



# NVIDIA DGX Station A100

## システム アーキテクチャ

*AI時代のワークグループ アプリケーション*

テクニカル ホワイトペーパー

# 目次

1. はじめに
2. NVIDIA DGX Station A100 システム アーキテクチャ
  - 2-1. NVIDIA A100 GPU エラスティック コンピューティング時代向けの第8世代データセンター GPU
  - 2-2. 第3世代 Tensor コア
  - 2-3. TensorFloat-32 (TF32) による Tensor コアの活用
  - 2-4. 細粒度構造化スパース性
  - 2-5. マルチインスタンス GPU (MIG)
  - 2-6. 4 基の NVIDIA A100 Tensor コア GPU を搭載
  - 2-7. 大規模な複合ワークロードを高速化する第 3 世代の NVLink
  - 2-8. PCIe Gen4 による高速化性能
  - 2-9. 冷却方法の改善
  - 2-10. サーバークラス AI アプライアンス向けのサーバークラス CPU
  - 2-11. リモート管理による柔軟性の強化 (BMC)
  - 2-12. グラフィックスを適宜使用する
  - 2-13. その他のシステム機能
3. セキュリティ
  - 3-1. ドライブの暗号化
  - 3-2. システム メモリの暗号化
  - 3-3. トラステッド プラットフォーム モジュール (TPM) テクノロジー
4. 完全に最適化された DGX ソフトウェア スタック
5. 状況を一変させるパフォーマンス
  - 5-1. ディープラーニングのトレーニングと推論
  - 5-2. 機械学習とデータサイエンス
6. NVIDIA DGXperts への直接アクセス
7. まとめ

# 1. はじめに

データサイエンスチームは、反復の高速化（各実験に費やす時間を削減できれば実施できる実験の数が増える）、また従来よりも大規模かつ高品質なデータセットの使用により、ワークフローとモデル品質を向上させる方法を模索しています。こうした専門家たちはイノベーションの最先端に身を置き、自社組織に大きな影響を与えるプロジェクトの開発を進めています。しかしながら、GPU 対応のラップトップやワークステーション、利用可能な GPU クラウド インスタンス、データセンター AI サーバーの借用部分など、予備の AI コンピュート サイクルを探し続けることも少なくありません。

このようなチームには、組織内の他の領域に振り回されることのない専用の AI リソースが必要です。すなわち、多忙なデータサイエンティストが投げかける可能性のあるジョブをすべて処理可能な妥協のない専用の AI システム、パフォーマンスを最大化するためにハードウェアとソフトウェア間で完全に最適化されたアクセラレーテッド AI プラットフォームです。また、AI 時代向けのワークグループ サーバーが必要です。そのようなチームのために、NVIDIA DGX Station™ A100 が用意されています。



NVIDIA DGX Station A100 は、AI スーパーコンピューティングをデータサイエンスチームにもたらし、データセンターの設置や IT への追加投資なしにデータセンターテクノロジーを提供します。複数の同時接続ユーザー向けに設計されており、ワークステーション用フォームファクターでサーバーグレードコンポーネントを活用します。これは、マルチインスタンス GPU (MIG) 対応の NVIDIA A100 Tensor コア GPU を 4 基、完全に相互接続することで、合計最大 320 GB の GPU メモリを備えながら、オフィスや自宅の標準的な電源に接続可能な唯一のシステムであり、お客様が好きな場所に配置できる強力な AI アプライアンスとなります。

このホワイトペーパーでは、DGX Station A100 の設計とアーキテクチャについて説明します。

## 2. NVIDIA DGX Station A100 システム アーキテクチャ

DGX Station A100 の外側のシャーシは旧 DGX Station から変更されていませんが、内部は全面的に再設計されています。



図1 : DGX Station A100 のシャーシと側面パネルの図

## 2-1. NVIDIA A100 GPU エラスティック コンピューティング時代向けの 第8世代データセンター GPU

中心部では、NVIDIA DGX Station A100 システムの心臓部は、NVIDIA A100 GPU です (図2)。これは、数々の改良や新機能によって NVIDIA V100 GPU より高い性能を発揮し、大規模で複雑なAIワークロードや複数の小さなワークロードを効率的に高速化するように設計されています。A100 GPU には、40 ギガバイト (GB) の高帯域幅 HBM2 と、従来よりも大容量で高速なキャッシュが組み込まれており、AI および HPC のソフトウェアとプログラミングの複雑さを軽減します。



図2 : NVIDIA A100 Tensor コアGPU

NVIDIA A100 GPU には、AI ワークロードと HPC アプリケーションのパフォーマンスをさらに高速化するために次の新機能が搭載されています。

- 第3世代 Tensor コア
- 細粒度構造化スパース性
- マルチインスタンス GPU

## 2-2. 第3世代 Tensor コア

NVIDIA A100 GPU には、新しい第3世代の Tensor コアが搭載されています。Tensor コアは、混合精度行列の融合積和演算に特化した高性能演算器であり、AI ワークロードと HPC アプリケーションを高速化します。

NVIDIA V100 搭載の NVIDIA DGX Station で使用されていた第 1 世代の Tensor コアは、FP16 と FP32 の混合精度行列演算を高速化します。A100 Tensor コア GPU に搭載された最新の Tensor コアは、さらに大きな行列を使用して効率を高め、INT4 およびバイナリ データ型を含む性能を NVIDIA V100 Tensor コアに対して2倍に向上させます。また、A100 Tensorコア GPU では次の新しいデータ型に対応します。

- TF32
- IEEE 準拠の FP64
- Bfloat16 (BF16)

Tensor コアの BF16/FP32 混合精度演算は、Tensor コアの FP16/FP32 混合精度演算と同じ速度で実行され、ディープラーニングトレーニングの別の選択肢が得られます。

## 2-3. TensorFloat-32 (TF32) による Tensor コアの活用

これまで、AI トレーニングで通常利用される FP32 数値演算では Tensor コアによる高速化が行われませんでした。NVIDIA A100 アーキテクチャでは、FP32 数値演算で Tensor コアを使用する新しい TensorFloat-32 (TF32) 数値演算が導入されます。新しい TF32 演算は、前世代のデータセンター GPU で利用可能な FP32 FMA 演算の 10 倍の速さで実行されます。

新しい TensorFloat-32 (TF32) 演算では、8 ビットの指数部 (FP32 と同じ範囲)、10 ビットの仮数部 (FP16 と同じ精度)、および1ビットの符号部 (図3) を使用して計算を実行します。このように、TF32 は FP32 の範囲と FP16 の精度を組み合わせるものになります。計算実行後、標準の FP32 出力が生成されます。

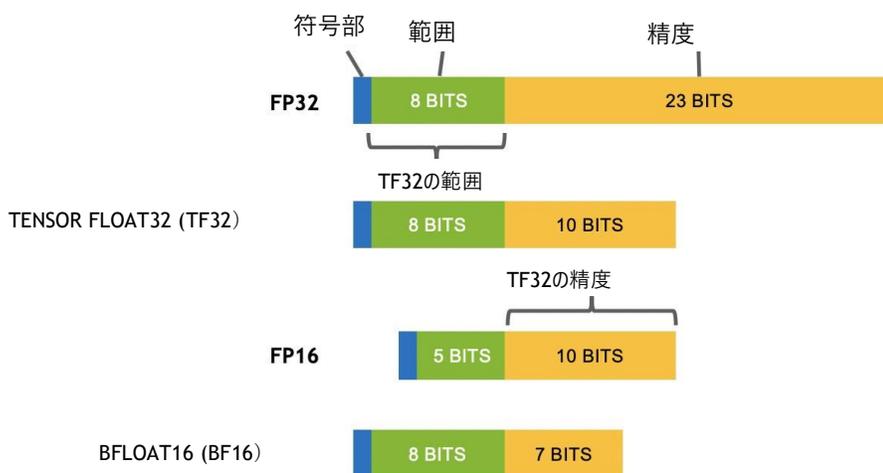


図3 : Tensor Float 32、FP32、FP16、BF16の説明

Tensor コアの外部では FP32 データパスを使用できるため、NVIDIA A100 は FP32 データの入出力に対して、TF32 で高速化された数値演算を実現できます。

TF32は、NGC ディープラーニング コンテナ 20.06 のリリース以降、TensorFlow、PyTorch、MXNet のデフォルトモードになっています。TensorFlow 1.15 については、[ソースコード](#)と Python パッケージもリリースされています。これらのディープラーニングフレームワークでは、コードを変更する必要はありません。V100 の FP32 と比較すると、A100 の TF32 は最も要求の厳しい[対話型 AI](#) モデルの1つである BERT-Large モデルのトレーニングを 6 倍超スピードアップします。

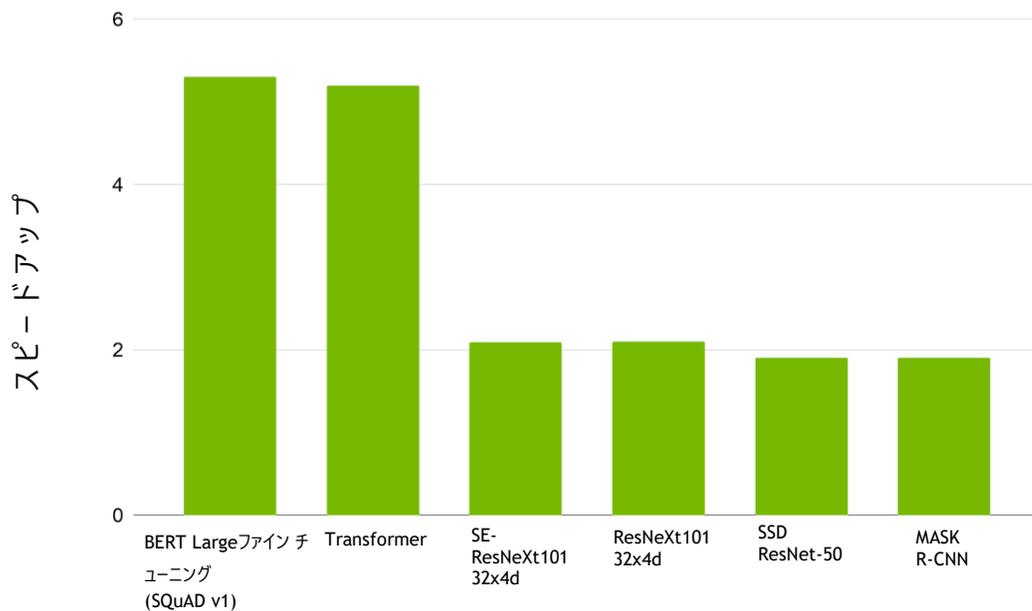


図4 : TF32 は、FP32、NGC pytorch:20.06-py3 コンテナのPyTorch 1.6、BERT-Large モデルのトレーニングに比べて5 倍を超えるスピードアップを実現できます。DGX A100 (A100 GPU の8 倍)での結果。モデル スクリプトはすべて、[Deep Learning Examples リポジトリ](#)に入っています。

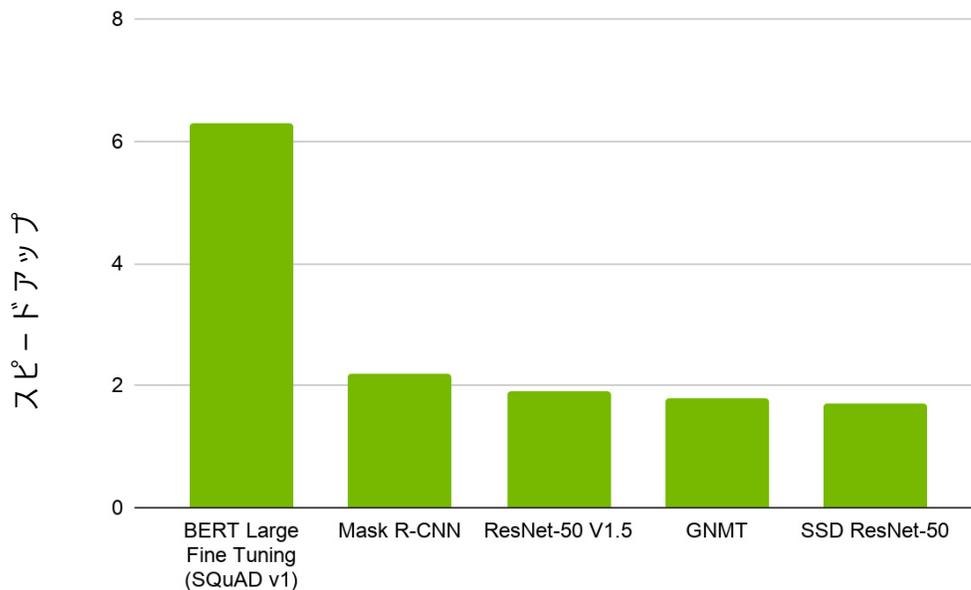


図5 : TF32は、FP32、tensorflow:20.06-tf1-py3 コンテナ-コンテナのTensorFlow 1.15、BERT-Large モデルのトレーニングに比べて6 倍を超えるスピードアップを実現できます。DGX A100 (A100 GPU の8 倍)での結果。モデル スクリプトはすべて、[Deep Learning Examples リポジトリ](#)に入っています。

## 2-4. 細粒度構造化スパース性

NVIDIA A100 GPU は、細粒度構造化スパース性をサポートしており、精度を損なうことなくニューラル ネットワークを高速化します。スパース性は、多くの場合、プルーニングによって発生します。これは、ネットワークの精度にほとんど寄与しない重みを削除する手法です。通常、これには「ゼロ化」とゼロまたはゼロに近い値を持つ重みの削除が含まれます。このように、プルーニングによって密なネットワークをスパース ネットワークに変換し、計算、メモリ、およびエネルギー要件を削減して同一水準の精度を実現することができます。ただし、現在まで、このタイプの細粒度スパース性は、モデル サイズの縮小とパフォーマンスの高速化という期待に沿っていませんでした。

細粒度構造化スパース性と A100 がサポートする2：4 パターン (図6) を使用すると、スパース ネットワークの各ノードが同じ量のメモリ アクセスと計算を実行し、ワークロードがバランスよく分散されて、コンピュータ ノードの使用率も上がります。さらに構造化スパース行列は効率的に圧縮できるため、スパース Tensor コア形式をハードウェアでサポートする A100 GPU では行列積和演算のスループットが倍増します。

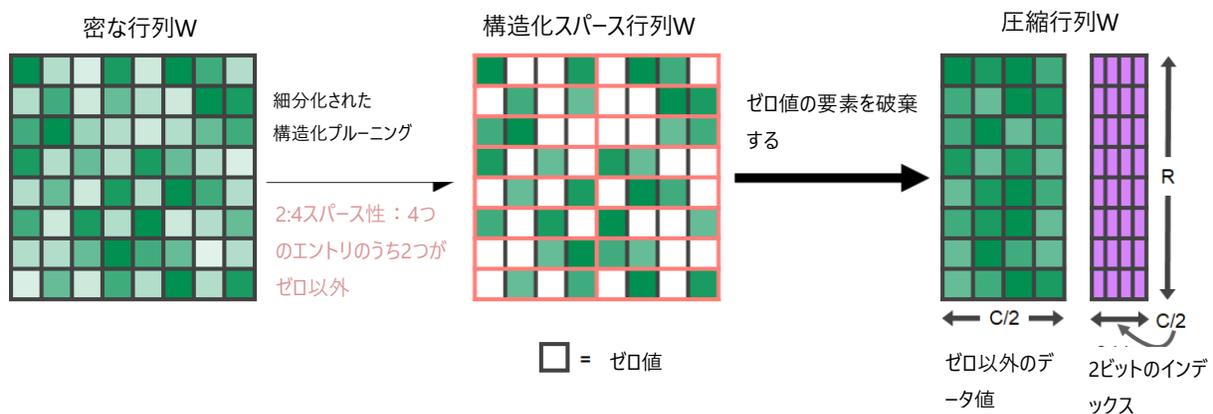


図6 : NVIDIA A100 GPU は、効率的な圧縮形式と 2 倍のインストラクション スループットで細粒度構造化スパース性をサポートします。その結果、さまざまなAI ネットワークにまたぐことでTensor コアの計算が高速化され、推論パフォーマンスが向上します。細粒度 構造化スパース性により、A100 のINT8 Tensor コア演算は V100 の20 倍のパフォーマンスをもたらし、FP16 Tensor コア演算速度は V100 の5 倍になります。

## 2-5. マルチインスタンス GPU (MIG)

NVIDIA A100 GPU には、GPU の使用率を高めるため、マルチインスタンス GPU (MIG) と呼ばれる新しいパーティショニング機能が組み込まれています。MIG は、空間パーティショニングを使用し、1つの A100 GPU の物理リソースを7つもの独立した GPU インスタンスに分割します。MIG を使用すると、NVIDIA A100 GPU が GPU ごとに同時インスタンスを使用し、V100 の最大 7 倍のスループットによってサービス品質を保証します (図7)。

NVIDIA A100 GPU で MIG が有効になっていれば、各 GPU インスタンスが独自のメモリ、キャッシュ、ストリーミング マルチプロセッサを備えているため、並列コンピューティング ワークロードは隔離された GPU メモリと物理 GPU リソースにアクセスできます。これにより、複数のユーザーが同じ GPU を共有してすべてのインスタンスを同時に実行できるため、GPU の効率が最大になります。

MIG は、DGX Station A100 システムの任意の数の GPU で個別に有効化できます。すべての GPU を MIG 対応にする必要はありません。一方、DGX Station A100 システムのすべての GPU で MIG が有効になっている場合は、最大 28 人のユーザーが同時かつ独立して GPU アクセラレーションを利用できます。

実際、DGX Station A100 は、MIG 対応のワークステーション用フォーム ファクターによる唯一のシステムです。DGX Station A100 の MIG を有効にすることがメリットになる典型的なユースケースを次に示します。

- DGX A100 サーバーで配備して実際に稼働させる前の、小規模かつ低レイテンシのモデルが関わる、GPU 全体のすべてのパフォーマンスを必要としないバッチ サイズ 1 の複数の推論ジョブの評価
- モデル探索用の Jupyter ノートブック
- 学生や大組織のデータサイエンスチームのメンバーなど、複数のユーザー間での GPU のリソース共有

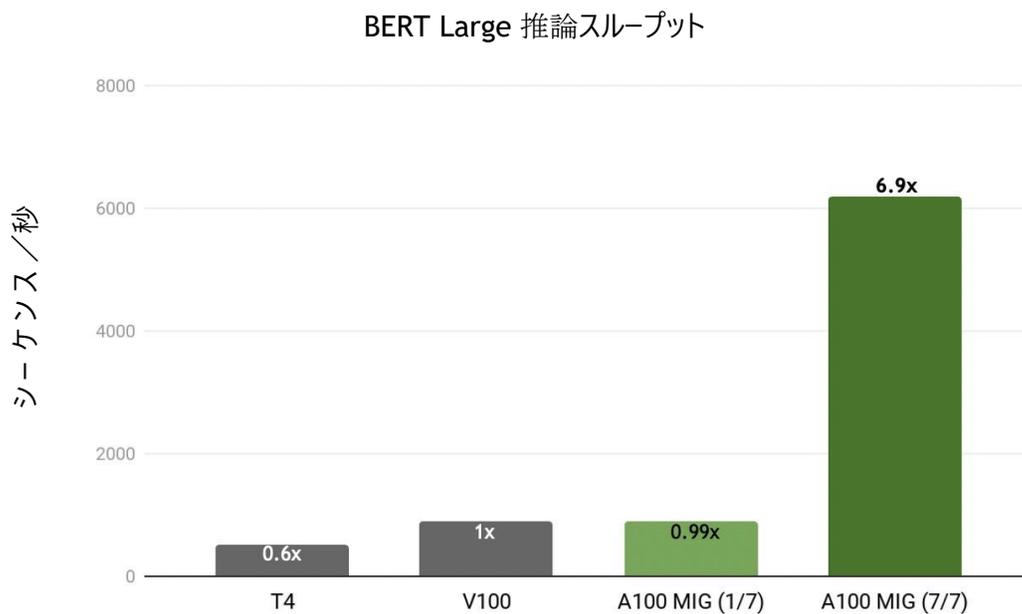


図7

： マルチインスタンス GPU (MIG) による最大で7 倍速い推論スループット。上記の結果は DGX A100 で実行したものです。

BERT Large 推論 (シーケンス長=128)

- T4 : TRT7.1、精度=INT8、バッチサイズ=256
- V100 : TRT7.1、精度=FP16、バッチサイズ=256
- 7 個の 1 g.5 gb MIG インスタンスを備える A100。TensorRT リリース候補版、バッチサイズ=94、精度=INT8、スパース性 (1g.5gb は A100 の最小インスタンスで、コンピュータの 117 と総メモリの 5 GB を指定)

さらに個の A100 GPU を搭載した DGX A100 で実行すると、次の例 (図8) に示すように、大きく異なるワークロードに対してユーザーが異なる GPU を構成することができます。

- 4 個の GPU を AI トレーニングに使用
- 2 個の GPU を HPC またはデータ分析に使用
- 2 個の GPU を MIG モードで使用、14 個の MIG インスタンスにパーティショニング、それぞれが推論を実行



図8： NVIDIA DGX Station A100 システムの各種 GPU で異なるワークロード

MIG はさまざまな利用法をサポートしているので、ユーザーはベアメタルまたはコンテナで CUDA アプリケーションを実行できます。MIG サポートは Docker 用の NVIDIA Container Toolkit (旧称：nvidia-docker2) を使用して利用できるため、ユーザーは GPU インスタンスで CUDA アクセラレーテッド コンテナを実行できます。詳しくは、[「CUDA アプリケーションをコンテナとして実行する」](#)を参照してください。

## 2-6. 4 基の NVIDIA A100 Tensor コア GPU を搭載

DGX Station A100 の最もユニークな機能の1つは、4-way NVIDIA HGX™ GPU ボードの組み込みです。ハイパフォーマンス アクセラレーテッド データセンター サーバー向けに設計された NVIDIA HGX™ プラットフォームは、NVIDIA GPU のフルパワーおよび NGC™ の完全に最適化された NVIDIA AI と HPC のソフトウェア スタックを組み合わせ、最高のアプリケーション パフォーマンスを提供します。エンドツーエンドのパフォーマンスと柔軟性を備えた NVIDIA HGX により、研究者や科学者はシミュレーション、データ分析、AI を組み合わせて科学の進歩を促進することができます。

4基の A100 Tensor コア GPU、総計 160GB の HBM2 メモリまたは 320GB の HBM2e メモリ、総帯域幅 2.4TB/s により、データサイエンティストは、DGX Station A100 の設置場所として選んだところなら、どこからでもデータセンターのテクノロジーとパフォーマンスにアクセスできます。

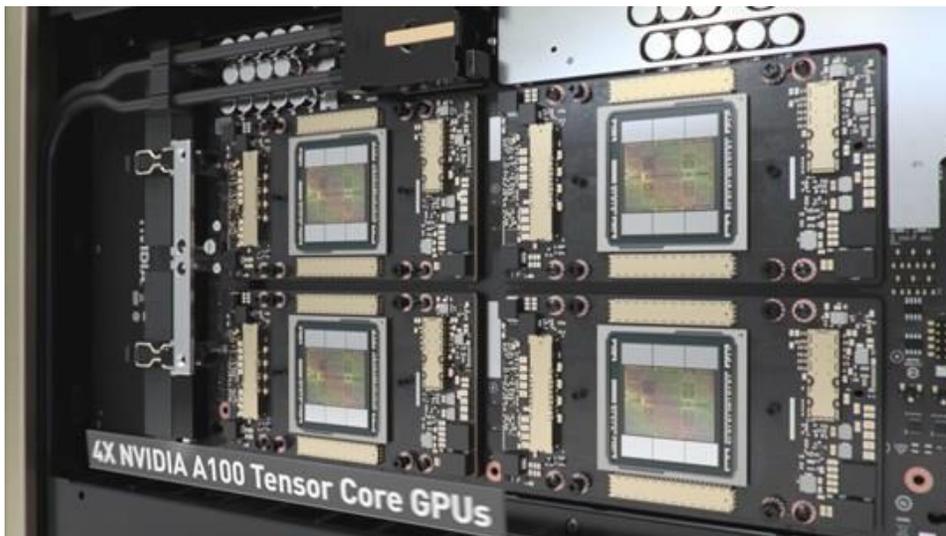


図9 : NVIDIA DGX Station A100 に内蔵されている4-way NVIDIA HGX ボード

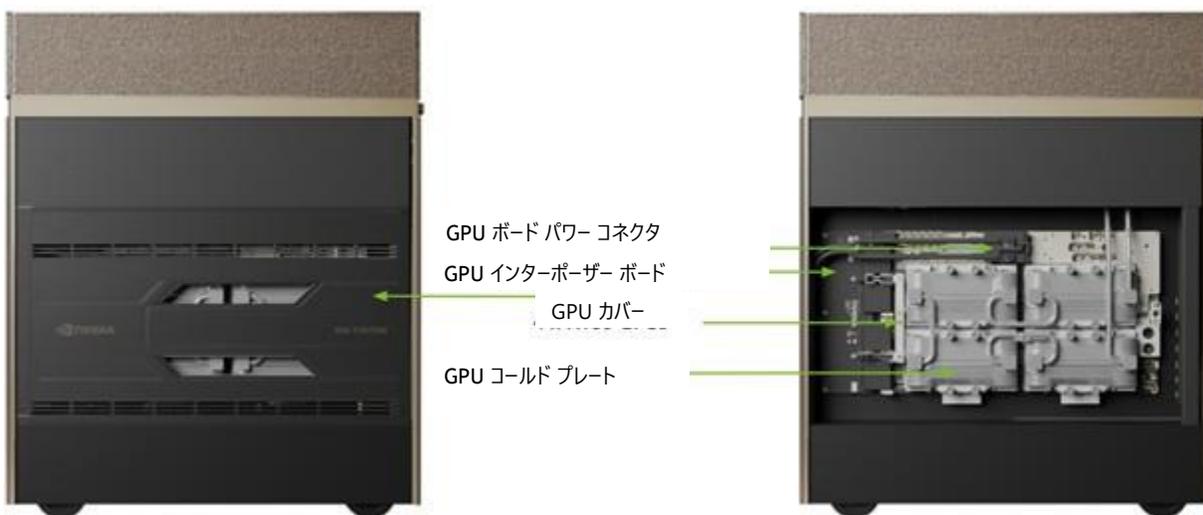


図10 : NVIDIA DGX Station A100 の図 (GPU 側、GPU コールドプレート付き)

## 2-7. 大規模な複合ワークロードを高速化する第 3 世代の NVLink

NVIDIA® NVLink® は、GPU 間的高速直接相互接続です。

HGX GPU ベースボードの 4 基の A100 GPU は第 3 世代の NVLink で直接接続され、完全な接続性を可能にしています。すべての A100 GPU が高速 NVLink ポートを使用して他の A100 GPU のメモリにアクセスできます。A100 間のピア帯域幅は 300GB/s の双方向であり、最速の PCIe Gen4 × 16 バスの 3 倍以上高速になります。

トポロジー図については、図 11 を参照してください。

## 2-8. PCIe Gen4 による高速化性能

NVIDIA A100 GPU は、16 レーンの PCI Express Gen 4 (PCIe Gen4) バスによって PCI スイッチインフラストラクチャに接続されており、それぞれが 31.5Gb/s を提供するので総計 252Gb/s となり、帯域幅が PCIe 3.0/3.1 の 2 倍になります。これらは、DGX ディスプレイ アダプター、NVMe ストレージ、および CPU へのアクセスを提供するリンクになります。

トレーニング ワークロードでは、通常、精度を向上させるために同じデータセットを何度も読み取る必要があります。このデータを何度も繰り返し転送してネットワーク帯域幅全体を使い果たすのではなく、ハイ パフォーマンス ローカルストレージを NVMe ドライブとともに実装してデータをキャッシュします。これにより、データをメモリに読み込むスピードが上がり、ネットワークとストレージ システムの輻輳も軽減されます。

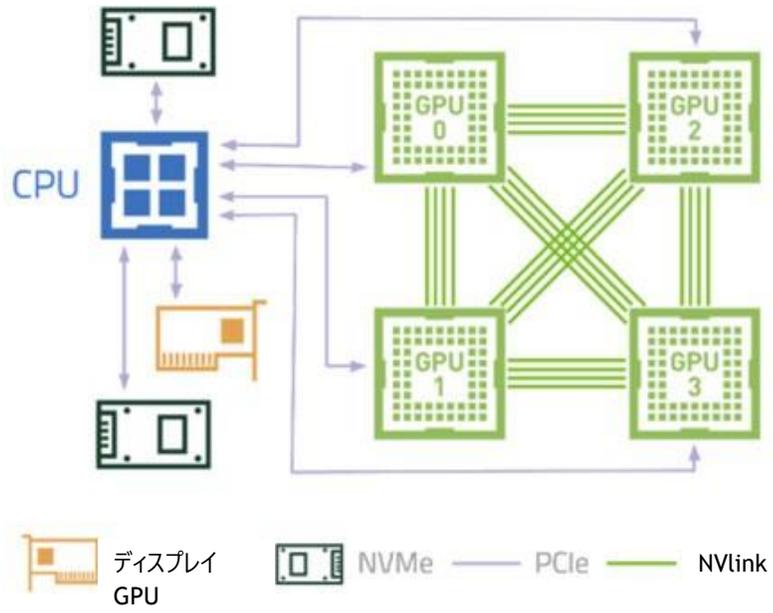


図11 : DGX Station A100 デバイストポロジー図

## 2-9. 冷却方法の改善

前世代の DGX Station と同様、DGX Station A100 は、公称動作温度 5～30°C のオフィス環境で静かに動作するように設計されています。ただし、前世代とは異なり、これまでより高い GPU/CPU コンポーネント温度に対処できるだけでなく、完全にメンテナンスフリーでそれを実行できる新しい改良型の冷媒ベース冷却システムを備えています。これにより、水位のチェックと補給が不要になります。冷却システムに漏れが発生してもシステムが損傷する可能性がなく、環境を破壊することもなく、ユーザーは整備部品について心配する必要はありません。



図12 : 図10 : NVIDIA DGX Station A100 冷却システムの図

冷媒システムは、1 台の循環ポンプ、GPU とシステム CPU に取り付けたコールドプレート、各種システムコンポーネントを相互接続するための配管、および熱を排出するための冷媒リザーバーを搭載した熱交換器ユニットで構成されています。3 つの低速ファンが、ささやくように静かに (<37dBm) 凝縮器への空気流を作り出し、収集した熱を周囲の外気中に排出します。

## 2-10. サーバークラス AI アプリケーション向けのサーバークラス CPU

DGX Station A100は、Zen 2 マイクロ アーキテクチャをベースとする最新の AMD EPYC 7742 エンタープライズクラスのサーバ プロセッサを搭載しています。AMD EPYC 7742 プロセッサは、最新の TSMC 7nm 製造プロセスを使用して HPC と AI のワークロード向けに最高のパフォーマンスを提供するものであり、すでに数多くの世界記録とベンチマークで実証されています。DGX Station A100 システムには、ブート、ストレージ管理、ディープラーニング フレームワークのスケジューリングと調整用のこれらの CPU の1つが搭載されています。CPU は最大速度 3.4GHz のブーストで動作し、1 コアあたり 2 スレッドの 64 コアを備えています。

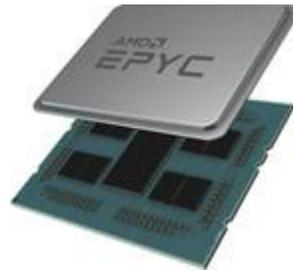


図13 : *コールドプレートで覆われた AMD EPYC7742 プロセッサの図*

この CPU は、大規模なメモリ容量と帯域幅を提供し、総計 204.8GB/s のメモリ帯域幅用に 8 個のメモリチャネルを備えています。DGX Station A100 のメモリ容量は 512GB で、8 個の DIMM スロットに 64GB DDR4-3200 ECC RDIMM メモリを標準搭載しています。

I/O については、AMD EPYC 7742 プロセッサは 128 個の PCIe Gen4 リンクを提供します。これにより、GPU およびその他の IO デバイスへの高速接続をサポートするプロセッサからシステムに最大帯域幅が提供されます。各 DGX Station A100 システムには、複数の 1.92TB NVMe M.2 ブート OS SSD、および 1 個の 7.68TB PCIe gen4 NVMe U.2 キャッシュ SSD が付属しています。

## 2-11. リモート管理による柔軟性の強化 (BMC)

リモート管理は、IT サービスがシステム管理を行えるようにしながら、DGX Station A100 のコンピューティング リソースを複数の研究者やチーム間で共有するための柔軟性を提供します。また、DGX Station A100 を他の IT インフラストラクチャとともに小部屋や机の下にも設置することができます。

リモート管理機能は、DGX ステーション A100 のマザーボードに組み込まれたフル装備のベースボード マネジメント コントローラ (BMC) を介して実行されます。システムの監視と管理をリモートで行う Web ベースのユーザー インタフェースと、既存のインフラストラクチャ ツールがシステムを管理および監視できるようにする IPMI インタフェースと RedFish インタフェースを提供します。



図 14 : DGX Station A100 BMC ログインページのスクリーンショット

Web ベースのユーザー インタフェースは、使いやすいメニューと全コンポーネントの詳細を提供する複数の画面により、全センサーの読み取りとシステムログの確認を行うための安全な方法を提供します。このインタフェースでは、GPU、メモリDIMM、CPU、ディスプレイカード、およびマザーボードの温度監視が提供されます。ファン回転数、消費電力、システム電圧も履歴グラフおよび現在の読み取り値とともに監視、表示されます。

これらの機能はすべて IPMI インタフェースからも利用できます。したがって、ログ、統計値、センサー読み取り値を収集する監視ソフトウェアは、ユーザーの介入なしに情報を自動取得できます。IPMI インタフェースには、システムのシリアル コンソールにアクセスしてシステムの BIOS 設定やインストールされているオペレーティング システムを管理するための Serial Over LAN (SOL) インタフェースも備わっています。

さらに Web ベースのインタフェースは、リモートのキーボード、ビデオ、マウス (KVM) 機能も提供しており、ユーザーはモニターの表示を正確に確認するとともに、離れた場所からシステムを管理できます。KVM 機能には仮想ストレージ機能も備わっているので、リモート ボリュームをマウント、かつリモート DGX Station A100 を ISO イメージから再インストールまたは起動することができます。

DGX Station A100 で利用できる便利な機能の一つとして、リモート管理ネットワーク インタフェースと通常のシステム LAN ケーブルを 1 つのネットワーク接続部で接続することができます。これにより、小部屋内部の 1 つのネットワーク ドロップを両方のネットワーク機能で使用することができます。これは BMC で簡単に構成可能であり、ネットワーク コントローラ サイドバンド インタフェース (NCSI) と呼ばれるテクノロジーを活用します。

BMC は、システムの背面パネルにあるボタンを点灯させることにより、システムをリモート識別できる LED 制御も提供します。この LED は、取り付けられたボタンで制御することもできます。システムを管理するリモート チームと、システムを保守するローカル チームを調整するのに優れたツールです。

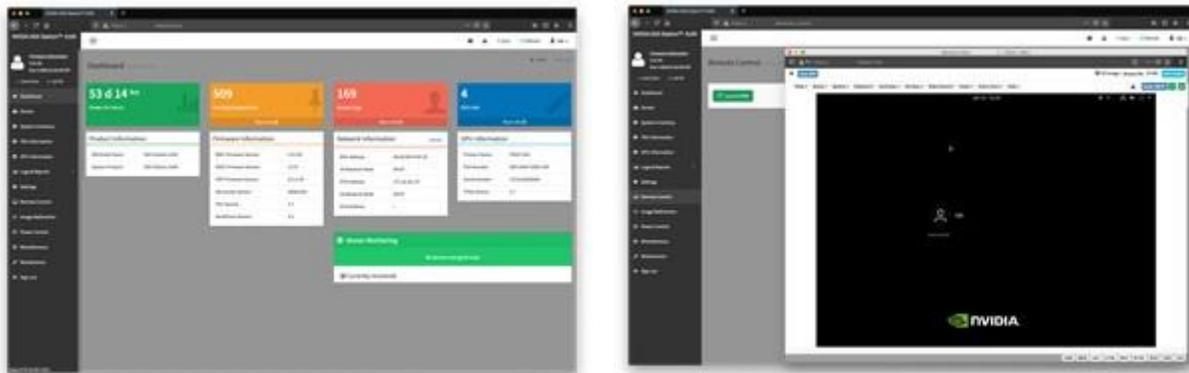


図15 : DGX Station A100 BMC Web UI のサンプル スクリーンショット

## 2-12. グラフィックスを適宜使用する

DGX Station A100 には、現在、独自の専用ワークステーション グレード グラフィックスアダプター、NVIDIA Turing™ グラフィックス プロセッシング ユニット (GPU) テクノロジーで動作する NVIDIA DGX ディスプレイ アダプター、および 4GB の GDDR6 メモリが搭載されています。柔軟なシングル スロットと薄型フォーム ファクターのカードは、最も要求の厳しいワークフローを満足するのに十分強力です。コンパクトで電力効率の高いフォーム ファクターによる最高の 3D パフォーマンスを活用します。

主要なパフォーマンス機能は次のとおりです。

- 4 GB GDDR6 グラフィックス メモリにより、データ サイエンティストや研究者が前例のない規模、複雑さ、および豊かさで結果をレンダリングできます。
- 896 個のストリーミング マルチプロセッサ (MX) コアにより、冷たくて静かなシングル スロット フォーム ファクターでピーク消費電力 50W の驚異的なパフォーマンスを実現します。
- VESA DisplayPort 1.4 Mini-DP 出力による 4 つの同時表示と最大 5K の解像度をサポートします。

DGX ディスプレイ アダプターに加えて、BMC には PCIe をベースとする独自の 2D/VGA ディスプレイ アダプター機能もあります。このアダプターは、解像度が最大 1920×1200@60Hz の標準的 (DB-15) なアナログ VGA ディスプレ

イをサポートしています。BMC 2D/VGA ディスプレイは BIOS で有効/無効にできます。デフォルトでは有効になっています。または、BIOS の「ディスプレイ アダプター」設定を「オンボード」モードに変更することにより、BMC 2D/VGA ディスプレイ アダプターをデフォルトの X Window System ディスプレイに設定できます。

## 2-13. その他のシステム機能

このシステムは、高速データ転送用に 2 つのオンボード 10 ギガビット イーサネット インタフェースで IO 接続を十分に提供します。これらのインタフェースを使用してネットワーク共有ストレージをマウントすることで、低遅延で高帯域幅の HPC または AI アプリケーションのデータを読み込みます。前述のように、ポートの1つは NCSI 機能を備えているので、ネットワーク ケーブルを統合し、システムに隣接するポートの必要数を減らすことができます。

システムは簡単に整備できます。障害発生時は、コンポーネントは明確にラベル付けされており、交換できるツールがほとんどないか、まったくないため、お客様がこのシステムを直接修理できます。整備可能な品目には、DIMM、NVMe ソリッド ステート ドライブ、ディスプレイ カード、トラステッド プラットフォーム モジュール、およびバッテリーなどがあります。

システム重量が 105lb なので、ホイールを組み込み、パッケージからフロアに移動し、オフィスや実験室で扱いやすいようにしました。ホイールには、ユニットが動かないように取り付け可能なロック一式が付属しています。これらのロックは、机や台の上に目立つように設置してユニットが転がらないようにする場合にも大変役立ちます。

セキュリティのため、意図せずに移動する可能性のある場所に DGX Station A100 を設置する場合に備え、Kensington ロックが追加されました。

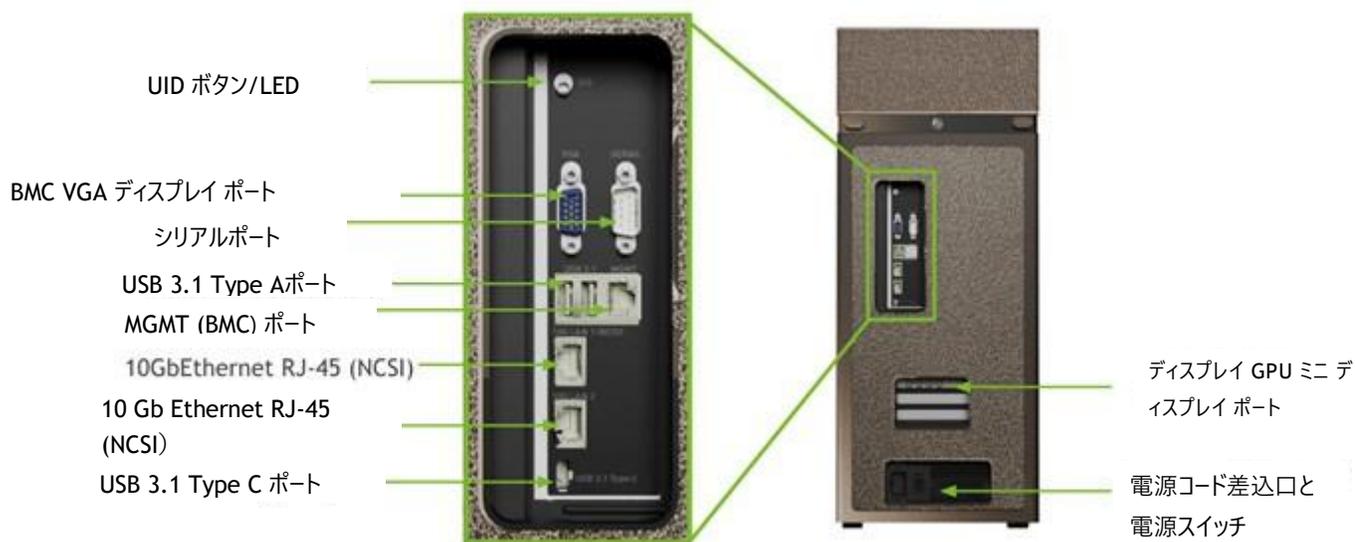


図16： DGX Station A100 の背面パネル図 (すべてのポートを含む)

## 3. セキュリティ

NVIDIA DGX Station A100 システムは、数多くの高度なセキュリティ機能をサポートしています。

### 3-1. ドライブの暗号化

NVIDIA DGX™ OS ソフトウェアは、ファイル システム レベルのシステム パーティションの暗号化、およびオプションの自己暗号化ドライブ (SED) を使用するデータ ドライブのフル ドライブ暗号化をサポートします。SED は、パフォーマンスに影響を与えることなく、ドライブ上のデータをオン ザ フライで自動的に暗号化します。ドライブは、TPM モジュールに格納されているキーから、または集中型キー サーバーを使用してロックを自動解除できます。

システム パーティションの暗号化はソフトウェア内部で実行します。手動でパスフレーズを入力して、または自動で集中型キー サーバーを使用して、起動するたびにファイル システムのロックを解除する必要があります。

### 3-2. システム メモリの暗号化

NVIDIA DGX Station A100 の 64 コア AMD EPYC CPU は、システム BIOS にてデフォルトで有効になっているセキュア メモリ暗号化 (SME) とセキュア 仮想マシン暗号化 (SEV) 機能により、厄介な攻撃に対する保護を顧客に別途提供します。CPU は、使用中のデータを保護するためのハードウェア アクセラレーテッド メモリ暗号化を提供し、プロセッサで利用可能な次の新しいセキュリティ コンポーネントを活用します。

- AES-128 暗号化エンジン：メモリ コントローラに組み込まれており、適切なキーを受け取ると、メイン メモリ内のデータを自動的に暗号化および復号化します。
- AMD セキュア プロセッサ：安全にキーを生成かつ管理するための暗号化機能を提供します。

AMD の SME 機能と SEV 機能の詳細については、<https://developer.amd.com/sev/>を参照してください。

### 3-3. トラステッド プラットフォーム モジュール (TPM) テクノロジー

NVIDIA DGX Station A100 システムには、

トラステッド プラットフォーム モジュール (TPM 2.0)<sup>1</sup>の業界標準に準拠したセキュアな暗号化プロセッサが内蔵されています。暗号化プロセッサは、DGX A100 のセキュリティ サブシステムの基盤であり、暗号化操作によってハードウェアを保護します。

有効にすると、DGX OS が完全に起動してアプリケーションが実行されるまで、TPM によって起動プロセスの完全性が保証されます。

TPM は、ボルト キーと SED 認証キーを安全に保管するための自己暗号化ドライブおよびドライブ暗号化ツールでも使用されます。

1. トラステッド コンピューティング グループのトラステッド プラットフォーム モジュールのホワイトペーパーを参照してください (<https://trust-edcomputinggroup.org/resource/trusted-platform-module-tpm-summary/>)。

## 4. 完全に最適化された DGX ソフトウェア スタック

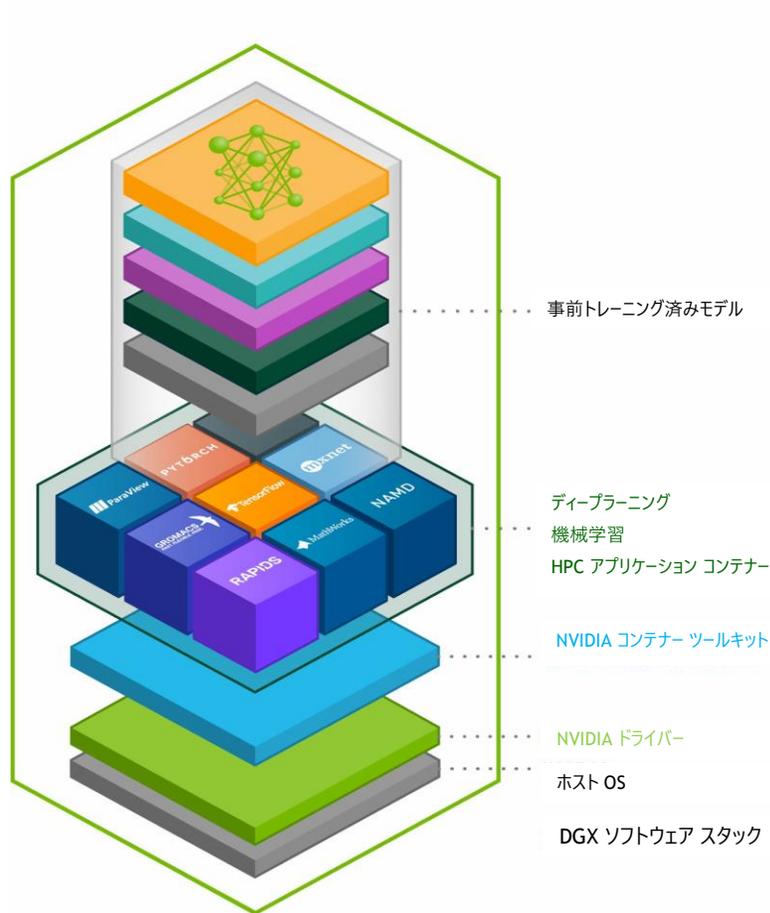
DGX Station A100 ソフトウェアは、AI ワークロードを大規模に実行するために構築されています。主な目標は、プラクティショナーが最小限のセットアップ作業でディープラーニング フレームワーク、データ分析、および HPC アプリケーションを DGX Station A100 に導入できるようにすることです。プラットフォーム ソフトウェアである DGX OS の設計は、Ubuntu に基づいて最適化された OS が中心になっており、必要なドライバー、ライブラリ、およびツールがすべてシステムにプリインストールされています。お客様は、マッチするソフトウェア ドライバーやライブラリを特定してインストールしなくても、すぐに成果を上げることができます。

DGX OS は、すぐに使えるようにすべての DGX システムにプリインストール済みです。このソフトウェアは、ソフトウェアの追加インストールおよび最新のセキュリティ パッチとバグ フィクスによる既存のソフトウェアのアップグレード用に Canonical と NVIDIA が共催するリポジトリからも入手できます。このリポジトリには NVIDIA の追加ドライバーと [CUDA ツールキット](#) の各バージョンが入っているので、お客様は新機能が必要になったときにソフトウェア スタックを簡単にアップグレードできます。

DGX OS は、コンテナ化されたアプリケーションと SDK ソフトウェアを使用するために必要なツールとライブラリも提供しており、GPU のサポートも含まれています。コンテナ化されたソフトウェアを使用すれば、十分にテストされたディープラーニング フレームワークやその他のアプリケーションへのアクセスがさらに速くなります。プラクティショナーは、[NGC レジストリ](#) の事前トレーニング済みモデル、モデル スクリプト、Helm チャート、およびソフトウェア開発キット (SDK) とともに、ディープラーニング (DL)、機械学習 (ML)、ハイパフォーマンス コンピューティング (HPC) アプリケーション用として GPU に最適化されたコンテナを取り出せます。コンテナはすべて、DGX システムで開発、テスト、チューニングされており、すべての DGX 製品と互換性があります (DGX-1、DGX-2、DGX A100、DGX Station、DGX Station A100)。

DGX のお客様は、プライベート レジストリにもアクセスできます。プライベート レジストリは、企業内や組織内で他のユーザーと共有できるカスタム コンテナ、モデル、モデル スクリプト、および Helm チャートを安全に保管できる場所を提供します。[NGC プライベート レジストリ](#) の詳細については、こちらの [ブログ投稿](#) をご覧ください。

図17 は、これらすべての要素が DGX ソフトウェア スタックの一部としてどのように組み合わされているのかを示して



います。

図17： NVIDIA DGX ソフトウェア スタック

DGX ソフトウェア スタックには、次の主要コンポーネントが含まれています。

- **NVIDIA コンテナ ツールキット**を使用すれば、ユーザーは GPU で高速化された Docker コンテナを構築して実行できます。ツールキットには、コンテナを自動構成して NVIDIA GPU を活用するためのコンテナ ランタイム ライブラリとユーティリティが入っています。
- **GPU アクセラレーテッド コンテナ**は下記をサポートするソフトウェアを備えています。
  - > トレーニング用のディープラーニング フレームワーク([PyTorch](#)、[MXNet](#)、[TensorFlow](#)など)
  - > 推論プラットフォーム([TensorRT](#)など)
  - > データ分析([RAPIDS](#)、エンドツーエンドのデータサイエンスと分析のパイプラインを完全に GPU で実行するためのソフトウェア ライブラリスイートなど)

- > ハイパフォーマンス コンピューティング (HPC) ([CUDA-X HPC](#)、[OpenACC](#)、[CUDA®](#)など)
- [NVIDIA CUDA ツールキット](#) (各 GPU アクセラレーテッド コンテナに組み込まれている)は、ハイパフォーマンスGPUアクセラレーテッド アプリケーション作成用の開発環境です。

CUDA 11 では、ソフトウェア開発者と DevOps エンジニアが、次のような新しい NVIDIA A100 GPU の主要革新技術の恩恵を享受できます。

- > 新しい入力データ型フォーマットのサポート、線形代数用 CUDA ライブラリでの Tensor コアとパフォーマンスの最適化、FFT、および行列乗算
- > Linux オペレーティング システムでの MIG インスタンスの構成と管理 (DGX ソフトウェア スタックの一部)
- > タスク グラフ、非同期データの移動、細分化された同期、L2 キャッシュ常駐制御のためのプログラミングと API

新着情報の詳細については、[CUDA 11 公開機能開発者ブログ](#)をご覧ください。

## 5. 状況を一変させるパフォーマンス

### 5-1. ディープラーニングのトレーニングと推論

革新的機能、データセンターグレードのテクノロジー、バランスの取れたシステム設計が詰め込まれた DGX Station A100 は、当該フォームファクターにおいてディープラーニングのトレーニングと推論で前例のないパフォーマンスを提供します。

大きなコンピューティングパワーを備えた画期的な A100 GPU、大容量 DRAM への高帯域幅アクセス、および高速ローカルストレージを組み合わせることにより、NVIDIA DGX Station A100 システムが BERT のような複雑なネットワークを飛躍的に高速化するのに最適なものになります。

DGX Station A100 システムは、それぞれが複雑なモデルを処理するために 5petaFLOPS の AI コンピューティング機能を備えています。BERT の大規模モデルは大量のメモリを必要としますが、DGX Station A100 はそれぞれ 160GB または 320GB の高帯域幅 GPU メモリを提供します。NVIDIA 相互接続テクノロジーの NVLink がすべての GPU をまとめ、効率的にスケーリングを行うために高帯域幅通信を備える大規模 AI モデルにおいて 1 つの GPU として機能します。

たとえば、下図に示すとおり、DGX Station (NVIDIA V100 GPU を 4 個搭載) に比べると、BERT Large 事前トレーニング フェーズ 1 が格段に速くなります。

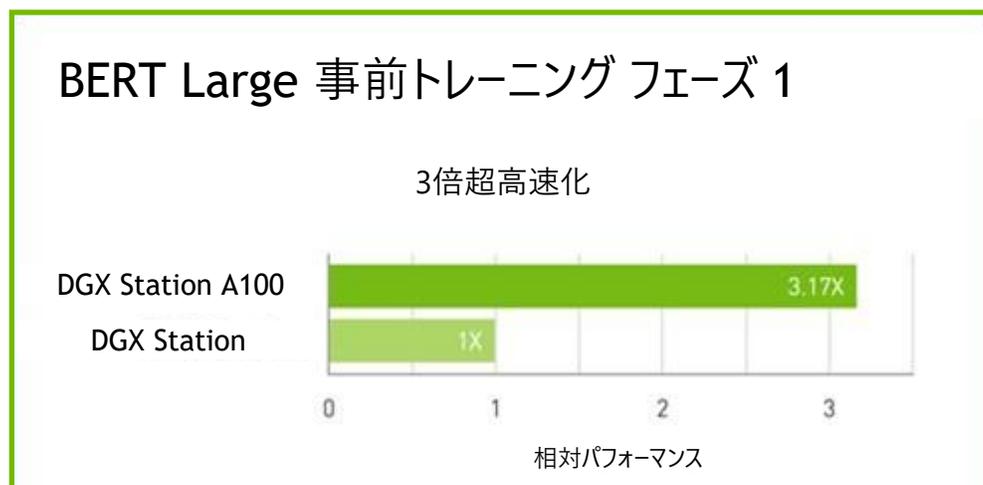


図18： BERT Large 事前トレーニング (DGX Station に対する DGX Station A100 の速度向上結果)：バッチサイズ=64；混合精度；AMP 搭載；実データ；シーケンス長=128

DGX Station A100 は、通常はディープラーニング推論のユースケースに使用されませんが、NVIDIA HGX ベースボードに 4 個の A100 Tensor コア GPU を搭載しており、マルチインスタンス GPU (MIG) テクノロジーをサポートしま

す。前述のとおり、GPU を小さなインスタンスにスライスして推論ワークロードを実行すると、特定の状況下で意味をなし、次のパフォーマンスが現行世代で大幅に向上しました。

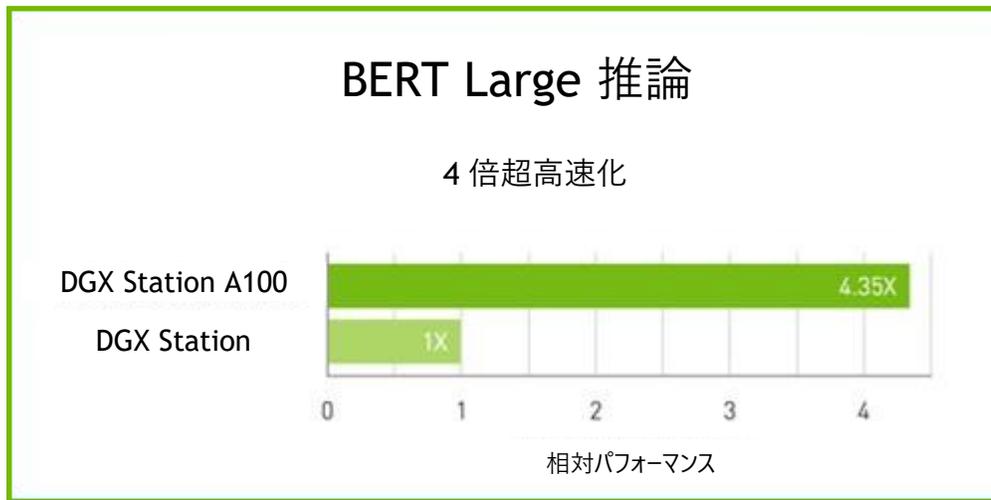
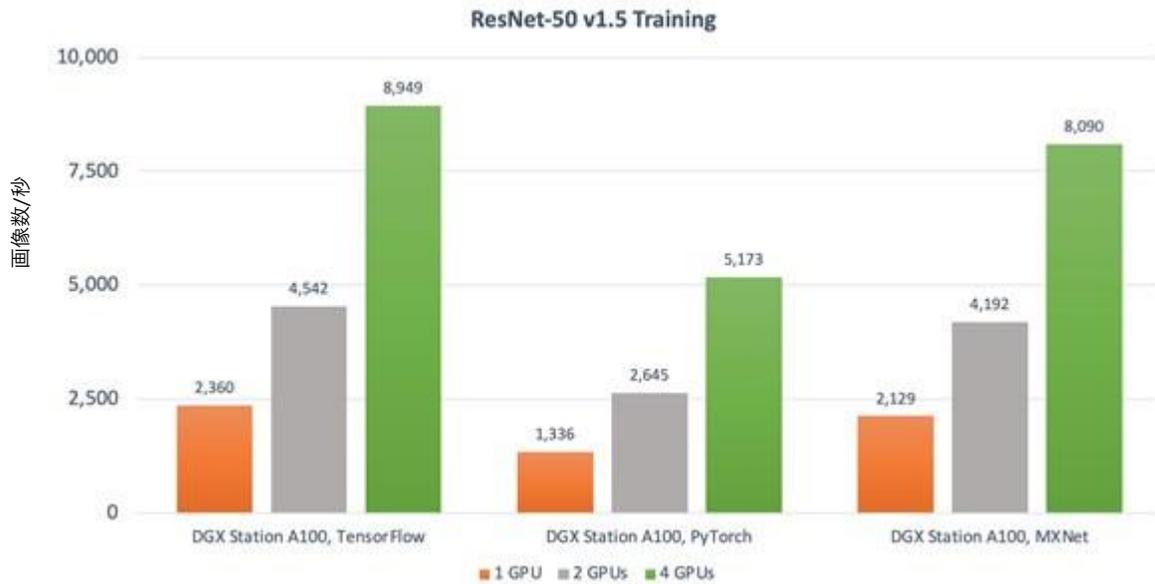


図19 : BERT Large 推論 (DGX Station に対する DGX Station A100 の速度向上結果) : バッチ サイズ =256 ; INT8 精度 ; 合成データ ; シーケンス長=128 ; cuDNN 8.0.4



機械学習とディープラーニングのワークロードが膨れ上がるにつれて、マルチGPUトレーニングの重要性が増します。DGX Station A100 では、1 個、2 個、4 個の GPU 間でほぼ線形のスケラビリティが得られます。これは、画像関連ディープラーニングのトレーニング ワークロードの典型例です。

#### ResNet-50 v1.5 トレーニング

図20 : DGX Station A100 (4x A100 SXM 80GB), ResNet-50 V1.5 トレーニング, 混合精度, TensorFlow (21.02-

*tf1\_py3, bs 256), PyTorch (21.02\_py3, bs 256), MXNet (21.02\_py3, bs 192)*

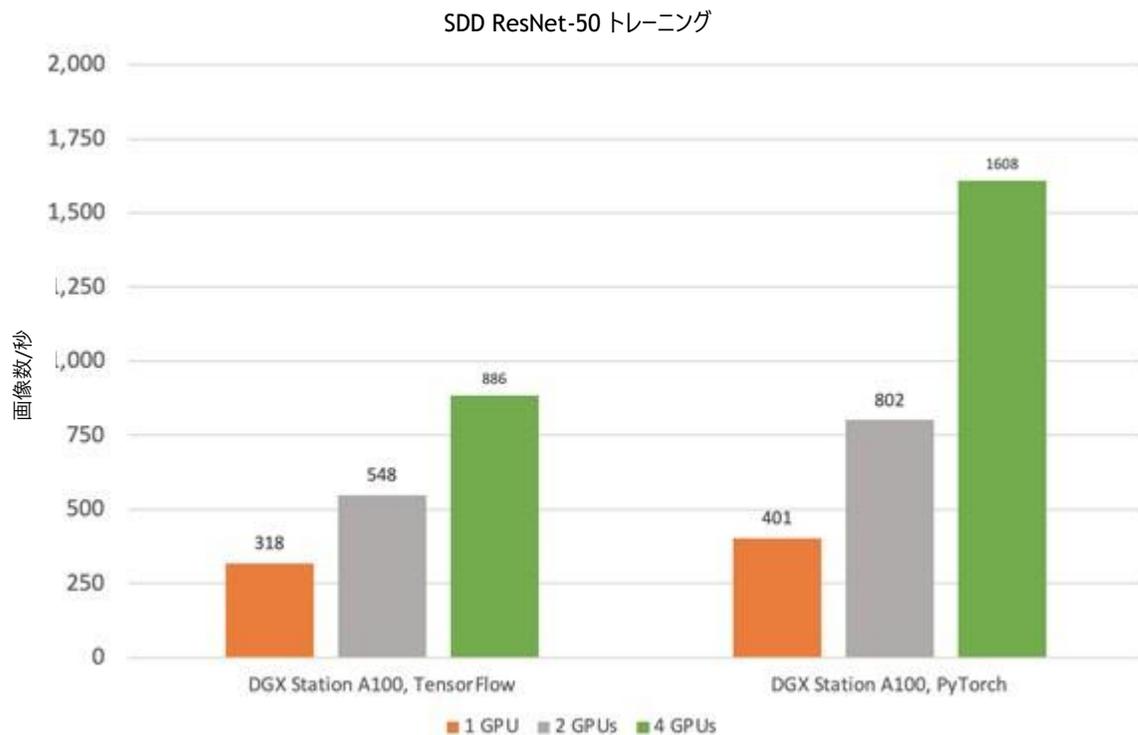


図21 : DGX Station A100 (4xA100 SXM 80GB), SSD ResNet-50 トレーニング、混合精度、TensorFlow (21.02-tf1\_py3コンテナ、bs 32)、PyTorch (21.02\_py3、bs 128)

## 5-2. 機械学習とデータサイエンス

### DGX Station A100 + RAPIDS を使用したほぼリアルタイムの実験

NVIDIA は、自社製品をテストするため、標準ベンチマークとして GPU ビッグデータ ベンチマーク (GPU-BDB) を構築し、典型的な大手小売会社や金融会社における実際の運用を模倣しました。30 のクエリで構成されており、大規模な抽出、変換、読み込み (ETL) 操作、自然言語処理、および 1TB または 10TB スケールの構造化データと非構造化データを組み合わせて使用する機械学習を必要とするものです。ベンチマークは「エンドツーエンド」で評価されます。すなわち、データをディスク上で開始および終了し、その間のすべてを測定します。

RAPIDS を使用して DGX Station A100 で 1TB ベンチマークを実行したところ、各クエリが平均 13.7 秒で完了し、総実行時間は 5 分ででした。NVIDIA DGX-1® サーバーで同じワークフローを実行すると、平均クエリ完了時間は 68.7 秒、総実行時間は 35 分ででした。DGX-1 サーバーと比較すると、DGX Station A100 は効率が 5 倍超アップしています。

DGX Station A100 は、コンパクトでオフィスフロアにぴったりのパッケージでありながらサーバーグレードのパフォーマンスを提供しますが、さらに難易度の高いビジネス上の課題を解決するデータサイエンスチームにほぼリアルタ

イムの双方向性を提供します。このレベルの双方向性により、チームが従来より短い時間で多くのことを容易に実行できるようになるため、企業はそれらのデータの価値を最大化することができます。

## 6. NVIDIA DGXperts への直接アクセス

[NVIDIA DGXperts](#) は、数千の DGX システムを導入した経験とフルスタック AI 開発の専門知識を有する、AI に精通した 16,000 人超の専門家によるグローバル チームです。メンバーのスキルセットには、システムの設計と計画、データセンターの設計、ワークロード テスト、ジョブ スケジューリング、リソース管理、持続的な最適化などがあります。

NVIDIA DGX A100 またはその他の DGX システムを所有していれば、NVIDIA エンタープライズ サポート サービスの一部としてこの専門家たちに直接連絡できます。NVIDIA DGXperts は、お客様の社内の AI に関する専門知識を補充し、エンタープライズ グレードのプラットフォームと AI に精通した人材の強化を組み合わせ、お客様が組織内の AI プロジェクトの目標を達成できるようにします。

## 7. まとめ

NVIDIA DGX Station A100 システムの革新的技術により、開発者、研究者、IT マネージャー、ビジネス リーダーなどが可能性の限界を押し広げ、プロジェクトや組織全体において AI のメリットを最大限活用できるようになります。

データサイエンティストのチーム全体で共有できるパワフルなシステムです。AI 時代のワークグループ サーバーとして設計されており、トレーニング、推論、分析のワークロードを MIG で並列実行可能で、最大 28 台の独立した GPU デバイスを個々のユーザーとジョブに提供できるので、アクティビティが抑えられ、システム全体の性能に影響を与えません。

このようなデータ サイェンス チームは、企業のオフィス、実験室、研究施設、さらには自宅で仕事をするこもあるでしょう。大規模な AI インフラストラクチャを設置するには多大な IT 投資と工業用の強力な電源供給と冷却機能を備えた大型データセンターが必要ですが、DGX Station A100 なら、お客様の作業スペースがどこにあっても標準的な壁のコンセントに接続するだけです。さらに、その革新的な冷却ベースのデザインがひんやりした手触りを与えます。

NVIDIA DGX Station A100 は、ワークステーションのフォームファクターでデータセンター級の AI サーバーを実現し、専用の電力供給と冷却なしに標準的なオフィス環境で使用するのに向いています。4 個の超強力な NVIDIA A100 Tensor コア GPU、最上位のサーバー グレード CPU、超高速 NVMe ストレージ、および最先端の PCIe Gen4 バスを備える設計になっています。DGX Station A100 にも NVIDIA DGX A100 と同じベースボード マネジメント コントローラ (BMC) が搭載されており、システム管理者はリモート接続によって要求タスクを実行できます。DGX Station A100 は、オフィス環境向けの最もパワフルな AI システムであり、データセンターなし

にデータセンター テクノロジを提供します。

詳しくは、下記をご覧ください。

- [NVIDIA DGX Station A100のWebページ](#)
- [NVIDIA DGX Station A100のデータシート](#)

[NVIDIA Ampere アーキテクチャの詳細に関する開発者ブログ](#)

## 注意事項

本書は情報提供のみを目的とするものであり、製品の特定の機能、状態、または品質を保証するものではありません。NVIDIA Corporation ("NVIDIA") は、本書に記載されている情報の正確性または完全性について、明示的または黙示的を問わず、一切の表明も保証もしておらず、いかなる誤りについても責任を負いません。NVIDIA は、そのような情報による結果または使用、あるいは特許など、その使用に起因する可能性のあるサードパーティの権利の侵害について一切責任を負わないものとします。本書は、任意の材料（下記に定義）、コード、または機能を開発、公開、または提供することを確約するものではありません。

NVIDIA は、予告なしに、本書に対して訂正、修正、拡大、改善、およびその他の変更を加える権利を留保します。

お客様は、注文する前に最新の関連情報入手し、当該情報が最新かつ完全であることを確認することが推奨されます。

NVIDIA 製品は、NVIDIA とお客様の正式な代表者が署名した個別の販売契約書（「販売条件」）において別途合意されていない限り、注文確認時に提供された NVIDIA 標準販売約款に従って販売されます。NVIDIA は、本書で参照している NVIDIA 製品の購入に関するお客様の一般条件を適用することについて、ここで明示的に否認します。本書によって直接的または間接的に契約上の義務が生じることはありません。

NVIDIA 製品は、医療、軍用、航空、宇宙、または生命維持装置、あるいは NVIDIA 製品の故障または誤動作が人身傷害、死亡、または物的損害や環境被害につながると合理的に予想されるアプリケーションでの使用に合わせて設計、認可、または保証されていません。NVIDIA は、そのような機器またはアプリケーションへの NVIDIA 製品の組み込みおよび/または使用について一切の責任を負いませんので、当該組み込みおよび/または使用はお客様の責任で行ってください。

NVIDIA は、本書に基づく製品が特定用途に適しているかどうかを表明または保証していません。各製品のパラメーターのテストは、必ずしも NVIDIA が実施しているとは限りません。本書に記載されている情報の適用可能性を評価および判断し、お客様が計画したアプリケーションに当該製品を確実に適合させ、アプリケーションに必要なテストを実施してアプリケーションまたは製品の不備を回避するのはお客様単独の責任となります。お客様の製品設計に支障があると NVIDIA 製品の品質と信頼性に影響を与える恐れがあり、本書に記載のない付加的または異なる条件および/または要件が生じる可能性があります。NVIDIA は、(i) 本書に反する方法での NVIDIA 製品の使用、あるいは (ii) お客様の製品設計に基づく起因する可能性のある不備、損害、費用、または問題に関して一切の責任を負いません。

明示的または黙示的を問わず、本書に伴う NVIDIA の特許権、著作権、またはその他の NVIDIA の知的財産権にかかるライセンスは付与されません。サードパーティの製品またはサービスについて NVIDIA が公開した情報は、それを以て当該製品またはサービスを使用するためのライセンスを NVIDIA が付与するものではなく、それを保証または承認するものではありません。そのような情報を利用するには、サードパーティの特許またはその他の知的財産権に基づいてサードパーティが付与するライセンス、あるいは NVIDIA の特許またはその他の知的財産権に基づいて NVIDIA が付与するライセンスが必要になる場合があります。

本書に記載されている情報の複製は、NVIDIA が書面で事前に承認し、変更せずに複製し、該当するすべての輸出法と規制に完全に準拠し、かつ関連するすべての条件、制限、および注意事項が添えられている場合のみ許可されます。

本書およびすべての NVIDIA 設計仕様書、参照ボード、ファイル、図面、診断データ、リスト、およびその他の文書（一括して「資料」）は「現状のまま」で提供されます。NVIDIA は、明示的、黙示的、または法定か否かを問わず、あるいは素材についていかなる保証も行わず、非侵害、商品適格性、および特定用途への適合性に関するすべての黙示的な保証を明示的に否認します。法律で禁止されていない範囲において、NVIDIA はいかなる場合も直接的、間接的、特別、偶発的、懲罰的、または必然的な損害を含むがそれらに限定されない損害については、本書の使用に伴って責任の理論に関係なく発生し、NVIDIA がそのような損害の可能性について知らされていたとしても、一切の責任を負わないものとします。いかなる理由であっても、お客様が被る可能性のある損害にかかわらず、本書に記載されている製品について NVIDIA がお客様に対して負う責任の総額および累計額は、製品の販売条件に従って限定するものとします。

## 商標

NVIDIA、NVIDIA のロゴ、DGX、DGX-1、DGX-2、DGX Station、NVIDIA Turing、および NVIDIA Volta は、NVIDIA Corporation の米国およびその他の国における商標および/または登録商標です。その他の会社名および製品名は、それらが関係するそれぞれの会社の商標または登録商標の可能性があり得ます。

## Copyright

©2021 NVIDIA Corporation. All rights reserved.

