

# Dell EMC PowerEdgeサーバー上で のVMware vSphere 7とBitfusionの 連携

## 概要

組織は新たな商機を切り開き、ビジネスを加速するために、人工知能（AI）や機械学習（ML）を急速に取り入れています。これらのアプリケーションの開発と導入には、大規模な計算処理能力を備えた弾力性のあるプラットフォームが必要です。これは、従来のCPUに加えてGPUが急増する原因になっています。ただし、このような構成は、使用率、パフォーマンス、拡張性にかかわらず、ユーザーまたはフレームワークごとにGPUが静的に割り当てられることを想定している現行の手法や導入モデルにはうまく適合しません。大規模な導入環境では、GPUサーバーの共有やプール化が行えず、通常は、GPU使用率が平均15%未満のローカライズされたアプリケーションがサーバーで実行されます。これは、上級ユーザーの場合も同じです。結果として生じるサイロ化は、データセンターのモダナイゼーションを低迷させると同時に、設備投資と運用コストの増加を招きます。

vSphere Bitfusionは、GPUを従来のコンピューティング リソースのように抽象化、分割、自動化、共有できるファーストクラス リソースとして扱うことで、この問題を解決します。GPUには、共通インフラストラクチャの一部としてデータセンターのどこからでもアクセスできます。また、vSphere Bitfusionでは、複数のサーバーからのオンデマンドでネットワークに接続されたGPUを使用して、単一GPUの一部から複数のGPUまでコンピューティング リソースを動的に調整することもできます。これにより、使用率が2~4倍向上し、ビジネスの敏捷性とコスト削減に直接つながります。

## 対象者

このホワイト ペーパーは、コストを最適化した統合型vSphere Bitfusionソリューションの設計と導入を考えている方を対象としています。このホワイト ペーパーの情報は、GPUアクセラレーターが必要なAI/MLワークロードをサポートしているアーキテクトやエンジニアに最適です。ここでは、Dell

---

EMC PowerEdgeサーバー上で実行される、新たに統合されたvSphere Bitfusionソリューションについて詳しく説明します。

2020年6月



## 目次

|                      |   |
|----------------------|---|
| ソリューション概要 .....      | 1 |
| 実行フロー .....          | 3 |
| ソリューション アーキテクチャ..... | 4 |
| テクノロジーの概要 .....      | 5 |
| ネットワーク設計の考慮事項 .....  | 7 |
| まとめ .....            | 8 |
| 関連資料 .....           | 8 |
| 著者.....              | 8 |

本書の情報は作成時点のものです。Dellは本書の情報について、いかなる表明または保証もせず、その商品性、特定用途への適合性に関するいかなる保証も拒否します。

この資料に記載される、すべてのソフトウェアの使用、複製、および頒布には、当該のソフトウェア ライセンスが必要です。

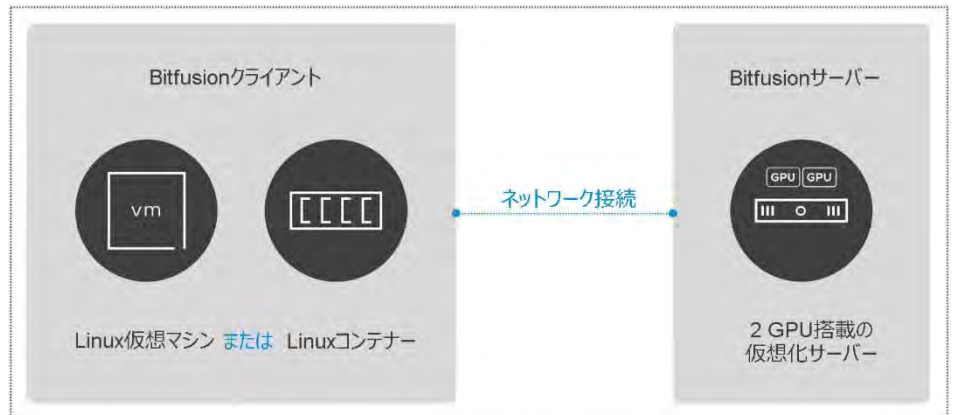
Copyright © 2020 Dell Inc. その関連会社。All Rights Reserved. (不許複製・禁無断転載) Dell, Dell Technologies, EMC、ならびにこれらに関連する商標およびDell又はEMCが提供する製品およびサービスにかかる商標はDell Inc.またはその関連会社の商標又は登録商標です。Intel、Intel ロゴは米国およびその他の国における Intel Corporation の商標です。および/またはその他の国におけるIntel Corporation またはその子会社の商標です。商標です。

本書に掲載されている情報は、発行日の時点で正確な情報であり、予告なく変更される場合があります。

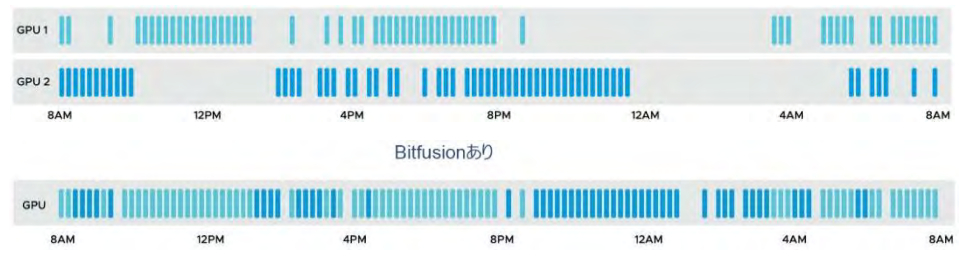
Published in the USA 6/20.

## ソリューション概要

VMware vSphere Bitfusionは、ハイパーバイザーやAI/MLアプリケーションの変更や特別なソフトウェアを必要とせずに簡単に導入できるクライアント/サーバー モデル アーキテクチャに基づいています。仮想マシンまたはコンテナであるクライアントは、AIワークロードをローカルで実行し、ネットワーク全体でGPUへのアクセスを提供する仮想サーバーに接続します。



Bitfusionを使用すると、GPUを任意のマシンにリモートで接続して総所有コストを削減できるだけでなく、単一のGPUを任意のサイズの複数の仮想GPUに分割して、複数のワークロードを同じGPUで並列実行することもできるため、使用率が高まります。



Bitfusionでは、GPUのリアルタイムでの集約と分割が可能のため、十分に活用されていないGPUコンピューティング サイクルをより効率的に利用できます。たとえば、通常はCPUマシンでワークロードを維持し、ワークロードにGPUが必要な場合にのみリモートでGPUを接続することで、GPUの使用率を2~4倍に高めることができます。

Bitfusionにより、機械学習アルゴリズムの初期のテストや検証など、ノードとGPUをすべて利用できるとは限らないユースケースのユニット エコノミクスが向上します。断片化されたGPU（GPUの20分の1程度）を実行時に割り当てることで、同じ物理ハードウェアで以前よりもはるかに多くのユーザーをサポートできます。これにより、インフラストラクチャ管理の範囲と負担を増加させるさまざまな低処理能力デバイスに頼らなくても、きめ細かなリソース制御を行えるようになります。Bitfusionは、ハイ パフォーマンスのGPUインスタンスと大幅なコスト削減を同時に実現して、開発とテストのさまざまな段階に合わせて支出と容量のサイズを適切に調整できるようにします。

## BITFUSIONクライアント

Bitfusionクライアントは、ネットワークを介してGPUまたはGPUの一部にリモートで接続する仮想マシンまたはコンテナです。この機能は、GPU言語呼び出しをインターセプトしてリモートのBitfusion GPUサーバー インスタンスにリダイレクトする軽量エージェントによって実現されています。クライアント サービスは、GPUリソースの割り当てとジョブの完了時のリソース解放を行う実行コマンドを実行します。

エンド ユーザーは、GPUリソースを予約形式で要求し、システムがシャットダウンするかサービスが停止するまでリソースを保持することもできます。また、アイドル時間を設定して、未使用のGPUリソースを解放してプールに戻すことも可能です。

クライアントのセットアップはシンプルかつ迅速にできます。コマンド ラインを使用してエージェントをダウンロードして、ゲストOS（CentOS 7またはUbuntu 16.04/18.04）にインストールします。エージェントはAI/MLアプリケーションと基盤となる仮想インフラストラクチャの両方に対して透過的であるため、vMotion、High-Availability、Distributed Resource SchedulerなどのvSphere機能は影響を受けず、他のGPUアプローチにはない大きな利点です。エージェントがインストールされたら、リモートのBitfusion GPU クラスタに対する可視性とアクセスを確保するために、vCenter UIでBitfusionクライアントを有効にする必要があります。環境内のさまざまなBitfusionクライアントに関する情報は、vCenterユーザー インターフェイス（UI）のBitfusionプラグインの [クライアント] タブに表示されます。

Bitfusionクライアントは、vSphere 6.7 Standard以降を実行している仮想化サーバーに導入できます。

## BITFUSIONサーバー

Bitfusionサーバーは、vSphere 7 Enterprise Plusを実行している仮想化されたGPUサーバーに仮想アプライアンス（OVA）として導入されます。Bitfusionサーバーは、ホストのGPUリソースを公開し、Bitfusionクライアントからの接続を受信し、着信AI/ML GPU言語リクエストを物理GPUに渡して、結果を返します。

セットアップは簡単で、インストーラーに表示される手順に従って、インターフェイスの数、IPアドレス情報、DNSなどを指定します。サーバー アプライアンスが正常に導入されたら、vCenter UIでアプライアンスを有効にする必要があります。次のステップは、DirectPath I/O（パススルー）経由でサーバー アプライアンスをGPUに接続し、サーバー アプライアンスのvmx構成ファイルに変数を追加して、サーバー アプライアンスをvCenterに登録することです。

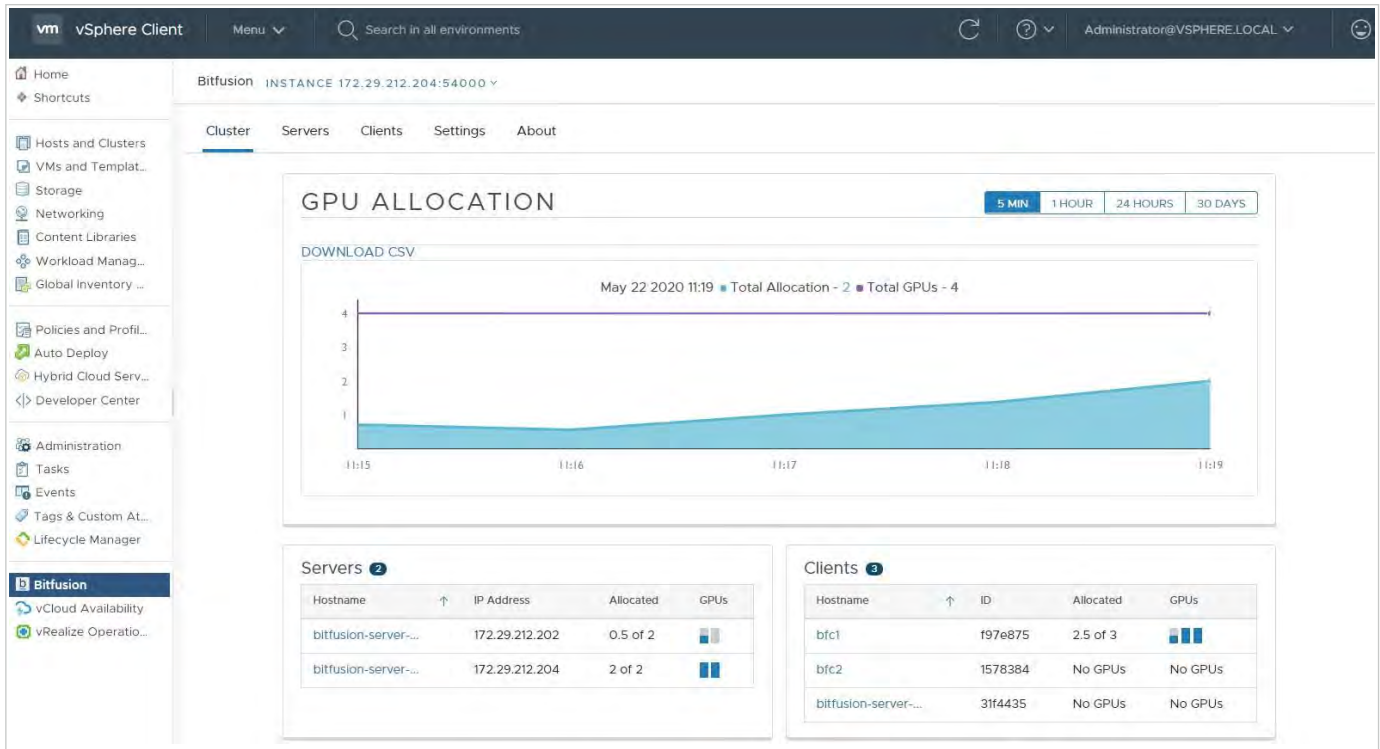
注：Bitfusionクラスタは複数のBitfusionサーバーで構成されていますが、これらのサーバーが同じ物理クラスタ内または同じvSphereクラスタ内にある必要はありません。クラスタは、データセンター全体に分散されたサーバーから集約できます。

## BITFUSIONプラグイン

vCenterサーバー アプライアンスに統合されたBitfusionプラグインは、Bitfusionサーバーとクライアントがオンラインになったときにその登録を追跡します。また、クリーンなBitfusion UIを提供して、GPUクラスタ、サーバー、クライアント、ユーザーを監視および管理できるようになるだけでなく、GPUクォータなどのグローバル クライアント デフォルトを定義し、アイドル状態のGPUを自動接続解除して他のクライアントのために解放するなどのさらに詳細な設定が可能になります。

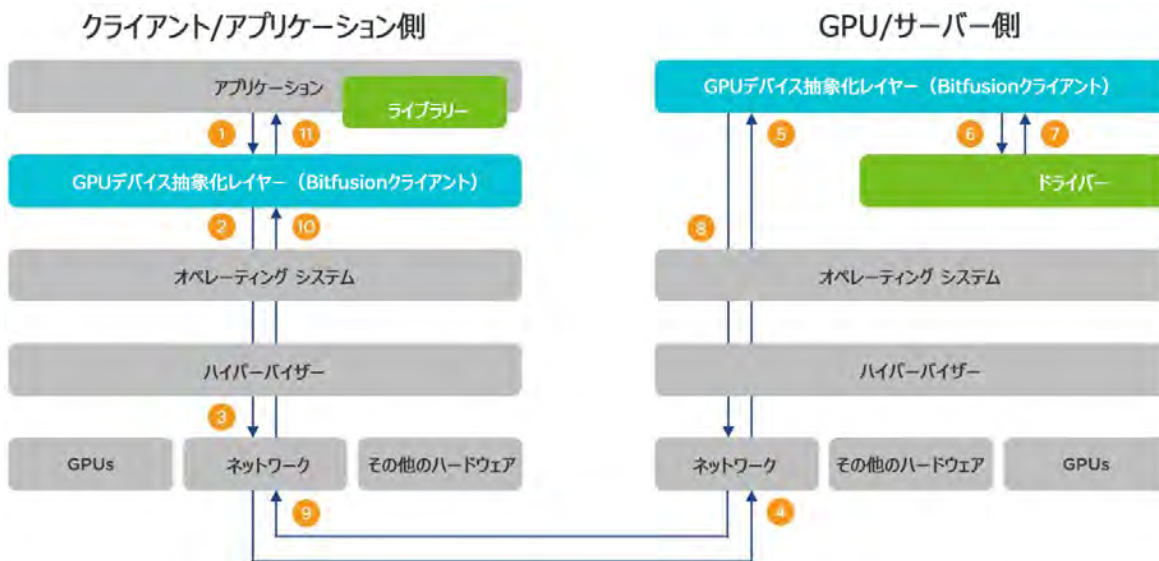
ダッシュボードには、GPUの割り当てと使用状況が複数の観点から可視化されます。クラスタ ビューには、Bitfusionサーバー、各サーバーのGPUの数、使用中のGPU、GPUの使用状況の推移が一覧表示されます。サーバー ビューには、接続されているクライアントとそのGPU使用率、GPUメモリー消費量、ネットワークトラフィックなど、各サーバーの情報がより詳細に表示されます。また、最適なパフォーマンスを確保するためにサーバーのヘルス チェックも実行されます。

割り当てられたGPUの数と使用状況の推移を表示するクライアントビューもあります。



## 実行フロー

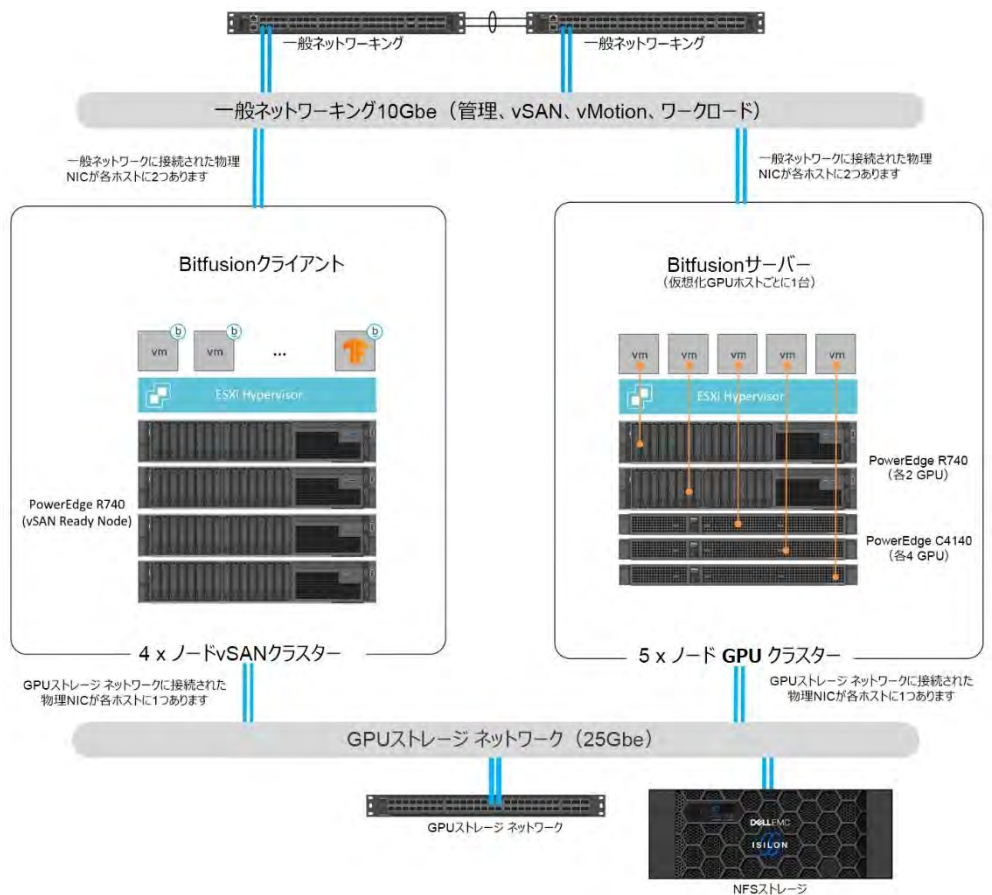
Bitfusionは、データセンター全体の仮想マシンが消費するプールされたリソースとしてホストの物理GPUを公開する、透過的なソフトウェアレイヤーとして実行されます。AI/MLジョブが開始されると、vSphere BitfusionがGPUリソースを割り当て、ネットワーク上で動的に接続します。次の図は、Bitfusionクライアント内のエージェントとBitfusionサーバー内のGPUとの間のエンド ツー エンドの実行フローを示しています。



1. APIストリームがアプリケーションからキャプチャーされ、Bitfusionランタイムにリダイレクトされて、デバイスID番号Nを要求します。
2. Bitfusionランタイムが、ローカルおよびリモートのすべての使用可能なGPUに基づいて、デバイスID NをN'に変換します。
3. BitfusionランタイムがGPU関数をマーシャリングし、ネットワーク経由でリクエストを送信します。
4. Bitfusionサーバーをターゲットとするリモートシステムにリクエストが到着します。
5. Bitfusionサーバーで受信されたリクエストが、マーシャリング解除されて変換されます。
6. 変換されたリクエストが、通常の関数呼び出しとしてGPUツールキットに送信されます。
7. ツールキットとライブラリーが、通常のフローとしてドライバーAPIとGPUドライバーを呼び出します。
8. Bitfusionサーバーランタイムが、GPUリターンコードと結果の出力をマーシャリングします。
9. Bitfusionサーバーをターゲットとするリモートシステムにリクエストが到着します。
10. リクエストがBitfusionクライアントで受信されてマーシャリング解除され、デバイスIDがN'からNに再変換されます。
11. コードがアプリケーションに返されます。アプリケーションが実行を再開します。

## ソリューションアーキテクチャ

次の図は、Dell Technologiesハードウェア上の新しいvSphere Bitfusion統合を検証するための環境設定を示しています。左側では、4つのPowerEdge R740 vSAN Ready Nodeで構成されるvSANクラスターが複数のクライアントVMをホストしており、各VMにBitfusionエージェントがインストールされています。右側には、2つのPowerEdge R740ノードと3つのPowerEdge C4140ノードで構成されるGPUクラスターがあり、サーバーの内部GPUに直接接続されたBitfusionサーバーアプライアンスが各ノードに搭載されています。このソリューションには、管理、vSAN、vMotion、VMトラフィックのための冗長10 Gbeネットワークに加え、GPU通信およびNFSストレージ専用の25 GbEネットワークが含まれています。



## テクノロジーの概要

ここでは、このソリューションで使用されている以下のテクノロジーの概要について説明します。

- VMware vSphere 7 Ent +
- VMware vSAN 7
- VMware vCenter 7
- Dell EMC PowerEdge R740 vSAN Ready Node AF6
- Dell EMC PowerEdge C4140
- Dell EMC PowerEdge R740

### VMWARE VSPHERE 7 ENTERPRISE PLUS

vSphereはVMwareの仮想化プラットフォームであり、CPU、ストレージ、ネットワークの各リソースを含む集約型コンピューティング インフラストラクチャにデータセンターを変換します。vSphereは、これらのインフラストラクチャを統合型の運用環境として管理し、その環境に参加しているデータセンターを管理するためのツールを提供します。vSphere BitfusionはvSphere 7 Enterprise Plusの機能ですが、BitfusionクライアントはvSphere 6.7以降の任意のエディションに導入できます。

### VMWARE VSAN 7

vSANは、VMwareのソフトウェア定義 ストレージ ソリューションです。vSphereハイパーバイザーおよびvCenter Server統合管理ソフトウェアとネイティブに統合されており、ストレージの導入と管理、およびストレージ サイロの削除を大幅にシンプルにします。AI/MLワークロードに関連付けられた大規模なデータセットは共有ストレージに保存する必要がありますが、ハイパーコンバージドvSANストレージは、VMディスク（VMDK）を保存するためのハイ パフォーマンスで拡張性とコスト パフォーマンスに優れたソリューションを提供します。

### VMWARE VCENTER 7

vCenter Serverは、データセンター全体でESXiホスト、仮想マシン、コンテナ、Kubernetesクラスターを管理するために使用される、vSphereの一元管理コンポーネントです。また、Distributed Resource Scheduler、High-Availability、vMotion、Storage vMotionなどの多くの高度なvSphere機能を管理することに加え、Bitfusionプラグインを使用した仮想GPUクラスター、サーバー、クライアントの可視化と管理のためのダッシュボードも提供します。

### DELL EMC POWEREDGE R740 VSAN READY NODE

Dell EMC vSAN Ready Nodeは、認定された構成で導入リスクを低減し、ストレージ効率を最大50%向上させて、vSANクラスターの構築または拡張を迅速化するのに役立つ、事前構成済みの構成要素です。各Ready Nodeは、必要な量のCPU、メモリー、ネットワーク、I/Oコントローラーおよびストレージ（SSD、HDD、またはフラッシュ デバイス）を備え、vSAN用に最適に構成されています。

- ✓ プリインストールされたハイパーコンバージド インフラストラクチャ
- ✓ バンドルされたハイパーコンバージド インフラストラクチャ ソフトウェア ライセンス
- ✓ ハードウェアとハイパーコンバージドインフラストラクチャ ソフトウェアの両方をカバーする、サーバー ベンダーによる単一ベンダー サポート





## DELL EMC POWEREDGE C4140サーバー

PowerEdge C4140は、要求度の最も高いテクニカル コンピューティング ワークロードを処理するように設計された、高密度の専用ラック サーバーです。C4140は、第2世代インテル® Xeon®スケーラブル・プロセッサを搭載したPowerEdgeサーバー ポートフォリオの主要なGPU高速化プラットフォームであり、異機種混在データセンター環境で拡張性の高いビジネス アーキテクチャを実現します。

わずか1Uのスペースに4つのダブルワイド アクセラレーターを備えたC4140は、スペース、コスト、管理要件を軽減しながら卓越したパフォーマンスを実現し、密度を最大限に高める、vSphere Bitfusion サーバー アプライアンスに最適な選択肢です。

提供される主な機能は次のとおりです。

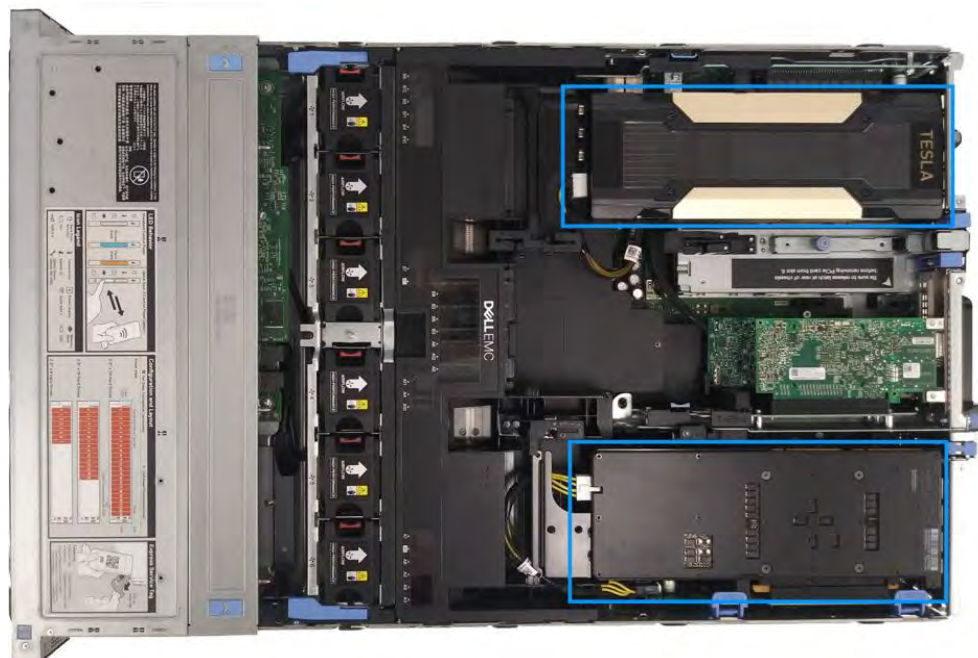
- 異機種混在データセンター環境で拡張性の高いビジネス アーキテクチャを実現する、PowerEdge サーバー ポートフォリオの主要なGPU高速化プラットフォーム
- 1Uフォーム ファクターに4つのGPUアクセラレーターを搭載してパフォーマンス密度を最大限高めることによって、テクニカル コンピューティング ワークロードを実行するために必要なサーバーの数を削減
- 優れた熱効率と高いアクセラレーター柔軟性を生み出す革新的なプラットフォーム設計



## DELL EMC POWEREDGE R740サーバー

Dell EMC PowerEdge R740は、2ソケットの2Uラック サーバーであり、メモリー、I/O容量、ネットワークの拡張性に優れたオプションを使用して、複雑なワークロードを実行できるように設計されています。R740は、第2世代インテル® Xeon® スケーラブル・プロセッサ・ファミリー、最大24個のDIMM、PCI Express® (PCIe) 3.0対応の拡張スロット、さまざまなネットワーク インターフェイス テクノロジーを備えています。

PowerEdge R740は、データ ウェアハウス、eコマース、データベース、ハイパフォーマンス コンピューティング (HPC) など、要求度の高いワークロードとアプリケーションを処理できる汎用プラットフォームです。PowerEdge R740の動的構成により、最大3つのダブルワイド300W GPUまたは最大6つのシングルワイド150W GPUをサポートできます。



## GPUサポート

vSphere Bitfusionはビジネス向けGPUをサポートしています（実際、vSphere BitfusionはGPU言語コードの新しいリリースと迅速に統合し、AI/MLアプリケーションとGPUのマトリックス上でテストしています）。

## ネットワーク設計の考慮事項

クライアントVMネットワークは、2つのvNICで構成されています。1つは従来の企業ネットワークを介してVMインスタンスにアクセスするため、もう1つはデータセットまたはホーム ディレクトリをホストしている必要なストレージ リソースとGPUにアクセスするためです。GPUネットワークは少なくとも10 GbEにすることを勧めますが、必要なパフォーマンス レベルはユースケースによって決まります。プールされたGPUへの接続のレイテンシーを最小限に抑え、パフォーマンスを可能な限り高めるため、PVRDMAの使用をお勧めします。

サーバー アプライアンスVMネットワークも2つのvNICで構成されています。1つは、vSphere Bitfusionサーバー アプライアンスを管理するためのvCenterアクセス用にVMware管理ネットワークとの接続に割り当てられており、2つ目のインターフェイスは、GPUとストレージへのアクセスに使用されます。どちらのインターフェイスでもVMXNet3が使用されていますが、レイテンシーを可能な限り低く抑える必要がある場合は、PVRDMAを使用してネットワーク パフォーマンスを向上させることもできます（これには、PVRDMAがクライアントでも有効になっていることが必要です）。

物理GPUサーバーには3つの物理インターフェイスがあり、それぞれが標準のvSwitchにマッピングされています。2つのインターフェイスはVMwareサービスの冗長リンクとして使用され、1つの追加インターフェイスはGPUにアクセスするために使用されます。重要な指針としては、スイッチ ホップ数を最小限に抑えるGPUアクセス専用の高速ネットワークを使用すると、レイテンシーが削減され、vSphere Bitfusionクライアントとサーバー間の通信に不要なジッターの影響が及ばなくなります。

## まとめ

VMware vSphere 7とBitfusionをDell EMCハードウェアで実行すると、データセンター全体のAI/MLワークロードにGPUアクセラレーションを提供するための、拡張性とコストパフォーマンスに優れたハイパフォーマンスソリューションが実現します。このソリューションは導入が簡単で、AI/MLアプリケーションと基盤となる仮想インフラストラクチャの両方に対して透過的であり、vMotion、High-Availability、Distributed Resource Schedulerなど、他のGPUアプローチでは必ずしも実現できないvSphere機能を維持します。

このソリューションは、データセンター全体のサーバーからのオンデマンド ネットワーク接続GPUを提供し、単一GPUの一部から複数のGPUまでコンピューティング リソースを動的に調整する機能によって、GPUリソースの使用率を最大限に高めて、必要な場所に処理能力が確実に割り当てられるようにします。これにより、使用率が2~4倍向上し、ビジネスの敏捷性とコスト削減に直接つながるため、組織は新たな商機を切り開き、ビジネスを加速できるようになります。

## 関連資料

[Dell EMC vSAN Ready Node](#)

[仮想SANクラスターの設計とサイズ設定](#)

[GPUおよびDell EMCサーバーのワークロード別推奨事項](#)

## 著者

VMware、CTOオフィス、Dell Technologiesソリューション、スタッフ ソリューション アーキテクト、  
**Chris Gully**

VMware、スタッフ ソリューション エンジニア、**Shawn Kelly**

[Dell Technologies Customer Solution Center](#)

詳細は[delltechnologies.com/poweredge-vmware](https://delltechnologies.com/poweredge-vmware)をご覧ください。