



Dell AIポートフォリオで AIワークロードの課題に対応

DellのAIポートフォリオとHPEの類似 製品の比較

人工知能 (AI) がビジネス環境を激変させたことに疑いの余地はありません。あらゆる規模の業界と企業が、保有するデータから深いインサイトを獲得し、ビジネスプロセスを自動化し、カスタマイズされたエクスペリエンスを顧客やユーザーに提供しており、これによって業界内での競争優位性を高めています。企業がAIの力を効果的に活用するには、包括的な統合ソリューションポートフォリオでAIライフサイクル全体を網羅するインフラストラクチャプロバイダーが必要です。

増大するAIニーズに対処し、AIニーズ固有の複雑さを乗り越えるため、デル・テクノロジーズやHewlett Packard Enterprise (HPE) といったインフラストラクチャベンダーが顧客をサポートしています。AI対応ポートフォリオが充実しているこれらのベンダーは、オンプレミスとクラウドの高性能インフラストラクチャソリューションに、戦略的パートナーシップ、サポートやコンサルティングの各種サービスを組み合わせた、様々なレベルのAIソリューションを提案しています。

このレポートはDellとHPEのAIポートフォリオ*に関する一般公開情報を調査したものです。レポートの目的は、AIニーズへの対応にデル・テクノロジーズを選択した場合に企業が得られるアーキテクチャ、パフォーマンス、サポート上のメリットを明らかにすることです。DellがAI導入向けに構築した各種サーバーの詳細情報を比較し、ML Commons®の業界ベンチマークテスト結果を参照しました。また、AI導入のあらゆる段階で顧客をサポートする追加のソフトウェアやサービスについても調査しました。

*注：PTによる調査はすべて2023年12月5日以前に完了しているため、この資料には、その日付以降にリリースされたDellまたはHPEの製品や変更は反映されていません。

AI 導入の課題

AI 戦略を採用することで、データセンターおよびその人員配置を行う IT 担当者は、数々の新しい課題に直面することになります。

- 社内 AI トレーニングまたは外部雇用を通じて、現行スタッフのスキルギャップを解消する
- 企業のデータの質、量、場所、現状など、AI のデータ準備ニーズを理解する
- AI について特定のビジネス目標を評価し、どの AI モデルと実装がメリットをもたらすかを適切に判断する
- 計画している AI システムについて、コンピューティング、ネットワーキング、ストレージのニーズを評価し、獲得計画を策定する

データセンターで AI 導入のメリットを享受するには、数々の高いハードルを乗り越える必要があり、ここに挙げた課題はその一例にすぎません。

Dell の AI ポートフォリオはこのような顧客の課題を解決することを目指しており、プロフェッショナル サービスとコンサルティング サービスの提供を通して、実装ロードマップの作成と AI モデル用のデータ準備を支援しています。¹ また、機械学習 (ML) の概念やその他の教育トピックを網羅するトレーニング コースや、AI 実装を確実に成功に導くための検証済み設計の提供も、ポートフォリオに含まれています。² さらに、Dell はサードパーティーとの提携を通じ、さらなる AI ツールを顧客に提供しています。Hugging Face コミュニティ内のカスタム Dell ポータルはその一例です。専用のコンテナとスクリプトを備えており、オープンソース AI モデルの導入³ や、Meta Llama 2 大規模言語モデル (LLM) の容易な導入を可能にします。⁴ モバイル ワークステーションから最大 8 基のハイエンド NVIDIA GPU をサポートするサーバーまで、Dell はコンピューティングとパソコンの選択肢を幅広く用意しています。また、ハイパフォーマンスのファイルおよびオブジェクトストレージレイのポートフォリオでは、AI で要求される非構造化データストレージを提供しています。これらのストレージ製品 (Dell PowerScale、ObjectScale、ECS、オンボード ストレージなど) は、AI ワークロードで頻繁に使用される非構造化データに対応します。⁵ また、Dell は Snowflake と提携しており、ハイブリッドクラウド ストレージ ソリューションを顧客に提供しています。⁶ Dell の分析によると、2023 年 8 月時点で同社は「最も広範な生成 AI ポートフォリオ」を提供しており、サーバーやストレージだけに留まらず、AI の実装過程に必要なリソースを提供しています。⁷

AI パフォーマンスと高速コンピューティングのオプション : Dell と HPE の比較

ワークロードのサイズやタイプに応じて、AI ワークロードの計算リソースには CPU、GPU またはその両方を使用できます。一部の CPU は AI に特化したアクセラレーターを提供しており、最新のインテル Xeon スケーラブル プロセッサに搭載されているインテル アドバンスド・マトリクス・エクステンション (インテル AMX) がその一例です。⁸ GPU は大規模なワークロードや要求が厳しいワークロードに向いていますが、GPU のフォーム ファクターがパフォーマンス レベルに影響することがあります。例えば、NVIDIA A100 および H100 モデルの GPU にはユニバーサル PCIe 版と SXM 版がありますが、SXM 版には、より高性能の NVIDIA SXM アーキテクチャが採用されています。⁹ また、冷却アーキテクチャや電力効率といったサーバー設計や大容量メモリーもパフォーマンスに影響します。多くのデータセンターでは現在も空冷が採用されています。これは、ハイパフォーマンス コンピューティング (HPC) ワークロードを処理するため、できるだけ空冷効率に優れたサーバーが必要になることを意味します。以下では、コンポーネントや冷却オプションなど、様々な特徴の観点から PowerEdge サーバー製品を説明します。また、公開されている MLCommons® MLPerf® スコアも紹介します。

AIモデル ベンチマーク性能 : MLPerf結果の比較

MLPerf®は、AIのトレーニングと推論性能を評価するベンチマークスイートです。組織がMLPerf®の公式結果を公開する場合、その結果はベンチマーク開発者MLCommons®によって設定された所定の条件に準拠している必要があります。¹⁰ これらのコンプライアンスガイドラインは、性能比較を容易にする基準を提供するものです。MLPerf®では、Datacenter、Edge、Mobile、Tinyデータセットを使用して推論テストを実施し、AIスコアとテスト中の電力消費ワット数を報告しています。推論ベンチマークスイートには、一般的なAI、ML、DLモデルのテストが多数含まれています（表1を参照）。

表1 : MLPerf®テストに含まれるAI、ML、DLモデルと、各モデルの一般的なユースケース。出典 : Principled Technologies

一般的な AI モデル	標準的な用途
ResNet	画像分類モデル。コンピューターに様々な画像を学習、記憶、識別させる場合に役立つ。ユースケースとしては、医療画像、ソーシャルメディアコンテンツのモデレーション、顔認識など
RetinaNet	オブジェクト検出の一種で、ResNet よりもさらに複雑なものに対応できる。画像やビデオフレームに含まれているオブジェクトの識別と検出に役立ち、それらのオブジェクトを重要度によって分類できる。自動運転、車両自動アシスト技術、監視、顔認識などに使用される
3D-UNet	医療画像のセグメント化に特化
RNN-T	音声認識。言語の自動翻訳などのユースケースで使用される
BERT	自然言語処理。テキストの要約、言語翻訳、タスクの自動完了などのユースケースで使用される
DLRM-v2-99.9	推奨モデル。ターゲットを絞った広告やパーソナライズされた製品推奨などのユースケースで使用される
GPTJ-99と99.9	自然言語処理用 LLM。テキスト生成に優れており、チャットボットやチャットベースの AI ツールといったユースケースで使用される

MLPerf

MLPerf®の結果には、AIモデル自体のパラメーターとは別にいくつかの追加パラメーターが含まれています。これらを利用して、大量のデータを1つのグラフまたは表内で解析できます。パラメーターについて簡単に説明します。

- 99.0 および 99.9 : モデルのトレーニング精度。より正確な出力を必要とする場合、モデルはより複雑になり、データの処理時間も長くなる
- オフライン サンプル / 秒 : テスト開始時にすべてのクエリーを送信するモード。システムにデータがすでに存在している状態をシミュレートする
- サーバー クエリー / 秒 : テスト中にクエリーを送信し続けるモード。データのライブストリームを分析している状態をシミュレートする

MLCommons® および MLPerf® の結果の詳細については、<https://mlcommons.org/benchmarks/inference-datacenter/>を参照してください。

このレポートに記載されている結果は、2023年11月から MLCommons® の Web サイトで公開されている MLPerf®v3.1 Inference Datacenter の結果です。¹¹ これらの結果にはテクノロジー メーカーやクラウド サービス プロバイダーから提出された結果が含まれており、幅広い構成を網羅しています。一般公開されている HPE の提出結果と比較して、Dell のサーバーは特定の AI モデルで高い結果を出しました（注：サーバー間で GPU 構成が異なるため、サーバーを直接比較することは困難です）。詳細については、表 2 を参照してください。

表 2：2023年11月29日時点で公開されている MLCommons® MLPerf® 3.1 結果に含まれている Dell と HPE のサーバー。
出典：Principled Technologies

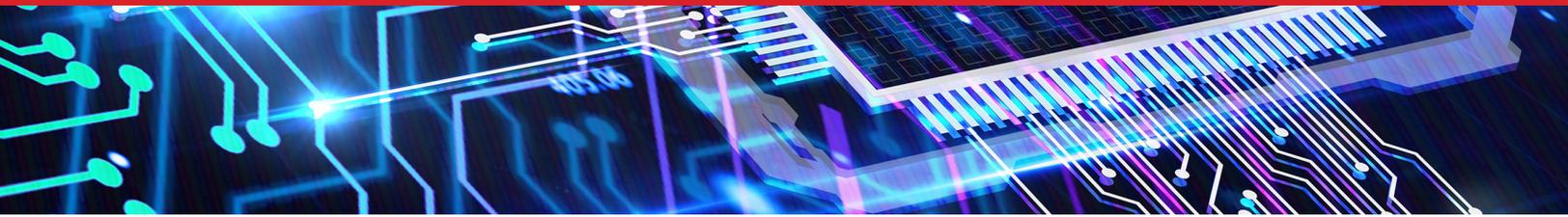
送信者	サーバーのモデル	GPU の数とモデル	説明
Dell ¹²	PowerEdge XE9680	8x NVIDIA H100 SXM	大規模ワークロード（大規模言語モデルなど）を使用した AI トレーニングおよび推論向け
	PowerEdge XE9640	4x NVIDIA H100 SXM	高密度の水冷却データセンターにおける大規模 AI モデルのトレーニング向け
	PowerEdge XE8640	4x NVIDIA H100 SXM	空冷データセンターの 4U フォーム ファクターで従来型の AI トレーニング、HPC、データ分析アプリケーションを駆動するユースケース向け
	PowerEdge R760xa	4 x NVIDIA H100 PCIe	様々なハイコンピューティング ワークロード向け。最上位性能の GPU を使用するほどではない AI-ML/DL のトレーニングや推論用
HPE ^{13,14}	ProLiant XL675d Gen10 Plus	8x NVIDIA A100 SXM	ハイパフォーマンス コンピューティングおよび AI 向け
	ProLiant DL380a Gen11	4 x NVIDIA H100 PCIe	中程度 AI ワークロード向けの 2U サーバー

Dell と HPE サーバーの直接的な比較

包括的な AI 戦略にはハードウェア以外の要素も多数含まれますが、AI ワークロードの成功を導く重要な要素の 1 つは高いハードウェア性能を確保することです。新しい GPU やその他のテクノロジーが利用可能になるにつれ、AI ワークロード機能も進化します。MLPerf®v3.1 の結果が最初に発表された時点で、入手可能な最高の NVIDIA GPU は H100 Tensor コアでした。Dell はこの GPU を使用した数種類のサーバー（PCIe と SXM5 の両フォーム ファクター）について、MLPerf® 結果を公開しています。¹⁵ 公開された HPE の結果には H100 の結果 1 件しか含まれておらず、フォーム ファクターも PCIe のみでした。当社の調査結果から、入手可能な GPU 対応 HPE サーバーには NVIDIA GPU の最高パフォーマンスを引き出す SXM5 H100 フォーム ファクターに対応しているものは少なく、HPE ProLiant サーバーでは対応しているものが 1 つもないことがわかりました。¹⁶ 以下に示すように、通常 GPU の性能が高くなると、AI ワークロードのパフォーマンスも向上します。

8GPU の MLPerf 結果

Dell PowerEdge XE9680 は、AI アクセラレーション用の NVIDIA H100 SXM5 GPU を最大 8 基、第 4 世代 Intel® Xeon® スケーラブル・プロセッサ・ファミリーを最大 2 基サポートします。PowerEdge XE 製品ファミリーはモジュラー型アーキテクチャを採用しており、SXM4 または SXM5 NVIDIA GPU または AMD の Open Compute Project Accelerator Module (OAM) GPU アセンブリをサポートしています。そのため標準の PCIe GPU よりもパフォーマンスを向上させることができます。¹⁷ ラックスペースがわずかに 6U の PowerEdge XE9680 は、コンパクトな 8 ウェイ NVIDIA H100 SXM5 サーバーです。最新の HPE ProLiant Gen11 サーバーは現時点で H100 SXM フォーム ファクターをサポートしておらず、¹⁹ 一部の HPE Cray Supercomputing サーバーのみがこれをサポートしています。¹⁸ HPE は Cray サーバーに関する MLPerf® 結果を提出しておらず、同社の AI ポートフォリオ ページでは ProLiant サーバーを大々的にアピールしているため、この資料では ProLiant に焦点を当てて説明します。（図 1 を参照）。



Featured AI products and services

PRODUCT

HPE Ezmeral Unified Analytics Software

Unlock data and insights faster by helping you develop and deploy data and analytic workloads. Provides fully managed, secure, enterprise-grade versions of the most popular open-source frameworks with a consistent SaaS experience.

[Explore more →](#)

PRODUCT

HPE Machine Learning Development Environment

Uncover hidden insights from your data by helping engineers and data scientists collaborate, build more accurate ML models and train them faster.

[Explore more →](#)

PRODUCT

HPE Machine Learning Data Management Software

Uncover hidden insights with a data pipelining and versioning solution that automates data pipelines and accelerates time to ML model production by processing petabyte-scale workloads.

[Explore more →](#)

PRODUCT

HPE ProLiant Servers

Speed time to value with systems that are optimized for computer vision inference, generative visual AI, and end-to-end natural language processing.

[Explore more →](#)

図 1 : 「Featured AI products and services」 ページ (<https://www.hpe.com/us/en/solutions/ai-artificial-intelligence.html>) のスクリーンショット。HPE ProLiant サーバーを大きく紹介している (2023 年 12 月 5 日時点)

2023 年 11 月に最初に発表された 8GPU サーバーの MLPerf[®]v3.1 結果では、NVIDIA SXM5 H100 GPU 搭載 Dell PowerEdge XE9680 のパフォーマンスが、NVIDIA SXM4 A100 GPU 搭載 HPE ProLiant XL675d Gen10 Plus のパフォーマンスを最大 4.25 倍上回りました (図 2 参照)。

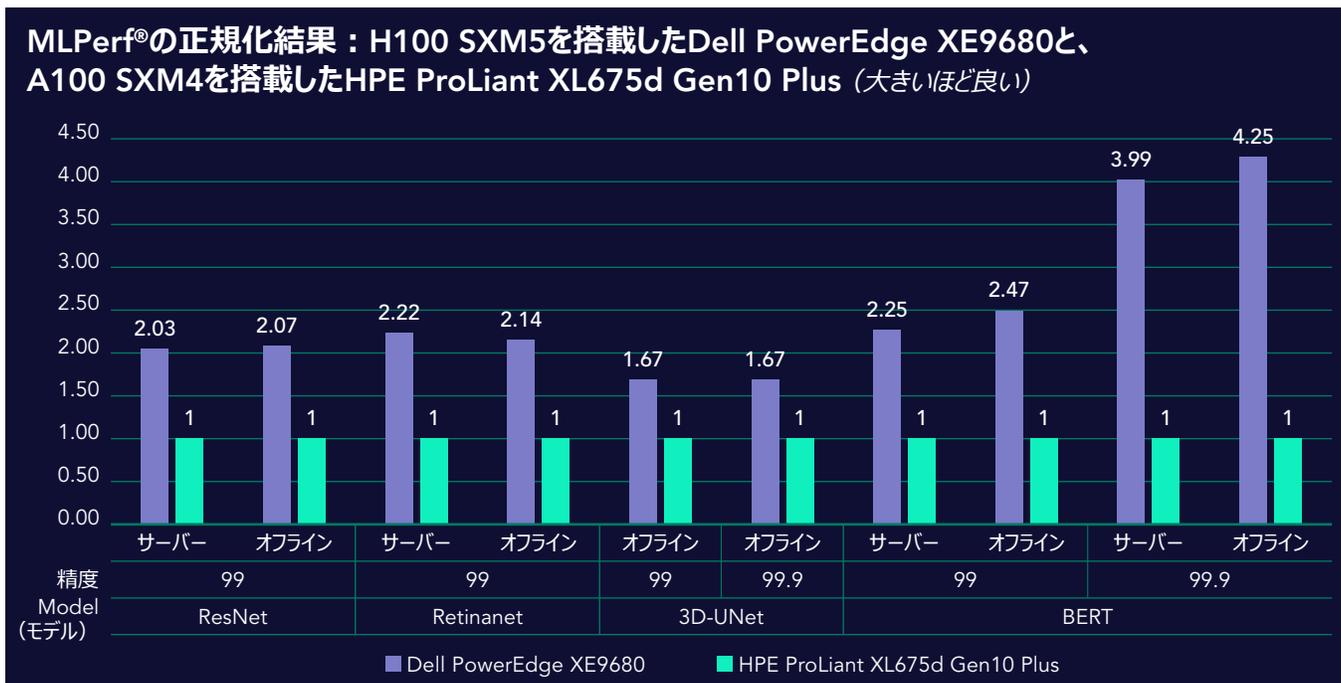


図 2 : 公開された Dell PowerEdge XE9680 と HPE ProLiant XL675d Gen10 Plus の MLPerf[®] 結果 (2023 年 11 月 29 日時点)。Dell システムは NVIDIA H100 GPU を使用しているが、HPE システムの GPU は 1 世代古い。出典 : Principled Technologies。MLCommons[®] からのデータを使用。^{20,21}

比較を容易にするために、図 2 から図 5 までのテスト結果を正規化しました。つまり、HPE ProLiant DL380a Gen 11 の各結果を 1 とし、対応する Dell PowerEdge R760xa の相対的な結果を表示しました。この結果から、GPU モデルが 1 世代違うだけで、多数の AI ワークロードにわたって大きなパフォーマンスの差が生じることを推測できます。

4GPU の MLPerf 結果

データセンターの省電力化と省スペース化が主な懸念事項である場合、2U の Dell PowerEdge XE9640 が解決策になる可能性があります。最大 4 基の NVIDIA H100 SXM GPU を搭載する PowerEdge XE9640 は、XE9680 の半分の GPU 計算能力を提供し、必要なスペースは XE9680 の 3 分の 2 です。²² Dell Smart Cooling テクノロジーを内蔵する高密度設計の Dell PowerEdge XE9640 は、CPU や GPU に対する直接水冷など、一連のサーマル テクノロジーを備えています。²³

PowerEdge XE9640 の 2U シャーシは、大型ファンやヒートシンクなど向上したエアフロー メカニズムに対応しており、PCIe カードやメモリーといったその他の重要コンポーネントの冷却に役立ちます。²⁴ 現在、4 ウェイ HGX H100 GPU を搭載する 2U 製品は、Dell でも HPE でも PowerEdge XE9640 以外にありません。HPE AI ポートフォリオには、1U と 2U の Gen11 ProLiant サーバーがありますが、PCIe フォーム ファクターの GPU に限定されます。²⁵

Dell PowerEdge XE9640 サーバーは Intel Max GPU シリーズ 1550 OAM GPU もサポートしています。これは低消費電力の高密度 GPU オプションで、PCIe カードと OpenCompute Accelerator Module (OAM) が含まれています。²⁶ 2023 年 12 月 5 日時点で、このような GPU を搭載したサーバーを HPE が提供しているかどうかは確認できませんでしたが、HPE は Intel データセンター GPU Max 1100 GPU を搭載する HPE ProLiant DL380 Gen11 と DL380a Gen11 を提供しています。²⁷ つまり、Dell PowerEdge XE9640 は、4 基の Intel Max 1550 OAM GPU を搭載した 2U サーバーとしては、唯一の製品であると考えられます。データセンターのスペースと電力効率を懸念している企業にとって、4 基の Intel Max 1550 GPU を搭載する 2U サーバーは最適なソリューションであり、ハイパフォーマンスのコンピューティングと高いエネルギー効率をデータセンターのスペースを犠牲にすることなく実現できます。

MLPerf[®] 3.1 の公開結果では、4 基の HGX H100 GPU を搭載する Dell PowerEdge XE9640 が、4 基の PCIe H100 GPU を搭載する HPE ProLiant DL380a のパフォーマンスを最大 1.99 倍上回りました (図 3 参照)。

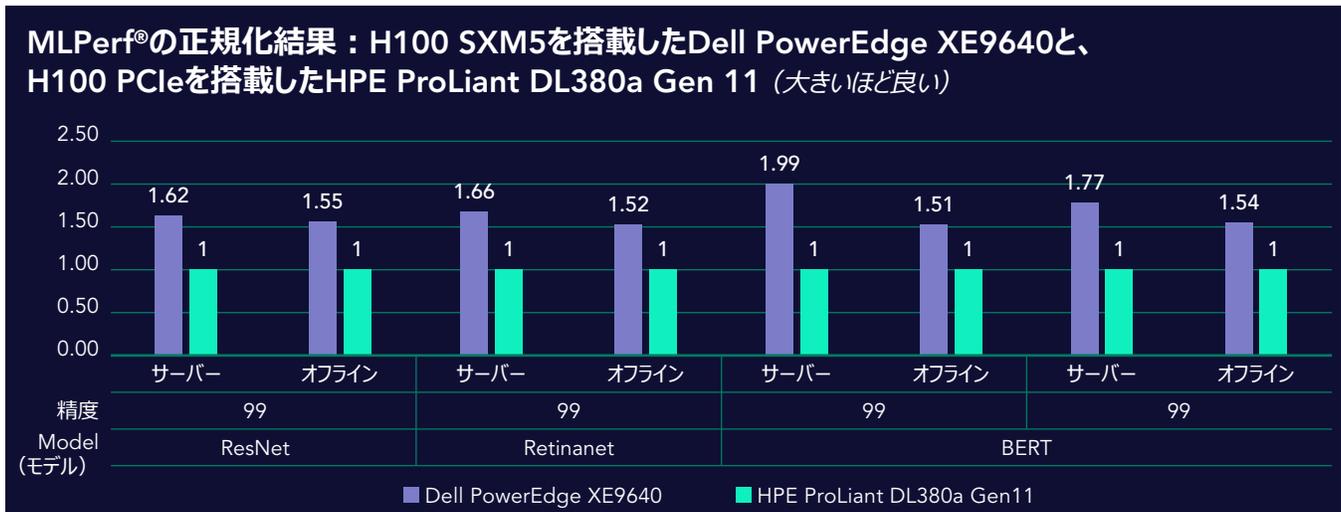


図 3: 公開された Dell PowerEdge XE9680 と HPE ProLiant XL675d Gen10 Plus の MLPerf[®] 結果 (2023 年 11 月 29 日時点)。Dell システムは NVIDIA H100 GPU を使用しているが、HPE システムの GPU は 1 世代古い。出典: Principled Technologies。MLCommons[®] からのデータを使用 ^{28,29}

PowerEdge XE8640 は 4 ウェイ GPU 構成で、プロセッサ用の空冷のほか、GPU 用に液体アシスト空冷ラジエーターを備えており、施設からラックへの水供給を必要としません。外部冷却剤を使用しない（できない）場合は、³⁰4 基の NVIDIA H100 SXM5 GPU をサポートする 4U の Dell PowerEdge XE8640 であれば、PowerEdge XE9640 と同等のコンピューティング能力を利用でき、直接水冷の必要もありません。³¹

Dell PowerEdge XE8640 の特長は、最新の第 4 世代インテル Xeon スケーラブル プロセッサと最大 4TB のメモリーを搭載している点です。³²AI やデータ分析によく見られる大規模データセットと複雑な計算に対応できます。ここでも、HPE Cray システムは NVIDIA H100 SXM5 GPU を搭載していますが、HPE ProLiant GPU 対応サーバーはこれをサポートしていません。

2023 年 11 月時点で公開されている MLPerf[®] のデータと比較すると、4 基の NVIDIA H100 SXM5 GPU を搭載した PowerEdge XE8640 サーバーは、9 つのカテゴリーで提出された 4GPU の全結果の中で最高の AI スループットを達成しました。図 4 に示すように、HPE ProLiant DL380a サーバーと比較して、最大 2.07 倍のスコアを記録しています。

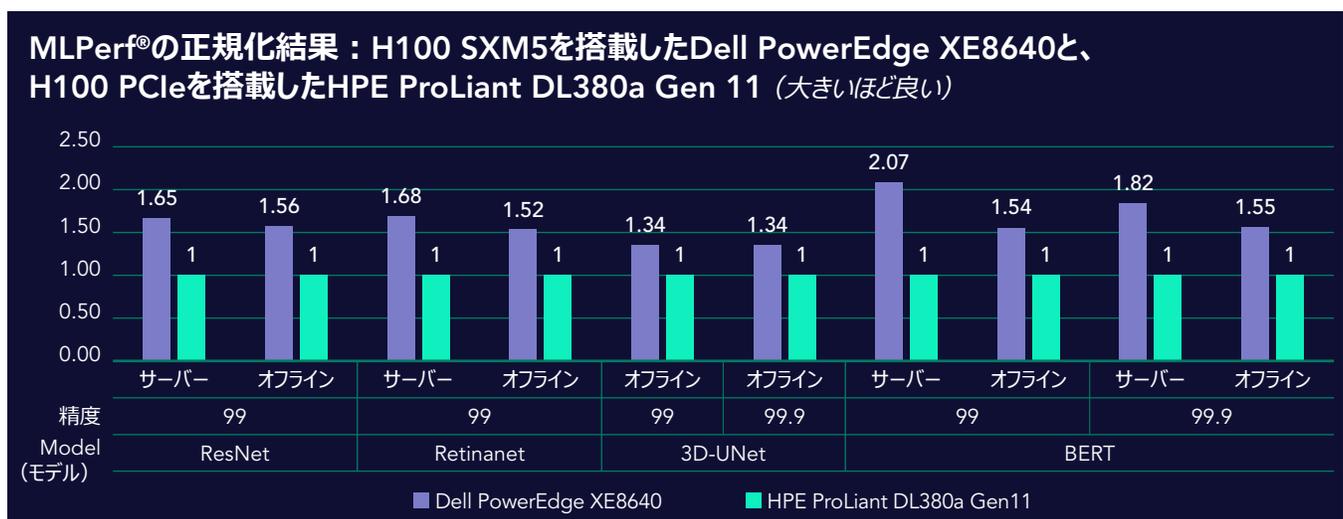


図 4：公開された Dell PowerEdge XE8640 と HPE ProLiant DL380a Gen11 の MLPerf[®] 結果（2023 年 11 月 29 日時点）。Dell システムは NVIDIA H100 SXM フォーム ファクターを使用し、HPE システムはパフォーマンスがやや劣る PCIe フォーム ファクターを使用している。出典：Principled Technologies、MLCommons[®] からのデータを使用。^{33,34}

最後にご紹介するのは、小規模に始めて必要なだけ拡大していきたい企業に適している、2U の Dell PowerEdge R760xa サーバーです。NVIDIA、AMD、インテルの様々な GPU に対応し、最大 4 基のダブル幅 PCIe Gen 5 GPU または 12 基のシングル幅 PCIe GPU をサポートしています。³⁵ 32 個の DIMM スロット、2.5 インチディスク用の 8 ドライブ ベイ、12 個の PCIe スロットを備えており、AI データ要件の増大に応じて拡張できる拡張性の高いストレージを提供し、最大 12 基のシングル幅 PCIe GPU または 4 基のダブル幅 PCIe GPU (NVIDIA H100 や L40S など) をサポートしています。³⁶ この拡張性は、機械学習モデルのトレーニングから高度なデータ処理まで、進化する AI タスクにサーバーが適応できることを意味します。

PowerEdge R760xa の空冷システムは、高密度コンピューティング環境をサポートし、最大 350W の高熱設計電力 (TDP) アクセラレーターにも対応します。³⁷ この機能により、コンピューティング負荷が高い状況でもパフォーマンスを維持できます。2023 年 11 月時点で公開されている MLPerf[®] の ResNet、RetinaNet、BERT の「サーバー」モードを使用したサーバー テスト結果では、4 基の NVIDIA H100 PCIe GPU を搭載した PowerEdge R760xa が、4 基の H100 PCIe GPU を搭載した HPE ProLiant DL380a Gen 11 のパフォーマンスを上回っています（図 5 を参照）。

MLPerf®の正規化結果：H100 PCIeを搭載したDell PowerEdge R760xaと、H100 PCIeを搭載したHPE ProLiant DL380a Gen 11（大きいほど良い）

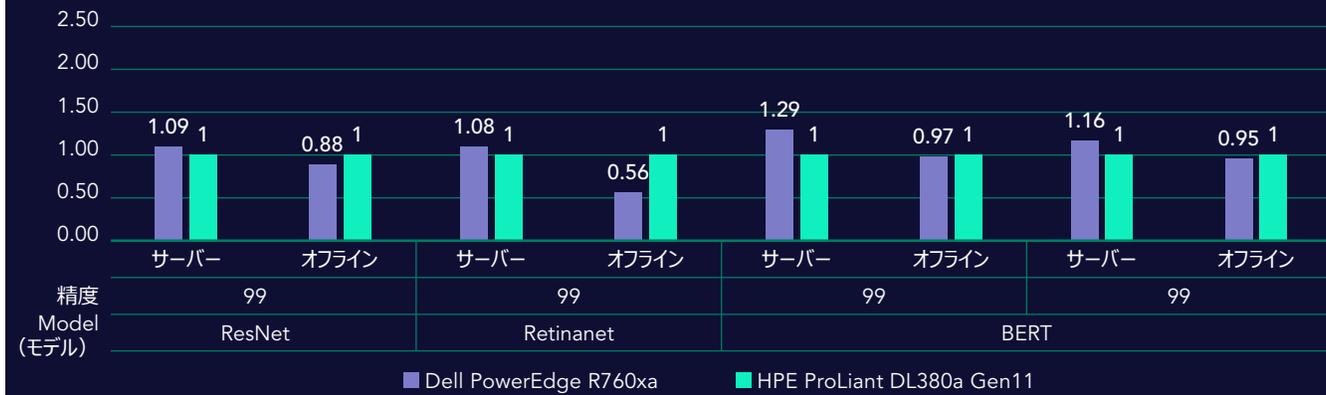


図 5：公開された Dell PowerEdge R760xa と HPE ProLiant DL380a Gen11 の MLPerf® 結果 (2023 年 11 月 29 日時点)。どちらのシステムも、NVIDIA H100 GPU の PCIe フォームファクターを使用している。出典：Principled Technologies。MLCommons® からのデータを使用 ^{38,39}

MLPerf® の結果を総合的に見ると、パフォーマンスはサーバーとコンポーネントによって大きく異なるため、ワークロードとそのパフォーマンス要求に対応できる適切なオプションを選択することが重要であることがわかります。AI ワークロード向けの Dell PowerEdge サーバーは冷却と密度のオプションを複数提供しているため、あらゆるデータセンターに対応でき、高い MLPerf® パフォーマンスを提供します。

Dell AI ポートフォリオの詳細

AI ワークロードの計画において、唯一の重要な考慮事項はコンピューティング性能です。しかし、AI の導入に着手する際は、ベンダーが提供するその他の AI ポートフォリオについても検討する必要があります。以下では、クライアントワークステーション、クラウドネイティブ製品、ストレージなど、AI ポートフォリオにとって重要なその他のカテゴリについて説明します。また、Dell の製品が HPE 製品よりも優れている領域についても注目していきます。

ワークステーション

AI 開発者やデータサイエンティスト向けの Dell Precision Data Science ワークステーションは、NVIDIA RTX™ GPU とインテル Xeon® CPU に加えて一連のデータサイエンス ツールを提供します。⁴⁰ これらのシステムはプロフェッショナルグレードのコンピューティング オプションを利用しており、100 以上のプロフェッショナル アプリケーションで認定されている NVIDIA GPU と、⁴¹ インテル Xeon スケーラブル プロセッサ アクセラレーター（インテル DL ブーストなど）を搭載しています。⁴² Dell Precision ワークステーションにはモバイル型、タワー型、ラック型があり、定置でのデータ分析や出先での科学分野モデリングまで、幅広いニーズに対応します。

HPE ワークステーションの製品ラインアップは Dell よりも狭く、主要製品は NVIDIA L4s を搭載するタワー型のシングルワークステーションです。モバイル型のワークステーション オプションはありません。⁴³ HPE のタワー型ワークステーション製品は様々なタスクに十分対応しますが、Dell の広範な製品ラインアップと比べると、柔軟性やワークロード対応範囲の点で劣ります。サイズと携帯性に関する選択肢が多彩な Dell Precision ワークステーションは、ラボ、オフィス、現場での作業など、様々な環境のニーズに対応し、カスタマイズ度の高いソリューションを提供します。

ストレージ

AI ワークロードの実行において、ストレージはコンピューティングと同じ位重要な要素です。データが増えるほど AI モデルの精度は向上しますが、大量のデータセットの保存と管理は、多くのデータセンターの能力にとって課題となりえます。また、モデルは一般的に非構造化データを使用してトレーニングされるため、AI 対応ストレージ システムは様々な種類のデータを簡単に処理できる必要があります。⁴⁴ AI、ML、DL のデータセットに対応する容量と拡張性を提供するため、Dell は PowerScale™ シリーズをファイル ストレージおよび Elastic Cloud Storage (ECS) 用として、またはオブジェクトストレージ用のソフトウェア定義 ObjectScale 用として提供しています。

PowerScale オールフラッシュ NAS ポートフォリオには、ノードあたり 3.84TB から最大 720TB の raw 容量をサポートする容量オプションがあり、クラスター化されたオールフラッシュ容量は最大 186PB に到達します。PowerScale の柔軟性と拡張性は、様々な顧客および AI のユース ケースに対応できます。⁴⁵ クラスター化された PowerScale F900 では、合計のローストレージの容量は最大 186PB に到達します。⁴⁶ 3 種類のオールフラッシュ PowerScale モデル (F200、F600、F900) すべてに、インライン データ圧縮と重複排除機能が搭載されており、ストレージ効率の向上に寄与しています。⁴⁷ 各 PowerScale ストレージ モデルでは Dell OneFS™ ファイル システムが使用されています。このファイル システムでは、ストレージを階層化するポリシーが採用されており、最速層にある最重要データを優先することでワークロードを最適化します。⁴⁸ Dell は OneFS ソフトウェアを Dell APEX File Storage for AWS として、AWS 市場でも提供しています。顧客は OneFS を自社の AWS コンピューティング インスタンスで使用できます。そのユーザー エクスペリエンスは、オンプレミスの OneFS アレイの同じ機能と一貫性があります。⁴⁹ HPE はパブリッククラウド統合を通じてハイブリッド ストレージ ソリューションを提供していますが、同社の製品で Dell APEX File Storage for AWS のようなクラウドネイティブのオプションは確認できませんでした。

Dell のオブジェクトストレージ オプションに含まれている Dell Enterprise オブジェクトストレージ (ECS) は、非構造化データをパブリッククラウド規模で保存するための特定用途向けのストレージです。⁵⁰ Amazon S3 オブジェクトストレージとの組み込みの互換性でハイブリッド クラウド機能を提供するほか、ECS ストレージ ノードはラックあたり最大 14PB の容量を提供します。⁵¹ HPE にもファイルおよびオブジェクトストレージのオプションはありますが、オブジェクトストレージ製品は Scality との提携を通じて提供されます。顧客は Scality の HPE ソリューションを HPE から購入できます。⁵²

プロフェッショナル サービス

Dell は多彩なプロフェッショナル サービスを幅広く提供しており、コンサルティング、データ準備、導入、サポート、マネージド サービスといったサービスで AI 導入を支援しています。検証済みのアーキテクチャやソリューションを求めている組織向けに、Dell は検証済みの AI 向け設計を提供しています。これは特定のユース ケースを対象にしたもので、AI リソースの設計と導入における推測的な作業を排除します。Dell の検証済み AI ソリューションには、ハードウェアとソフトウェアのバンドル、会話型 AI モデル、機械学習の運用などが含まれています。Dell は、事前構成済みの特定用途ソリューションと AI 関連サービスを組み合わせた包括的な AI ソリューションを通じて様々な AI ニーズに対応しています。これらのソリューションを利用することで、アドホック ソリューションを構築するよりも短期間かつ容易に AI の成功を実現できる可能性があります。

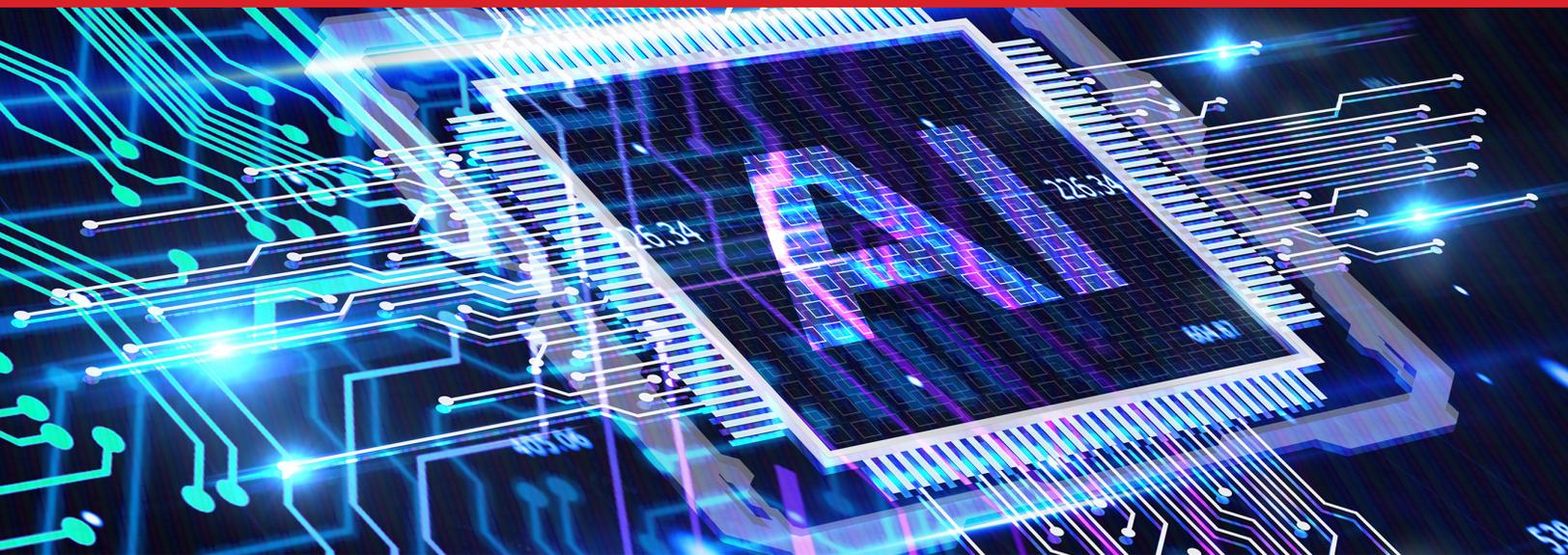
また、Dell のサービスでは、アドバイスから実装まで AI 導入に関するガイダンスを受けることができます。Dell ProConsult Advisory Services を利用することで、顧客は生成 AI プロセスの導入メリットを享受できる領域を特定でき、必要なソリューションや IT スキルが含まれたロードマップを作成することができます。Dell サービスには、大規模言語モデルの統合に必要なデータ準備や IT チーム向けの AI トレーニングを提供するサービスもあります。生成 AI の完全導入に向け、Dell のチームが顧客固有のユース ケースを精査し、ニーズに最適な AI モデルを構成します。HPE もプロフェッショナル サービスの提供を通じて、企業の AI 導入を支援しています。^{53, 54}

管理に関する考慮事項

サーバーには、管理者の時間を奪う継続的な管理が必要です。ファームウェア、ソフトウェア、ドライバーには定期的なアップデートが必要で、IT スタッフは性能と温度の維持管理、最適化などの作業を強いられます。以前の Principled Technologies (PT) テストでは、Dell サーバーの管理機能を Integrated Dell Remote Access Controller 9 (iDRAC9) で評価しました。⁵⁵ iDRAC9 OpenManage™ Enterprise (OME) では構成可能なスケジュール設定を使用する自動オンライン アップデートを利用できるため、管理者はサーバーを常に最新の状態に保ち、ワークロードの増加に合わせて、新しいサーバーのオンボードも、プロファイルを使用して迅速かつ簡単に行えます。iDRAC と OME を使用した場合、HPE OneView と HPE iLO を使用する場合よりも多くのリモート管理機能にアクセスでき、サーバーの導入もファームウェアのアップデートも簡単です。Dell PowerEdge サーバーには Dell 管理およびサービスが付属します。このサービスはシステム正常性の監視やファームウェアのアップデートなどのタスクにかかる時間と労力の削減に役立ち、IT サイクルを解放してイノベーションやその他のタスクに充てることができます。⁵⁶

表 3：2022 年 11 月の PT レポートによる、Dell と HPE の管理ツールの比較の概要。⁵⁷
出典：Principled Technologies

	Dell の管理ツールとの違い	どれだけ優れているか
リモート管理機能の数 iDRAC と iLO の比較	iDRAC の方が iLO よりも HTML5 コンソールおよび BIOS 構成機能が多く、より多くのリモート機能に対応する	HTML5 コンソール機能は 2.5 倍、BIOS 機能は 13 倍
サーバー導入の容易性 OME と OneView の比較	OME を使用した 1 対多プロファイル導入	OneView よりもサーバーの導入時間を 52% 短縮
ファームウェア アップデートの容易性 OME と OneView の比較	OME を使用した自動オンライン アップデート	Dell.com に接続して複数のサーバーをアップデートする。OneView でバンドルを手動でアップロードしてサーバーをアップデートする場合の所要時間を節約
アラートの容易性 OME と OneView の比較	OME でアラートポリシーを設定し、アラートに基づいて自動化されたアクションを実行する	OneView でアラートを受信するたびに手動でアクションを実行する場合と比較し、プロセスを自動化することで時間を節約でき、エラーが発生する可能性を低減できる
セキュリティ機能（システムロックダウンおよび動的 USB）の使いやすさ iDRAC と iLO の比較	iDRAC の方が時間もステップも少なく、再起動も不要	システム ロックダウンに要するステップは 1/4、時間を 91% 短縮
分析の堅牢性 CloudIQ for PowerEdge と InfoSight の比較	CloudIQ for PowerEdge にはカスタマイズ可能なレポートがあり、正常性の指標の数も多いため、管理能力に優れる	InfoSight と比較して 15 倍多い指標から選択可能



結論

AI の力を活用して業務の合理化と改善を図ることは、大変困難な作業ではあるものの、ビジネスにもたらす影響は甚大です。テクノロジーがこれまで以上に急速に進歩する中、AI に適したベンダーと提携することが重要になっています。包括的な AI ポートフォリオを提供するだけでなく、計画、準備、実装、管理サービスも提供できる Dell のような企業を選択することで、顧客はこれらの課題に正面から向き合うことができます。MLPerf® ベンチマーク テストから、Dell AI ポートフォリオの製品は、AI ワークロードに対して一貫した強力なパフォーマンスを提供することがわかりました。ハイパフォーマンスで柔軟性に優れたサーバー オプションに加え、複数のストレージ オプション、検証済みソリューション、AI 専用カスタマイズされたプロフェッショナル サービスを通じて、Dell は企業の AI 導入と AI のメリットを享受するための活動を支援しています。

1. Dell 「Increasing Your Data Value with Dell Generative AI Solutions」 (2023 年 12 月 19 日にアクセス) <https://www.dell.com/en-us/blog/increasing-your-data-value-with-dell-generative-ai-solutions/>
2. Dell 「Dell AI solutions」 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>
3. Dell 「Dell Technologies and Hugging Face to Simplify Generative AI with On-Premises IT」 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/dt/corporate/newsroom/announcements/detailpage.press-releases~usa~2023~11~20231114-dell-technologies-and-hugging-face-to-simplify-generative-ai-with-on-premises-it.htm#/filter-on/Country:en-us>
4. Dell 「Dell and Meta Collaborate to Drive Generative AI Innovation」 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/blog/dell-and-meta-collaborate-to-drive-generative-ai-innovation/>
5. Dell 「Dell AI-Ready Data Platform」 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/storage-for-ai.htm?hve=explore+unstructured+storage#tab0=0>
6. Dell 「Snowflake and Dell Partnership Gains Momentum」 (2023 年 12 月 19 日にアクセス) <https://www.dell.com/en-us/blog/snowflake-and-dell-partnership-gains-momentum/>
7. Robert McNeal 「Dell, VMware and NVIDIA Bring AI to Your Data」 (2024 年 1 月 17 日にアクセス) <https://www.dell.com/en-us/blog/dell-vmware-and-nvidia-bring-ai-to-your-data/>。上記リンク先：「2023 年の 8 月の Dell の分析に基づくと、デル・テクノロジーズは、ワークステーション パソコン（モバイルおよび固定型）からサーバーまで AI ワークロードをサポートする各種ソリューションを揃え、ハイパフォーマンス コンピューティング、データストレージ、クラウド ネイティブ ソフトウェアデファインド インフラストラクチャ、ネットワークング スイッチ、データ保護、HCI、各種サービスを提供しています」
8. Intel 「Accelerate Artificial Intelligence (AI) Workloads with Intel Advanced Matrix Extensions (Intel AMX)」 (2023 年 12 月 12 日にアクセス) <https://www.intel.com/content/www/us/en/content-details/785250/accelerate-artificial-intelligence-ai-workloads-with-intel-advanced-matrix-extensions-intel-amx.html>

9. Viperia 「NVIDIA's H100 and A100 GPU Cards: Exploring the Intricacies of SXM and PCI-E Connections」 (2023年12月12日にアクセス) <https://www.viperatech.com/unraveling-the-mysteries-sxm-vs-pci-e-connections-in-nvidias-high-end-h100-and-a100-gpus/>
10. GitHub 「MLPerf® Results Messaging Guidelines」 (2024年1月16日にアクセス) https://github.com/mlcommons/policies/blob/master/MLPerf_Results_Messaging_Guidelines.adoc
11. MLCommons® 「MLPerf® Inference: Datacenter Benchmark Suite Results」 (2023年12月12日にアクセス) <https://mlcommons.org/en/inference-datacenter-31/>
12. Dell 「PowerEdge XE Servers」 (2023年12月12日にアクセス) <https://www.dell.com/en-us/dt/servers/specialty-servers/poweredge-xe-servers.htm?hve=explore+poweredge+xe#tab0=0>
13. HPE 「HPE ProLiant XL675d Gen10 Plus Configure-to-order Server」 (2023年12月12日にアクセス) <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.1013142988.html>
14. HPE 「HPE ProLiant DL380a Gen11」 (2023年12月12日にアクセス) <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.proliant-dl380-server.1014696168.html>
15. MLCommons® 「MLPerf® Inference: Datacenter Benchmark Suite Results v 3.1」 (2023年12月12日にアクセス) <https://mlcommons.org/benchmarks/inference-datacenter/>
16. HPE 「NVIDIA Accelerators for HPE ProLiant Servers」 (2023年12月12日にアクセス) https://www.hpe.com/psnow/doc/c04123180.html?jumpid=in_pdp-psnow-qs
17. Dell 「PowerEdge XE9680 Specification Sheet」 (2024年1月19日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9680-spec-sheet.pdf>
18. HPE 「HPE & NVIDIA financial services solution sets new records in performance」 (2023年12月12日にアクセス) <https://community.hpe.com/t5/alliances/hpe-amp-nvidia-financial-services-solution-sets-new-records-in/ba-p/7197388>
19. HPE 「QuickSpecs: HPE Cray Supercomputing XD670」 (2023年12月12日にアクセス) <https://www.hpe.com/psnow/doc/a50004292enw>
20. 検証済み MLPerf® スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023年12月5日、エントリー 3.1-0069 から抽出)。MLPerf® の名前とロゴは、米国およびその他の国における MLCommons® Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
21. 検証済み MLPerf® スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023年12月5日、エントリー 3.1-0085 から抽出)。MLPerf® の名前とロゴは、米国およびその他の国における MLCommons® Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
22. Dell 「PowerEdge XE9640 Rack Server」 (2023年12月12日にアクセス) <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>
23. Accelsius 「Enabling the AI Revolution with Liquid Cooling」 (2023年12月12日にアクセス) <https://www.accelsius.com/blog/enabling-the-ai-revolution-with-liquid-cooling>
24. Dell 「Dell PowerEdge XE9640 Technical Guide」 (2023年12月12日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9640-technical-guide.pdf>
25. HPE 「HPE ProLiant DL380a Gen11」 (2023年12月12日にアクセス) https://www.hpe.com/psnow/doc/PSN1014696168WWEN.pdf?jumpid=in_pdp-psnow-dds
26. Intel 「Intel® Data Center GPU Max Series Technical Overview」 (2023年12月12日にアクセス) <https://www.intel.com/content/www/us/en/developer/articles/technical/intel-data-center-gpu-max-series-overview.html#gs.088741>
27. HPE 「Intel Data Center GPU Max 1100 48GB Accelerator for HPE Data sheet」 (2023年12月12日にアクセス) <https://www.hpe.com/psnow/doc/PSN1014779728WWEN>
28. 検証済み MLPerf® スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023年12月5日、エントリー 3.1-0066 から抽出)。MLPerf® の名前とロゴは、米国およびその他の国における MLCommons® Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
29. 検証済み MLPerf® スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023年12月5日、エントリー 3.1-0084 から抽出)。MLPerf® の名前とロゴは、米国およびその他の国における MLCommons® Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください

30. Dell 『AI and HPC —With Air or Liquid Cooling』 (2023 年 12 月 12 日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/servers/briefs-summaries/poweredge-xe9640-and-xe8640-infographic.pdf>
31. Dell 『PowerEdge XE8640 : Drive AI, HPC modeling and simulation workloads with superior performance』 (2023 年 12 月 12 日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe8640-spec-sheet.pdf>
32. Dell 『PowerEdge XE8640 Rack Server』 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>
33. 検証済み MLPerf[®] スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023 年 12 月 5 日, エントリー 3.1-0067 から抽出)。MLPerf[®] の名前とロゴは、米国およびその他の国における MLCommons[®] Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
34. 検証済み MLPerf[®] スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023 年 12 月 5 日, エントリー 3.1-0084 から抽出)。MLPerf[®] の名前とロゴは、米国およびその他の国における MLCommons[®] Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
35. Dell 『PowerEdge R760xa Rack Server』 (2023 年 12 月 12 日にアクセス) https://www.dell.com/en-us/shop/dell-poweredge-servers/poweredge-r760xa-rack-server/spd/poweredge-r760xa/pe_r760xa_16902_vi_vp#features_section
36. SANStorageWorks 『Dell EMC PowerEdge R760xa: Powerful and scalable for GPU workloads』 (2023 年 12 月 12 日にアクセス) <https://www.sanstorageworks.com/PowerEdge-R760xa.asp>
37. Dell 『Dell PowerEdge Servers and NVIDIA GPUs』 (2023 年 12 月 12 日にアクセス) <https://infohub.delltechnologies.com//design-guide-generative-ai-in-the-enterprise-inferencing/dell-poweredge-servers-and-nvidia-gpus-1/>
38. 検証済み MLPerf[®] スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023 年 12 月 5 日, エントリー 3.1-0064 から抽出)。MLPerf[®] の名前とロゴは、米国およびその他の国における MLCommons[®] Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
39. 検証済み MLPerf[®] スコア v3.1 推論, Closed. <https://mlcommons.org/benchmarks/inference-datacenter/> (2023 年 12 月 5 日, エントリー 3.1-0084 から抽出)。MLPerf[®] の名前とロゴは、米国およびその他の国における MLCommons[®] Association の登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.org を参照してください
40. Dell 『Workstations for AI』 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/dt/ai-technologies/index.htm?hve=explore+dell+precision+for+ai#pdf-overlay=https://www.delltechnologies.com/asset/en-us/products/workstations/briefs-summaries/ai-industry-brochure.pdf>
41. NVIDIA 『NVIDIA RTX in Professional Workstations』 (2023 年 12 月 12 日にアクセス) <https://www.nvidia.com/en-us/design-visualization/desktop-graphics/>
42. Intel 『Intel[®] Deep Learning Boost (Intel[®] DL Boost)』 (2023 年 12 月 12 日にアクセス) <https://www.intel.com/content/www/us/en/artificial-intelligence/deep-learning-boost.html>
43. HPE 『HPE ProLiant ML350 Gen11』 (2023 年 12 月 12 日にアクセス) <https://buy.hpe.com/us/en/compute/tower-servers/proliant-ml300-servers/proliant-ml350-server/hpe-proliant-ml350-gen11/p/1014696172>
44. ComputerWeekly.com 『Storage requirements for AI, ML and analytics in 2022』 (2023 年 12 月 12 日にアクセス) <https://www.computerweekly.com/feature/Storage-requirements-for-AI-ML-and-analytics-in-2022>
45. Dell 『PowerScale AI-Ready Data Platform』 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale>
46. Dell 『Compare PowerScale』 (2023 年 12 月 12 日にアクセス) <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale#compare-module>
47. Dell 『Dell PowerScale All-Flash』 (2023 年 12 月 12 日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h15963-ss-powerscale-all-flash-nodes.pdf>
48. Dell 『Dell PowerScale OneFS Software Features』 (2023 年 12 月 12 日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h18275-onefs-software-features-data-sheet.pdf>
49. Dell 『Dell APEX File Storage for AWS』 (2023 年 12 月 12 日にアクセス) <https://www.delltechnologies.com/asset/en-us/products/storage/briefs-summaries/h19575-so-apex-file-storage-for-aws.pdf>

50. Dell 「Dell ECS Enterprise Object Storage」(2023年12月12日にアクセス)
<https://www.dell.com/en-us/dt/storage/ecs/index.htm?hve=explore+ecs#tab0=0&tab1=0>
51. Dell 「Dell ECS Enterprise Object Storage」(2023年12月12日にアクセス)
<https://www.dell.com/en-us/dt/storage/ecs/index.htm#tab0=0&tab1=0&accordion0>
52. HPE 「Storage Solutions for Scality」(2023年12月12日にアクセス)
<https://www.hpe.com/us/en/storage/file-object/scality.html>
53. HPE 「Make AI work for you」(2024年1月16日にアクセス)
<https://www.hpe.com/us/en/solutions/ai-artificial-intelligence.html>
54. HPE 「HPE AI Services – Generative AI Implementation」(2024年1月16日にアクセス)
<https://www.hpe.com/us/en/services/generative-ai-implementation-service.html>
55. Principled Technologies 『Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE』(2023年12月12日にアクセス)
<https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>
56. Principled Technologies 『Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE』(2023年12月12日にアクセス)
<https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>
57. Principled Technologies 『Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE』

MLPerfの名前とロゴは、米国およびその他の国におけるMLCommons Associationの登録商標および未登録商標です。All rights reserved. (不許複製・禁無断転載) 無断での使用は固く禁じられています。詳細については、www.mlcommons.orgを参照してください

▶ オリジナルの英語版のレポートは
<https://facts.pt/zPmSx4c>でご覧いただけます

このプロジェクトは、デル・テクノロジーズの委託を受けて作成されています。



Facts matter.®

Principled Technologies は、Principled Technologies, Inc. の登録商標です。
他のすべての製品名は各社の商標です。

保証の免責事項、責任の制限：

Principled Technologies, Inc. はそのテストの精度と妥当性を確保するために適切な努力を行っていますが、テストの結果と分析、それらの精度、完全性、または品質に関して、特定の目的に対する適合性の黙示保証を含め、明示または黙示にかかわらず、いかなる保証も放棄します。すべての個人または事業体は自己の責任においてテストの結果に依存し、Principled Technologies, Inc. およびその従業員、その請負業者が、テスト手順や結果における疑わしいエラーや欠陥による損失や損害についてのいかなる主張に対しても、何ら責任を負わないことを認めるものとします。

Principled Technologies, Inc. は、そのテストに関連する間接的、特別的、付随的、結果的な損害に対して、当該損害の可能性について知らされていた場合でも、一切責任を負わないものとします。いかなる場合も Principled Technologies, Inc. は、直接的損害を含め、Principled Technologies, Inc. のテストに関連して支払われた金額を超える責任を負わないものとします。お客様の唯一の救済手段は、ここに示すとおりです。