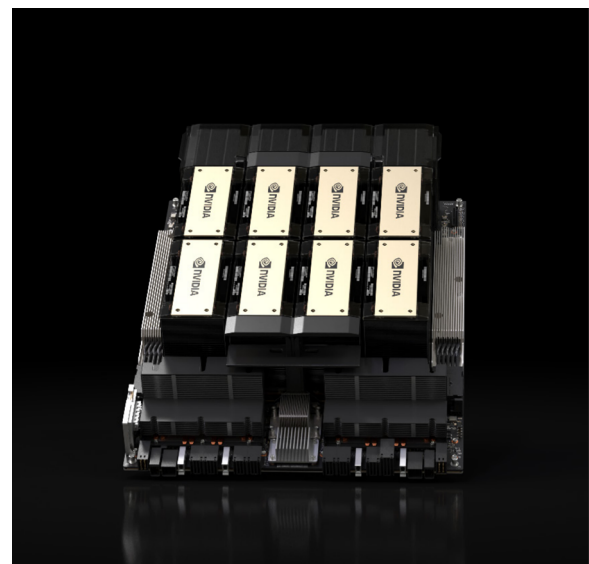


NVIDIA HGX H100

世界をリードするAIコンピューティングプラットフォーム



AIおよび高性能コンピューティング向けの専用設計

AI、複雑なシミュレーション、大規模なデータセットには、複数のGPUを超高速接続で結び、完全に加速されたソフトウェアスタックが必要です。NVIDIA HGX™ AIスーパーコンピューティングプラットフォームは、NVIDIA GPU、NVIDIA NVLink™、NVIDIAネットワークング、そして完全に最適化されたAIおよび高性能コンピューティング（HPC）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、迅速なインサイトの導出を可能にします。

比類なきエンドツーエンドの加速コンピューティングプラットフォーム

NVIDIA HGX H200は、H200 Tensor Core GPUと高速インターコネクトを組み合わせ、世界で最も強力なサーバーを形成します。最大8基のGPUを構成し、最大1.1テラバイト（TB）のGPUメモリと、毎秒38テラバイト（TB/s）の総メモリ帯域幅を提供し、前例のない加速性能を実現します。これにより、32ペタFLOPSという驚異的なパフォーマンスが組み合わさり、AIおよびHPC向けの世界で最も強力な加速スケールアップサーバープラットフォームが生まれます。

HGX H200とHGX H100の両方には、最大400ギガビット毎秒（Gb/s）の速度を可能にするNVIDIA Quantum-2 InfiniBandおよびSpectrum™-X Ethernetを活用した高度なネットワークングオプションが含まれています。また、HGX H200およびHGX H100には、クラウドネットワークング、コンポザブルストレージ、ゼロトラストセキュリティ、そしてハイパースケールAIクラウドにおけるGPUコンピュータの弾力性を可能にするNVIDIA® BlueField®-3データ処理ユニット（DPU）も搭載されています。

ディープラーニング推論：パフォーマンスと多様性

多種多様なビジネス課題において、多様なニューラルネットワークが活用される中で、汎用アクセラレーターは最高のパフォーマンスを提供するだけでなく、あらゆるアプリケーションでこれらのネットワークを加速するために必要な多様性も備えている必要があります。HGX H200とHGX H100はその両方で卓越しており、NVIDIAがAIパフォーマンスを大規模に提供するリーダーとなり、市場のリーダーシップを維持するために、世界で最も強力な加速AIコンピューティングプラットフォームを生み出しています。

技術仕様

	HGX H200 4-GPU	HGX H200 8-GPU	HGX H100 4-GPU	HGX H100 8-GPU
フォームファクター	4x NVIDIA H200 SXM	8x NVIDIA H200 SXM	4x NVIDIA H100 SXM	8x NVIDIA H100 SXM
FP8テンソルコア	16 PFLOPS	32 PFLOPS	16 PFLOPS	32 PFLOPS
INT8テンソルコア	16 POPS	32 POPS	16 POPS	32 POPS
FP16/BFLOAT16テンソルコア	8 PFLOPS	16 PFLOPS	8 PFLOPS	16 PFLOPS
TF32テンソルコア	4 PFLOPS	8 PFLOPS	4 PFLOPS	8 PFLOPS
FP32	270 TFLOPS	540 TFLOPS	270 TFLOPS	540 TFLOPS
FP64	140 TFLOPS	270 TFLOPS	140 TFLOPS	270 TFLOPS
FP64テンソルコア	270 TFLOPS	540 TFLOPS	270 TFLOPS	540 TFLOPS
メモリ	564GB HBM3e	1.1TB HBM3e	320GB HBM3	640GB HBM3
GPU総帯域幅	19GB/s	38GB/s	13GB/s	27GB/s
NVLink	Fourth generation	Fourth generation	Fourth generation	Fourth generation
NVSwitch™	N/A	Third generation	N/A	Third generation
NVSwitch GPU間帯域幅	N/A	900GB/s	N/A	900GB/s
総帯域幅	3.6TB/s	7.2TB/s	3.6TB/s	7.2TB/s

* With sparsity.

始める準備はできましたか？
NVIDIA HGXの詳細については、
www.nvidia.com/hgxをご覧ください。

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, BlueField, HGX, Magnum IO, NVLink, NVSwitch, and Spectrum are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 3361709. JUL24