



サーバー アクセラレーター

アプリケーションを飛躍的に加速化



インサイトとイノベーションの促進

デジタル エンタープライズの成功には、Big Fast Dataの活用が不可欠です。ところが、データセットが増加するに伴い、従来のデータセンターはパフォーマンスと拡張性の限界に達しつつあります。特に、リアルタイム データ ソースの取得とクエリーに関しては状況が差し迫っています。

10人中7人

今後5～10年間の最重要テクノロジー3つのうちの1つとしてAIを挙げているテクノロジー部門経営陣¹

最大100倍

GPUあたりのCPU数²

600個を超えるアプリケーション

高速化とすべてのディープ ラーニング フレームワーク³

一部のアプリケーションでは、可視化、モデリング、シミュレーションを高速化する目的でアクセラレーターがかなり前から利用されていますが、最近では、さらに多くの主流アプリケーションにおいてインサイトとイノベーションの促進にアクセラレーターを活用できるようになりました。グラフィックス プロセッシング ユニット (GPU) やフィールド プログラマブル ゲート アレイ (FPGA) などのアクセラレーターは、並列処理を利用して一度に大量のデータを迅速に処理することによって、CPUを補完し、高速化しています。高速化されたデータセンターでは、経済性も高まります。つまり、従来よりも少ない台数のサーバーでパフォーマンスを飛躍的に向上させることができるため、インサイトの迅速な取得とコスト削減につながります。

現在、複数の業界において組織が競合他社に差をつけるため、サーバー アクセラレーターの採用を進めています。そのねらいは、データから得られるインサイトで製品とサービスを強化し、アプリケーション パフォーマンスの向上によって生産性を高め、高速でパワフルな分析で運用を最適化して、すべてを以前よりも迅速に処理することで市場投入までの期間を短縮することにあります。

Dell EMCは、お客様がアプリケーションを高速化できるよう、Dell EMC PowerEdgeサーバーで使用可能なサーバー アクセラレーターの選択肢をご用意しています。

アクセラレーターのユースケース

- **機械学習とディープ ラーニング** — アクセラレーターは、トレーニングと推論の両方のワークロードを高速化するために必要な並列処理を実現することにより、AIを理論からメインストリームへ移行しました。
- **予測分析** — アクセラレーターによって実現するAIは、分析を高速化できます。これにより、動的な関連づけが可能になり、驚異的なスピード、精度、規模で予測可能な成果がもたらされます。
- **データベースの高速化** — アクセラレーターで集計、ソート、グループ化の処理を高速化して、従来のデータベースを過負荷状態にしている複雑な分析の演算を解くことができます。
- **ストリーミング データ** — モノのインターネット (IoT) によって大量のデータが生成されています。アクセラレーターを使用すると、ストリーミング データの取得、探索、可視化を同時に行ってリアルタイムで分析できます。

¹ 『2019 Edelman AI Survey』 (2019年3月)

² 『NVIDIA Tesla V100 Tensor Core GPU』 (2019年5月)

³ 『NVIDIA Tesla P100』 (2019年5月)

GPU

グラフィックス プロセッシング ユニット (GPU) は、コンピューティング パフォーマンスを高速化するように設計されたコプロセッサです。通常、各GPUには、数学機能を効率的に実行するように設計された数千個のコアが搭載されています。ワークロードの一部はCPUからGPUにオフロードされますが、コードの残りの部分はCPUで実行されるため、全体的なアプリケーション パフォーマンスが向上します。

Dell EMCでは、サーバーのPCIeスロットに挿入するPCIeカード タイプとサーバーのマザーボードにマウントできるSXM2モジュール タイプの両方のGPUを提供しています。

SXM2モジュール搭載のPowerEdge C4140サーバーの場合は、NVLINK高速インターコネクトをGPU間で利用できます。

FPGA

フィールド プログラマブル ゲート アレイ (FPGA) は、特定の種類のアルゴリズムを従来のソフトウェア ソリューションの最大1,000倍のスピードで実行するために必要な技術的特性を備えたアクセラレーターです。FPGAは、特定のタスクまたはアプリケーションの要件を正確に満たすよう構成できます。

Dell EMCでは、サーバーのPCIe スロットに挿入するPCIeカードとして、さまざまなFPGAを提供しています。

並列処理

並列処理は、複数のマイクロプロセッサ上でプログラム タスクを分割して同時に実行することで、処理時間を短縮する手法です。並列処理は、2つ以上のプロセッサを搭載したコンピューター、またはコンピューター ネットワークを使用して実行できます。並列処理は並列コンピューティングとも呼ばれます。

コードの最適化

サーバー アクセラレーターをフルに活用するには、ソフトウェア コードを最適化する必要があります。

コンピューティング ユニファイド デバイス アーキテクチャ (CUDA) は、コンピューティング カーネルの実行のために、GPU仮想命令セットと並列コンピューティング要素への直接アクセスを可能にします。

FPGA構成は、通常、ハードウェア記述言語 (HDL) を使用して指定されます。HDLを使用すると、特定用途向け集積回路 (ASIC) を模倣し、特定のタスクまたはアプリケーションの要件に合わせて基礎を構成できます。

- **可視化** — アクセラレーターは、CADなどの3D可視化アプリケーションのパフォーマンスを向上させます。これにより、ソフトウェアはユーザーの操作に合わせてリアルタイムでモデルを描画できるようになります。
- **モデリングとシミュレーション** — アクセラレーターによってリアルタイムよりも高速なモデリング シミュレーションが可能になります。その結果、設計を早期に評価し、設計変更を迅速にテストして、設計の反復数を増加させることができます。
- **財務モデリング** — 高速化されたHPC/人工知能 (AI) ソリューションにより分析ツールに革新がもたらされることで、業界全体が大量のデータ ポイントを活用して、リスクとリターンをより的確に理解できるようになっています。
- **地底探査データの処理** — 石油/ガス企業は、大規模な地底探査データ ストアから情報を抽出するための効果的で新しい方法を見出し、アクセラレーターを活用して、結果を得るまでの期間の短縮とコスト削減を実現しています。
- **信号処理** — アクセラレーターによってプロバイダーは、コンピューター、ラジオ、ビデオ、携帯電話などからの大量の信号データ ストリームをリアルタイムでモデリングして分析できます。

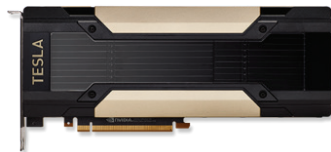
Dell EMC PowerEdgeサーバー用のGPUとFPGA

一部のDell EMC PowerEdgeタワーおよびラック サーバーで利用可能なパフォーマンス アクセラレーターを使用して、アプリケーションを高速化できます。次の表は、[PowerEdgeサーバー](#)に適合するアクセラレーター (GPUとFPGA) の数とモデルを示しています。Xilinx® Alveo™ U200およびインテル®プログラマブル アクセレーション カード (PAC) (Arria® 10 GX搭載) はFPGAで、それ以外はNVIDIA® GPUです。

2Sタワー	1U、2Sラック		2U、2Sラック		4U、4Sラック	
T640	R640	C4140	R740	R7425	R840	R940xa
4 アクセラレーター	1 アクセラレーター	4 アクセラレーター	3 アクセラレーター	3 アクセラレーター	2 アクセラレーター	4 アクセラレーター
			P4			
	T4		T4	T4		
P40		P40	P40	P40	P40	
M10			M10	M10	M10	
V100 16GB または32GB		V100 16GB または32GB PCIeまたは SXM2	V100 16GBまたは 32GB	V100 16GB または32GB	V100 16GBまたは 32GB	V100 16GBまたは 32GB
			U200	U200	U200	U200
	10GX		10GX		10GX	10GX

U = ラック ユニット、S = ソケットまたはCPU
表に記載されているアクセラレーターの数値は、ダブル幅のフル ハイト カードの場合の数です。T4は、ロープロファイルのハーフ ハイト カードです。

GPUは、CUDAコア数、メモリーの量、電力/冷却要件によって異なります。たとえば、NVIDIA T4 GPUは2,560個のCUDAコアと16GBのメモリーを搭載し、最大70ワットの電力を使用します。NVIDIA Tesla® V100は5,120個のCUDAコアと最大32GBのメモリーを搭載し、最大300ワットの電力を使用します。インテルとXilinx®FPGA仕様は、35~64MBの組み込みおよびオフチップメモリー容量を含み、電力使用量は10~225ワットです。



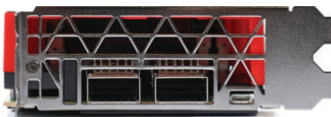
47倍向上

ディープラーニングの推論のスループットをCPUと比較⁴



1.5 TFLOP

(1秒あたりの浮動小数点数演算回数)と最大40 GFLOP/ワットの電力効率⁵



最大90倍

機械学習の推論、ビデオトランスコーディング、データベースの検索および分析のパフォーマンスをCPUと比較⁶

NVIDIA Tesla GPU

NVIDIA Tesla GPUには、可視化やその他のワークロードをこれまで以上に高速に実行するために必要となる処理能力があります。さらに、Teslaは、仮想デスクトップ インフラストラクチャ (VDI) 向けのハイ パフォーマンスとユーザー密度を実現します。

- [T4 Tensorコア GPU](#)
- [Tesla V100 GPUアクセラレーター](#)
- [NVIDIA NVLink™ ファブリック](#) インターコネクト
- [NVIDIA GPU CLOUD™](#) コンテナ
- [ソフトウェア アプリケーション カタログ](#)
- [NVIDIA開発者](#) 向けリソース

NVIDIA GPUは、[Dell EMC PowerEdgeサーバー](#) (R640、R740、R740xd、R7425、R840、R940xa、C4140を含む) と[Dell EMC Ready Solutions for HPC and AI](#)で利用できます。

インテルFPGA

インテルFPGAは、データ分析、画像推論、暗号化、圧縮などのワークロードと正確に一致するデータパスで動的に再プログラミングできます。

- [インテルArria 10 GX FPGA搭載のインテルPAC](#)
- [ソフトウェア アクセラレーション スタック](#)
- [インテルFPGAアクセラレーション ハブ](#)

[Dell EMC PowerEdgeサーバー](#) (R640、R740、R740xd、R840、R940xaを含む) と[HPCおよびAIソリューション](#)で利用できます。[構成ガイド](#)をダウンロードしてください。

Xilinx FPGA

Xilinx 16nm UltraScale™アーキテクチャを基盤として構築されたXilinx Alveoアクセラレーター カードは、高速化要件とアルゴリズム規格の変化に適應でき、ハードウェアを変更することなく任意のワークロードを高速化できるほか、総所有コストを削減します。

- [Alveo U200](#)
- [ソフトウェア エコシステム](#)
- [開発者ツール](#)

[Dell EMC PowerEdgeサーバー](#) (R740、R740xd、R7425、R840、R940xaを含む) と[HPCおよびAIソリューション](#)で利用できます。

高速化されたDell EMCソリューション

Dell EMCとパートナー各社が提供しているGPUおよびFPGA内蔵ソリューションで時間を節約できます。

Dell EMC Readyソリューション

[Dell EMC Ready Solutions for HPC](#)により、高度なコンピューティングを迅速かつシンプルに採用できます。Dell EMCは、柔軟で拡張性に優れたハイ パフォーマンス コンピューティング ソリューションの選択肢を、さまざまな業界のユースケースに対応するよう最適化された、サーバー、ネットワーキング、ストレージ、ソリューションおよびサービスとともに提供します。

[Dell EMC Ready Solutions for AI](#)には、AIイニシアティブを加速させるために必要なすべてのものが含まれています。AIをシンプルにするこれらの統合システムは、機械学習やディープ ラーニングに最適です。お客様組織は、顧客とそのビジネスに関するより深いインサイトをより迅速に獲得できるようになります。

⁴『NVIDIA Tesla V100 Tensor Core GPU』(2019年5月)

⁵『Intel Arria 10 FPGAs Features』(2019年5月)

⁶『Alveo U200 Data Center Accelerator Card』(2019年5月)



Dell EMC パートナーが提供するソリューション

[Amulet Hotkey®仮想デスクトップ ソリューション](#)は、エンタープライズ クラスのサーバーを仮想GPU アクセラレーターと組み合わせて、データ センター向けに最適化された高密度ソリューションを提供することによって、Windows® 10への移行をシンプルにします。さらに、仮想GPUは、Windows 10、Microsoft® Office 365®、YouTube®などの日常的に使うプログラムについて、高速なグラフィックス処理の仮想化ニーズの高まりに対応して、優れた仮想デスクトップ エクスペリエンスをもたらします。[Amulet Hotkeyのお客様の成功事例をご確認ください。](#)

[BittWare Stratix® 10ベースのFPGA Accelerated Compute Node](#)では、FPGAアクセラレーター向けに最適化された高密度ラックマウント サーバーを使用して、要求度の高いデータ センターワークロードを実行できます。1Uあたり最大4個のインテルStratix 10 FPGAを使用できます。システムは、BittWareまたは Dell EMCから直接購入できます。[ビデオを視聴する](#)

[Kinetica®](#)は、GPUによるデータベース高速化、ビジュアル ディスカバリー、機械学習、高速の並列コンピューティングの各機能を含んだインサイト エンジンです。Dell PowerEdgeサーバーとNVIDIA GPUでKineticaを実行すると、複雑で予測不能な大量データに付随する課題に対処できます。記事『[Explaining GPUs to Your CEO: The Power of Productization](#)』をご確認ください。

[Tracewell Systems®](#)は、企業、政府機関、OEMにパワフルな既製のコンピューティング テクノロジーを提供しています。環境要因によりコンピューティングに特殊な課題が生じる場所（空中、海上、地表など）や、据え置きとモバイルの導入環境、さらには専門的なハードウェアやソフトウェアとの統合が必要となる状況にも対応します。[データシート、ビデオ、その他のリソースはこちらから入手できます。](#)

Dell EMC高速化ソフトウェア パートナー

[NVIDIA GRID™](#)仮想PCとGRID®仮想アプリケーションは、NVIDIA® Tesla® GPUでの実証済みパフォーマンスによって、仮想デスクトップと仮想アプリケーションを向上させます。

[Kinetica](#)ソフトウェアは、GPUを並列コンピューティングに使用して、従来のオンライン分析処理（OLAP）ワークロードを飛躍的に高速化します。

[BitFusion](#)ソフトウェアは、GPU、FPGA、ASICを分離して、データ センター内の任意の場所に動的に接続します。

[SQream Technologies®](#) GPU高速化データ ウェアハウスは、テラバイト規模からペタバイト規模に拡張して、任意の規模およびワークロードに適応できます。

[FASTDATA.io](#) PlasmaENGINE® GPUネイティブ ソフトウェアでは、複数のGPUを使用して、移動中の無限のデータを複数のノードにわたってリアルタイムで処理できます。

[RAPIDS](#)は、エンド ツー エンドのデータ サイエンス トレーニング パイプラインをNVIDIA GPUで実行するためにNVIDIA CUDA-X上に構築されたデータ サイエンス ライブラリのスイートです。

Dell Technologiesのパートナーになる

[Dell Technologiesパートナー プログラム](#)に参加すると、デジタル、IT、人材、セキュリティのトランスフォーメーションを世界中の組織で日々実現するパートナー エコシステムに加わることになります。Dell Technologiesパートナー プログラムは、エッジからコア、クラウドに至るまで、業界で最も堅牢なポートフォリオを支えています。設計理念はSimple. Predictable. Profitable.（シンプル、予測可能、高い収益性）です。



リソース

PowerEdgeサーバーでデータセンターをあらゆるワークロードに対応させる

PowerEdgeタワーサーバーは、組織の拡大に応じてお客様のペースで拡張できるように設計されています。PowerEdgeラックサーバーは、拡張性の高いアーキテクチャに加え、コンピューティングとメモリーの最適なバランスも兼ね備えているため、幅広いアプリケーションにわたってパフォーマンスを最大限に高めることができます。

パフォーマンスの結果を確認する

エンジニアリング部門のワークロード別ベンチマーキングデータ、リファレンスアーキテクチャ、ブログを hpcatdell.com で確認し、[サーバーソリューションリソースのサポート](#)をご利用いただけます。次のホワイトペーパーをダウンロードできます。

- 『[Deep Learning Performance Comparison — PowerEdge C4140 Scale-up vs. Scale-out](#)』
- 『[GPU Database Acceleration on PowerEdge R940xa](#)』
- 『[Deep Learning Inference on PowerEdge R7425](#)』
- 『[Deep Learning Inferencing with Intel Programmable Acceleration Card on Dell EMC Infrastructure](#)』

エデュケーション サービスにアクセスする

必要なスキル、トレーニング、認定をeducation.emc.comで取得できます。[ディープラーニングで問題を解決する方法](#)をDell EMCのディープラーニング研究所で習得できます。

HPCおよびAIイノベーションラボにアクセス

世界最高クラスのソリューション設計、テクノロジーテスト、コラボレーションがdell EMC.com/ja-jp/innovationlabで実現します。

Center of Excellenceにつながる

イノベーションと専門技術のためのグローバルハブ：dell EMC.com/ja-jp/coe

お問い合わせ

詳細については、dell EMC.com/ja-jp/acceleratorsまたはdell EMC.com/ja-jp/serversにアクセスするか、Dell EMCセールス担当者または認定リセラーまで[お問い合わせ](#)ください。

リソースセンター

[サーバーソリューションリソース](#)
をダウンロード

技術文書

HPCエンジニアリング部門からのパフォーマンス結果、リファレンスアーキテクチャ、ブログを見る (hpcatdell.com)

仮想ラック

仮想ラック内のサーバーとソリューションを確認する (esgvr.dell.com)

Dell EMC HPCコミュニティに参加する

アイデアの交換を促進する世界的なテクニカルフォーラム
dellhpc.org

© 2019 Dell Inc. その関連会社。All rights reserved. (不許複製・禁無断転載) Dell, EMC, およびDellまたはEMCが提供する製品及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。その他の商標は、各社の商標又は登録商標です。参照番号：05/19 DELL-EMC-BRO-Accelerators-USLET-101

Intel®は、アメリカ合衆国およびその他の国におけるIntel Corporationの商標です。NVIDIA®、Tesla®、およびGRID™は、NVIDIA Corporationの商標です。Microsoft®、Windows®、およびOffice 365®は、米国およびその他の国におけるMicrosoft Corporationの登録商標または商標です。YouTube®は、Google Inc.の登録商標です。Arria®およびStratix®は、Altera Corporationの商標であり、米国特許商標庁およびその他の国で登録されています。Xilinx®、UltraScale™、およびAlveo™は、Xilinx, Inc.の商標です。SQream Technologies®は、SQream Technologies Ltd.の登録商標です。Kinetica®は、米国およびその他の国におけるKineticaとその子会社の商標です。FASTDATA.io®は、米国およびその他の国におけるFASTDATA.ioの登録商標です。Amulet Hotkey®は、Amulet Hotkey Ltd.の商標であり、英国、米国、およびその他の国で登録されています。Tracewell Systemsは、米国およびその他の国におけるTracewell Systemsの登録商標です。

掲載される情報は、発信現在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。