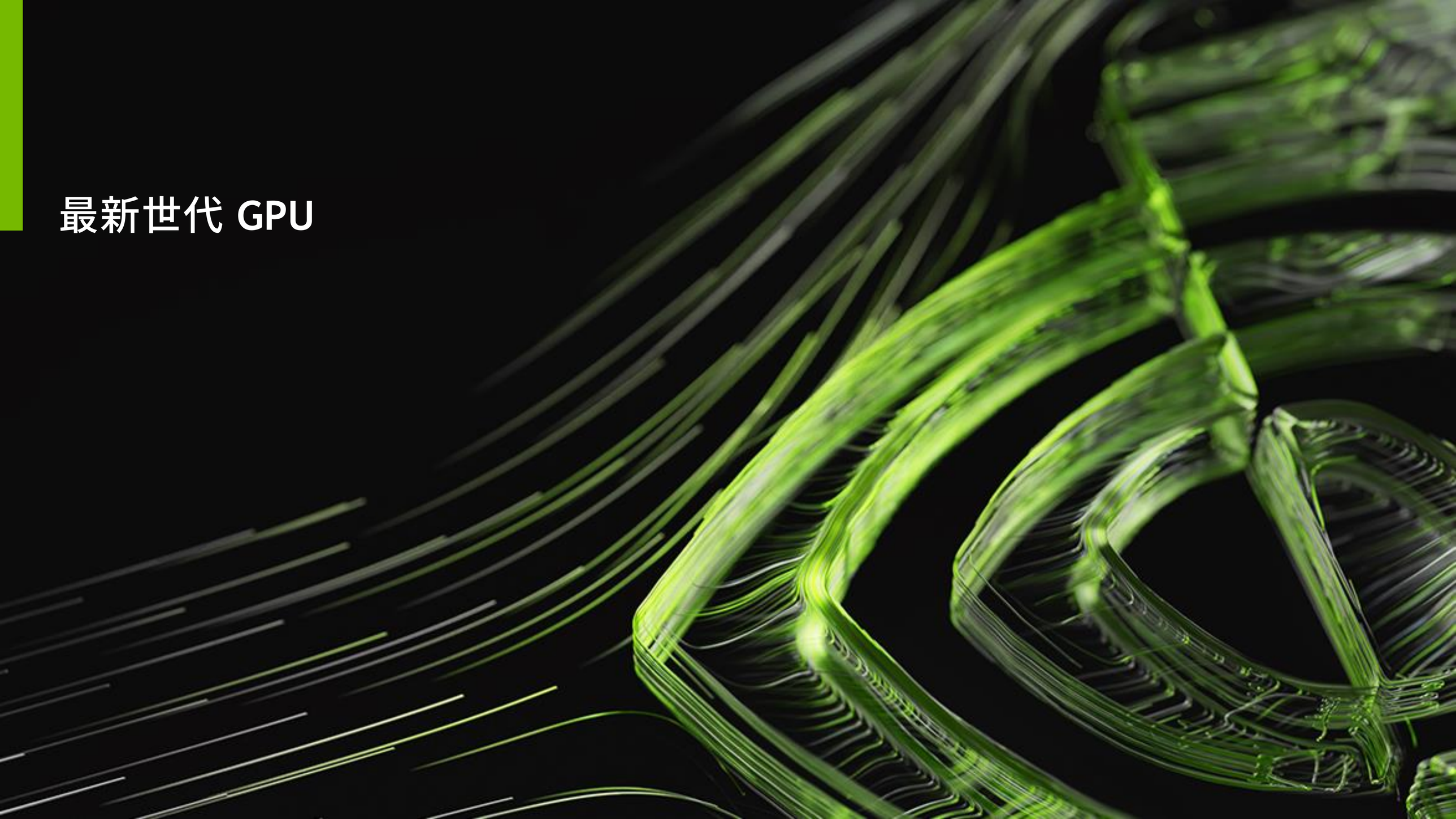
NVIDIA コンピューティング & ネットワーキング最新情報

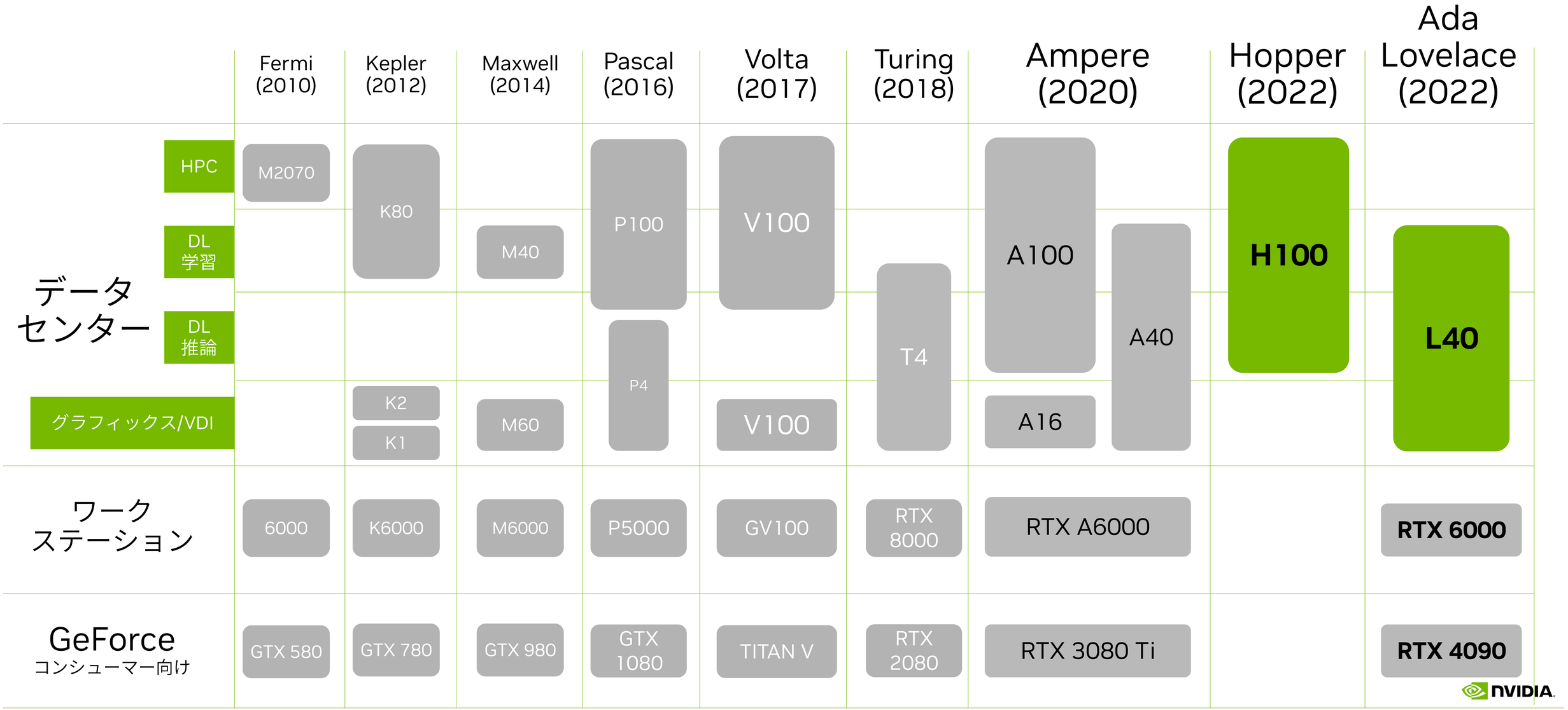
2022/12/5

佐々木 邦暢 (@_ksasaki)
エヌビディア合同会社

最新世代 GPU



NVIDIA GPU 製品のおおまかな一覧

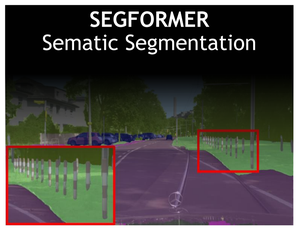


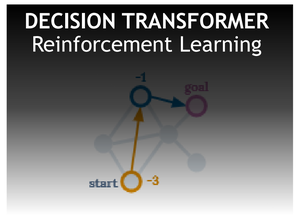
さらなるパフォーマンスとスケーラビリティを要求する次世代の AI

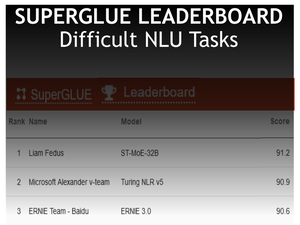
TRANSFORMERS TRANSFORMING AI

70%

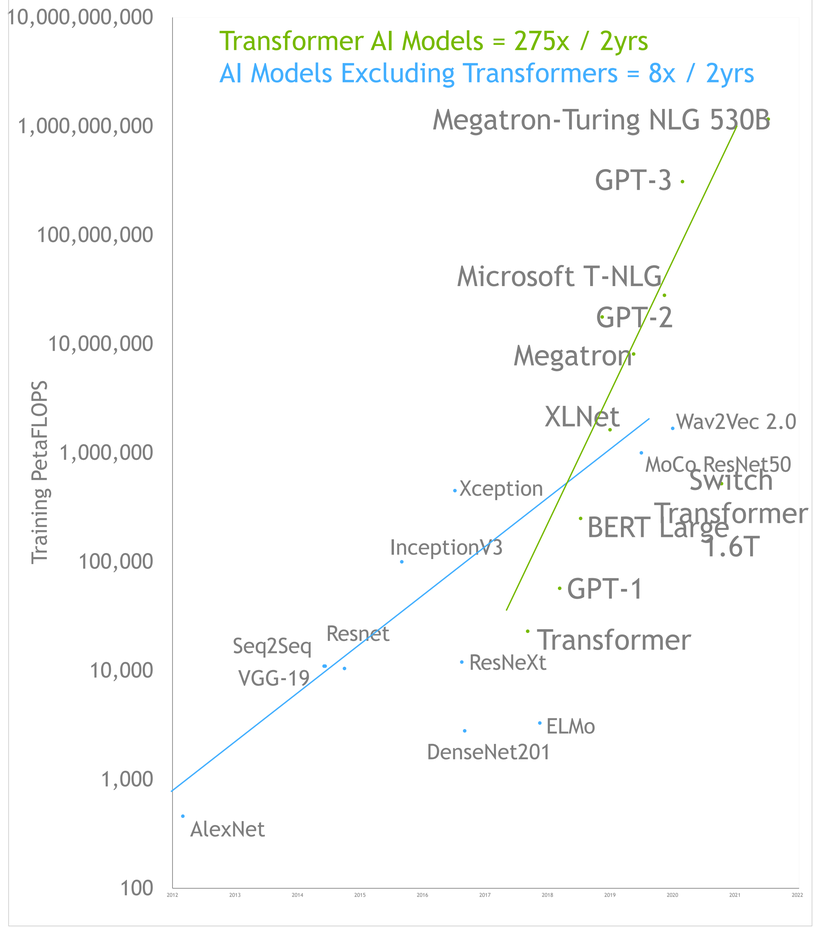
AI Papers
In last 2 years discuss Transformer Models



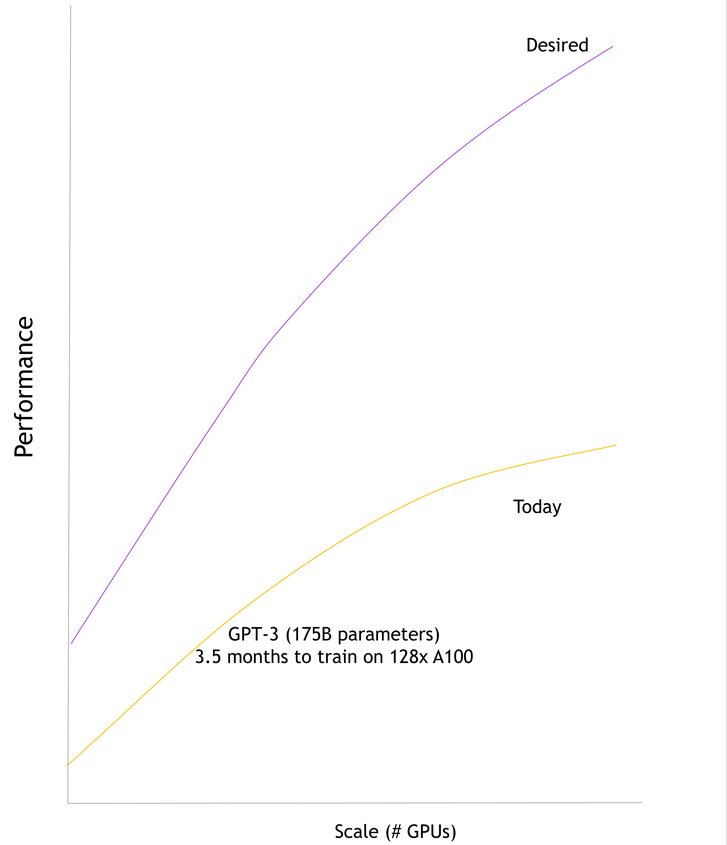




EXPLODING COMPUTATIONAL REQUIREMENTS



MONTHS TO TRAIN MODELS



MegaMolBART: https://catalog.ngc.nvidia.com/orgs/nvidia/teams/clara/models/megamolbart | SegFormer: https://arxiv.org/abs/2105.15203 | Decision Transformer: https://arxiv.org/pdf/2106.01345.pdf | SuperGLUE: https://super.gluebenchmark.com/leaderboard
Exploding Computational Requirements, source: NVIDIA Analysis and https://github.com/amirgholami/ai_and_memory_wall

NVIDIA

NVIDIA H100

世界の AI インフラを支える新しいエンジン

最高の **AI/HPC** 性能

4PF FP8 (6X)| 2PF FP16 (3X)| 1PF TF32 (3X)| 67TF FP64 (3X)

3.35TB/s (1.5X), 80GB HBM3 memory

TRANSFORMER モデルへの最適化

6X faster on largest transformer models

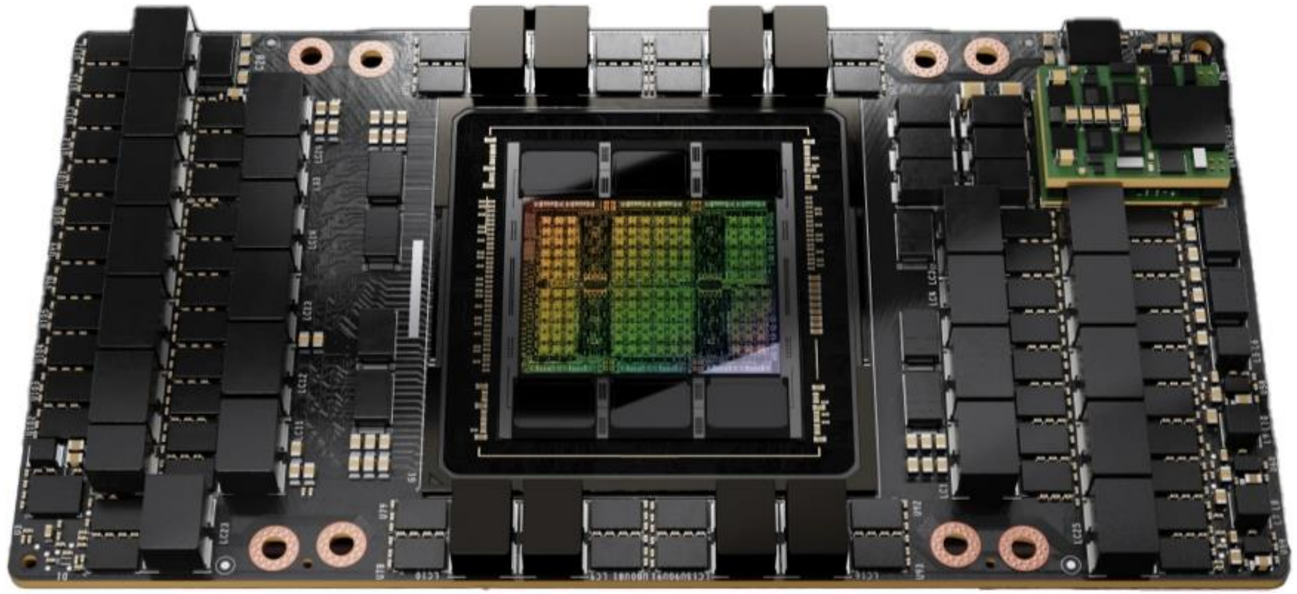
高い稼働率とセキュリティ

7 Fully isolated & secured instances, guaranteed QoS

2nd Gen MIG | Confidential Computing

史上最速でスケーラブルなインターコネクト

900 GB/s GPU-2-GPU connectivity (1.5X) | 128GB/s PCI Gen5



Custom TSMC 4N Process | 4.9 TB/s Total External B/W

https://resources.nvidia.com/en-us-tensor-core/nvidia-tensor-core-gpu-datasheet

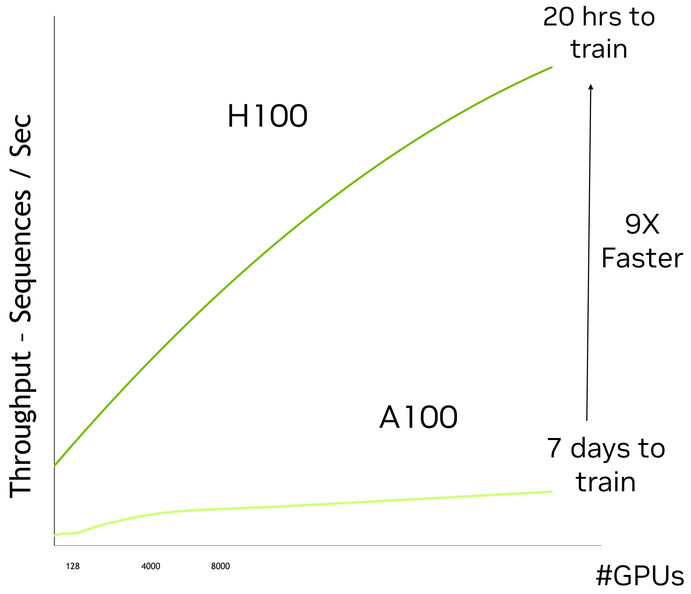
FP8, FP16, TF32 performance include sparsity. X-factor compared to A100

H100 の大幅な性能向上

次世代のブレークスルーを実現するパフォーマンスとスケーラビリティ

TRAINING

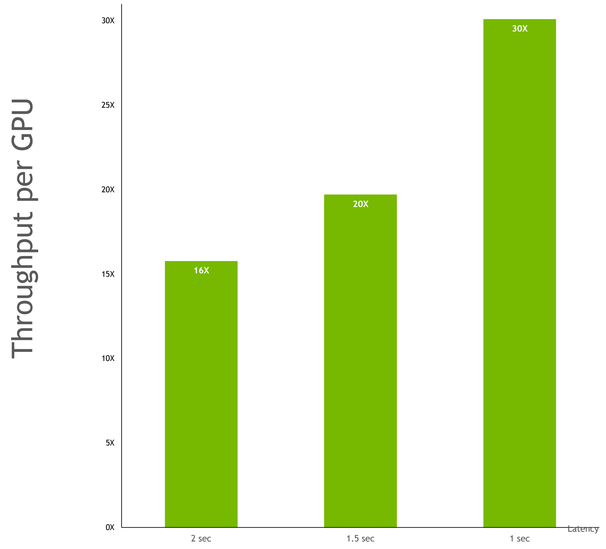
Up to 9X More Throughput



Mixture of Experts (395B) Training vs A100

INFERENCE

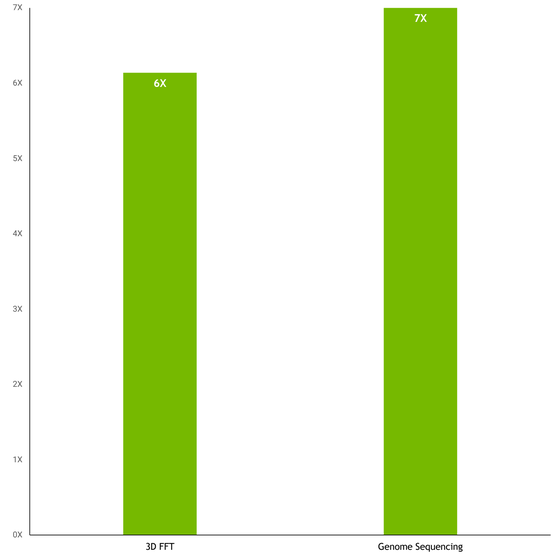
Up to 30X More Throughput



Megatron 530B Throughput vs A100

HPC

Up to 7X Higher Performance

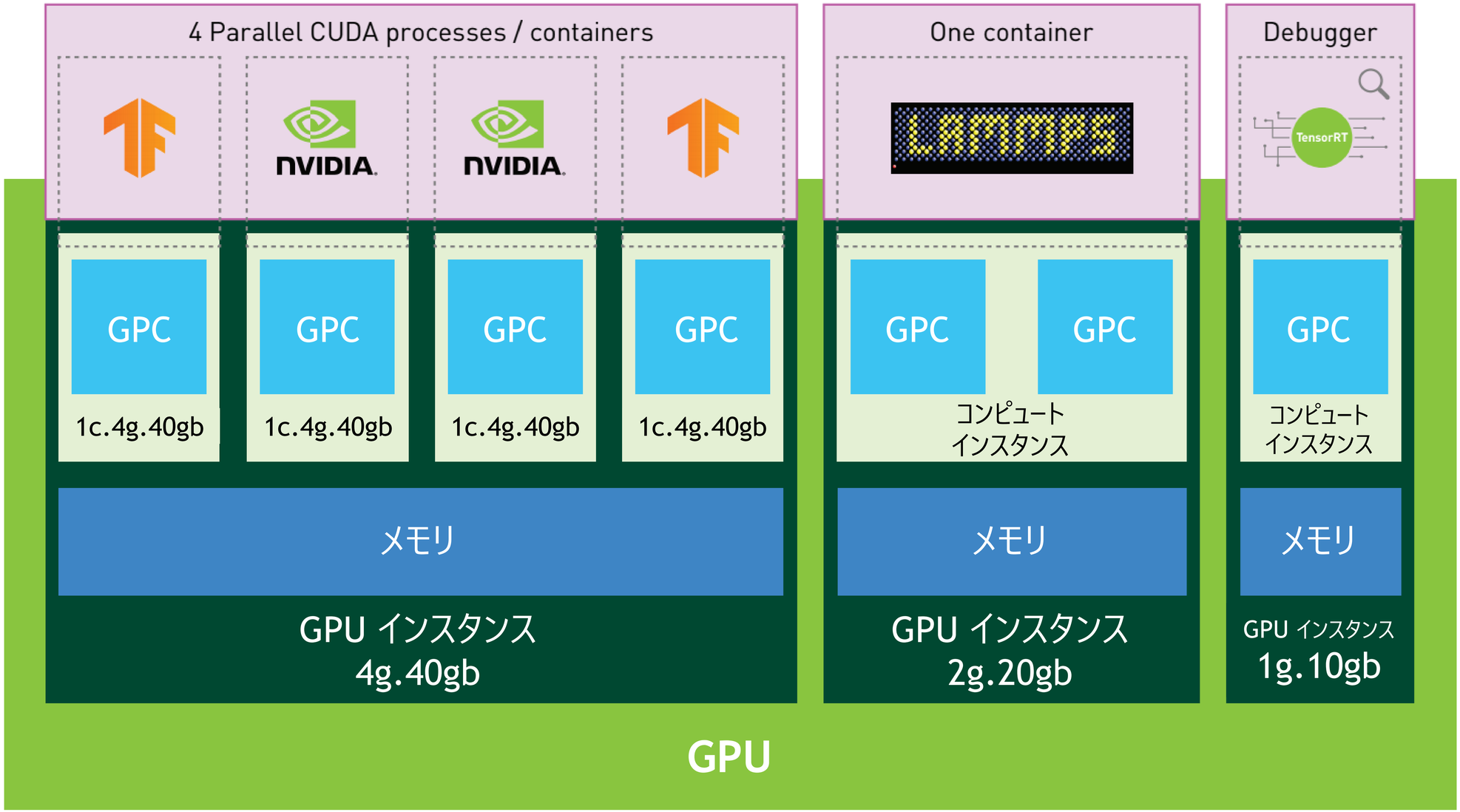


HPC App Performance vs A100

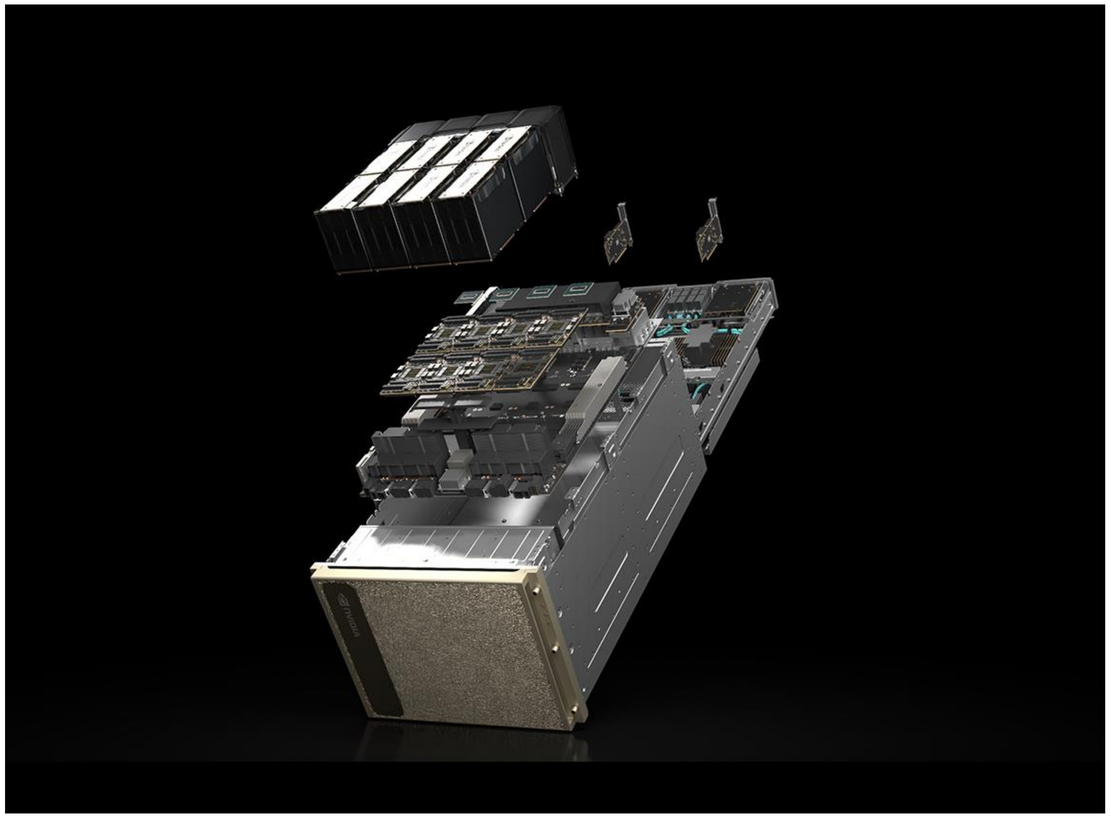
Projected performance subject to change
Training Mixture of Experts (MoE) Transformer Switch-XXL variant with 395B parameters on 1T token dataset | Inference on Megatron 530B parameter model based chatbot for input sequence length=128, output sequence length=20 | HPC 3D FFT (4K^3) throughput | A100 cluster: HDR IB network | H100 cluster: NVLink Switch System, NDR IB
HPC Genome Sequencing (Smith-Waterman) | 1 A100 | 1 H100

MULTI-INSTANCE GPU (MIG)

1 基の H100 GPU を最大 7 基に分割し有効活用



DGX H100 Hardware Feature Summary



NVIDIA DGX H100

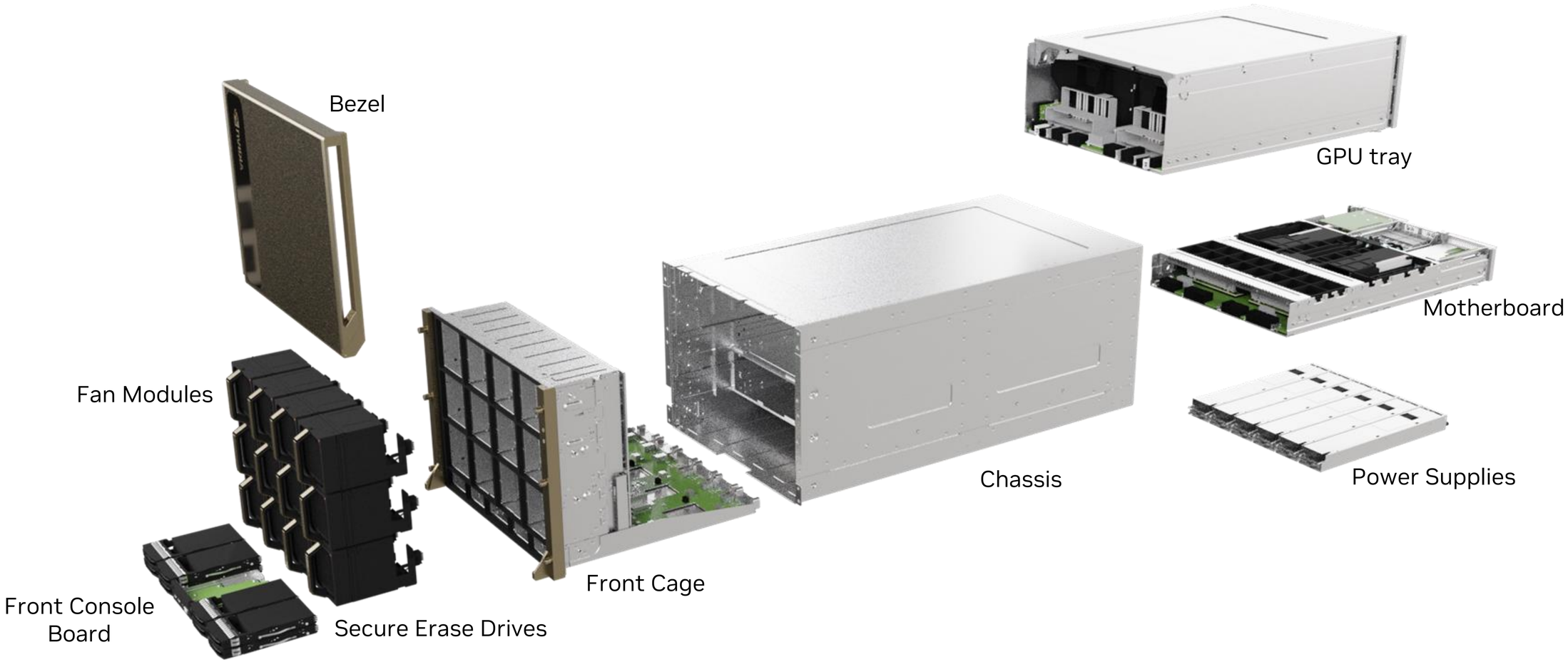
The world's first AI system with the NVIDIA H100 Tensor Core GPU

GPU	8 x NVIDIA H100 Tensor Core GPUs
GPU memory	640 GB total
Performance	32 petaFLOPs FP8
NVswitches	4
CPUs	2 x x86
System Memory	2 TB
Networking connectivity and speed	4 x OSFP ports provide connectivity to 8 x NVIDIA ConnectX-7s 8 x 400Gb/s InfiniBand/Ethernet* 2 x dual-port NVIDIA ConnectX-7 1 x 400Gb/s InfiniBand/Ethernet and 1 x 200Gb/s InfiniBand/Ethernet*
Cache Storage	8 x U.2 3.84TB NVMe Secure Erase Drives
Boot Storage	2 x 1.92TB M.2 NVMe (software encryptable)
Host Management	On-board 10Gb/s RJ-45 Ethernet Optional 50Gb/s QSFP Ethernet NIC
Remote System Management	Baseboard Management Controller (BMC) 1Gb/s RJ-45 network connectivity Remote Keyboard, Video, Mouse (KVM) Remote Storage RedFish, IPMI management
Operating System	DGX OS based on Ubuntu Additional support for Ubuntu and Red Hat Enterprise Linux

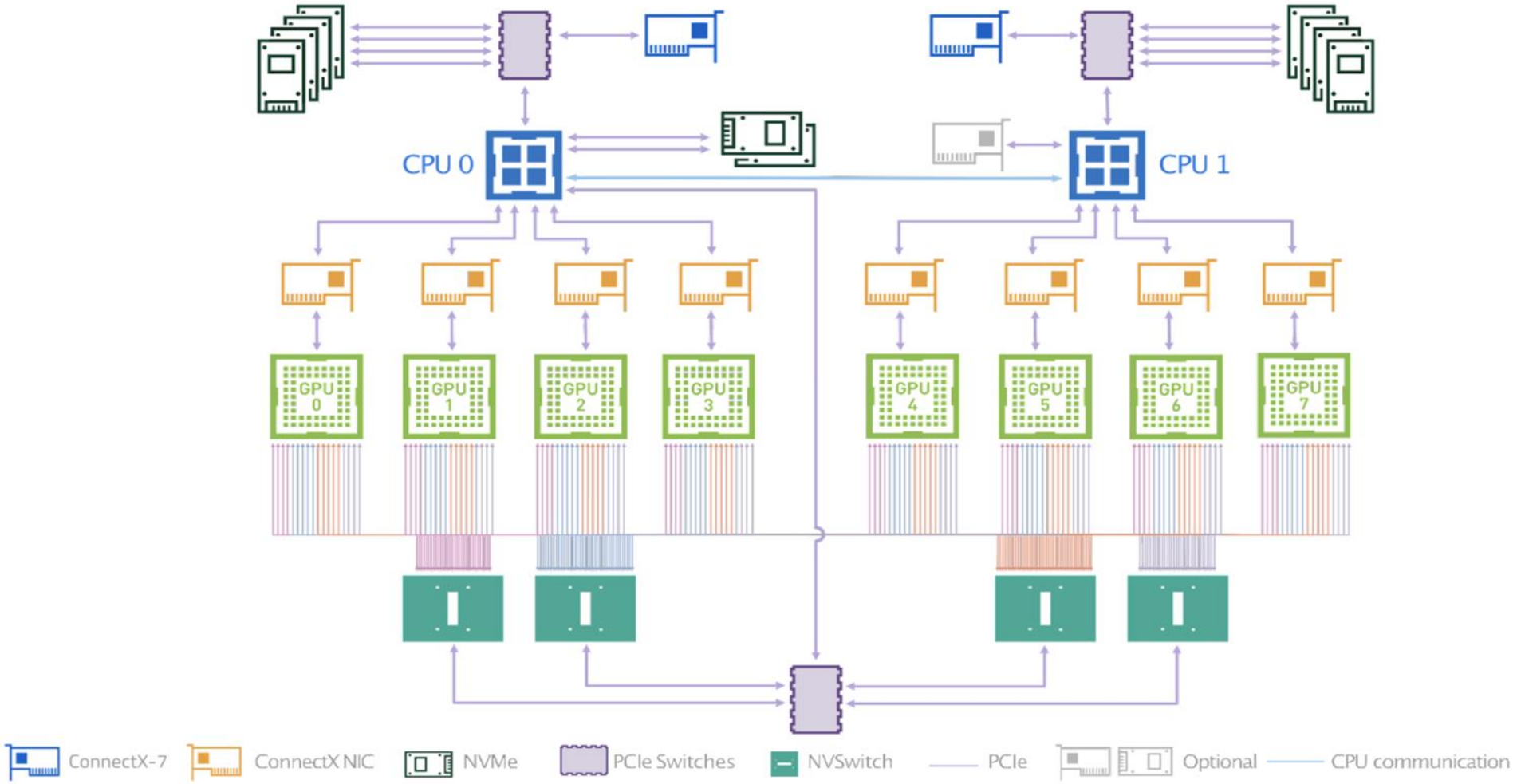
* Subject to Change

NVIDIA

DGX H100 Exploded View

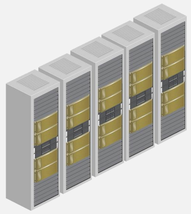


DGX H100 Block Diagram



Understanding DGX BasePOD vs DGX SuperPOD

NVIDIA DGX BasePOD



Scalable, foundational architecture

- Flexible reference architectures
- Powered by Base Command
- Sold through channel
- Partner professional services
- Validated against key **benchmarks**
- Foundation for **partner branded offerings**

Customer needs:

- Choice of flexible vs performance optimized designs
- Inclusion of non-SuperPOD certified storage

NVIDIA DGX SuperPOD

Digital twin of NVIDIA's infrastructure

- Turnkey data center **product**
- Powered by Base Command
- Sold through channel, but with mandatory…
  …white-glove NVIDIA PS – *order > install > commission*
- **Certified performance** for the most **complex workloads**
- **No customization, no partner re-branding**

Customer needs:

- a replica of NVIDIA's infrastructure
- a turnkey deployment
- full-stack support of the entire deployment

NVIDIA H100 PCIE

かつてないパフォーマンス、スケーラビリティ、
セキュリティをメインストリームサーバーで

高い **AI/HPC** 性能

3.2PF FP8 (5X) | 1.6PF FP16 (2.5X) | 800TF TF32 (2.5X) | 48TF FP64 (2.5X)
6X faster Dynamic Programming with DPX Instructions
2TB/s , 80GB HBM2e memory

最高のエネルギー効率

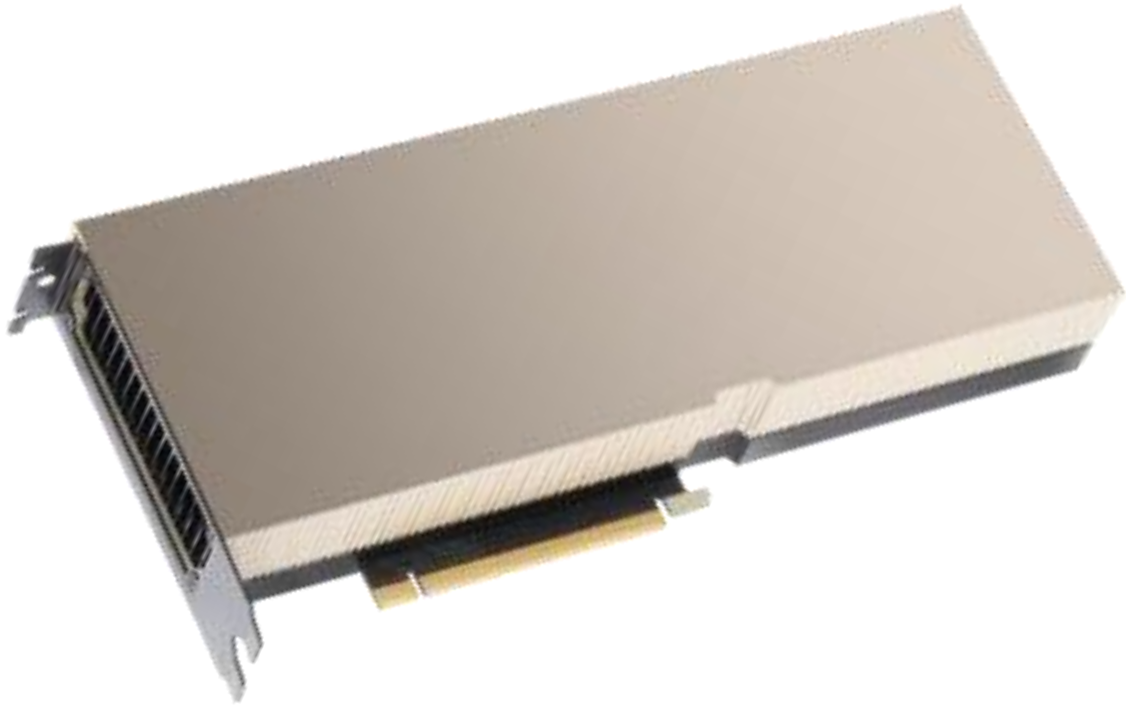
Configurable TDP - 200W to 350W
2 Slot FHFL mainstream form factor

高い稼働率とセキュリティ

7 Fully isolated & secured instances, guaranteed QoS
2nd Gen MIG | Confidential Computing

高速な接続

128GB/s PCI Gen5
600 GB/s GPU-2-GPU connectivity (5X PCIe Gen5)
up to 2 GPUs with NVLink Bridge



FP8, FP16, TF32 performance include sparsity. X-factor compared to A100

NVIDIA H100 GPU + NVIDIA AI ENTERPRISE

NVIDIA AI ENTERPRISE
5 年分が含まれます

Compute Accelerator for AI, HPC, Data Processing

Fastest Mainstream Compute
Ready for Enterprise AI

H100 PCIe
350W | 80GB
2-Slot FHFL

- PCIe 版 H100 を搭載する NVIDIA Certified サーバーには、H100 GPU 1基ごとに NVIDIA AI Enterprise サブスクリプションが付属します
- Business Standard Support が含まれます。
- こちらで有効化: nvidia.com/activate-h100

*NVIDIA AI Enterprise **not included** with HGX H100 (SXM form factor)*

H100 搭載システムが GREEN500 で首位を獲得

Green500 Data

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	405	**Henri** - Lenovo ThinkSystem SR670 V2, Intel Xeon Platinum 8362 2800Mhz (32C), NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States	5,920	2.04	31	65.091
2	32	**Frontier TDS** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	120,832	19.20	309	62.684
3	11	**Adastra** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Grand Equipement National de Calcul Intensif - Centre Informatique National de l'Enseignement Suprieur (GENCI-CINES) France	319,072	46.10	921	58.021

Hopperの Green500 デビュー

NVIDIA のテクノロジは、最新の Green500 リストの上位 30 のシステムのうち 23 に既に搭載されています。

例えば、ニューヨークのフラットアイアン研究所は、NVIDIA Hopper H100 GPU を搭載し、Lenovo によって構築された空冷式の ThinkSystem を使用したことにより、最も効率的なスーパーコンピューターとして Green500 リストのトップになりました。

Green500 によると、Henri と名付けられたこのスーパーコンピューターは、1 ワットあたり 650 億回の倍精度浮動小数点演算を行い、計算天体物理学、生物学、数学、神経科学、量子物理学の問題に取り組むために使用されます。

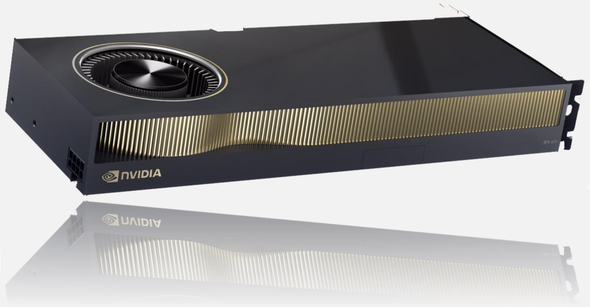
NVIDIA Hopper GPU アーキテクチャに基づく NVIDIA H100 Tensor コア GPU は、前世代の A100 GPU と比較して、最大 6 倍の AI パフォーマンスと最大 3 倍の HPC パフォーマンスを備え、驚異的な効率で動作するように設計されています。その第 2 世代である Multi-Instance GPU テクノロジは、GPU をより小さなコンピューティング ユニットに分割できるため、データセンターのユーザーが利用できる GPU クライアントの数が劇的に増加します。

https://blogs.nvidia.co.jp/2022/11/16/green/

NVIDIA

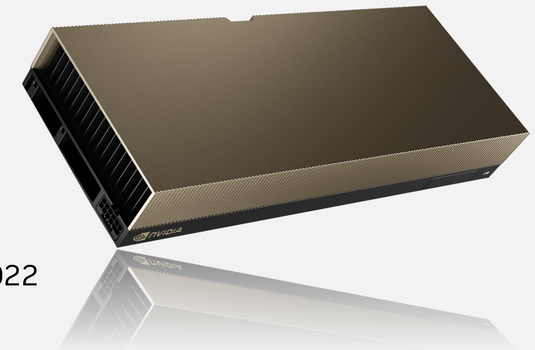
NVIDIA RTX 6000 Ada 世代
Ada for the Enterprise Desktop

- Up to 2x Faster than RTX A6000
- 48GB GDDR6 w/ECC Memory
- 4x DP1.4 Display Outputs
- PCIe Gen 4
- vGPU Support
- Available from channel partners starting in December 2022, OEM partners early 2023



NVIDIA L40
Ada for the Data Center

- Up to 2x Faster than A40
- 48GB GDDR6 w/ECC Memory
- 4x DP1.4 Display Outputs
- PCIe Gen 4
- vGPU Support
- Availability starting in December 2022



*Specifications, availability dates subject to change. vGPU support in Q1 2023

Ada Lovelace アーキテクチャ
エンタープライズ向け 次世代 RTX

NVIDIA GRACE CPU

Grace CPU スーパーチップ

AI と HPC インフラのための CPU

最高の CPU 性能

Superchip Design with 144 high-performance Neoverse V2 Cores
Estimated Specrate2017_int_base of over 740

最高のメモリバンド幅

World’s first LPDDR5x memory with ECC, 1TB/s Memory Bandwidth

最高のエネルギー効率

2X Perf/Watt, CPU Cores + Memory in 500W

2倍の搭載密度

2x density of DIMM based designs

NVIDIA のコンピューティングスタックが完全に動作

RTX, HPC, AI, Omniverse

提供開始: 2023 年前半

Grace Hopper スーパーチップ

大規模 AI と HPC を加速

最高の CPU+GPU 性能
Grace CPU plus Hopper GPU Acceleration

GPU から最大 600GB のメモリを利用可能
Enables Giant AI Models for Training & Inference

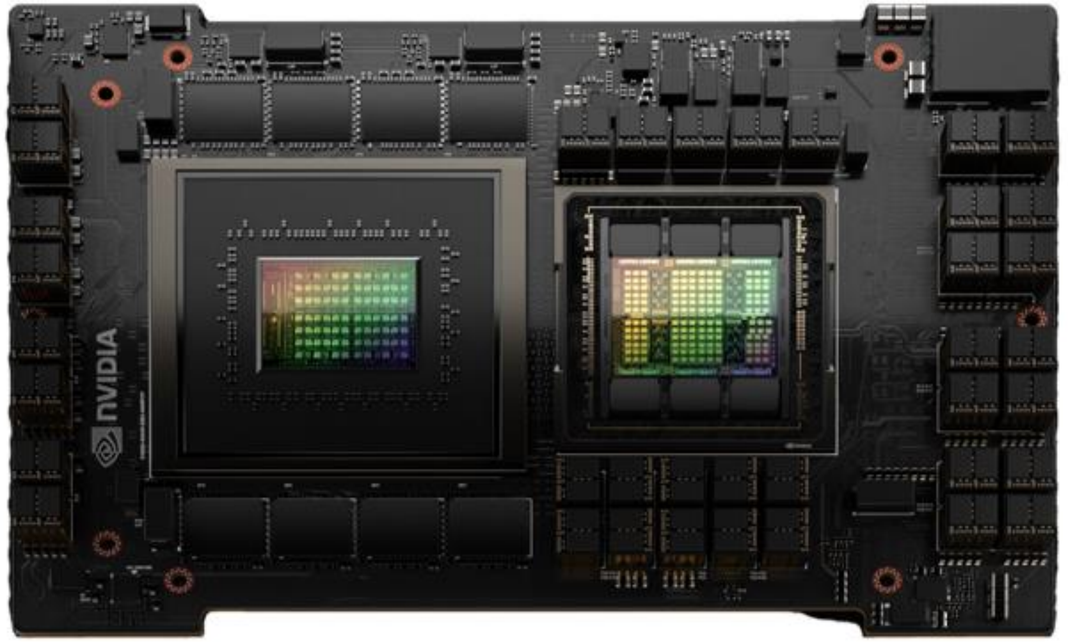
最高のメモリバンド幅 3.5TB/s
LPDDR5x and HBM3

新しい 900GB/S コヒーレント インタフェース
NVLink-C2C connecting Grace to Hopper

GPU とシステムメモリ間の帯域が15 倍
NVLink-C2C vs PCIe

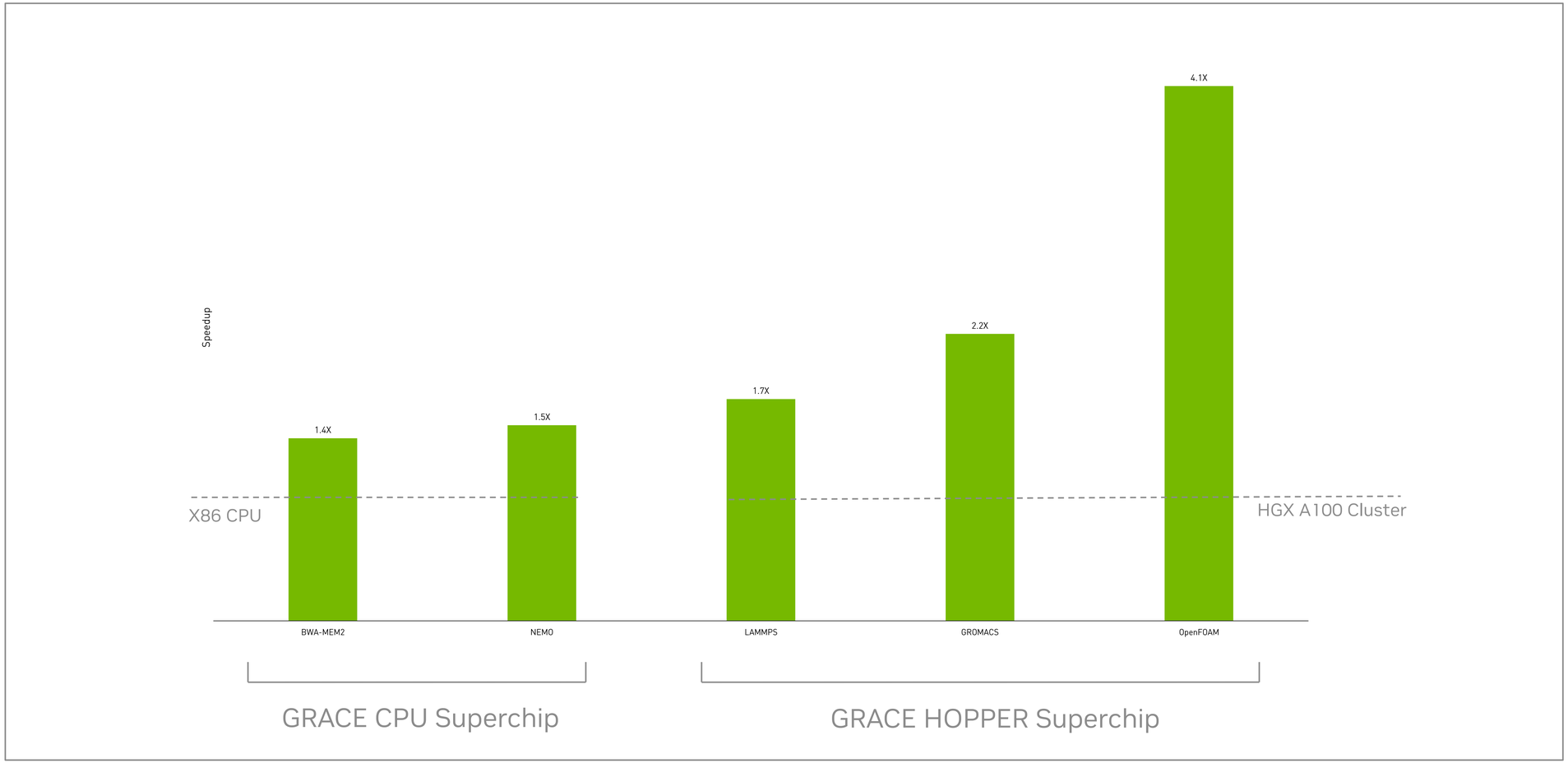
NVIDIA のコンピューティングスタックが完全に動作
RTX, HPC, AI, Omniverse

提供開始: 2023 年前半



Grace CPU および Grace Hopper スーパーチップ

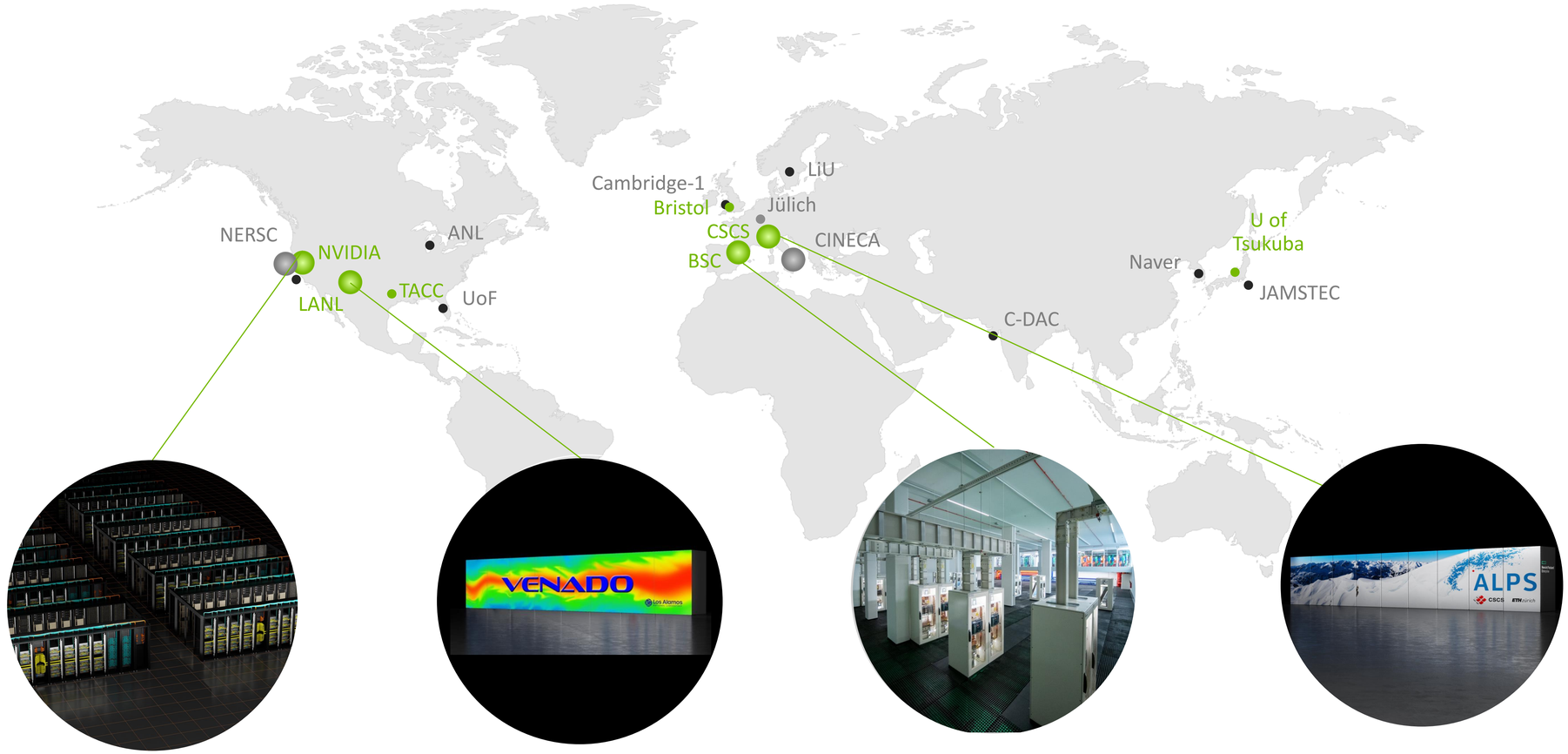
主要な HPC ワークロードを最大 4 倍高速化



Grace Hopper Superchip and Grace CPU Superchip Pre-silicon ESTIMATES ONLY: | Performance comparisons: (Left) AMD Milan CPU baseline| BWA-MEM2 (Single Socket/Chip) | NEMO (Dual Socket/Chip) | (Right) Cluster Comparison: 32 node HGX (4x A100 SXM4 80 GB & 4x HDR InfiniBand) baseline and 128x Single Grace-Hopper 66T / 80 GB nodes | LAMMPS datasets ReaxFF/c & SNAP (geomean) | OpenFOAM dataset Lid driven cavity | GROMACS dataset STMV 1066628 atoms

NVIDIA

エクサスケール AI スパコンが HPC を変革



NVIDIA (EOS)
18 EFLOPS AI Perf
Quantum-2 400Gb/s InfiniBand

LANL (Venado)
1 EFLOPS AI Perf

BSC (MareNostrum 5)
18 EFLOPS AI Perf
Quantum-2 400Gb/s InfiniBand

CSCS (ALPS)
20 EFLOPS AI Perf

Hopper + X86 Systems: University of Tsukuba, Bristol, and TACC | Grace Hopper / Grace CPU Superchips Systems: BSC, CSCS, and LANL

NVIDIA

NVIDIA ネットワーキング

AIアプリケーション
フレームワーク

プラットフォーム

NVIDIA HPC

NVIDIA AI

NVIDIA Omniverse

アクセラレーションライブラリ

GPU

CPU

Infrastructure Workload
- Communications
- Storage
- Security and Isolation

NIC

AIアプリケーション
フレームワーク

PLATFORMS

NVIDIA HPC

NVIDIA AI

NVIDIA Omniverse

アクセラレーションライブラリ

GPU

GPU

FREED

CPU

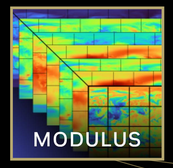
CPU

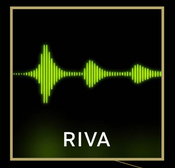
Cloud Native
Supercomputing

FREED

BlueField DPU

NIC





AI APPLICATION FRAMEWORK

PLATFORMS

NVIDIA HPC

NVIDIA AI

NVIDIA Omniverse

ACCELERATION LIBRARIES

GPU

In-Network Computing

コンピュテーショナルストレージ

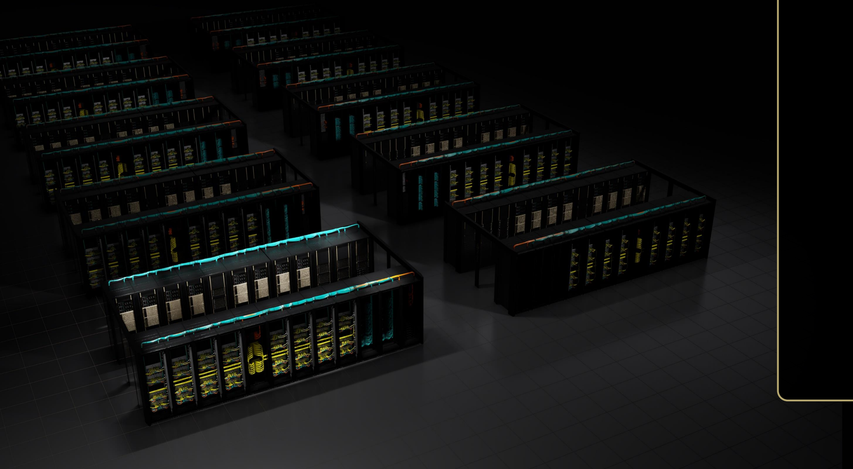
CPU

パフォーマンスの分離

強化されたテレメトリ

DPU

ゼロトラストセキュリティ



DPU で実現する新たな
データセンタープラットフォーム インフラストラクチャー

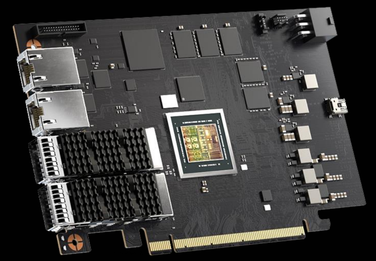
ソフトウェアデファインドネットワーク

ソフトウェアデファインドセキュリティ

ソフトウェアデファインドストレージ

Infrastructure Services

	BlueField-2	BlueField-3
ネットワーク バンド幅	200Gb/s	400Gb/s
RDMA メッセージ レート	215Mpps	370Mpps
Compute	SPECINT2K17: 9.8	SPECINT2K17: 42
メモリバンド幅	17GB/s	80GB/s



BlueField Infrastructure Compute Platform

プラットフォーム インフラストラクチャ

ストレージ

サイバー セキュリティ

ネットワーキング

DOCA

Orchestration

Security

Networking

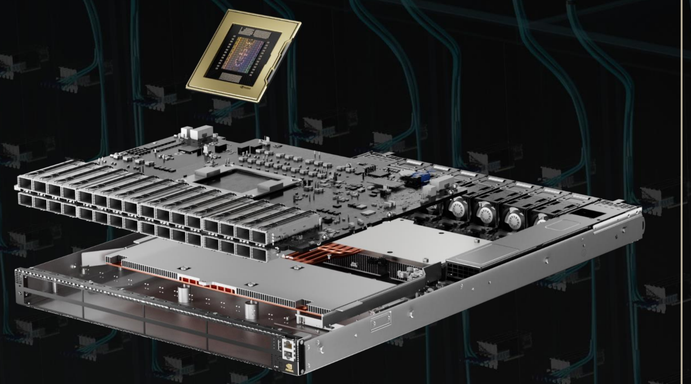
Storage

Management

Acceleration Libraries

Telemetry

NVIDIA Quantum-2 InfiniBand platform In-Network Computing



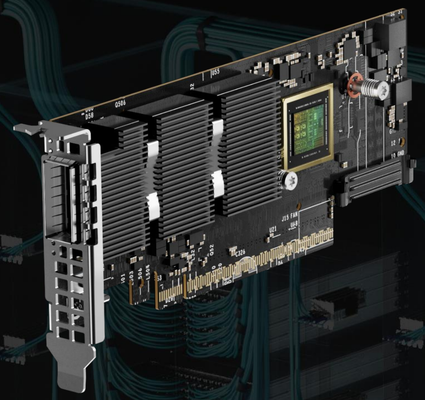
QUANTUM-2 SWITCH

64-Ports of 400 Gbps or 128-Ports of 200 Gbps

SHARPv3 Small Message Data Reductions

SHARPv3 Large Message Data Reductions

32X More AI Acceleration Engines



CONNECTX-7 INFINIBAND

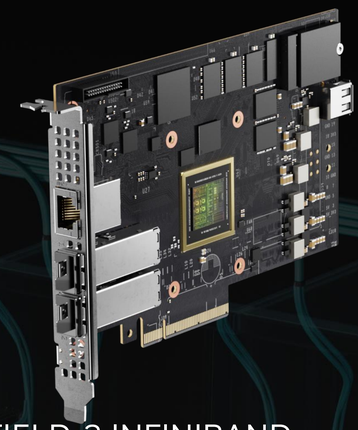
16 Core / 256 Threads Datapath Accelerator

Full Transport Offload and Telemetry

Hardware-Based RDMA / GPUDirect

MPI Tag Matching and All-to-All

PCIe Gen 5



BLUEFIELD-3 INFINIBAND

16 Arm 64-Bit Cores

16 Core / 256 Threads Datapath Accelerator

Full Transport Offload and Telemetry

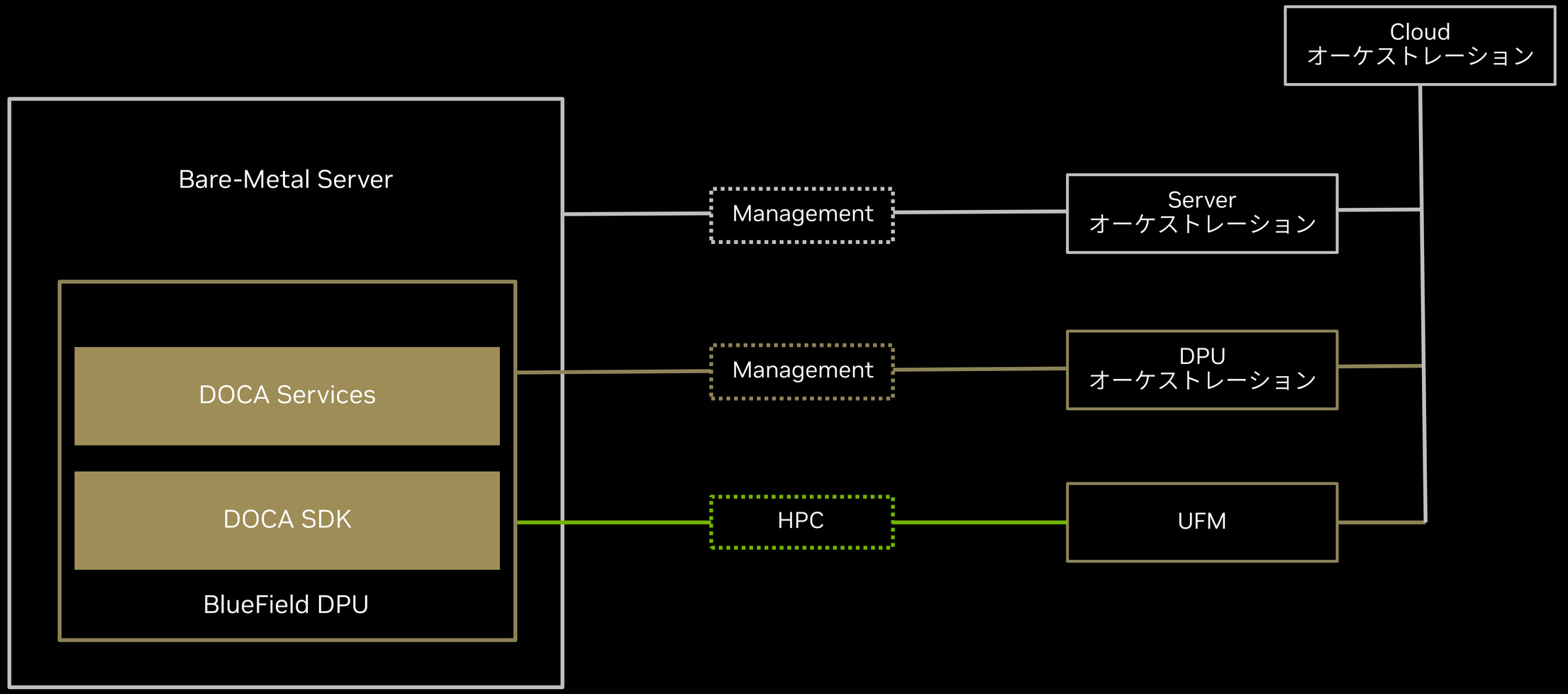
Hardware-Based RDMA / GPUDirect

Computational Storage

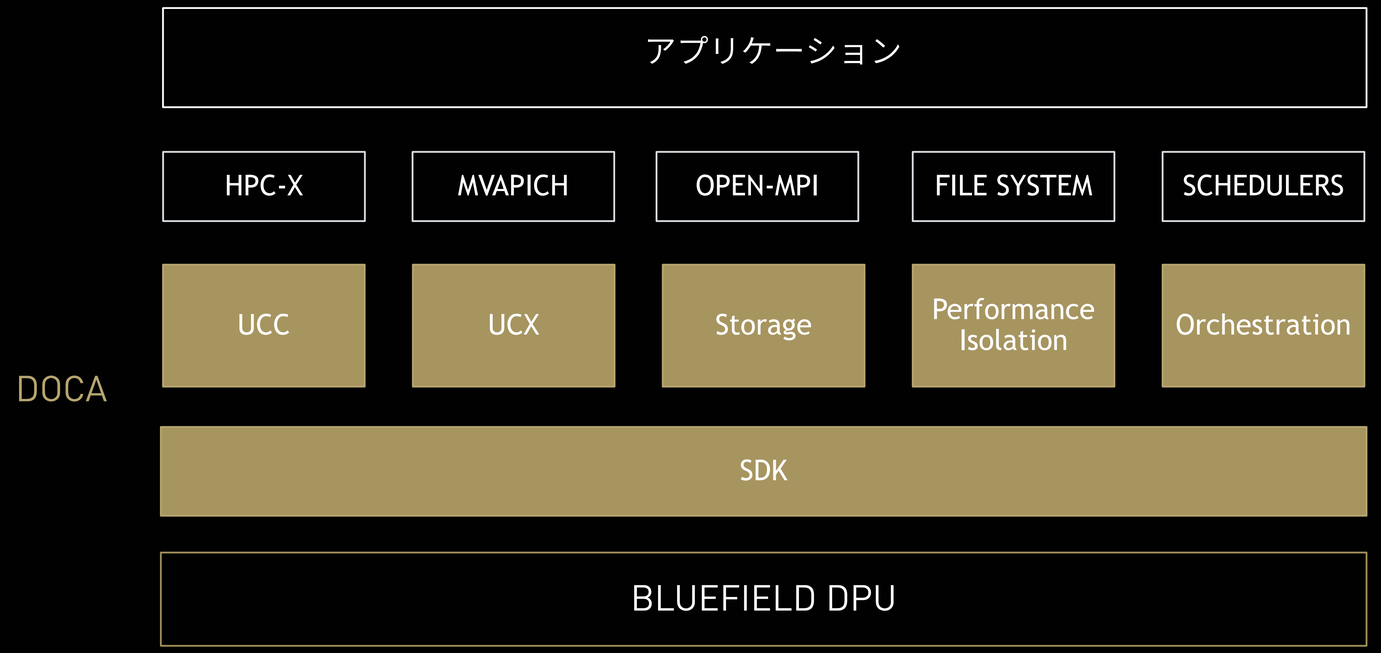
Security Engines

MPI Tag Matching and All-to-All

クラウドネイティブスーパーコンピューティング構成



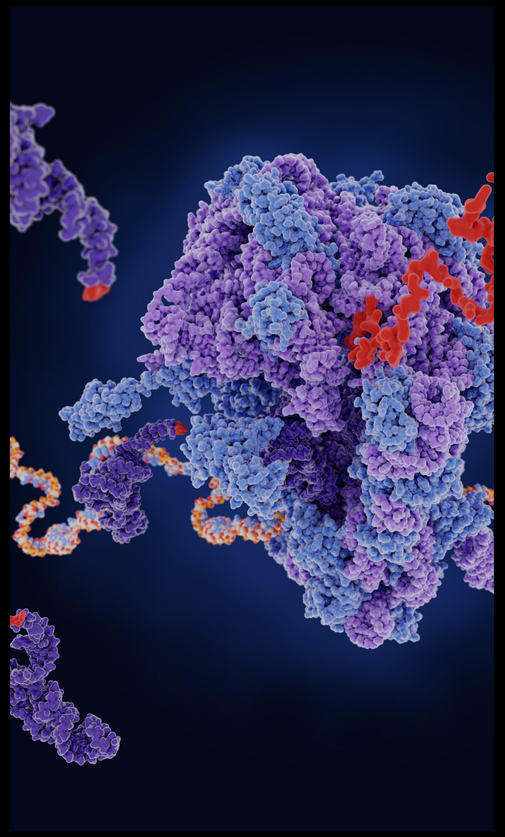
DOCA サービスによる HPC アプリケーションの高速化



より高いアプリケーションパフォーマンスの実現

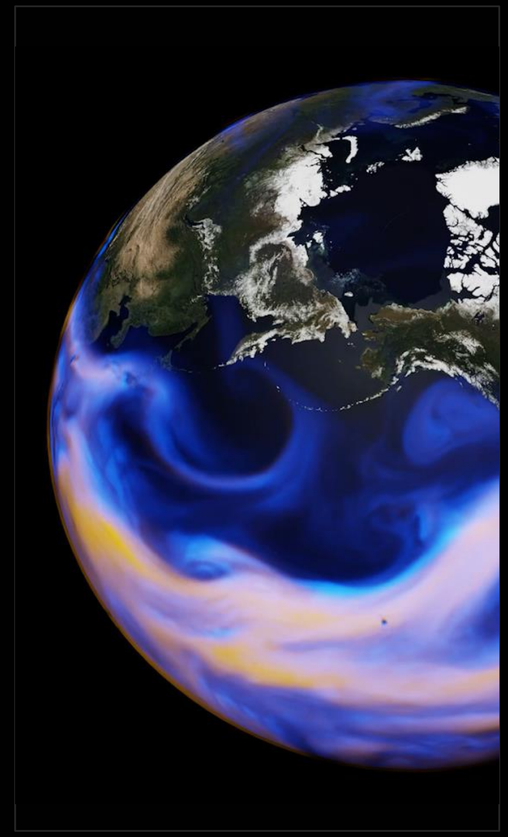
With BlueField DPUと Quantum In-Network Computing

分子動力学



数理モデリング

気象予測



Barcelona Supercomputing Center

Durham University

Georgia Institute of Technology

Japan Meteorological Agency

Los Alamos National Laboratory

Ohio State University

Sandia National Laboratory

Technical University Munich

University College London

NVIDIA コンピューティング プラットフォーム

ビジネス
アプリケーション

カスタマー エンゲージメント | 医療診断 | 異常検知 | 品質保証 | 自動化 | マーケティング | シミュレーション | ++

NGC

ソフトウェア ハブ

HELM

コンテナイメージ

学習済みモデル

ソフトウェア開発キット

アプリケーション
フレームワーク

スマートシティ **Metropolis** | 対話型 **AI** **Riva** | 自動運転 **Drive** | レコメンデーションシステム **Merlin** | ヘルスケア **Clara** | ++ ...

開発者
ツールキット

ML & DATA ANALYTICS: RAPIDS, dmlc XGBoost

AI TRAINING & INFERENCE: TensorFlow, TensorRT, PYTORCH, mxnet

HIGH PERFORMANCE COMPUTING: *NVIDIA HPC SDK*

RENDERING & VISUALIZATION: *IndeX*, *OptiX*

高速化
ライブラリ

COMPUTE: *CUDA-X*

NETWORKING, STORAGE & SECURITY: *DOCA*, *MAGNUM IO*

運用

VIRTUAL GPU SOFTWARE

TRITON INFERENCE SERVER

FLEET COMMAND

NVIDIA GPU Operator

Red Hat

vmware

NVIDIA CERTIFIED

VALIDATED SOLUTIONS

サーバー及びクラウド

DGX
HGX

ハイエンド

EGX

メインストリーム
とエッジ

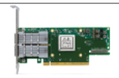
クラウドの
インスタンス

ハードウェア技術

GPU

NVSwitch

BlueField DPU



SMART NIC

Mellanox Switch

監視

DCGM | Grafana | Prometheus | UFM

NVIDIA


NVIDIA