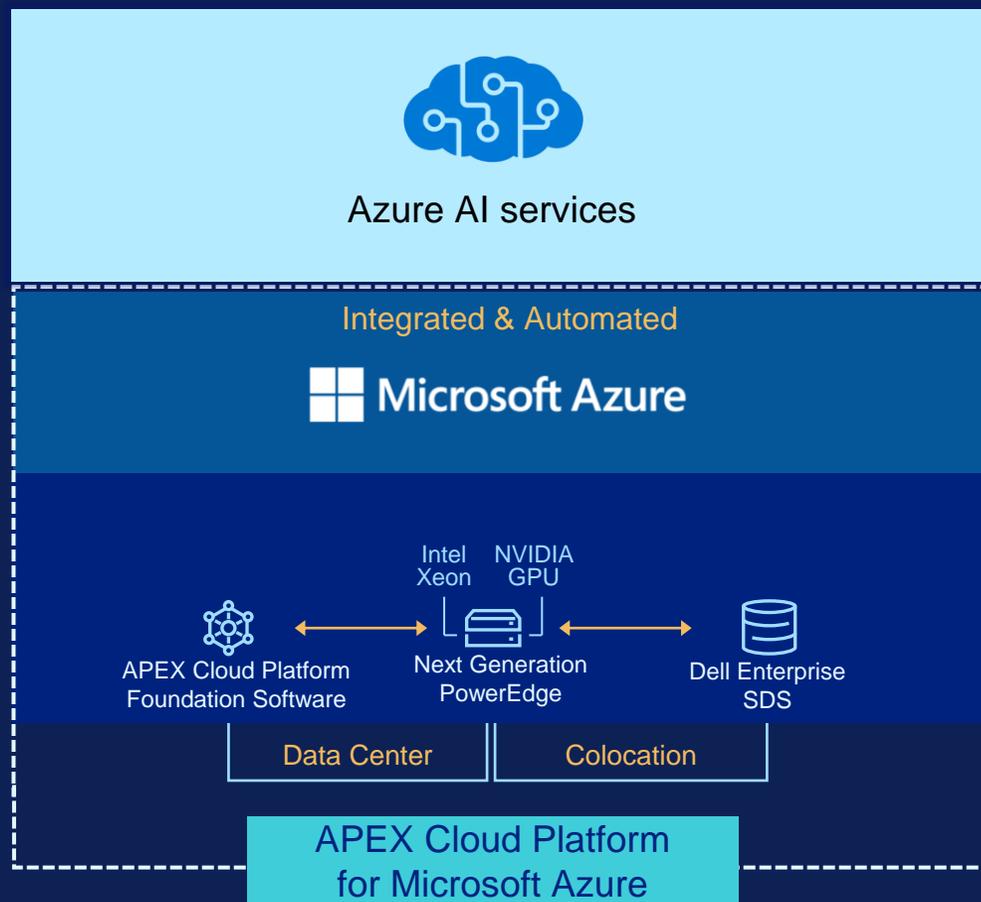
ISGポートフォリオ全体にわたる簡素化されたコンサンプション・モデルとデルのオプション・サービス\*\*

Configure the product → Select the services → Choose the consumption model

\* Data Center Utility can be considered a subscription but will not fall under the APEX Subscriptions external brand. No changes to DCU process/tools are being made with this release.

\*\* Services can be added to the subscription include sku'd infrastructure services and sku'd consulting services

# AI + APEX = APEX Cloud Platform大活躍！？



# 場所を問わないストレージとk8s管理 = APEX Navigator

## Dell APEX Navigator for Multicloud Storage

APEX File for AWS, APEX File for Azureのサポート *coming soon*

より多くのエンドポイント、より多くのクラウド、より多くの機能



## Dell APEX Navigator for Kubernetes

Dell Kubernetesデータストレージソフトウェアの導入、運用、管理を合理化

マルチクラウドのKubernetes環境全体の可視化とガバナンスで運用リスクを低減

リリース時点でPowerFlexをサポート。APEX Cloud Platform for Red Hat OpenShiftとPowerScaleのサポート予定 *coming soon*

最大 **95%**

クラウドインフラの設定にかかる時間を短縮<sup>2</sup>

**5分**

でKubernetesクラスタをオンボード<sup>3</sup>

**60秒** 以内

にアプリケーションモビリティジョブを開始<sup>4</sup>

<sup>1</sup> Subject to availability. Terms and conditions apply. For Dell APEX offers, see applicable offering agreements.

<sup>2</sup> Based on internal testing, January 2024, when comparing the manual configuration of cloud components vs. Navigator-driven cloud component orchestration and automated deployment of APEX Block Storage for AWS

<sup>3</sup> Based on internal testing, April 2024.

<sup>4</sup> Based on internal testing, April 2024.

**GA: United States only**

**DELL**Technologies

# AI-INのハイブリッドクラウド管理ツール：Dell APEX AIOps

可観測性とインシデント管理を一箇所にまとめる

GA済

インフラ可観測性

2024/10  
GA予定

アプリケーション  
可観測性

New!

GA済

インシデント管理

New!

AI/MLは、デジタル・ランドスケープ全体のITオペレーションを簡素化します。

強力なAIとノイズリダクションにより、問題を迅速に検出し、問題を軽減する

GenAIを搭載したAIOpsアシスタントを使用して、より速く答えを見つける



**10倍** インフラ問題解決までの時間短縮<sup>1</sup>

**70%** アプリケーションの問題解決時間の短縮<sup>2</sup>

**93%** お客様から報告された問題の削減<sup>3</sup>

<sup>1</sup>CloudIQユーザー調査、Dell Technologies、2021年。実際の結果は異なる場合があります。

<sup>2</sup>The Total Economic Impact™ Of IBM Instana Observability Cost Savings and Business Benefits Enabled By Instana Observability, Forrester, 2024。実際の結果は異なる場合があります。

<sup>3</sup>顧客調査、ムーグソフト、2023年。実際の結果は異なる場合がある。

An aerial night view of a city, likely Dubai, with numerous skyscrapers and illuminated streets. The image is overlaid with a complex network of glowing blue and white lines, suggesting a digital or data network. The text is prominently displayed in the upper half of the image.

**Why AI with Dell?**

**Because Dell AI is “End-to-End”**

**Unstoppable Hybrid Cloud**

Thanks for your  
great partnership.



DELL TECHNOLOGIES PARTNER PROGRAM  
**HEROES**



DELL TECHNOLOGIES PARTNER PROGRAM

**HEROES**