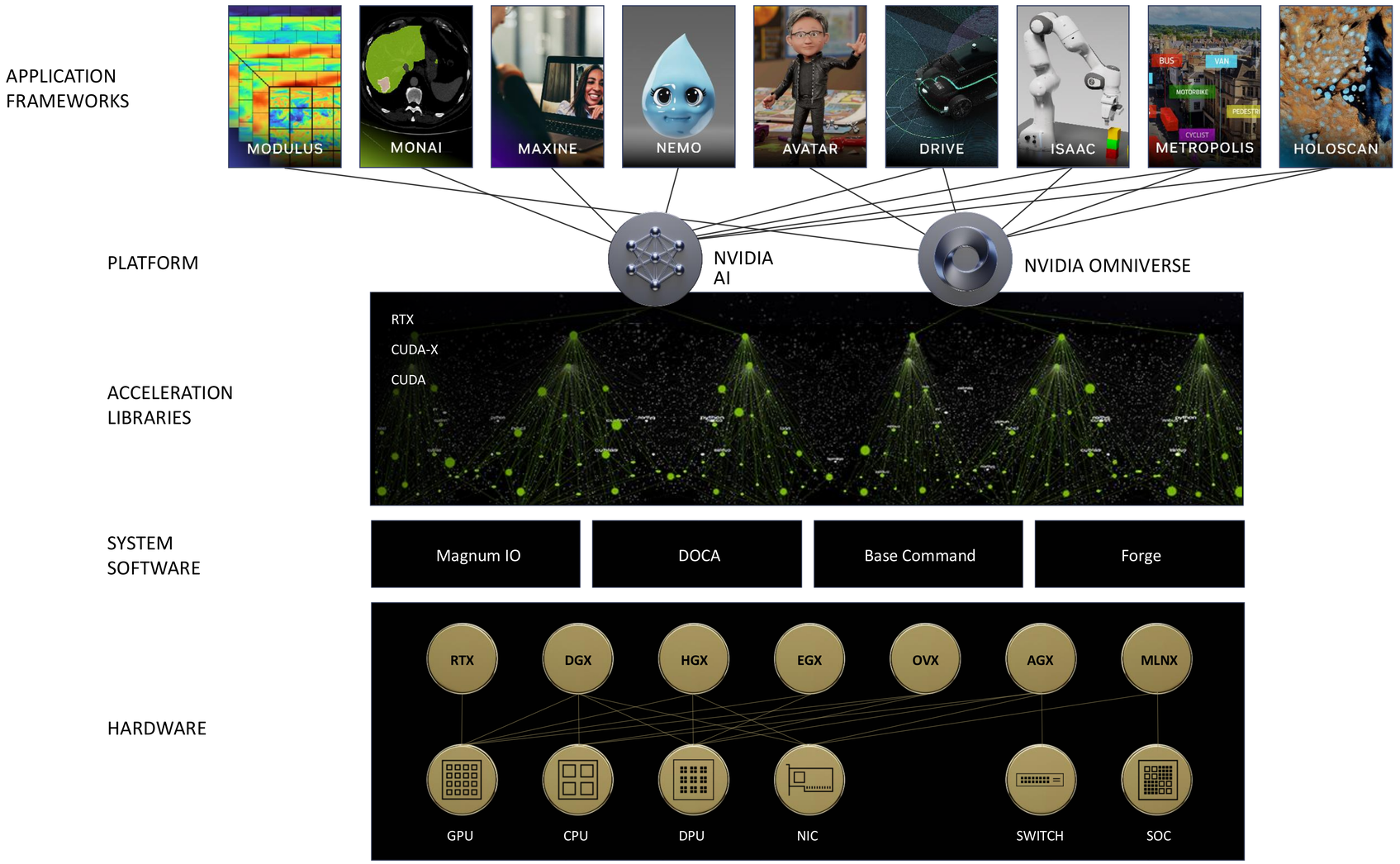
NVIDIA Fabrics for AI
@ IBM Storage Scale Users Forum

November 17, 2024

Bill Webb – bwebb@nvidia.com

NVIDIA Accelerated Computing

Accelerated Computing Services, Software and Systems Enabling New, Enhanced Business Models



AI Workloads Require an AI Fabric

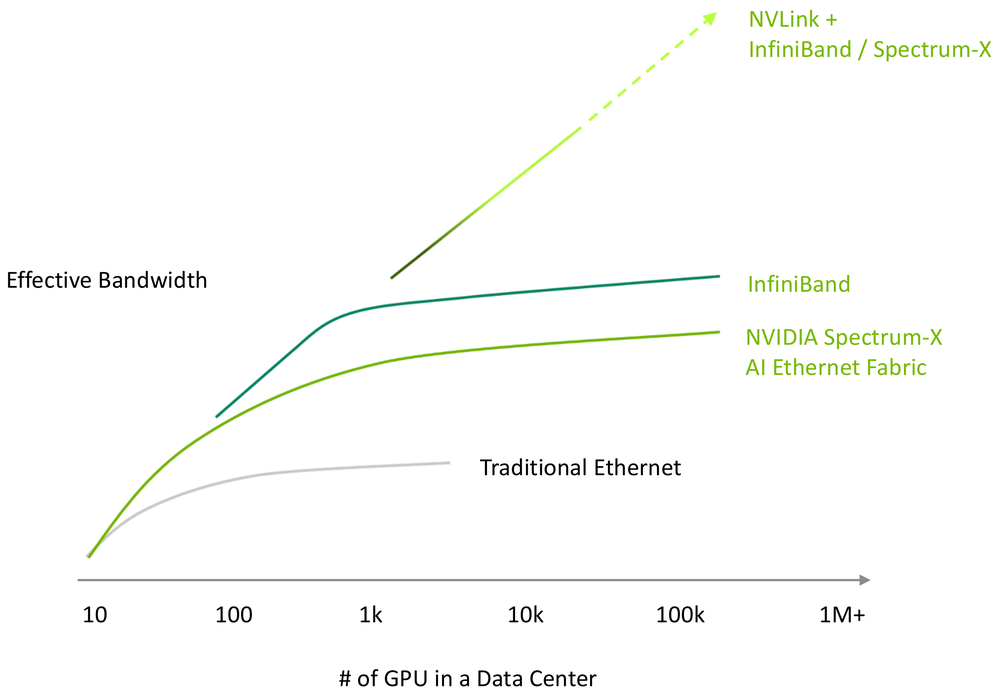
Full Stack Optimized

General Purpose Network		AI compute and storage Optimized Fabrics
Loosely-Coupled Applications		Tightly-Coupled Processes
TCP (Low Bandwidth Flows and Utilization)		RDMA (High Bandwidth Flows and Utilization)
High Jitter Tolerance		Low Jitter Tolerance
Oversubscribed Topologies		Nonblocking Topologies
Heterogeneous Traffic, Statistical Multi-Pathing		Bursty Network Capacity, Predictive Performance

Generative AI is a Data Center Scale Computing Problem

AI factories emerge as a new class of generative AI data centers

Purpose-built high performance networking is necessary to effectively scale AI



NVIDIA NVLink

Fastest Compute Fabric Connectivity

NVIDIA InfiniBand

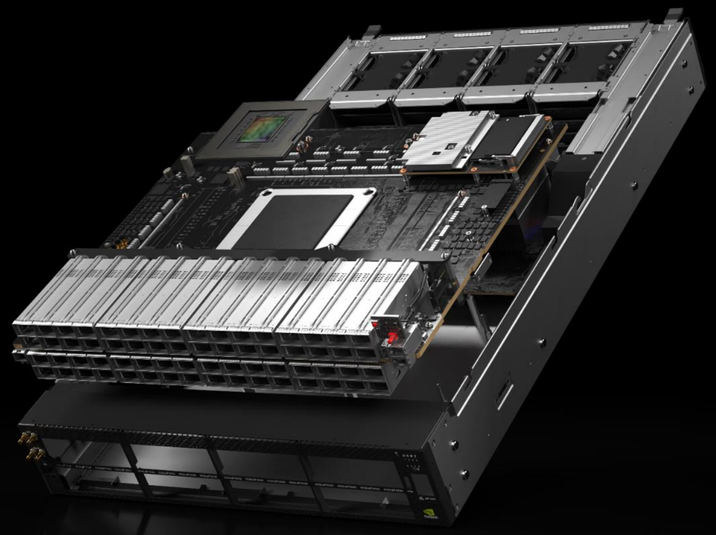
Gold Standard for Scale-Out AI Fabrics

NVIDIA Spectrum-X

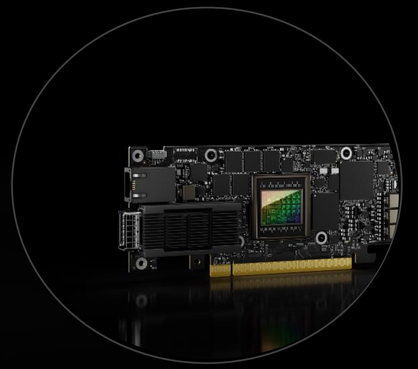
Ethernet Optimized for Multi-Tenant AI Factories

Purpose-Built Ethernet Fabric for AI

NVIDIA Spectrum-X Networking Platform



Spectrum SN5600 Ethernet Switch



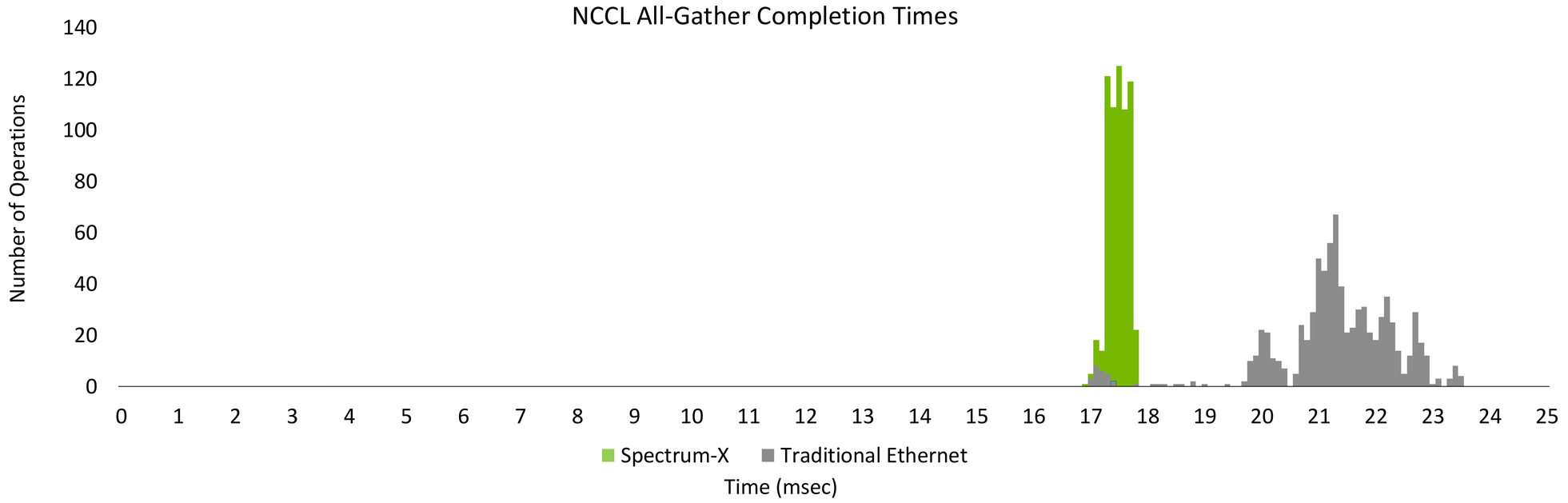
BlueField-3 SuperNIC

1.6X Higher AI Network Performance

Standard Ethernet Connectivity

What Makes Spectrum-X Special

Switch-to-SuperNIC, End-to-End Network Processing, Bringing High Performance to Ethernet



Schedules Data Transmission to Avoid Congestion

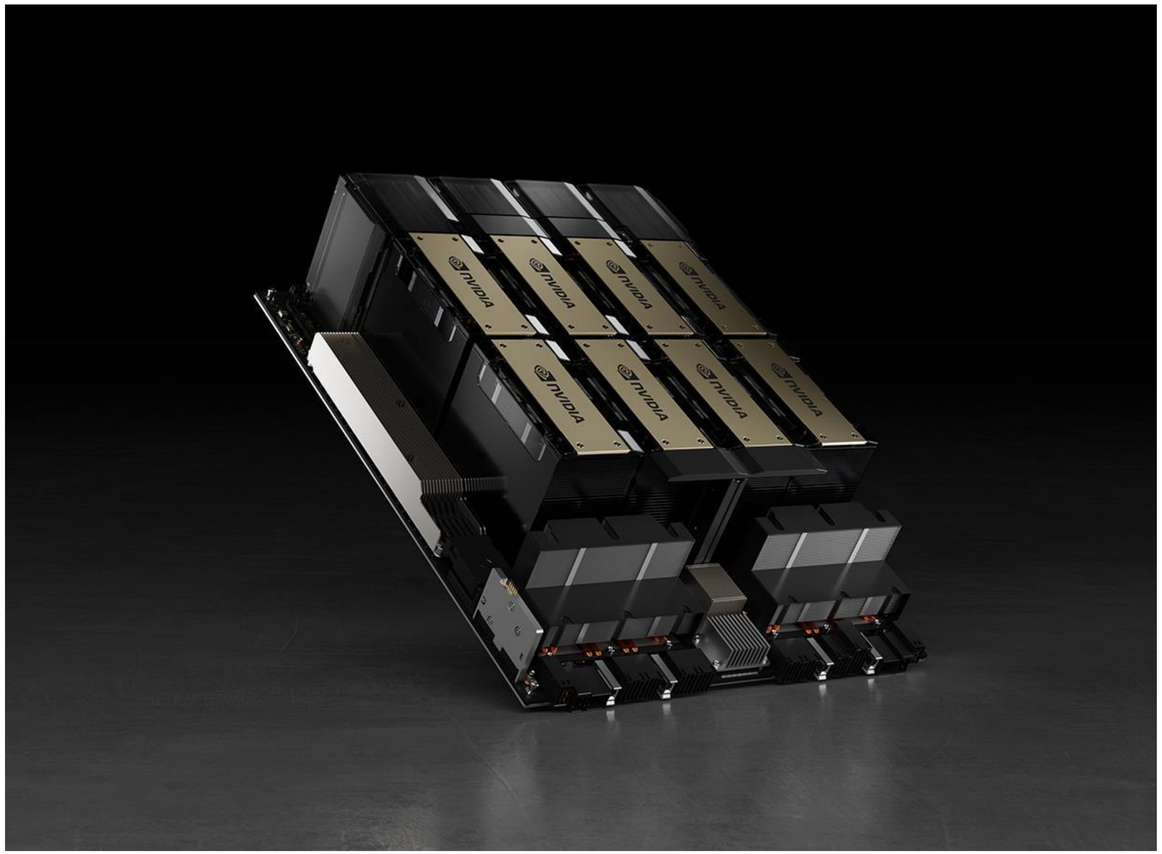
Ultra-High-Speed Traffic Monitoring Distributes Data Across All Switch Ports Ignoring Data Ordering

Reordering - Receive Data and Place it Back in Order

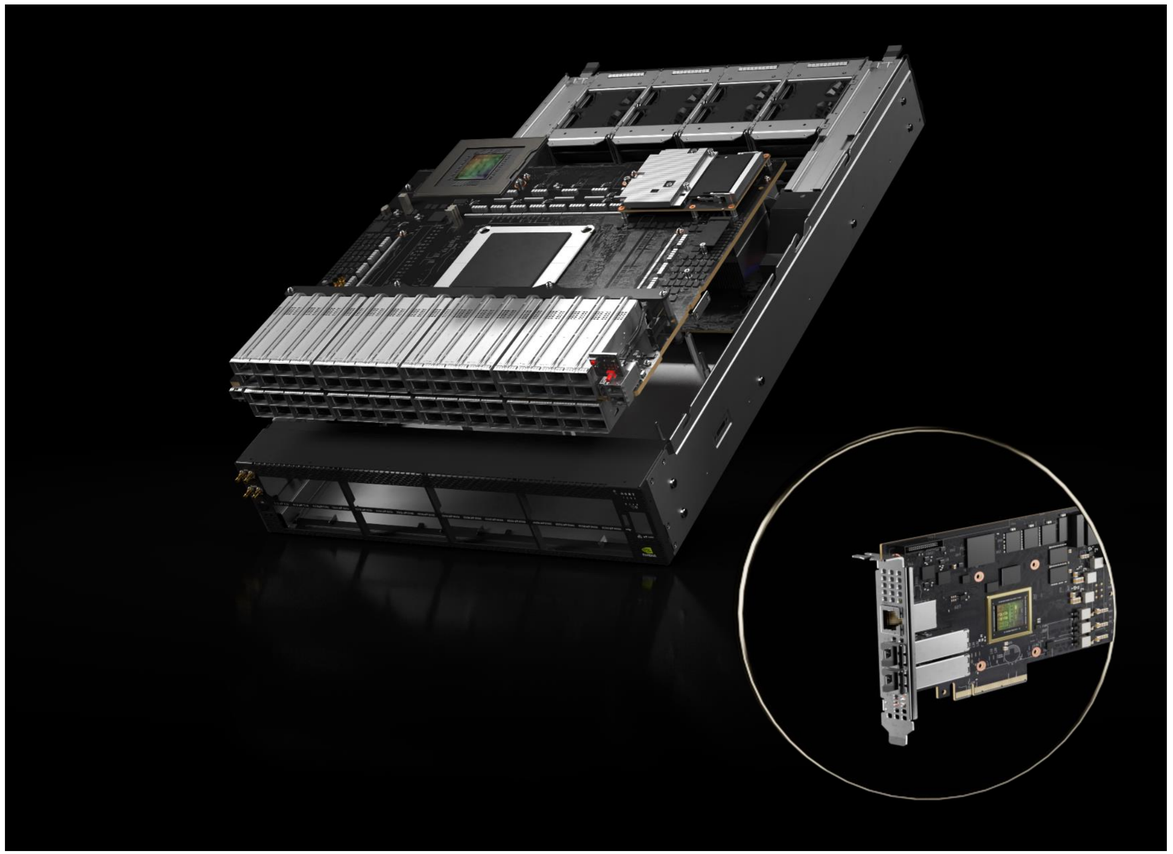
Spectrum-X provides 1.3X faster completion times for All-Gather Operations

Purpose-Built for Ethernet AI Clouds

NVIDIA Spectrum-X Combined with HGX Accelerates AI at Massive Scale



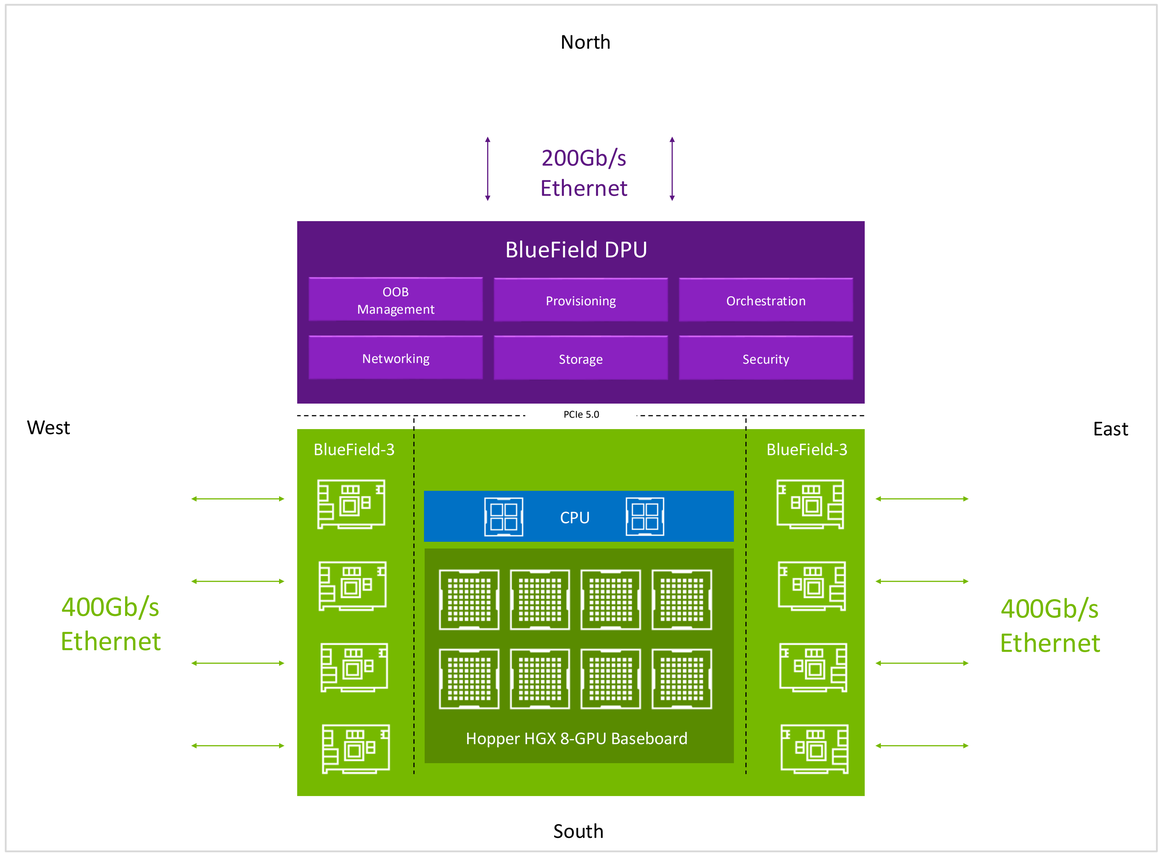
NVIDIA HGX H100/H200 AI Platform



NVIDIA Spectrum-X Networking Platform
Spectrum-4 Switch + BlueField-3 DPU

Spectrum-X and HGX H200 Accelerate AI at Massive Scale

BlueField-3 Delivers Key Features and Benefits Needed for AI Clouds



Hopper HGX 8-GPU Air-Cooled Platform w/ BlueField-3

Cloud Orchestration

Automated Provisioning
Elastic GPU Compute
Service Function Chaining
DOCA Services

Storage Acceleration

NVMe Storage Disaggregation
VirtIO Block Storage
Erasure Coding
Decompression

Secure Infrastructure

Isolated Control-Plane
Zero-Trust Security
Next-Generation Firewall
Deep Packet Inspection

Cloud Networking

Software-Defined Networking
VXLAN Overlay Networking
L3 Host-Based Networking (HBN)
VirtIO Full Emulation (VFE)

NVIDIA Spectrum SN5600

The First Ethernet Switch Purpose-Built for AI

FAST

4X bandwidth capacity increase

EFFICIENT

4X reduction in solution footprint

SECURE

In-flight and **at-rest** encryption

GREEN

50% Reduction in solution power

51.2Tbps
Aggregate bandwidth

400/800GbE
100G SerDes technology

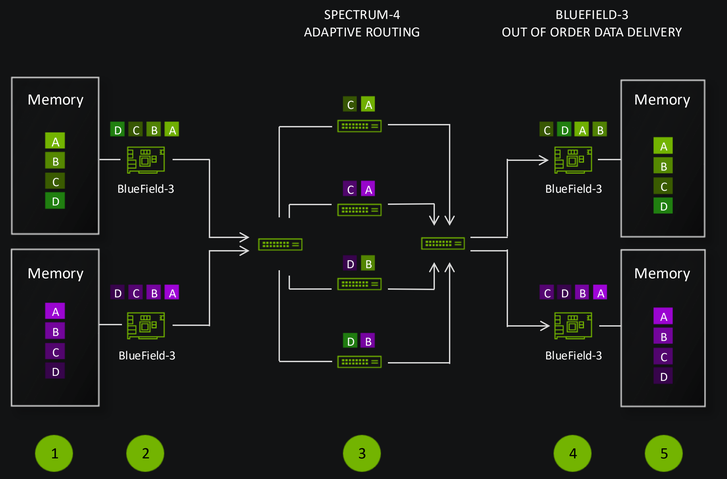
100B
Transistors

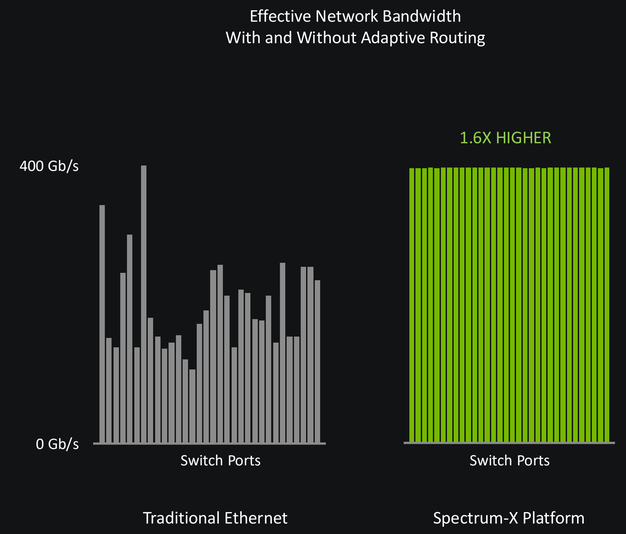
4N
NVIDIA design process

Enhanced Adaptive Routing

Increases effective data throughout by 1.6X

- BlueField-3 sends compute data into the network
- Spectrum-4 intelligently spreads data packets across **all available routes**
- BlueField-3 ensures in-order data delivery
- Increase from typical 60% to 95% effective bandwidth

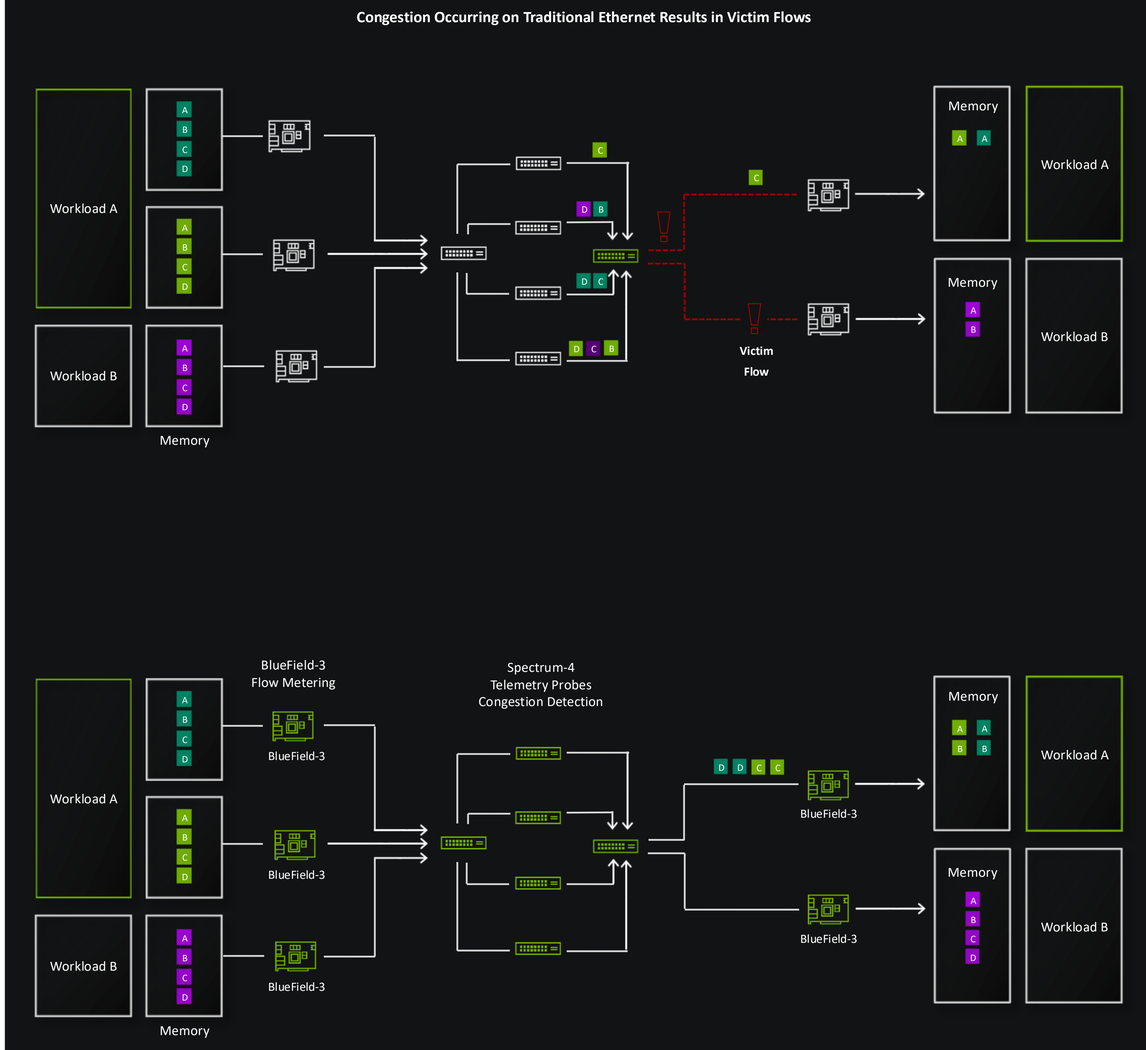




Programable Congestion Control

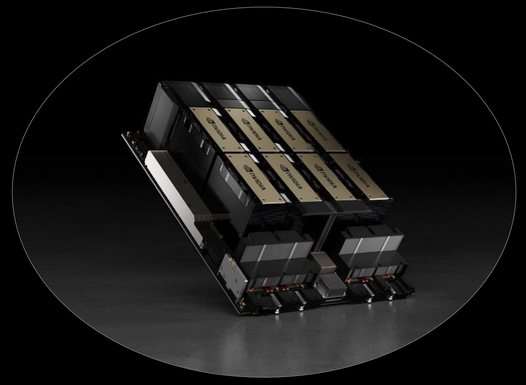
Noise Isolation for multi-tenant, diverse workflows

- Diverse workloads can impact each other’s performance
- Spectrum-X detects congestion spots in real time
- Programmable congestion control meters the data flow
- Results in performance isolation across workloads



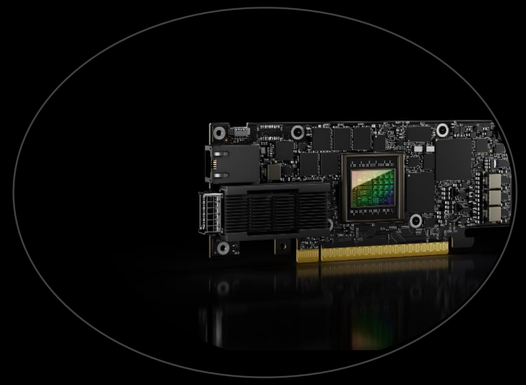
Israel-1 Spectrum-X Generative AI Cloud

Most Powerful Supercomputer in Israel with Peak Performance of 8 Exaflops



256 x HGX H100 Servers
2048 x H100 GPUs

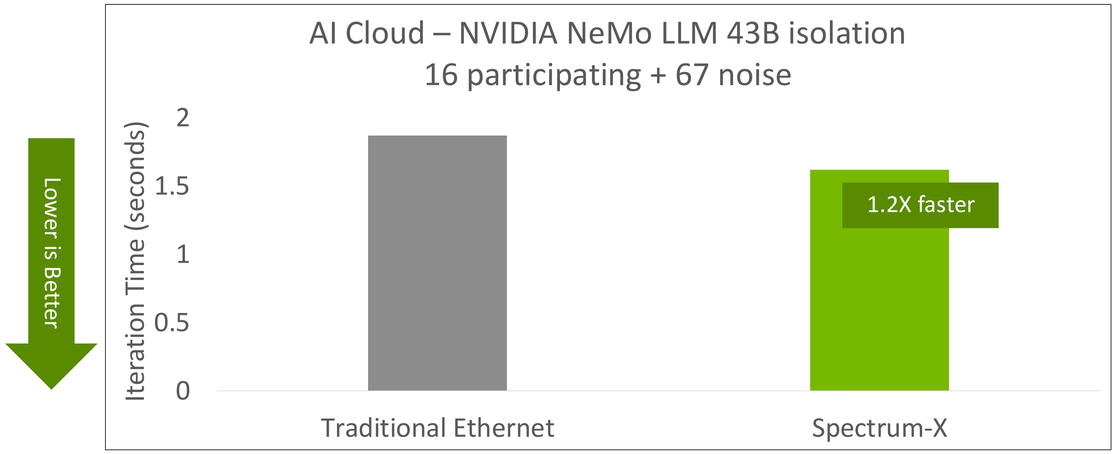
80 x Spectrum
SN5600 Switches

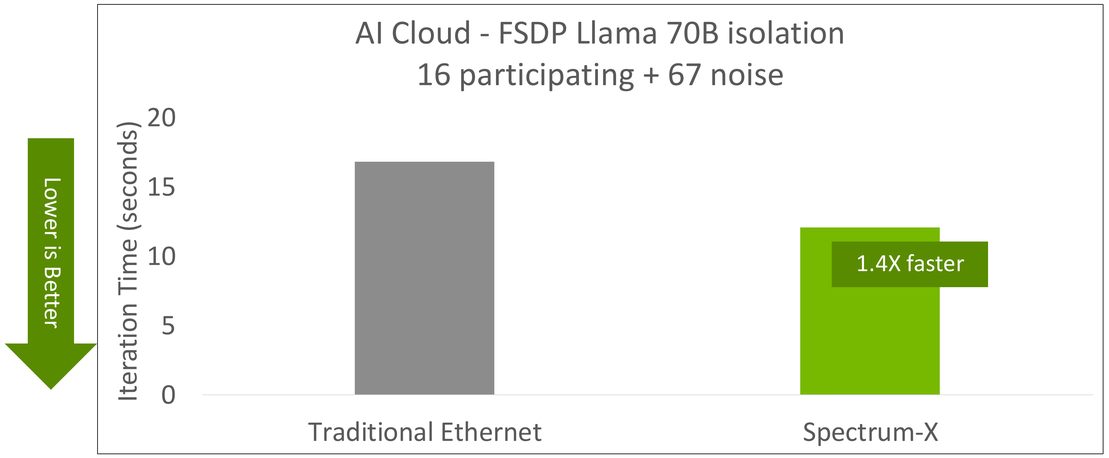


2048 x BlueField-3
SuperNICs

Spectrum-X Performance Update

AI Cloud Workload Performance Isolation





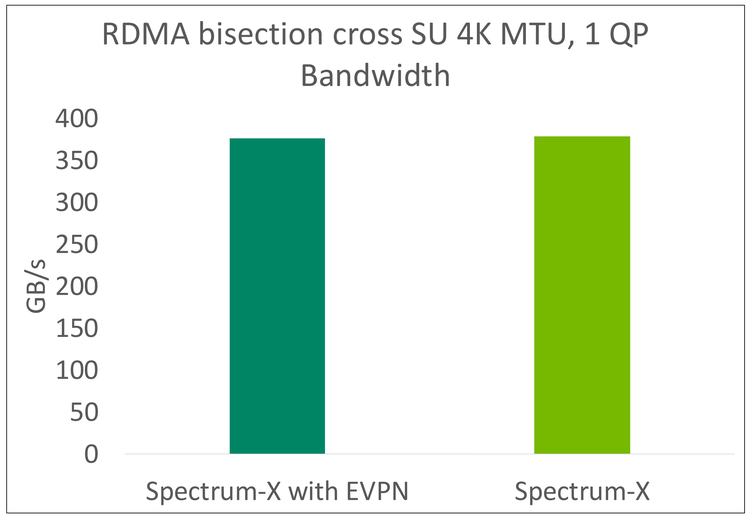
Spectrum-X accelerates iteration times for training the most common AI models such as Nemo and LLAMA.

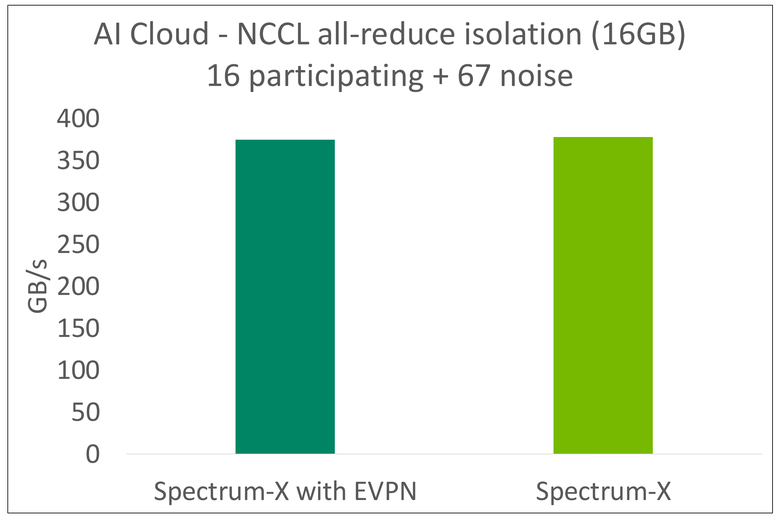
Faster training iterations lead to faster job completion times, accelerating time to insight.

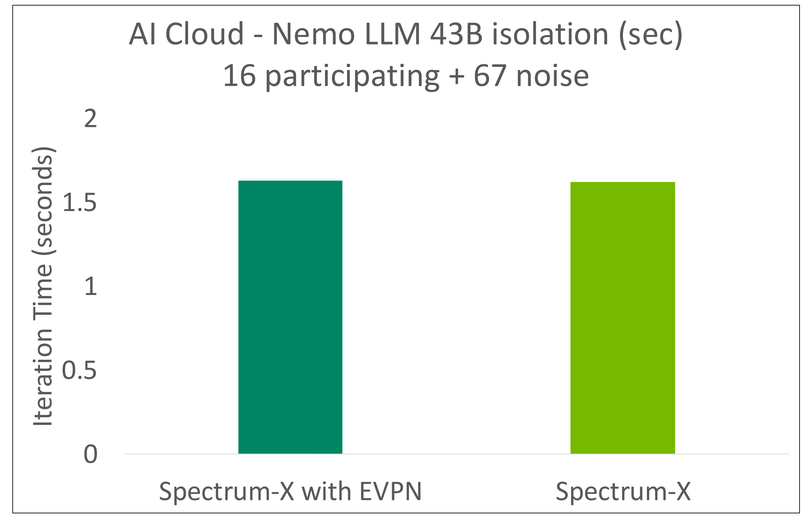
Spectrum-X accelerates training AI models in noisy AI Cloud environments

Spectrum-X Performance Update

Impact of Multi-Tenancy on AI Cloud Performance







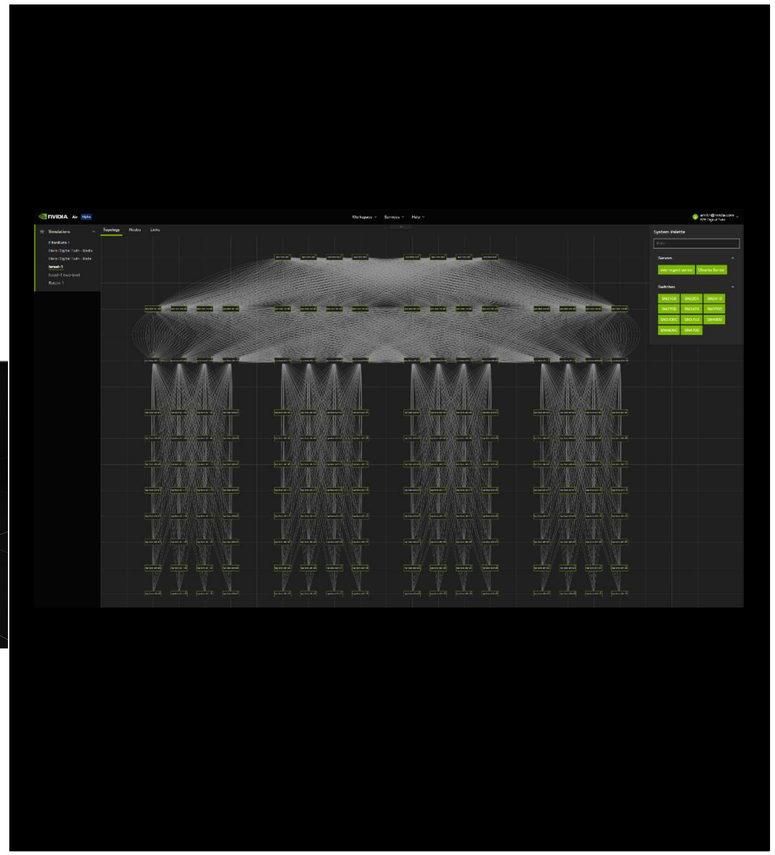
AI Clouds are multi-tenant architectures and require network virtualization.

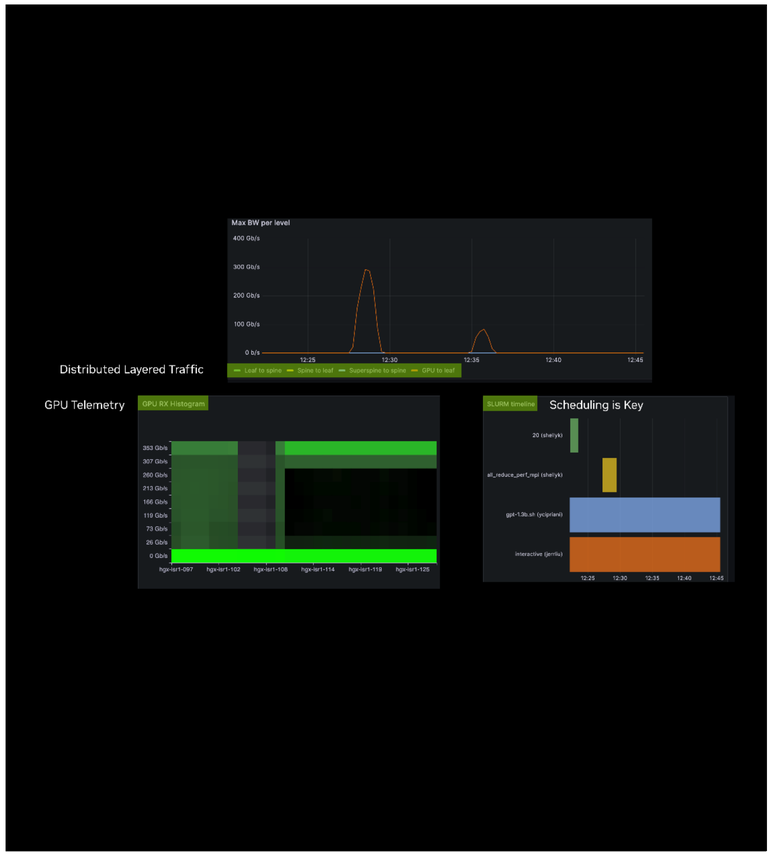
Across the application stack, Spectrum-X performance is essentially the same (within 1%) when implementing EVPN VXLAN.

Spectrum-X maintains highest performance in multi-tenant environments

The road to 100K’s of GPU’s and beyond

Fastest Time to Train



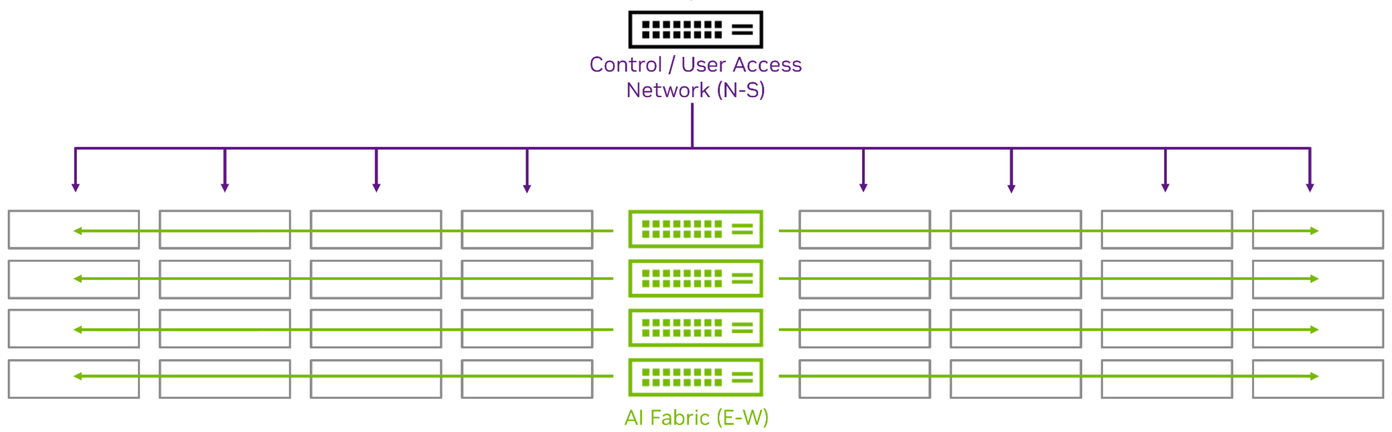


100,000 GPU commercial cluster built with SpectrumX

Validated, Optimized SpectrumX Reference Architecture running Training Jobs in Production

N-S Network is the Gateway to AI Success

Ensure Cloud-scale Efficiency with NVIDIA Spectrum Ethernet for N-S



North-South Networking Requirements:

Software stack built for automation and simulation

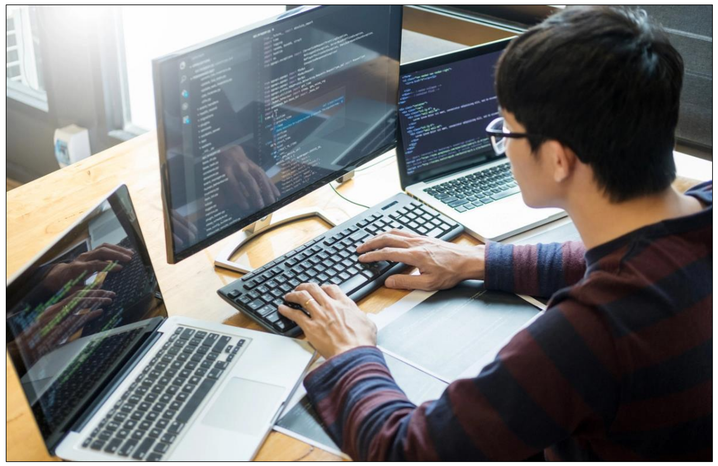
Accelerates GPU to AI storage traffic

Unified management across fabrics

NVIDIA Spectrum Software Stack

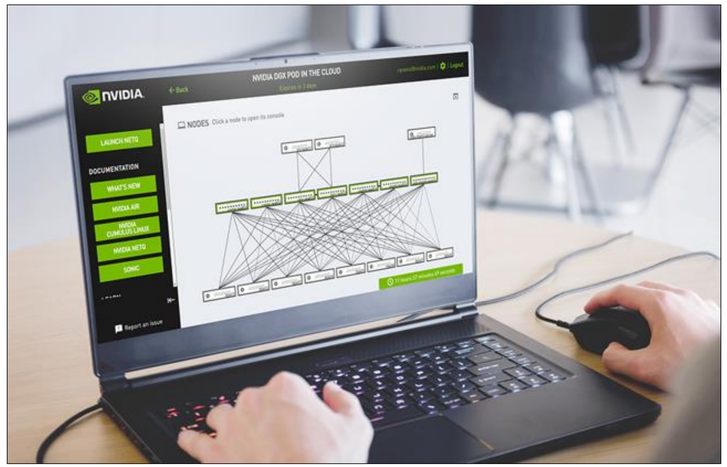
End-to-End Operational Efficiency

Cumulus Linux



Flagship Ethernet NOS
Built for Automation

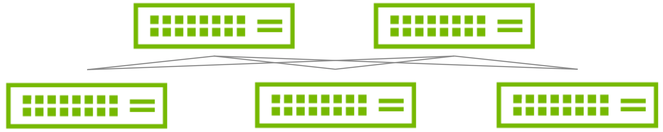
Air



DC Digital Twin Platform
Built for Network Simulation

NetQ

Fabric Validation Toolset
Built for Visibility

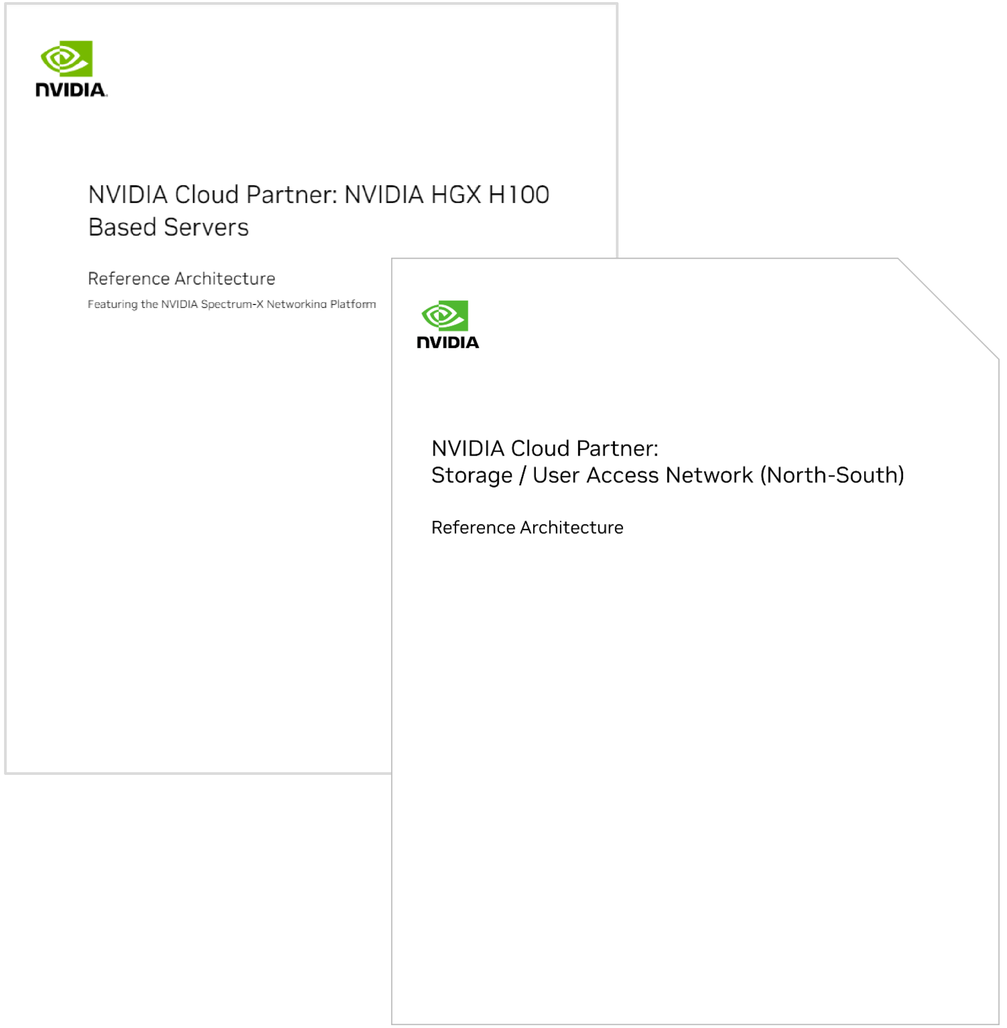


NVIDIA Cloud Partner Reference Architectures

Two RAs based on NVIDIA Spectrum and NVIDIA BlueField Networking

- Two Networking Reference Architectures for NVIDIA Cloud Partners
  - Spectrum-X for “E-W” AI Fabrics
  - Spectrum Ethernet for “N-S” Storage/User Access

- Both E-W and N-S networks are based on the same Ethernet stack
  - Cumulus Linux
  - NVUE Object Model
  - NetQ
  - NVIDIA Air
  - DOCA



- One OS
- One API
- One Protocol Stack

Leading OEMs and Ecosystem Partners Adopt Spectrum-X

- Dell, HPE, Lenovo, and Supermicro are integrating Spectrum-X into their data center offerings
- Spectrum-X will be available with HGX H100/200 and OVX L40S systems
- Systems feature the complete NVIDIA AI stack

DELL Technologies

Hewlett Packard Enterprise

Lenovo

SUPERMICRO

Key Takeaways

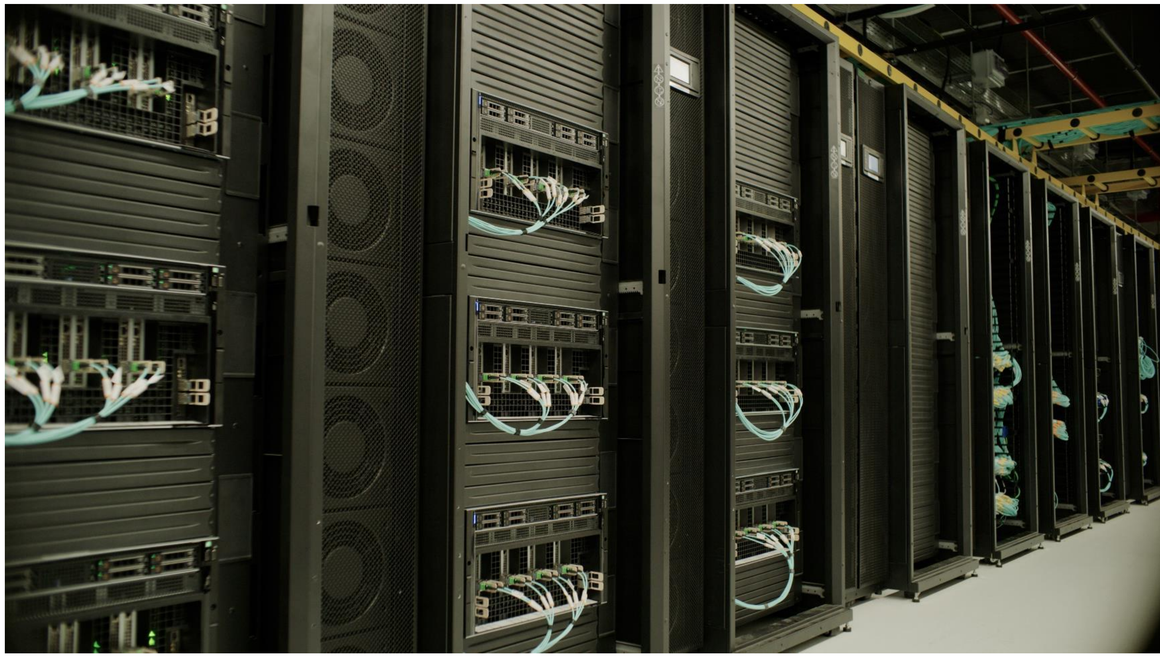
NVIDIA High Performance Platforms

- Learn more about NVIDIA SpectrumX and NVIDIA High Performance Networking – www.nvidia.com/en-us/networking/spectrumx

- Align your datacenter to the NVIDIA AI Reference Architecture. Future-proof your AI infrastructure.
  - Server: BlueField-3 DPU (for storage) and SuperNIC (Compute)
  - Fabric: Spectrum-4 switches

- IBM and NVIDIA are working to bring IBM as a certified storage partner to SpectrumX Reference Design. Talk to IBM and to NVIDIA to learn more.

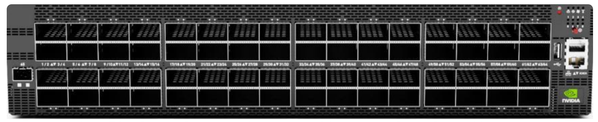
- Sign up for GTC updates - www.nvidia.com/gtc



Learn More About Spectrum-X

Available Resources

NVIDIA Spectrum-X Networking Platform



NVIDIA Spectrum SN5600
Ethernet Switch

NVIDIA BlueField-3
SuperNIC

Spectrum-X Intro Video

Data Center Digital Twin Israel-1 Video

Spectrum-X Datasheet

Spectrum-X Technical Whitepaper

Spectrum-X Webpage

Spectrum-X Technical Blog

Networking for AI Whitepaper

BlueField SuperNIC Whitepaper

Thank You

bwebb@nvidia.com