



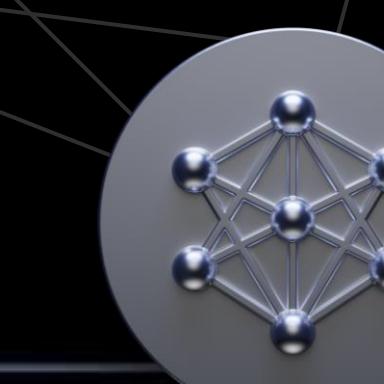
Entering A New Frontier of AI Networking Innovation

Gilad Shainer | GTC 2024

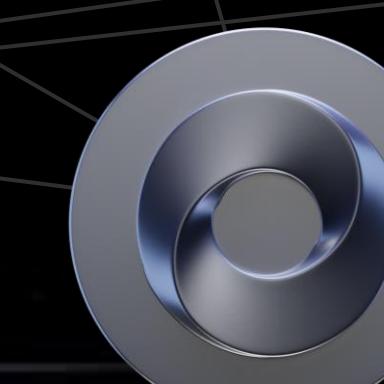
APPLICATION
FRAMEWORKS



PLATFORM

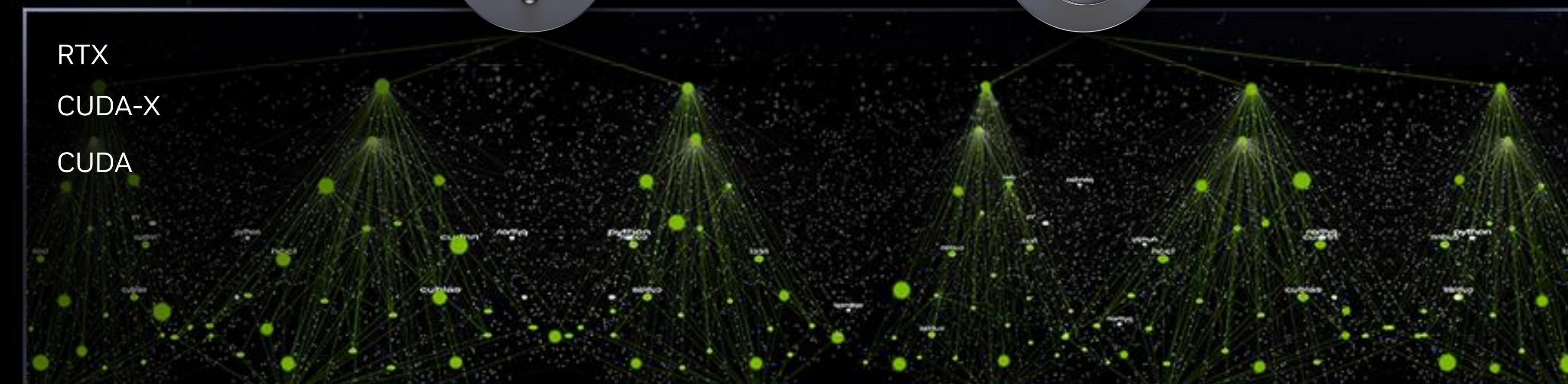


NVIDIA
AI



NVIDIA
OMNIVERSE

ACCELERATION
LIBRARIES



SYSTEM
SOFTWARE

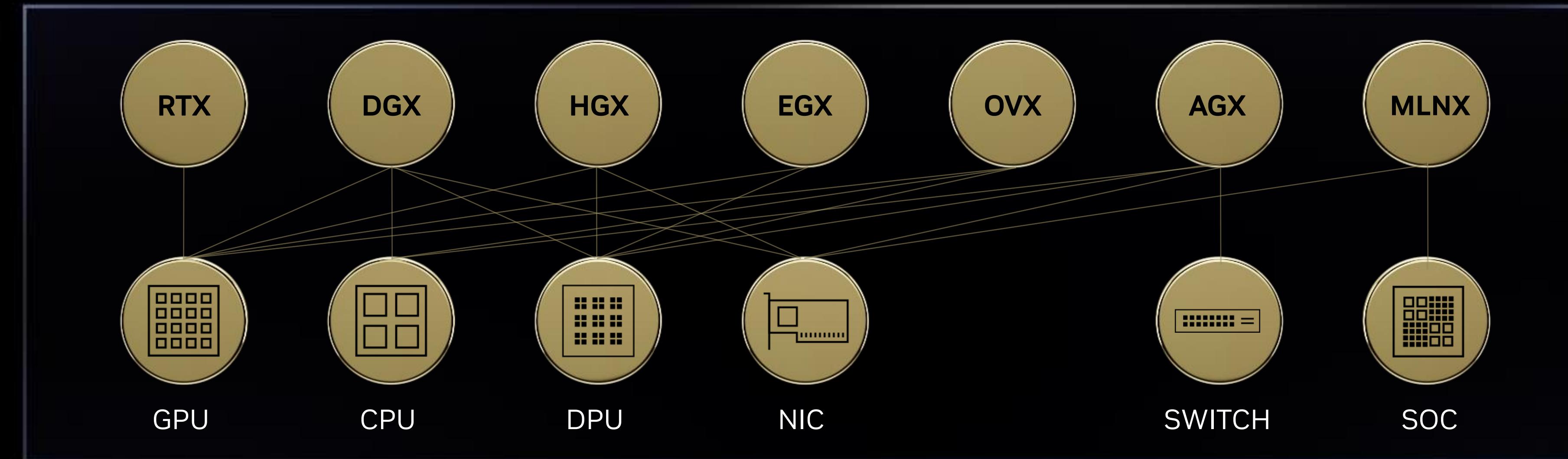
Magnum IO

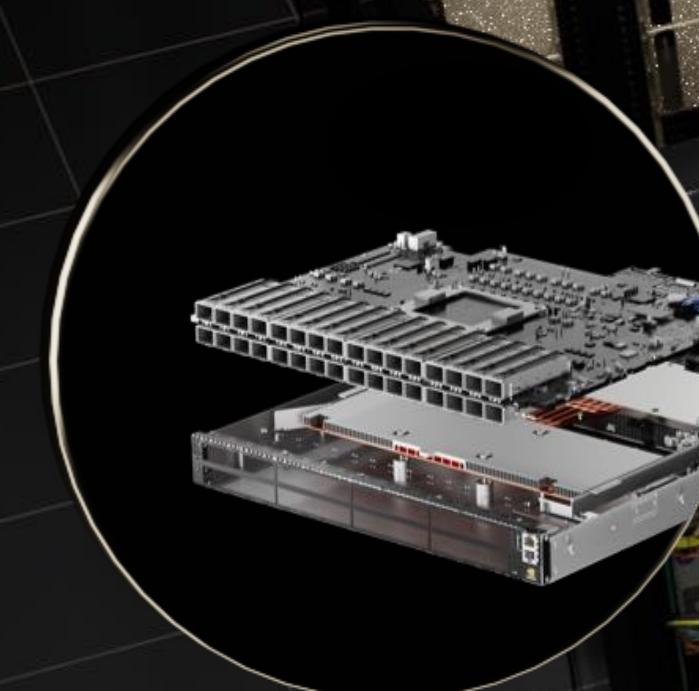
DOCA

Base Command

Forge

HARDWARE





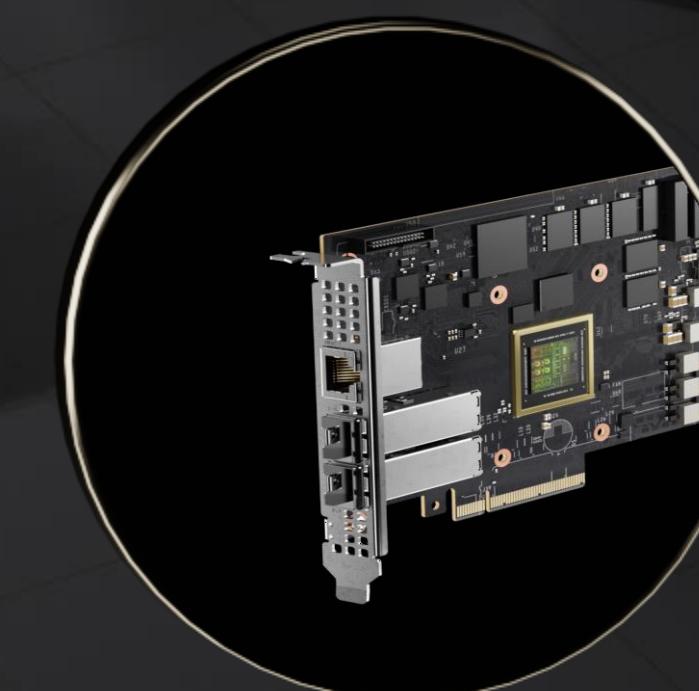
QUANTUM
INFINIBAND SWITCH



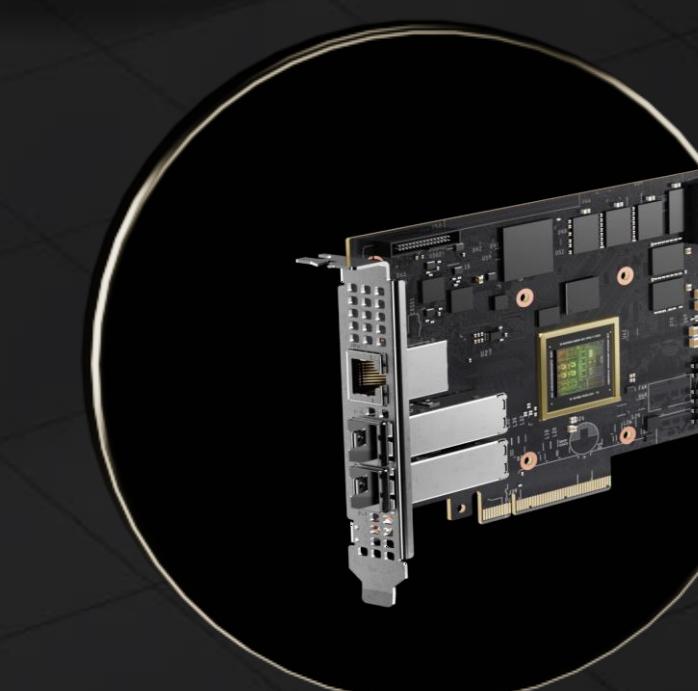
CONNECTX
SuperNIC



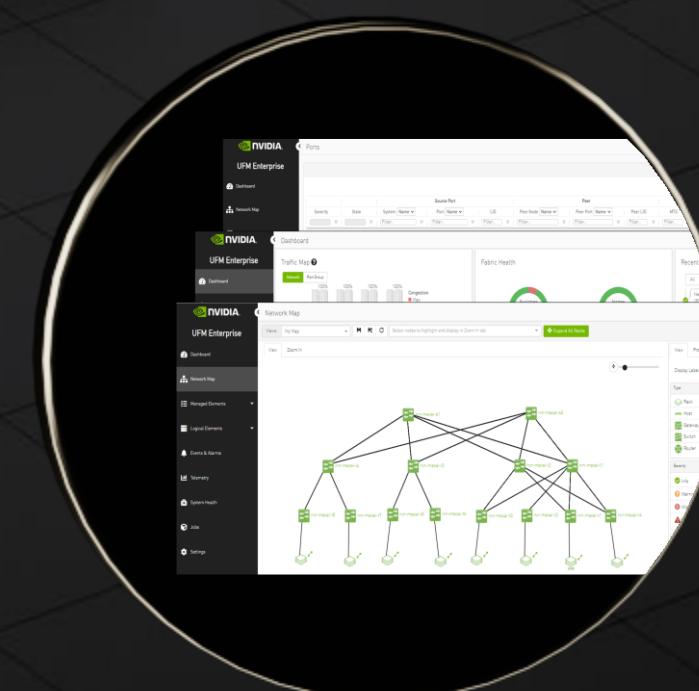
SPECTRUM
ETHERNET SWITCH



BLUEFIELD
SuperNIC



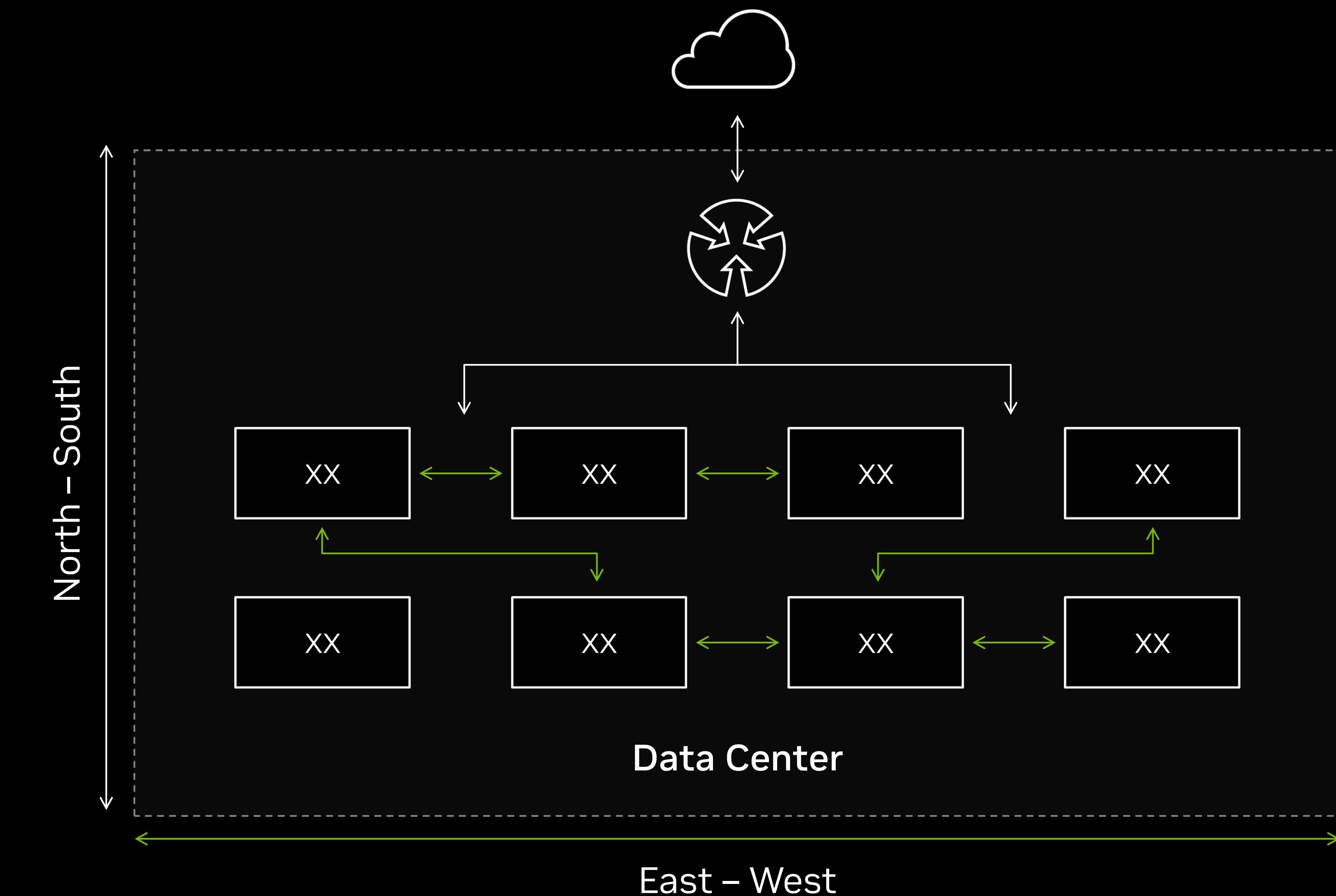
BLUEFIELD
DPU



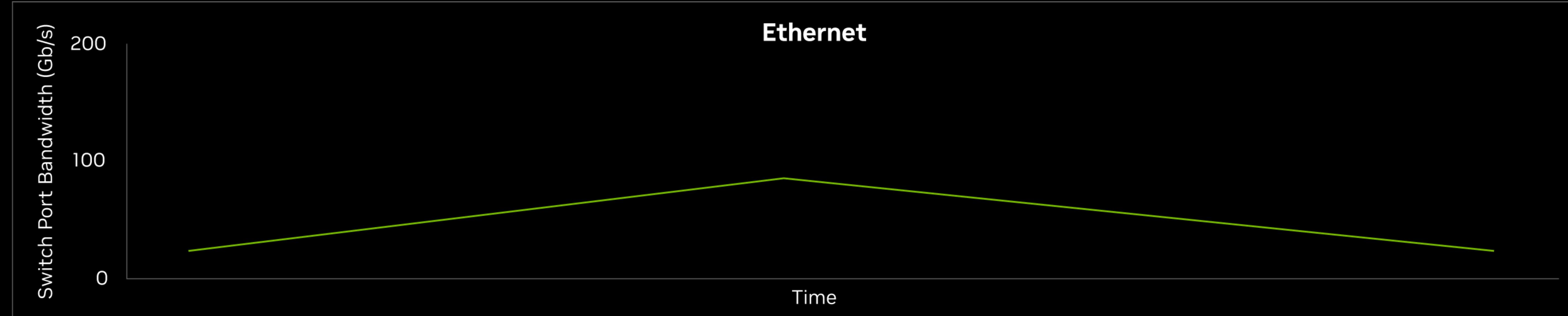
MANAGEMENT
& TELEMETRY

The Network Defines the Data Center

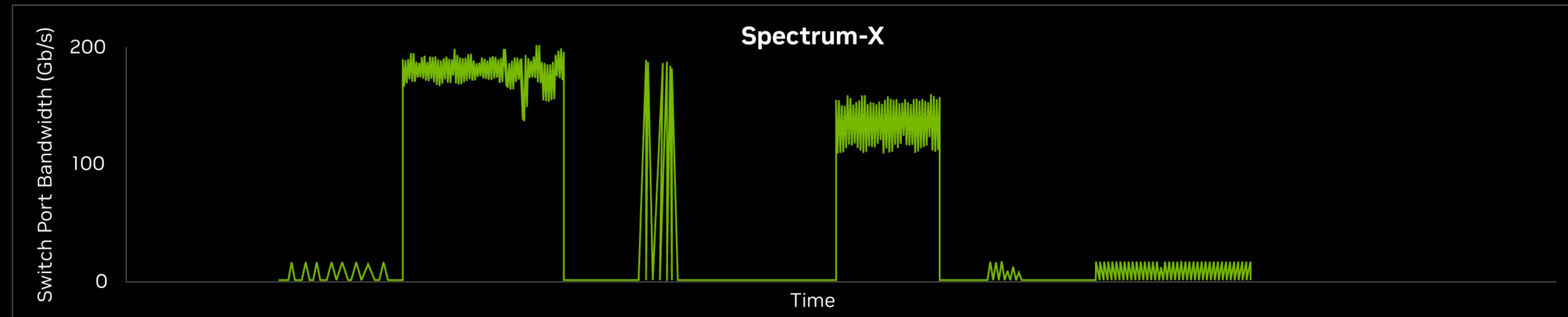
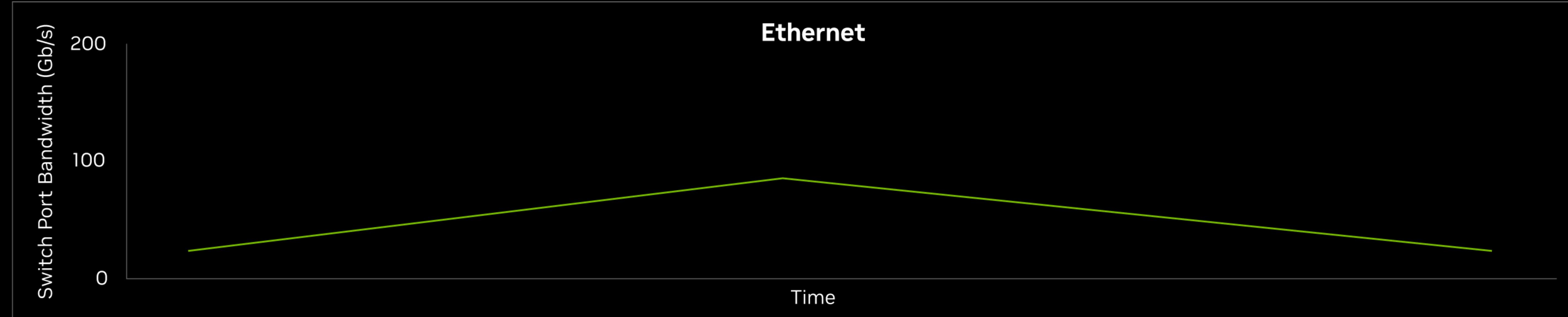
Control / User Access Network (North-South)	AI Fabric (East-West)
Loosely Coupled Applications	Distributed Tightly-Coupled Processing
TCP (Low Bandwidth Flows and Utilization)	RoCE (High Bandwidth Flows and Utilization)
High Jitter Tolerance	Low Jitter Tolerance (Long Tail Kills Performance)
Heterogeneous Traffic Average Multi-Pathing	Bursty Network Capacity Predictable Performance



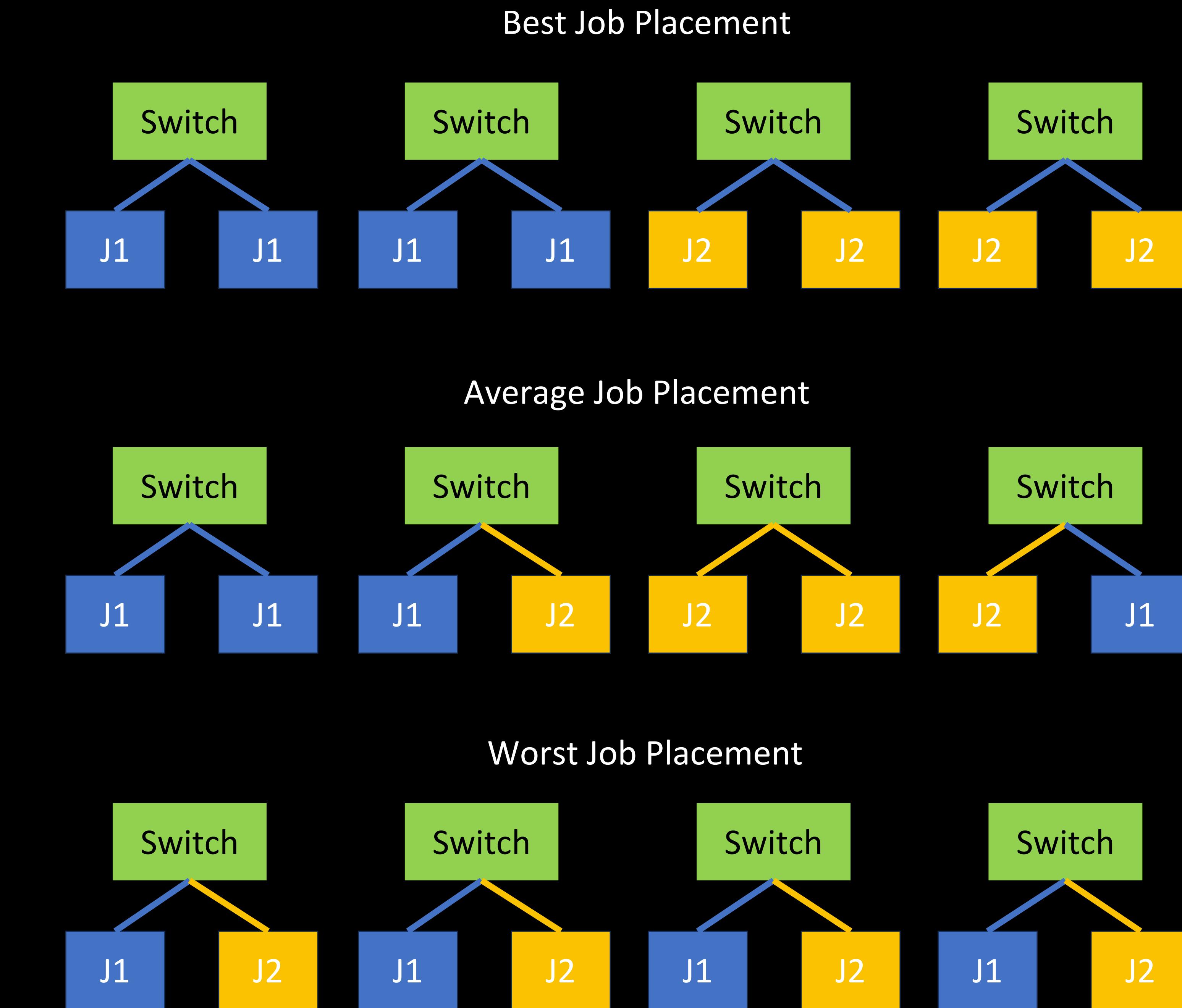
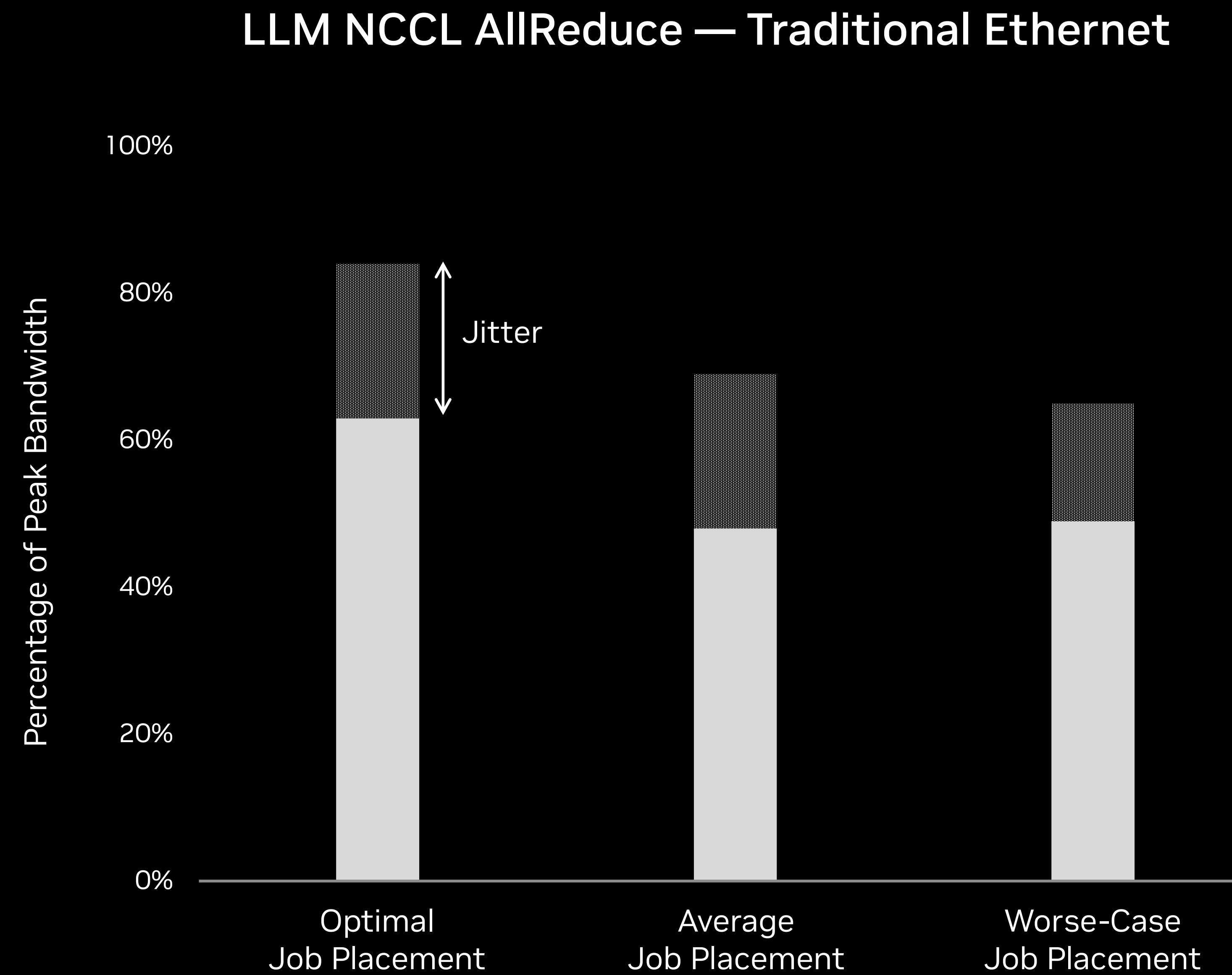
How Does GPT Workload Look Like?



How Does GPT Workload Look Like?



The Network Defines the Data Center

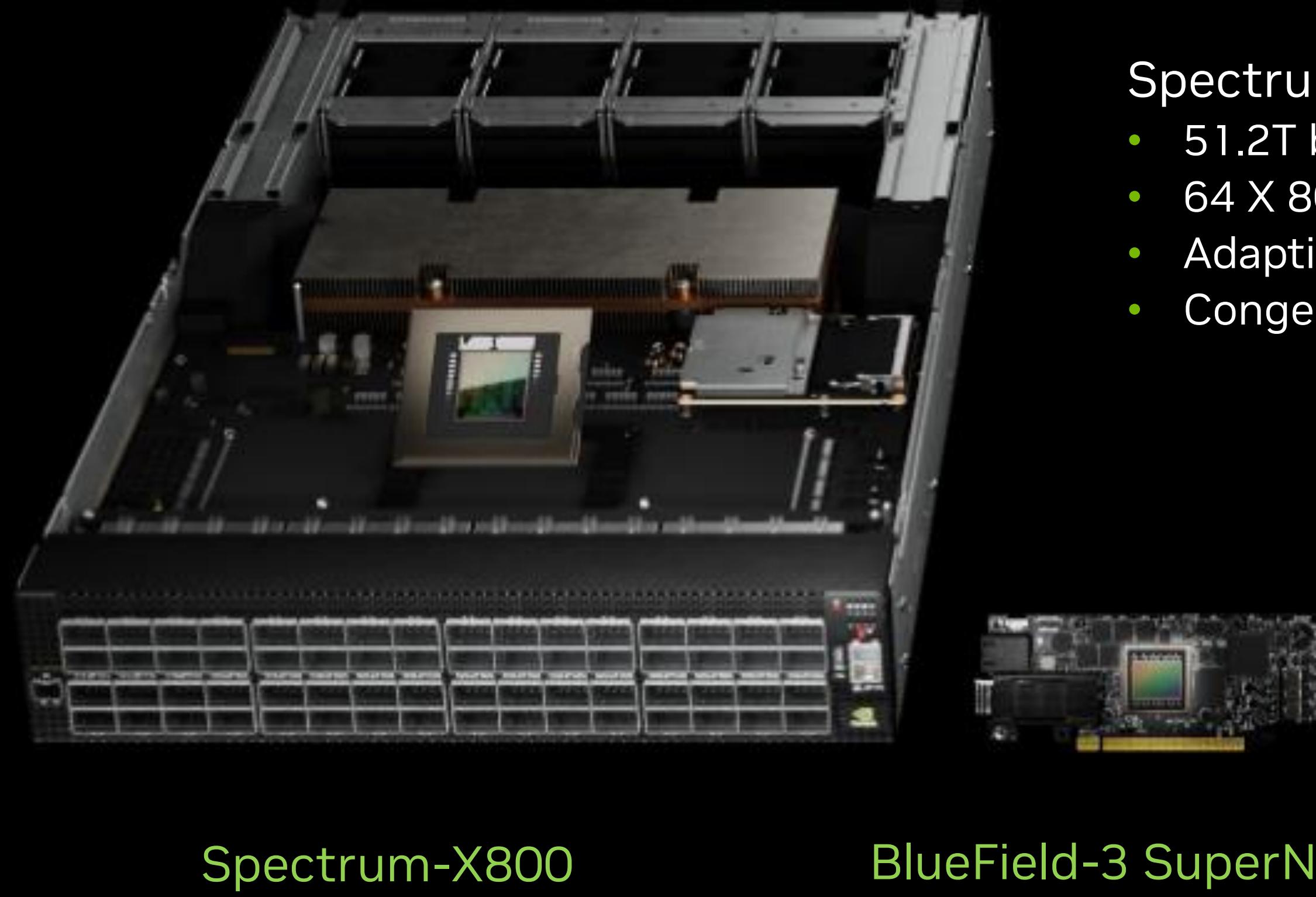
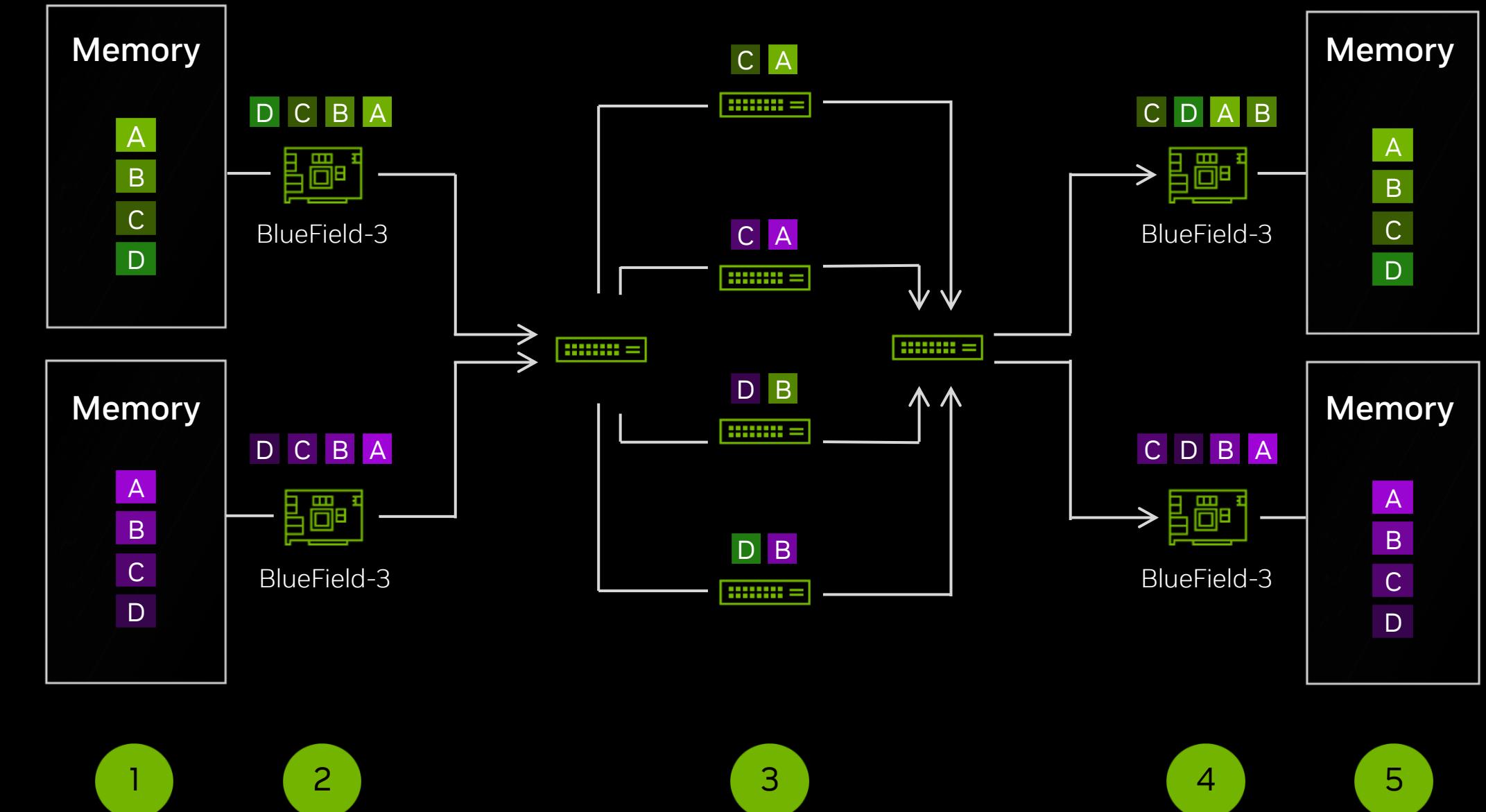


NCCL (NVIDIA Collective Communication Library) is the SDK library for AI communications - connects the GPUs and the network for the AI network operations

Spectrum-X800 Brings High-Performance AI to Ethernet

AI-optimized networking for every data center

- RoCE Adaptive Routing (local and remote information, a packet granularity)
- Congestion Control (telemetry probes)
- Noise Isolation (multi-jobs or a single large-scale job)
- High frequency telemetry (1000x)



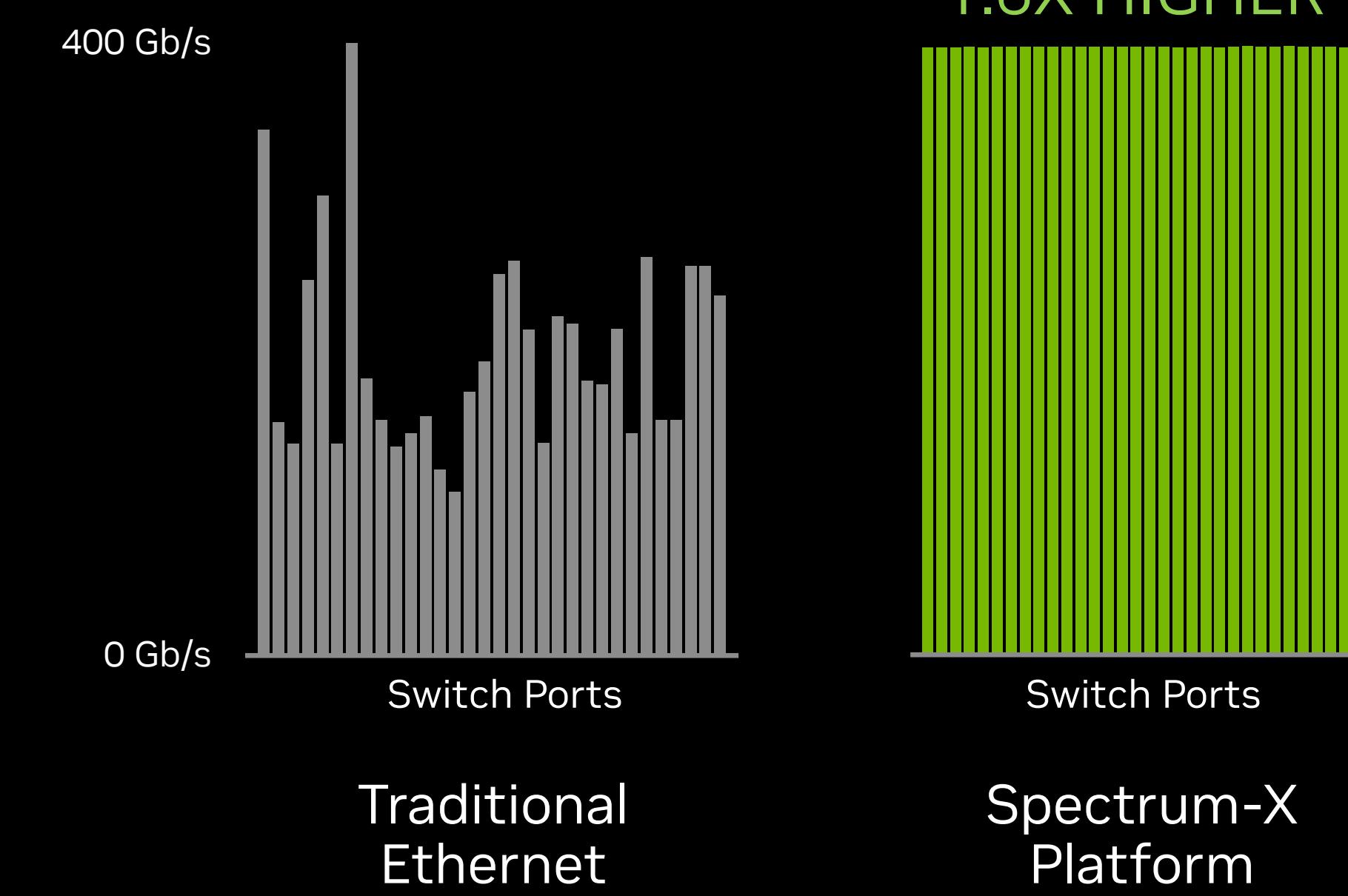
Spectrum-X800 Switch

- 51.2T bandwidth
- 64 X 800G Ports, 128 x 400G
- Adaptive routing
- Congestion control, noise isolation

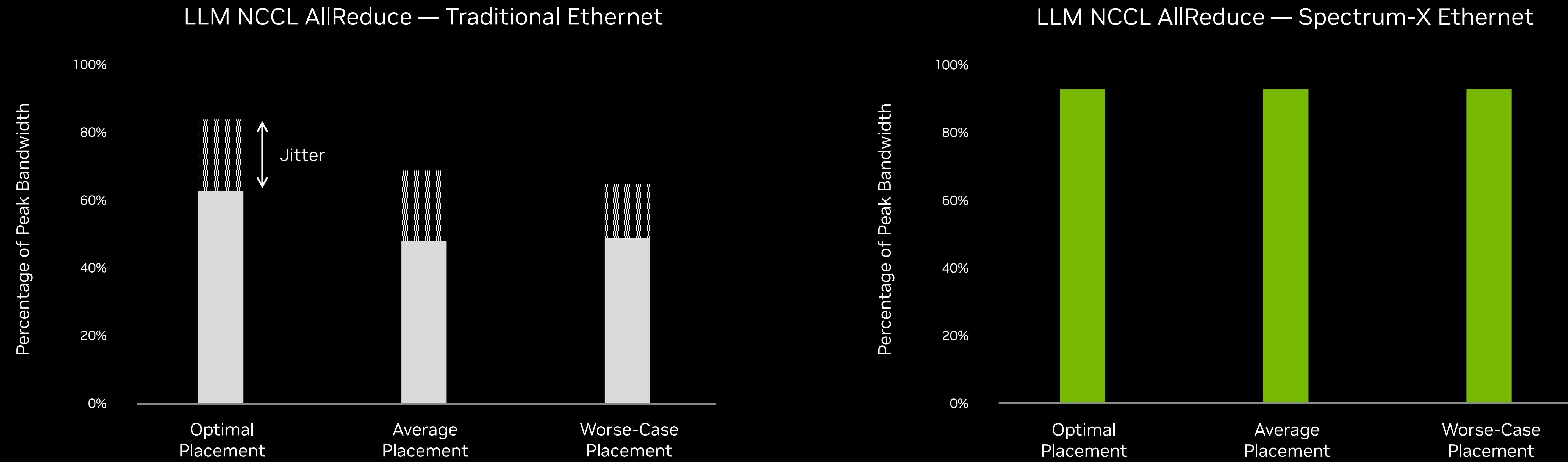
BlueField-3 SuperNIC

- 16 Arm 64-Bit Cores
- 16 Core / 256 Threads Datapath Accelerator
- DDR memory interface
- ConnectX NIC
- PCIe switch

Effective Network Bandwidth
With and Without Adaptive Routing



Spectrum-X New Class of Ethernet for AI



- Spectrum-X performance is consistent; Traditional Ethernet shows run-to-run bandwidth variability
- Results in 1.4x higher LLM performance (2K GPUs)

NVIDIA Spectrum-X Generative AI Cloud

Most powerful supercomputer in Israel



2,048 GPUs

2,560 BlueField-3 SuperNICs

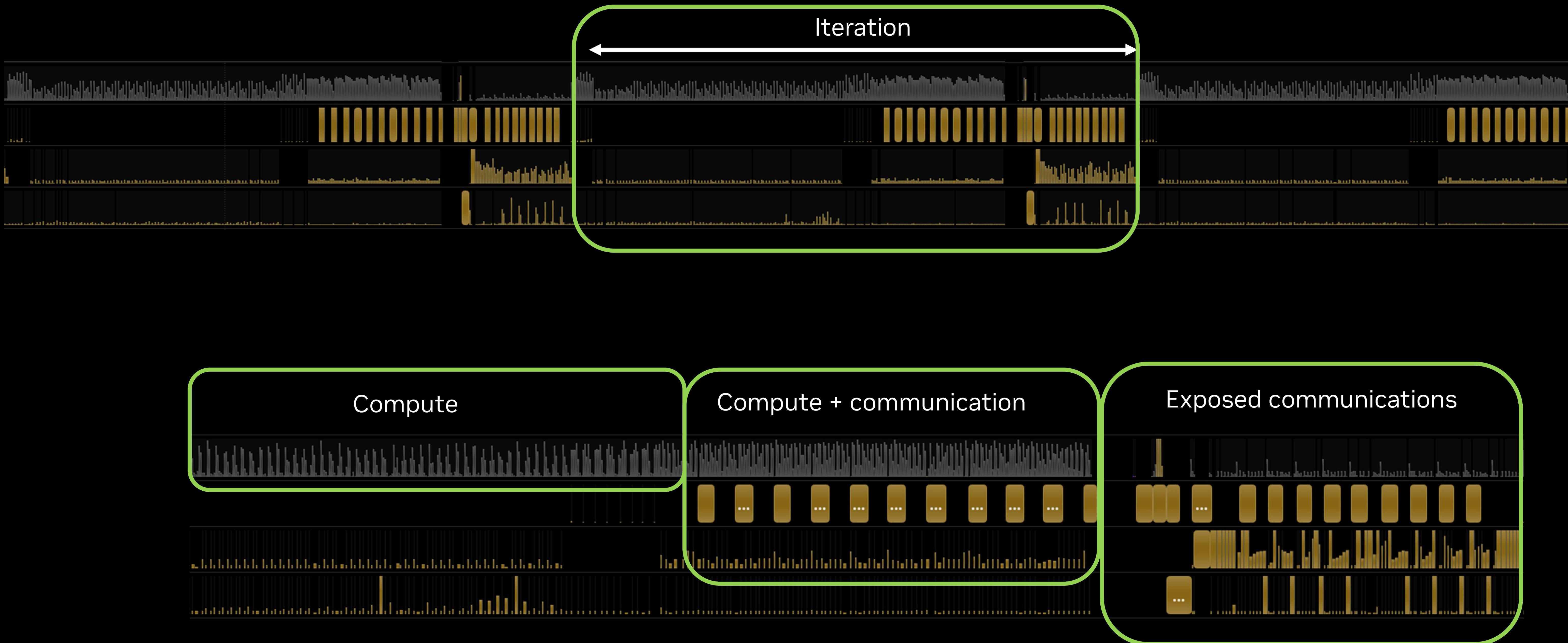
80+ Spectrum-x800 Ethernet switches

Peak AI performance of 8-Exaflops

Israel-1 AI Cloud Digital Twin with NVIDIA Air and Omniverse



LLM Compute and Communication Profiling

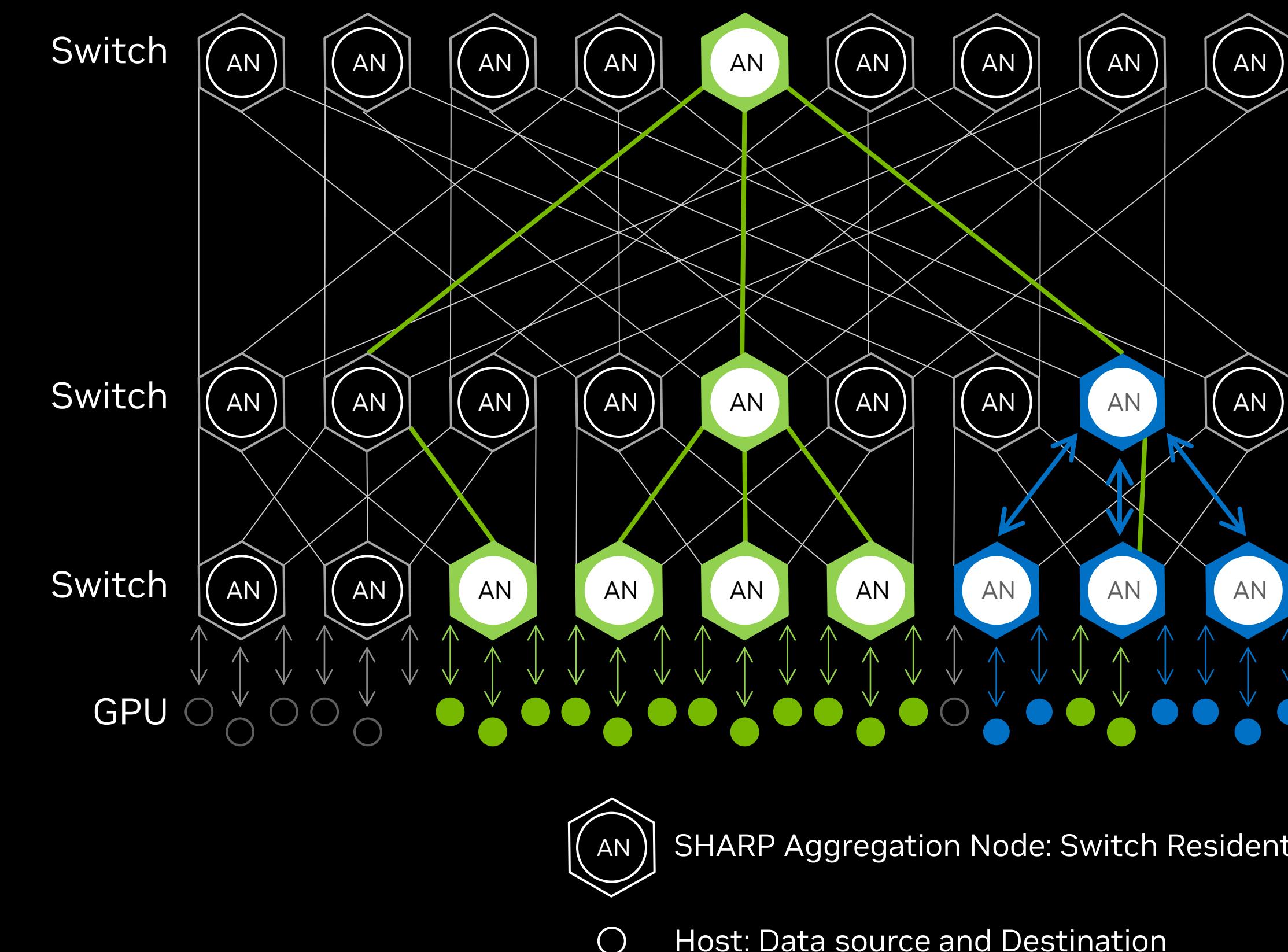


Representative profile from a large scale LLM training run
Communications is bursty in nature, an average bandwidth utilization is not a good network criteria

NVIDIA SHARP

Scalable Hierarchical Aggregation and Reduction Protocol Technology

- In-network data aggregation mechanism
- Multiple simultaneous outstanding operations
- Barrier, reduce, all-reduce, broadcast and more
- Sum, min, max, min-loc, max-loc, or, xor, and
- Integer and floating-point, 8/16/32/64 bits



Quantum-X800 InfiniBand Switch

Highest-Performance AI-Dedicated Infrastructure

- 144 ports of 800G, 5x higher switch capacity
- SHARP v4 with 14.4 TFlops of In-Network Computing, 9x higher
- Adaptive routing, congestion control



Quantum-X800

ConnectX-800 SuperNIC

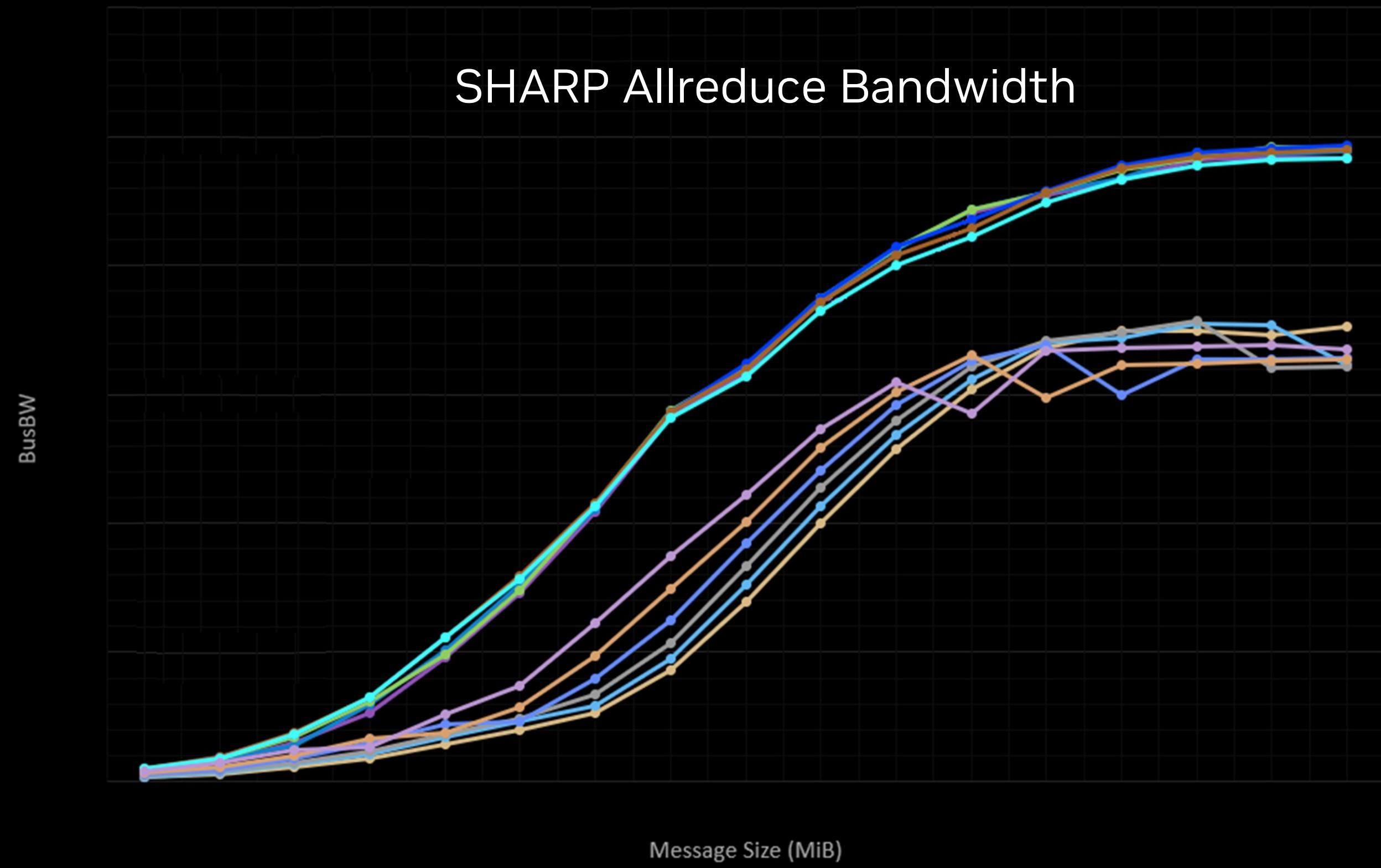
Quantum-X800 switch

- 144X 800G ports,
- SHARPv4 In-Network Computing
- Adaptive routing
- congestion control, and noise isolation

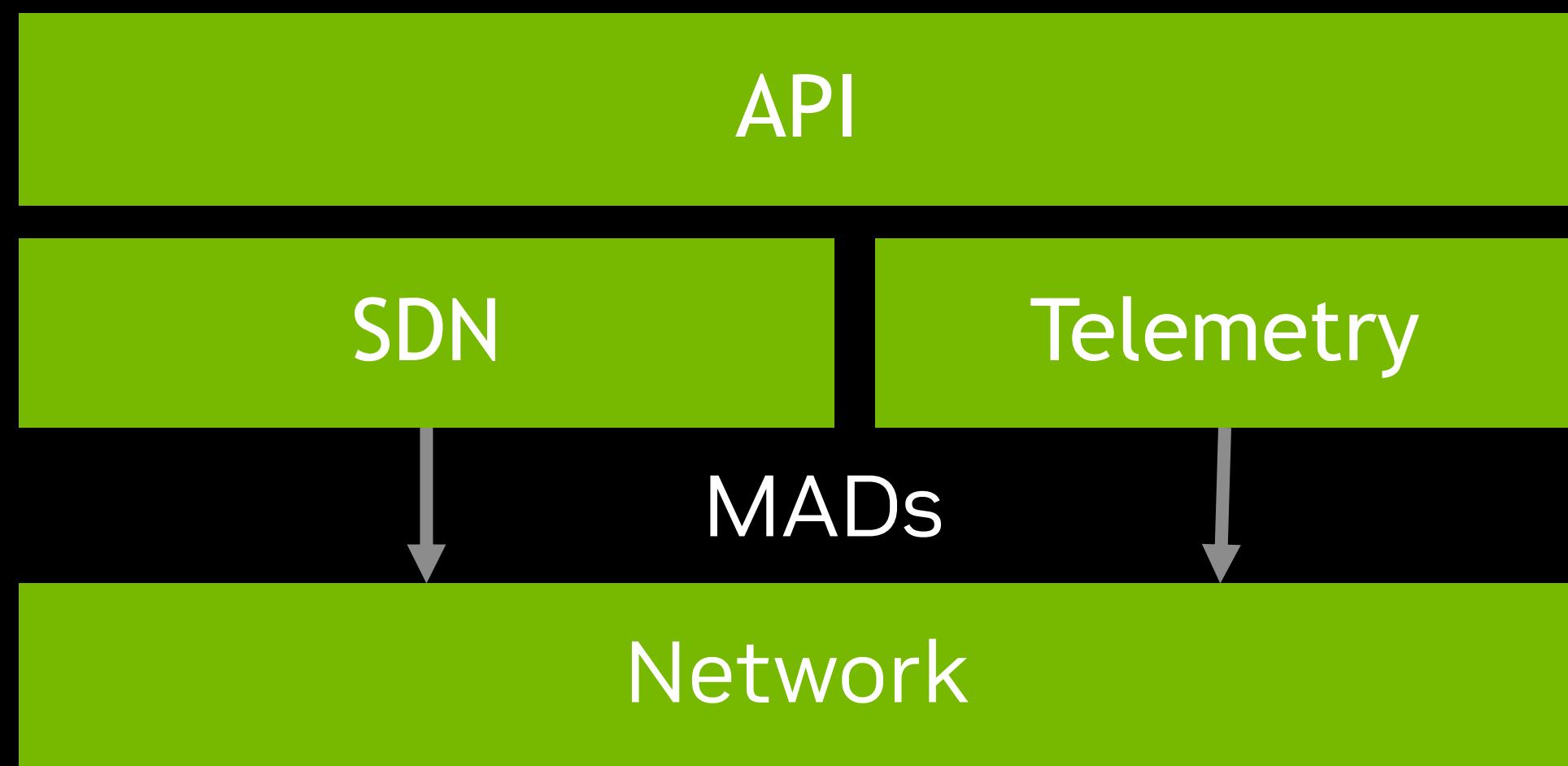


ConnectX-x800 InfiniBand SuperNIC

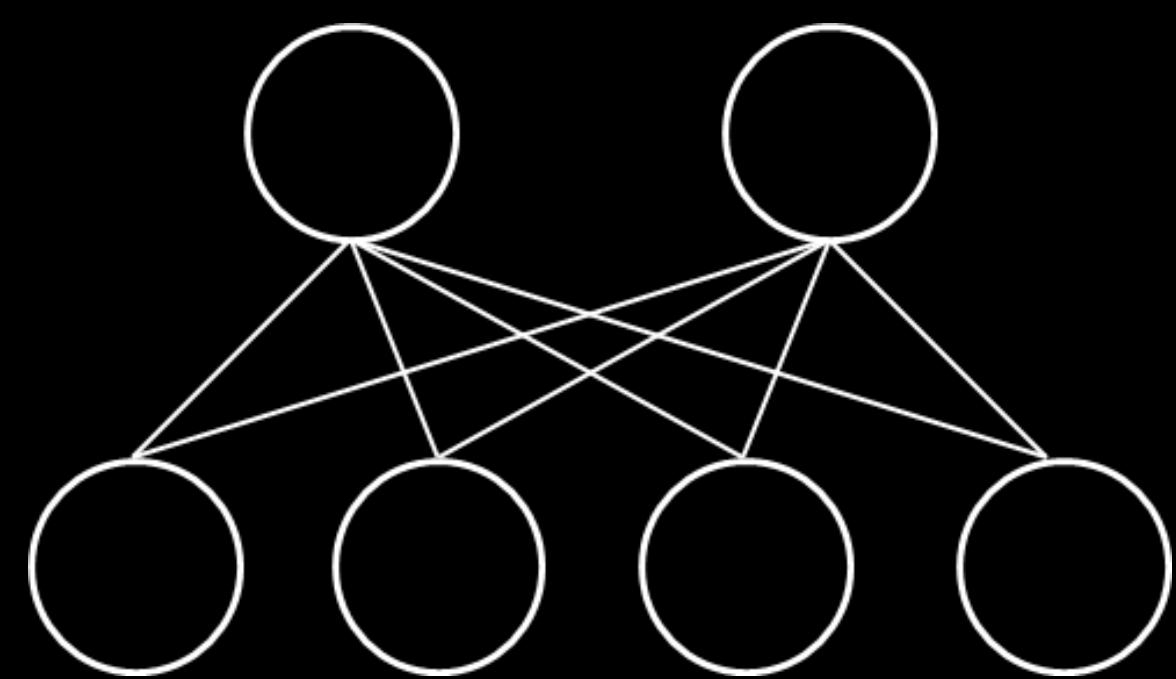
- PCIe Gen 6, PCIe switch
- Multi-host



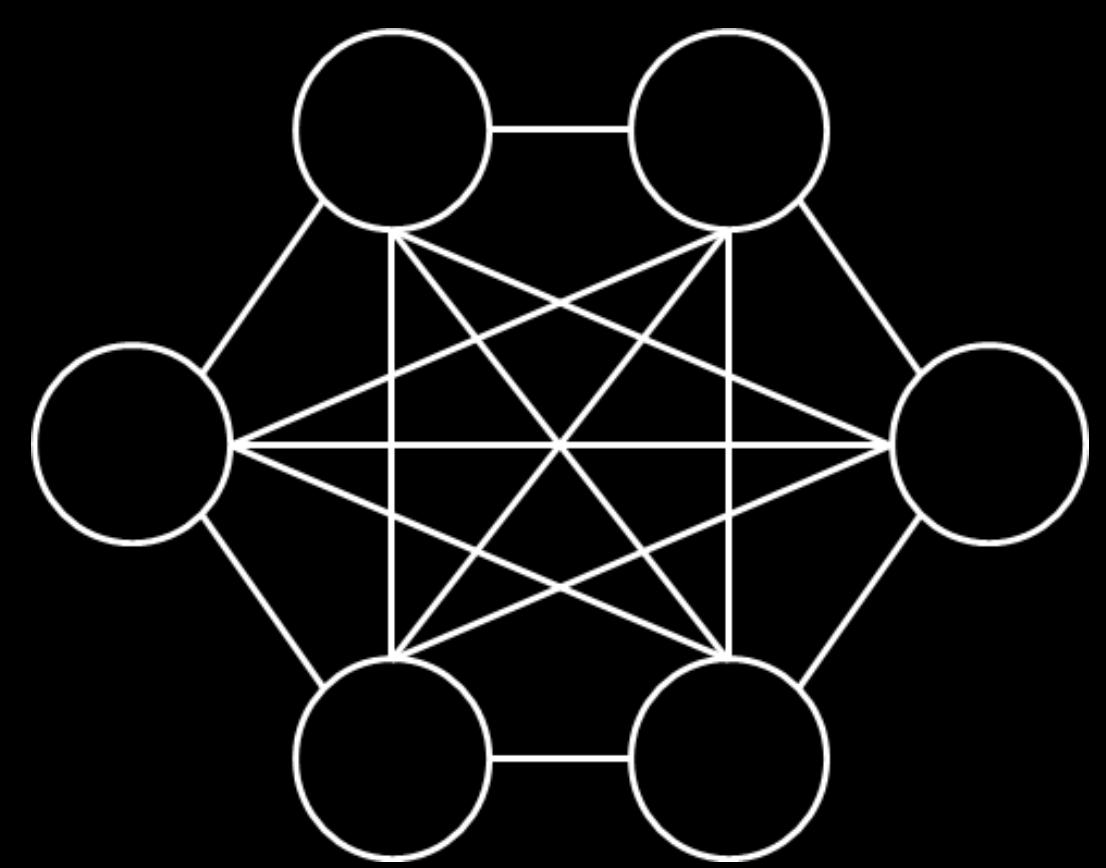
Exploring Topologies – Creating Routing Algorithms



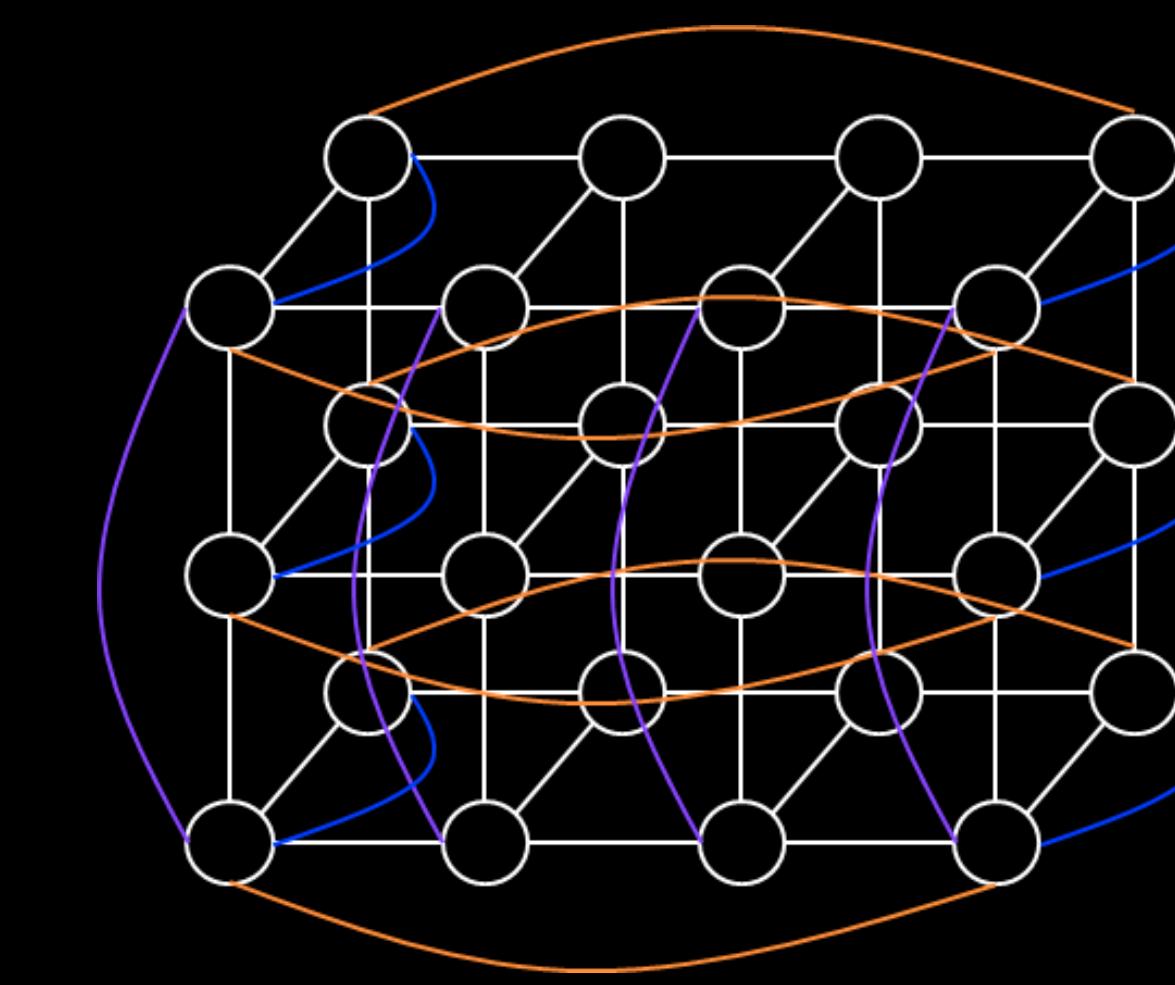
Existing routing algorithms support



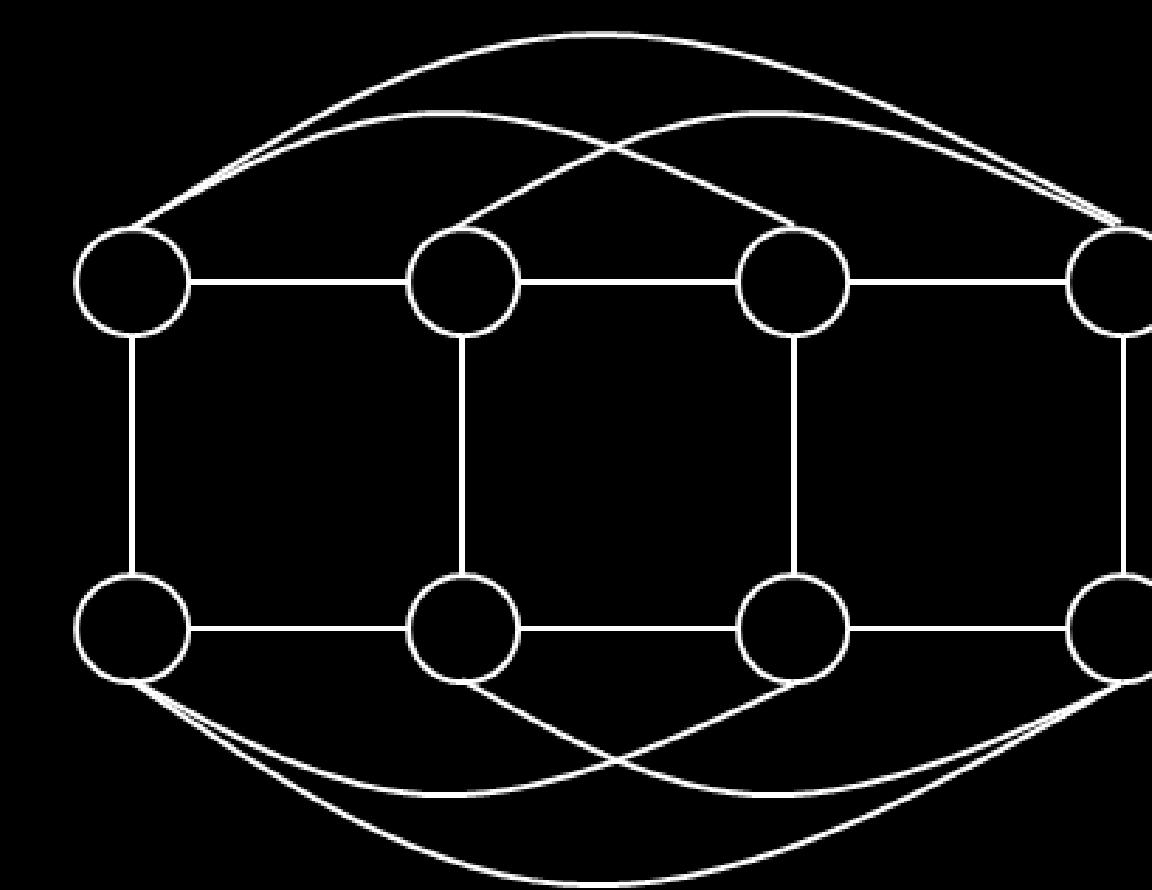
Fat-tree



Dragonfly

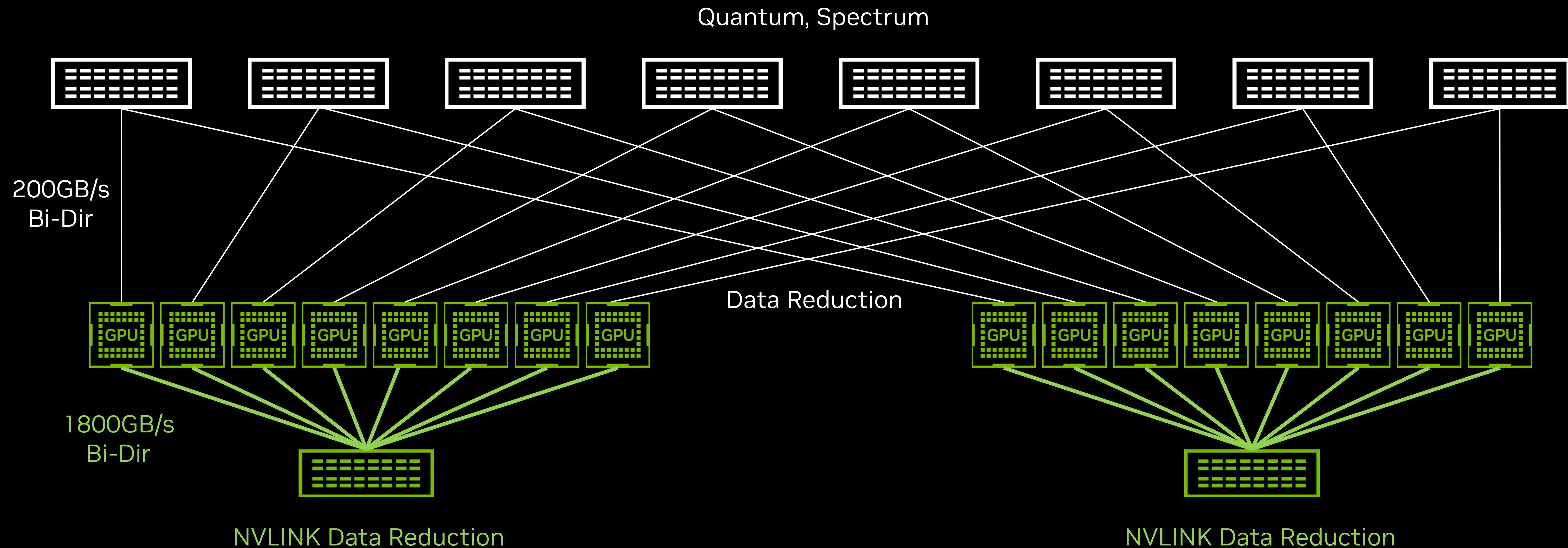


Torus



HyperX

Rail Optimized Topology



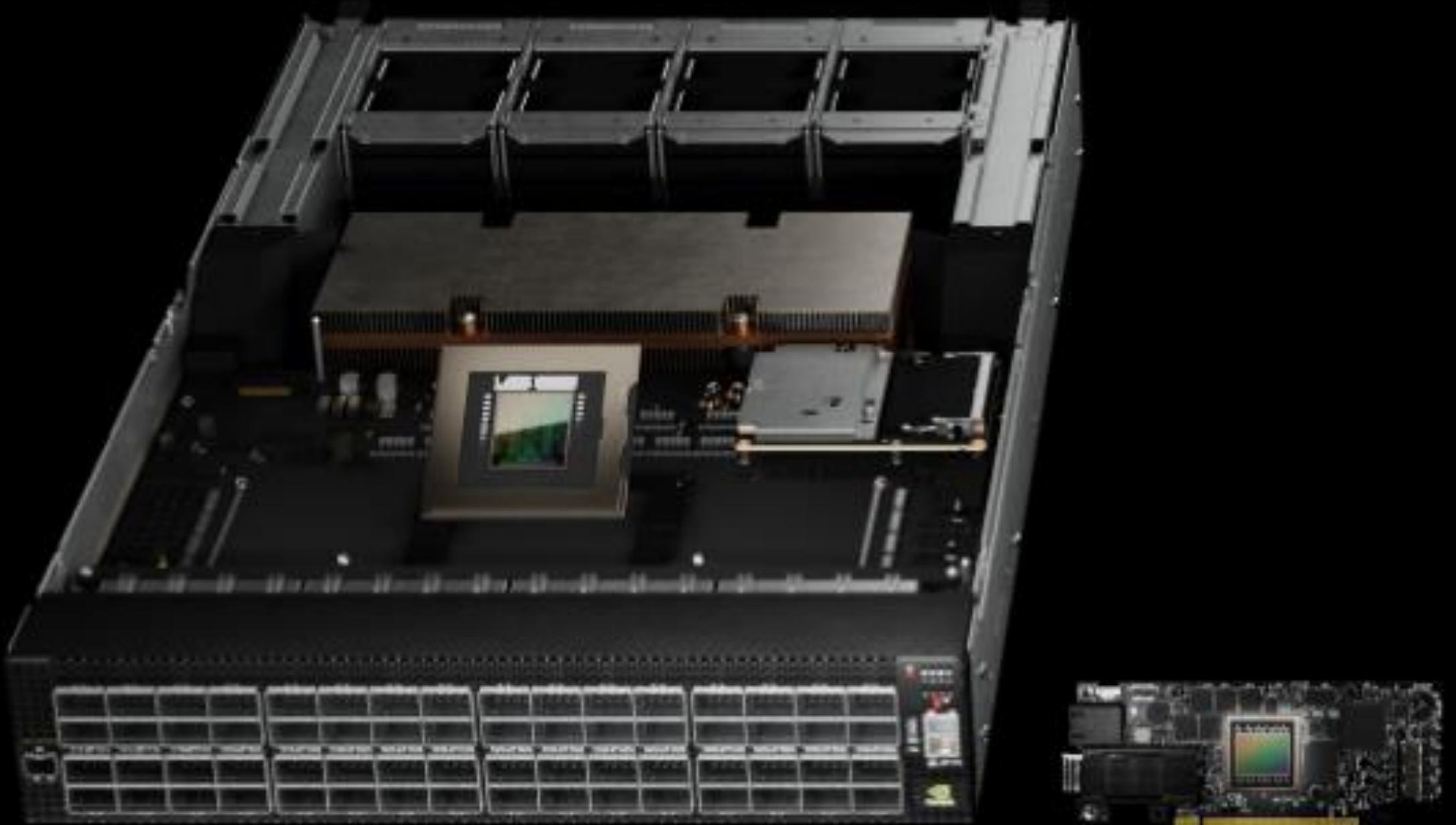


Quantum-X800 InfiniBand and Spectrum-X800 Ethernet

Optimized for Trillion-Parameter GPU Computing and AI Infrastructure

Bringing High-Performance AI to Ethernet

64 x 800GB/s | RoCE Adaptive Routing | Congestion Control | Noise Isolation

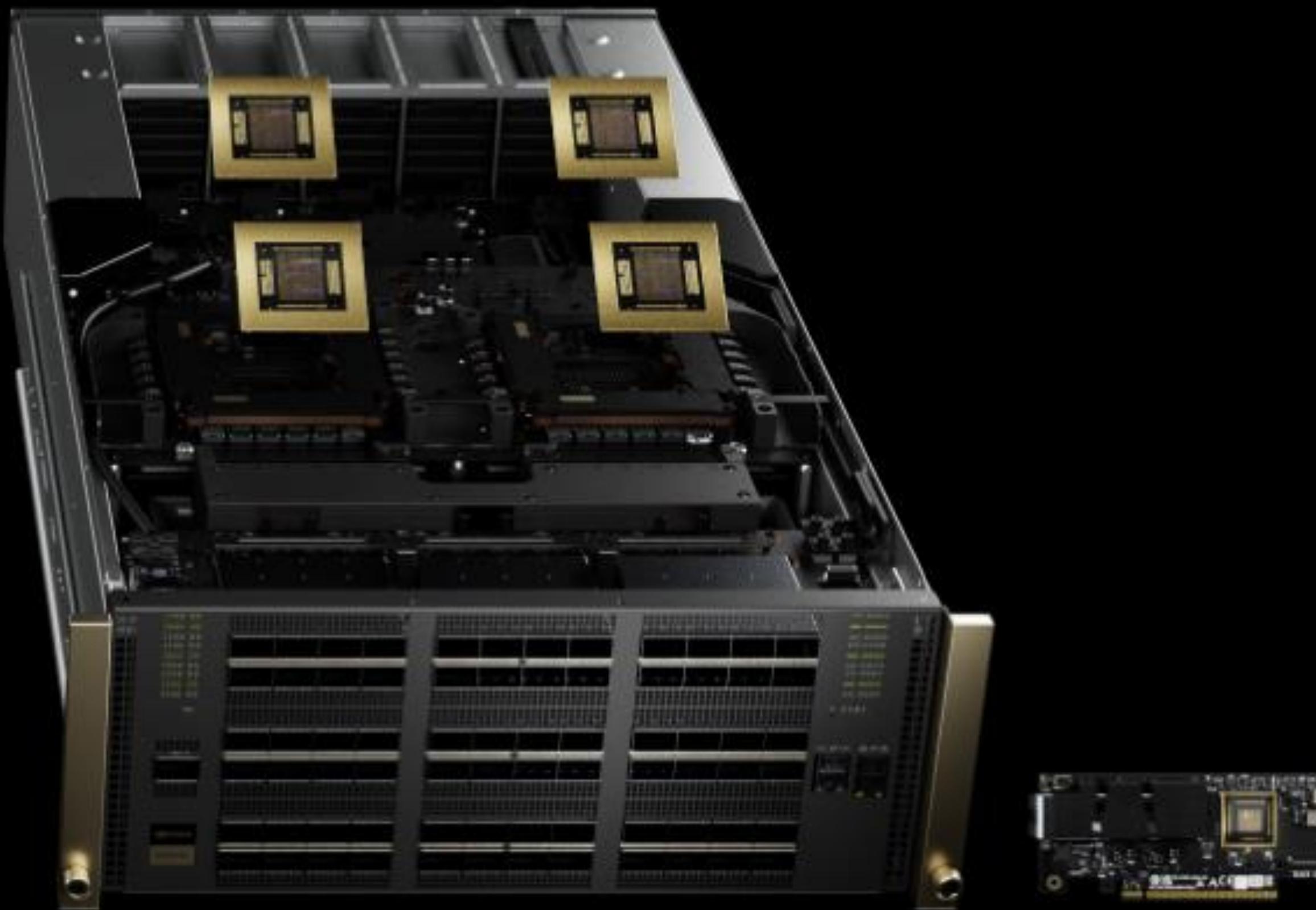


Spectrum-X800

BlueField-3 SuperNIC

Highest-Performance AI-Dedicated Infrastructure

144 x 800GB/s | 14.4Tflops In-Network Computing | Ultra-Low Latency



Quantum-X800

ConnectX-800 SuperNIC

