

NVIDIA DGX STATION A100

WORKGROUP APPLIANCE FOR THE AGE OF AI



データサイエンスチームはAIイノベーションの最先端を走り、企業や世界を変革するプロジェクトを開発しています。しかし、最も複雑なモデルを訓練するために必要なコンピュートサイクルを探すのに時間がかかることが多いのが現状です。その解決策として、ハードウェアとソフトウェアの両方で完全に最適化され、世界中のどこにいても同時に複数のユーザーに画期的なパフォーマンスを提供できる専用のAIプラットフォームが求められています。

NVIDIA DGX Station™ A100の概要

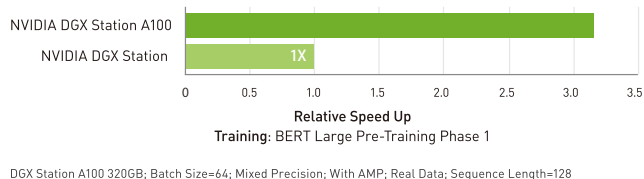
- > チームが制限なく使用できるトレーニング、推論、データ分析のための2.5ペタフロップスのパフォーマンスを提供する統合AIワークステーション。
- > サーバークレードで、プラグアンドゴーで、データセンターの電力と冷却を必要としません。
- > 世界クラスのAIプラットフォームで、複雑なインストールやITヘルプは必要ありません。
- > 4つのNVIDIA A100 Tensor Core GPUを完全に相互接続し、最大320ギガバイト(GB)のGPUメモリを搭載した世界で唯一のワークステーションです。
- > NVIDIAのノウハウと経験を活用して、AI変革を迅速に実現します。

NVIDIA DGX Station A100は、データサイエンスチームにAIスーパーコンピューティングをもたらし、データセンターや追加のITインフラストラクチャを必要としないデータセンター技術を提供します。複数同時にアクセスするユーザーのために設計されたDGX Station A100は、オフィスにやさしいフォームファクターでサーバークレードのコンポーネントを活用しています。これは、4つのGPUが完全に相互接続されたマルチインスタンスGPU(MIG)対応のNVIDIA A100 Tensor Core GPUを搭載した唯一のシステムであり、最大320GBの総GPUメモリを標準的なコンセントに差し込むことができます。

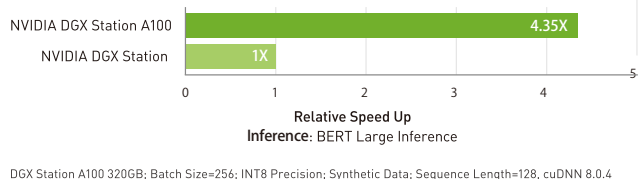
データサイエンスチームのための AIスーパーコンピューティング

DGX Station A100は集中化されたAIリソースを複数のユーザーに簡単に提供する、AI時代のワークグループ・アプライアンスです。トレーニング、推論、分析のワークロードを並行して実行することができ、MIを使用することで、最大28台のGPUデバイスを個々のユーザーに提供することができます。

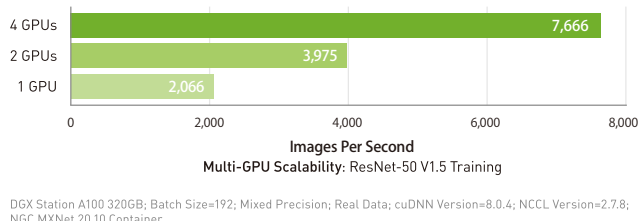
DGX Station A100 Delivers Over 3X Faster The Training Performance



DGX Station A100 Delivers Over 4X Faster The Inference Performance



DGX Station A100 Delivers Linear Scalability



DGX Station A100は、すべてのDGXシステムと同じ完全に最適化されたNVIDIA DGX™ソフトウェアスタックを搭載しており、最大のパフォーマンスと、個々のシステムからNVIDIA DGX POD™やNVIDIA DGX SuperPOD™に至るまで、DGXベースのインフラストラクチャとの完全な相互運用性を提供します。それにより、ユーザーやジョブの活動を抑制し、システム全体のパフォーマンスに影響を与えません。

データセンターを必要としない データセンター級のパフォーマンス

NVIDIA DGX Station A100は、特別な電源や冷却装置を必要とせず、標準的なオフィス環境での使用に適したワークステーション形式のデータセンタークラスのAIサーバーを提供します。

その設計には、4つの超強力なNVIDIA A100 Tensor Core GPU、最高級のサーバーグレードのCPU、超高速NVMeストレージ、および最先端のPCIe Gen4バスが含まれています。

DGX Station A100には、NVIDIA DGX A100と同じBMC(Baseboard Management Controller)も含まれており、システム管理者はリモート接続で必要なタスクを実行することができます。

DGX Station A100は、オフィス環境向けの最も強力なAIシステムであり、データセンターを介さずにデータセンター技術を提供します。

場所を選ばず設置できるAIアプライアンス

NVIDIA DGX Station A100は、企業のオフィス、研究室、研究施設、あるいは自宅で作業するといった昨今の仕事環境のデータサイエンスチームのために設計されています。

大規模なAIインフラストラクチャを設置するには、多額のIT投資と産業用の強力な電力と冷却を備えた大規模なデータセンターが必要ですが、DGX Station A100は、チームのワークスペースがどこにあっても、標準的なコンセントに差し込むだけです。

また、1人で簡単にセットアップを行うことができ、2本のケーブルで操作できるワールドクラスのAIプラットフォームを数分で稼働させることができます。

より大きなモデルを、より速い解決へ

NVIDIA DGX Station A100は単なるワークステーションではありません。机の下に置くことができる統合AIワークステーションです。64コアのデータセンター級CPUに加えて、NVIDIA DGX A100サーバーと同じNVIDIA A100 Tensor Core GPUを搭載しており、それぞれ40GBまたは80GBのGPUメモリを搭載し、高速なSXM4を介して接続されています。

NVIDIA DGX Station A100は、NVIDIA® NVLink®を活用し、4つのGPUを完全に相互接続してMIGをサポートし、システム性能に影響を与えずに、並列ジョブと複数のユーザーのために最大28台のGPUデバイスを提供する、オフィスに適した唯一のシステムです。

比類なきAI専門知識への統合的アクセス

NVIDIA DGX Station A100は、NVIDIAの何千人ものAI専門家に支えられた完全なハードウェアとソフトウェアのプラットフォームであり、世界最大のDGX試験場であるNVIDIA DGX SATURNVで得られた知識に基づいて構築されています。

NVIDIA DGX Station A100を所有することで、NVIDIA DGXpertsと直接アクセスすることができます。NVIDIAの10年以上のAIリーダーシップから得たノウハウと経験を活かして、AIの変革を迅速に進めるための処方的なガイダンスと設計の専門知識を提供する、AIに精通した専門家からなるグローバルチームです。これにより、ミッションクリティカルなアプリケーションを迅速に立ち上げて稼働させ、スムーズに稼働させ続けることができ、洞察に至るまでの時間を劇的に改善することができます。

SYSTEM SPECIFICATIONS

	NVIDIA DGX Station A100 320GB	NVIDIA DGX Station A100 160GB
GPUs	4x NVIDIA A100 80 GB GPUs	4x NVIDIA A100 40 GB GPUs
GPU Memory	320 GB total	160 GB total
Performance	2.5 petaFLOPS AI 5 petaOPS INT8	
System Power Usage	1.5 kW at 100–120 Vac	
CPU	Single AMD 7742, 64 cores, 2.25 GHz (base)–3.4 GHz (max boost)	
System Memory	512 GB DDR4	
Networking	Dual-port 10Gbase-T Ethernet LAN Single-port 1Gbase-T Ethernet BMC management port	
Storage	OS: 1x 1.92 TB NVME drive Internal storage: 7.68 TB U.2 NVME drive	
DGX Display Adapter	4 GB GPU memory, 4x Mini DisplayPort	
System Acoustics	<37 dB	
Software	Ubuntu Linux OS	
System Weight	91.0 lbs (43.1 kgs)	
Packaged System Weight	127.7 lbs (57.93 kgs)	
System Dimensions	Height: 25.1 in (639 mm) Width: 10.1 in (256 mm) Length: 20.4 in (518 mm)	
Operating Temperature Range	5–35 °C (41–95 °F)	



HPC システムズは NVIDIA Partner Network (NPN) 認定企業です。

NPNとは、NVIDIA社のSolution Provider向けのProgramです。Solution Providerは、VAR (Value Added Reseller: 付加価値再販業者) と呼ばれるパートナー企業が対象となります。当社はHigh Performance ComputingならびにDeep LearningのELITE Partnerに認定されています。



HPC システムズ株式会社
〒108-0022 東京都港区海岸 3-9-15 LOOP-X 8 階
TEL: 03-5446-5531 FAX: 03-5446-5550
Mail: hpcs_sales@hpc.co.jp

■ この内容は、2020年11月17日現在の内容です。

■ 価格、写真、仕様等は予告なく変更する場合があります。商品の色調は実際と異なる場合があります。

■ 社名、製品名などは、一般に各社の表示、商標または登録商標です。

■ NVIDIA、NVIDIAロゴ、CUDA、米国およびその他の国におけるNVIDIA Corporationの登録商標または商標です。