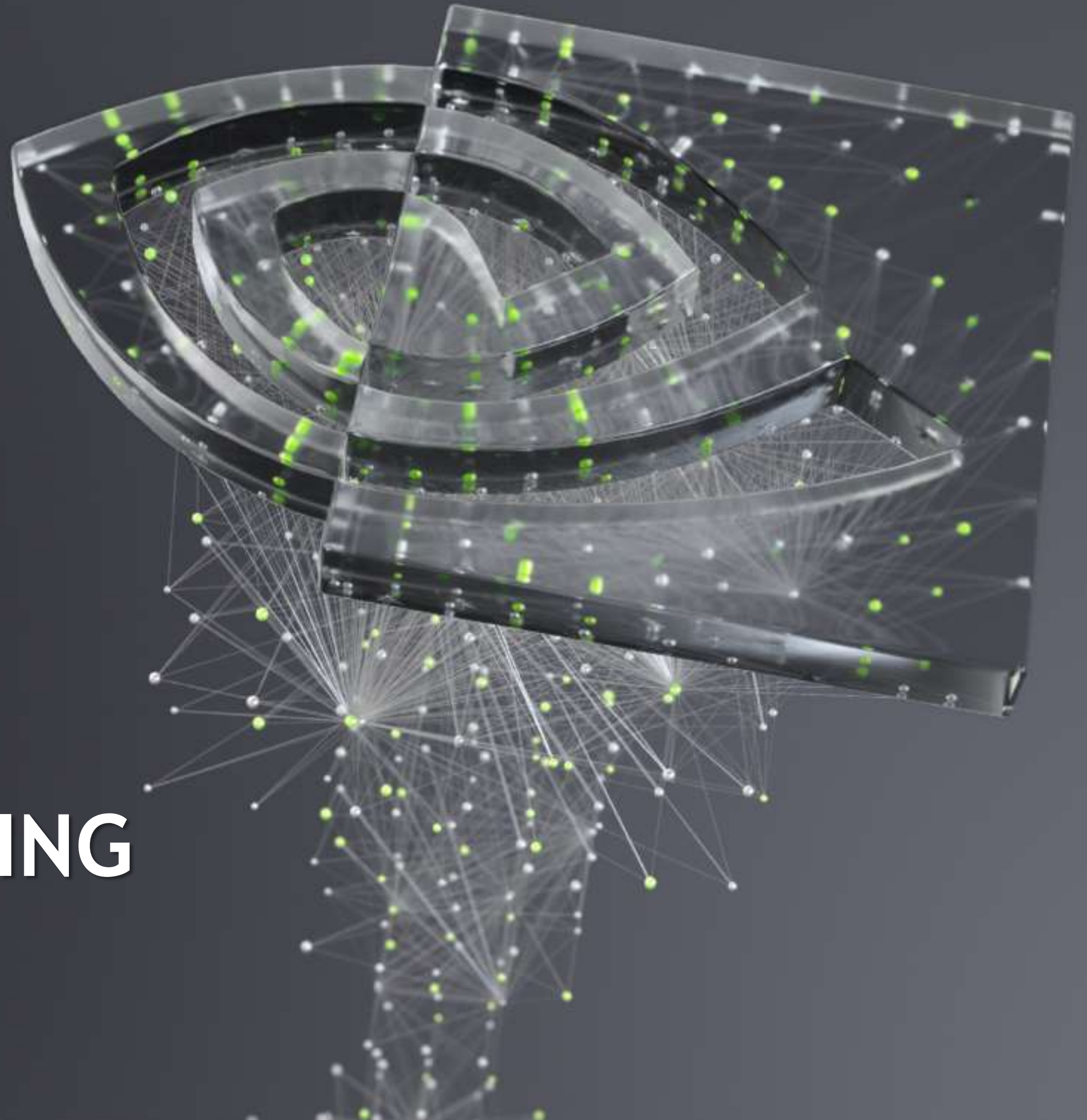


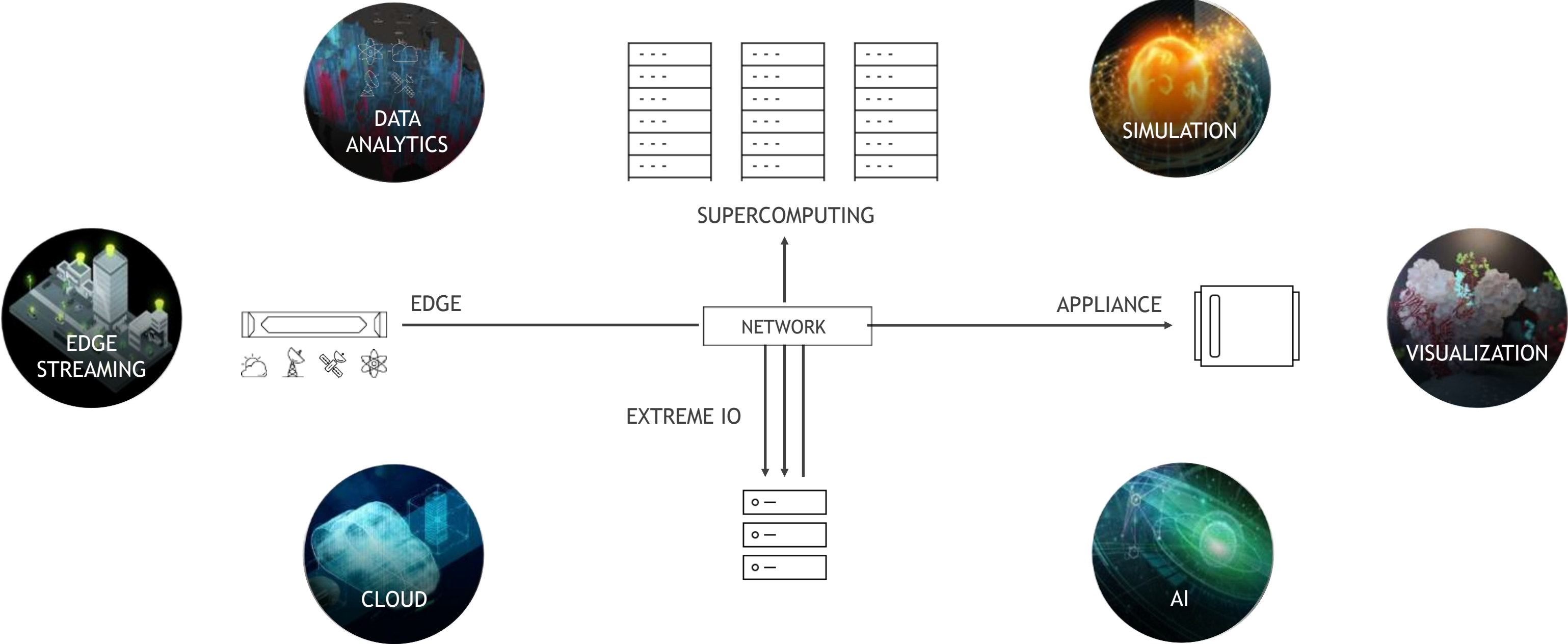


ACCELERATING SCIENTIFIC COMPUTING

DPU Accelerated Supercomputing



EXPANDING UNIVERSE OF SCIENTIFIC COMPUTING



HDR 200G INFINIBAND ACCELERATES NEXT GENERATION HPC AND AI SUPERCOMPUTERS (EXAMPLES)



9 PFlops
3K HDR Nodes
Dragonfly+ Topology



19.5 PetaFLOPS
2.5K HDR Nodes
Dragonfly+ and Fat Tree



16 PFlops
3K HDR Nodes
Dragonfly+ Topology



8K HDR Nodes
Dragonfly+ Topology



35.5 PFlops
2K HDR Nodes
Fat-Tree Topology



19.3 PFlops
5.6K HDR Nodes
Dragonfly+ Topology



63.5 PFlops
4.5K HDR Nodes
Fat-Tree Topology



HPC/AI Cloud
HDR InfiniBand

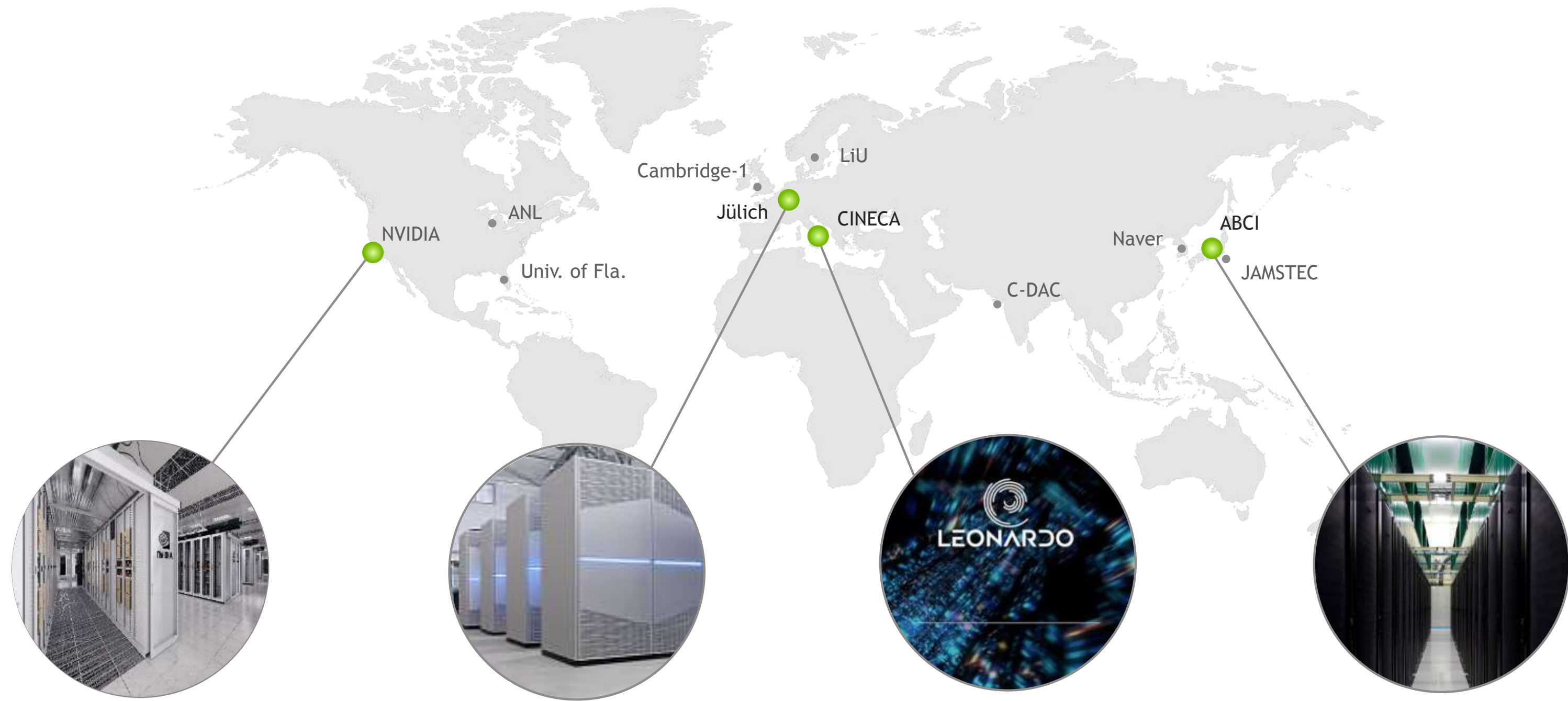


HDR InfiniBand
Supercomputers



23.5 PFlops
8K HDR Nodes
Fat-Tree Topology

NVIDIA PLATFORM POWERING THE EXASCALE AI SUPERCOMPUTERS



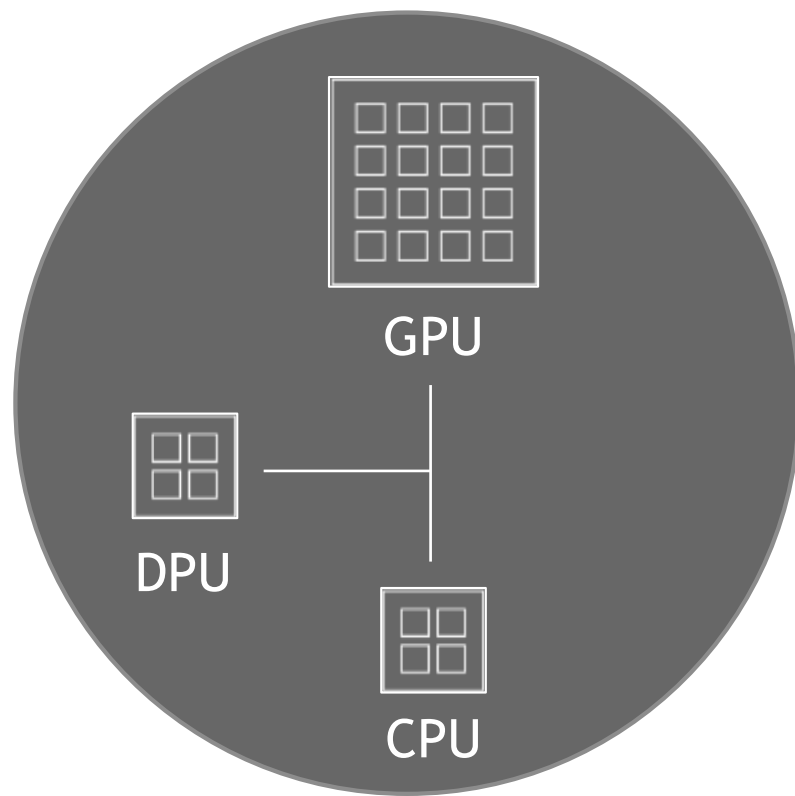
NVIDIA
2.8 EFLOPS AI Perf
HDR 200G InfiniBand

Jülich
2.3 EFLOPS AI Perf HDR
HDR 200G InfiniBand

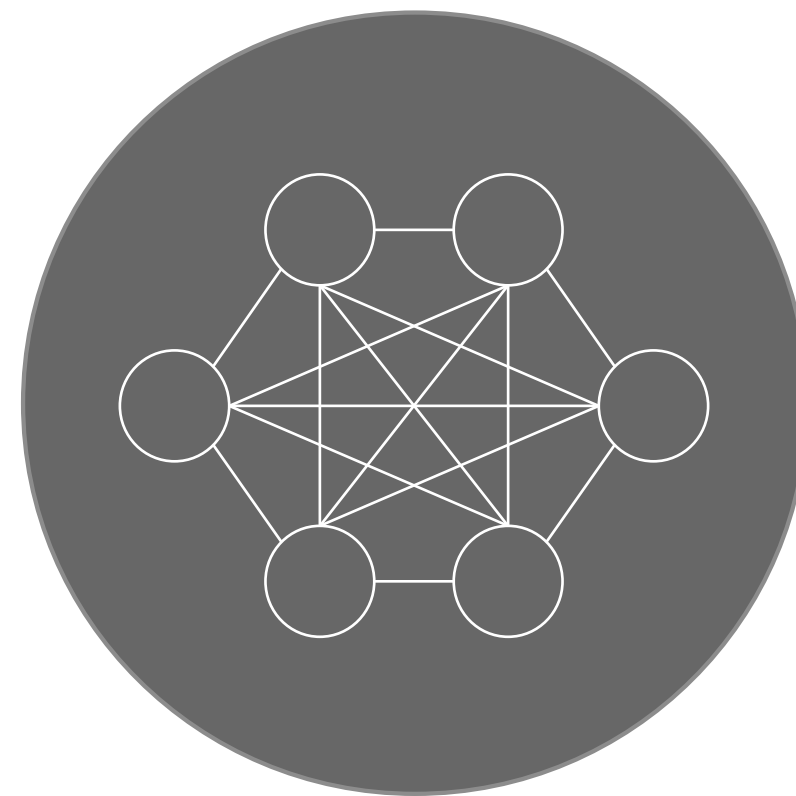
CINECA
10 EFLOPS AI Perf
HDR 200G InfiniBand

ABCI
600 PFLOPS AI Perf
HDR 200G InfiniBand

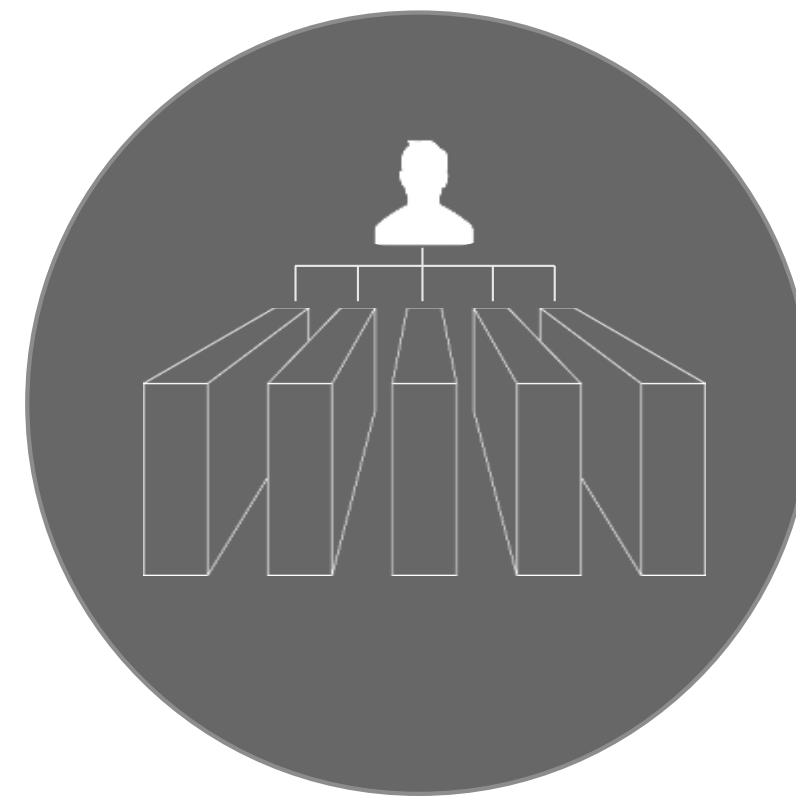
INFINIBAND TECHNOLOGY FUNDAMENTALS



Smart Networking



Architected to Scale

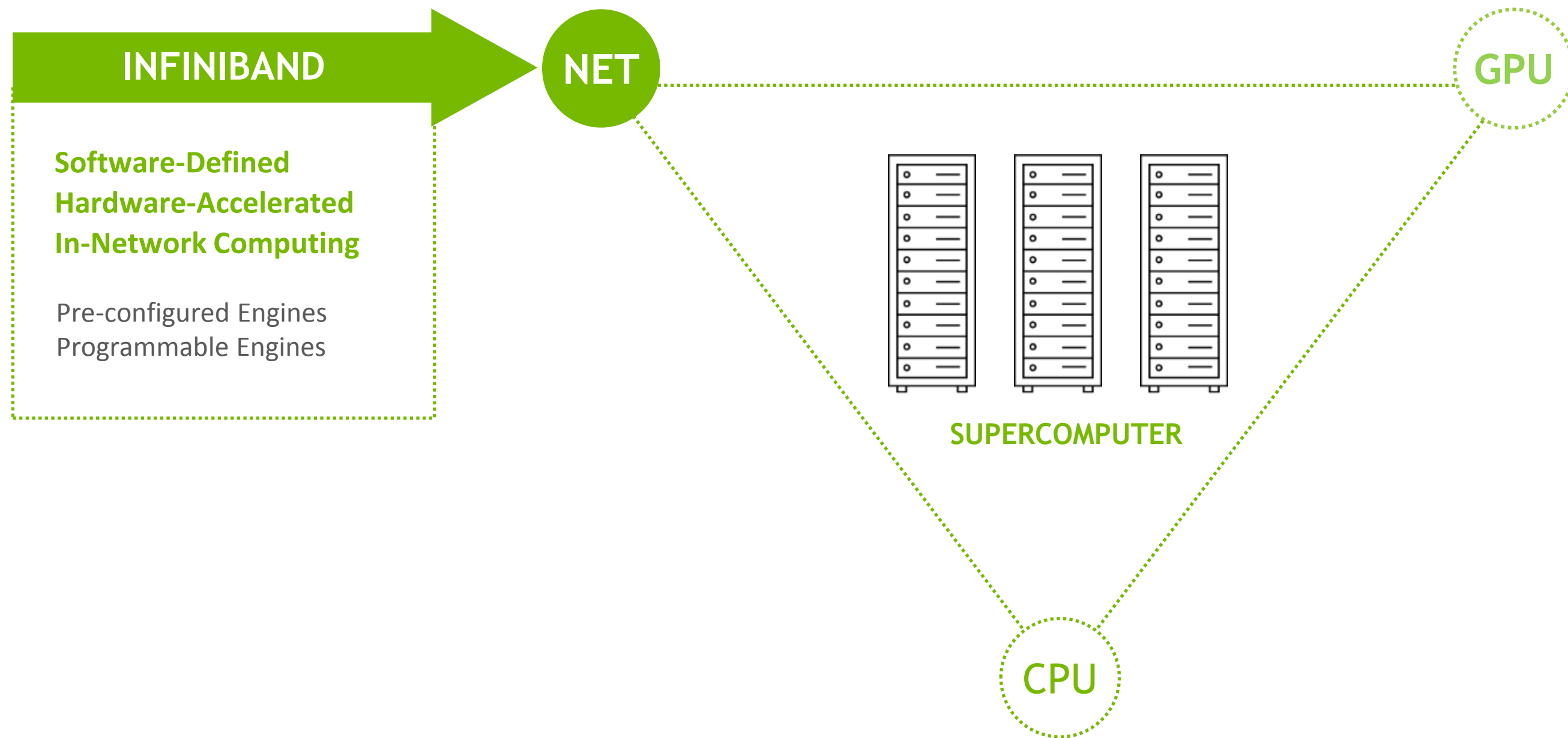


Centralized Management

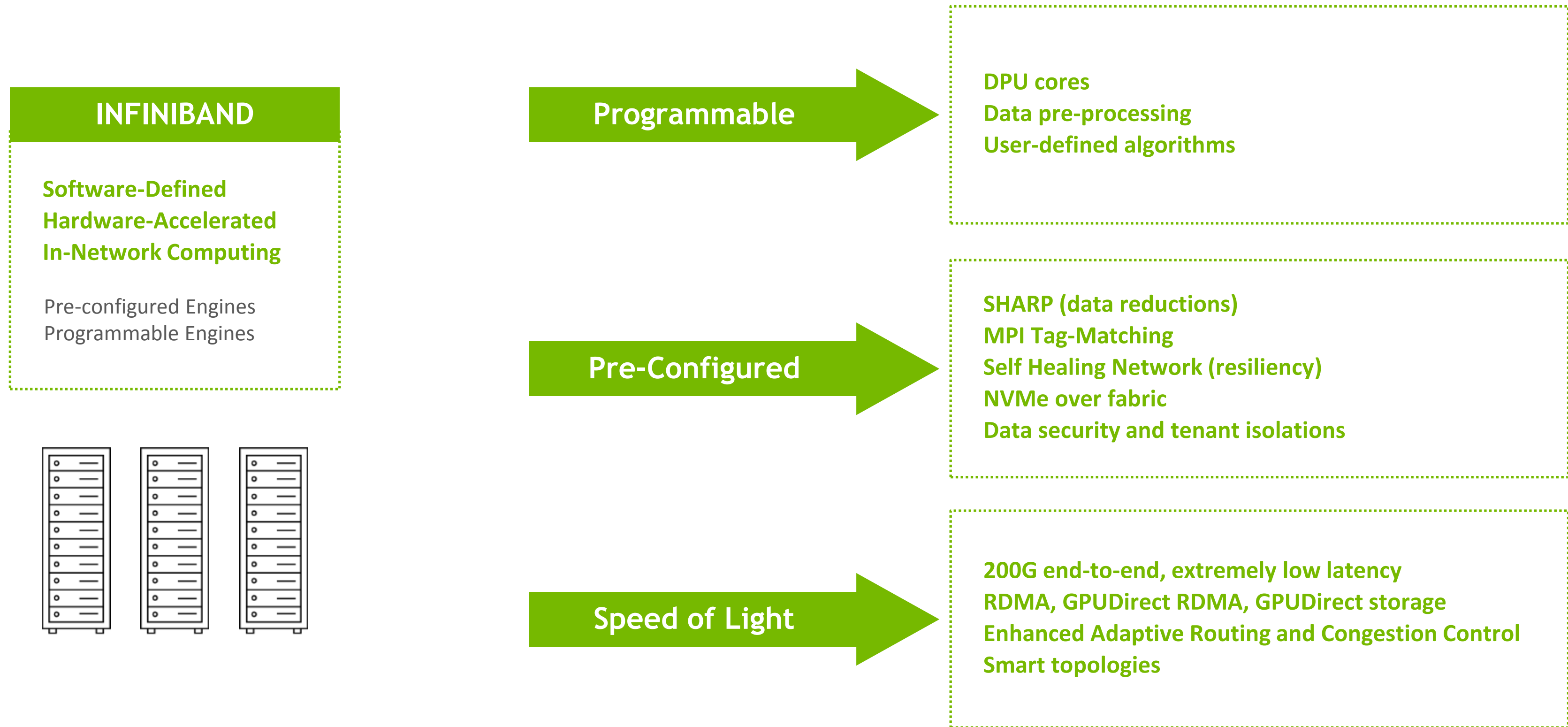


Standard

IN-NETWORK COMPUTING ACCELERATED SUPERCOMPUTING



IN-NETWORK COMPUTING ACCELERATED SUPERCOMPUTING

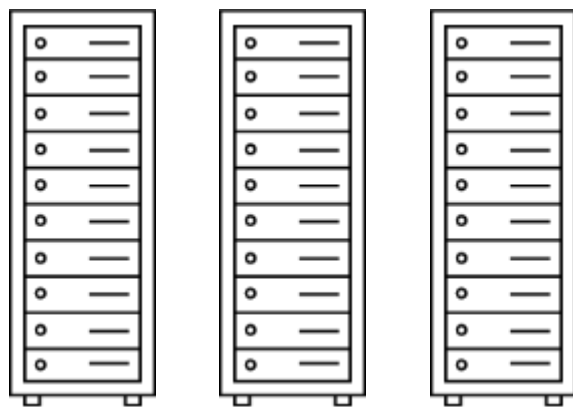


IN-NETWORK COMPUTING ACCELERATED SUPERCOMPUTING

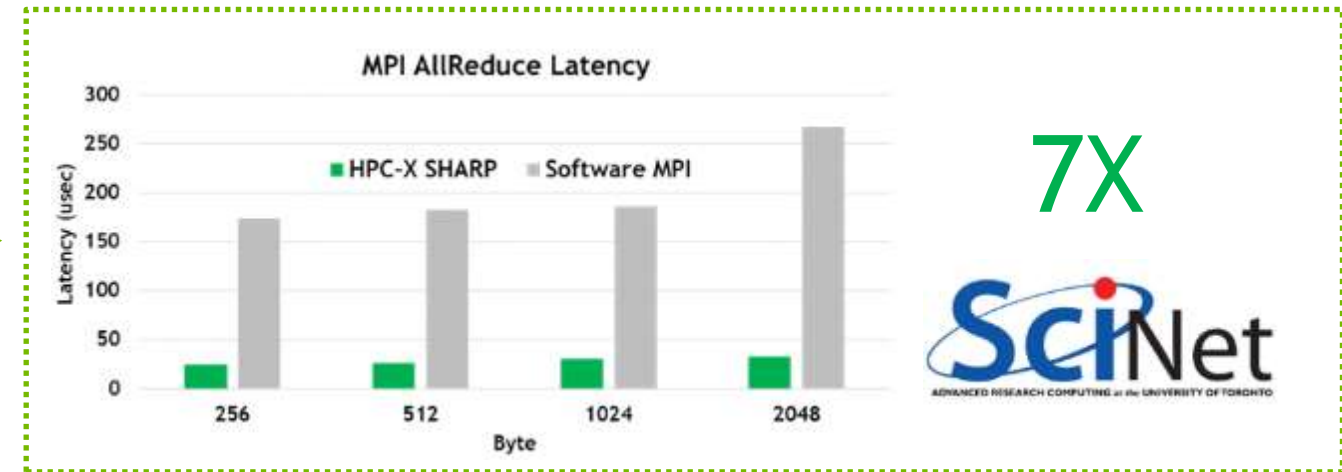
INFINIBAND

Software-Defined
Hardware-Accelerated
In-Network Computing

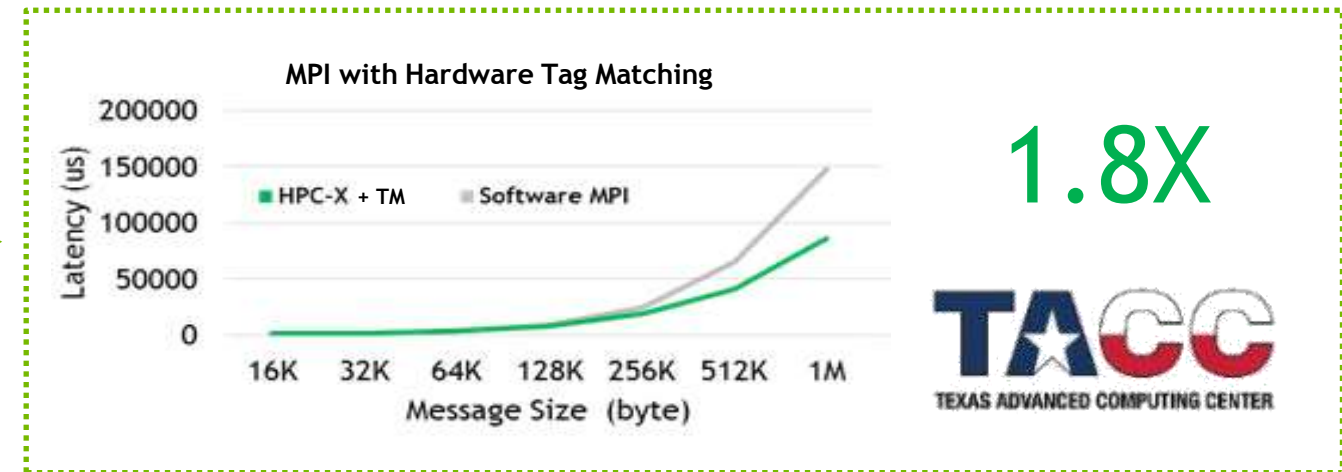
Pre-configured Engines
Programmable Engines



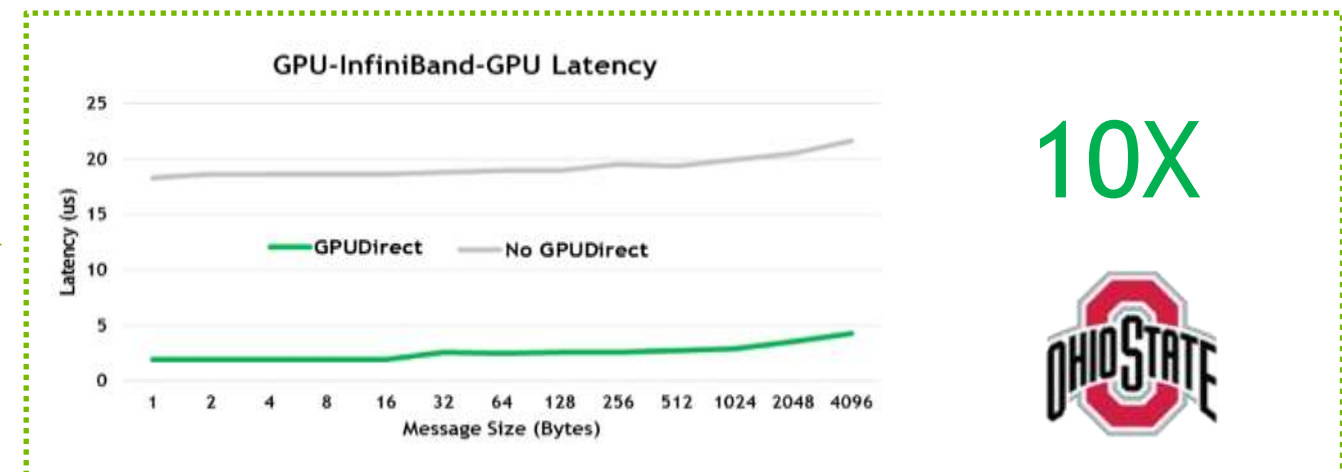
SHARP



MPI Tag-Matching



GPU Direct



HDR 200G INFINIBAND DEVICES



ConnectX-6 Adapter

HDR 200G InfiniBand
PCIe Gen3 and Gen4
In-Network Computing



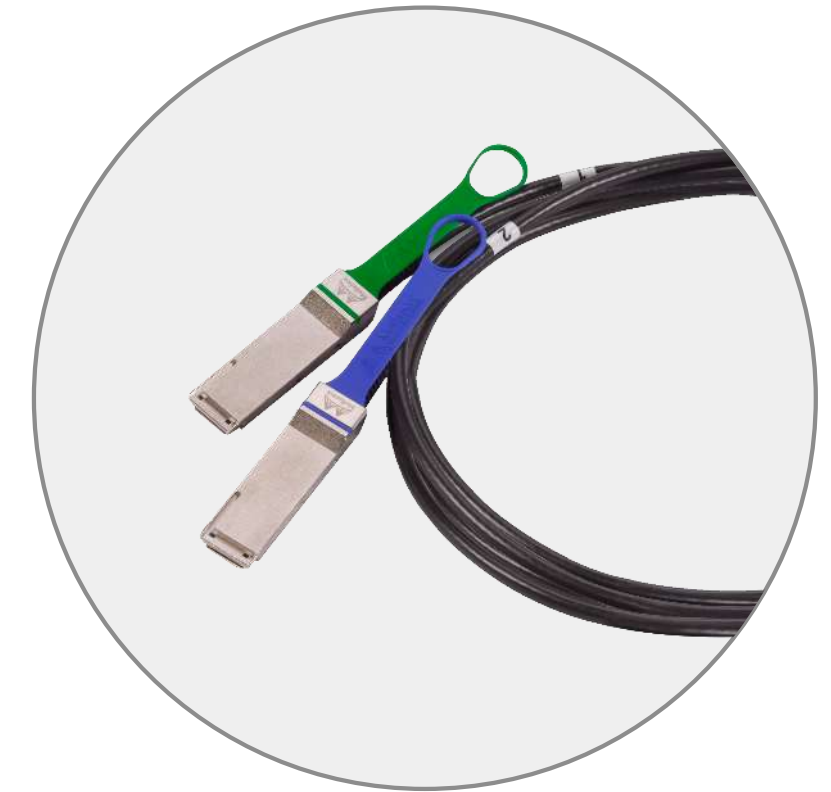
BlueField-2 DPU

HDR 200G InfiniBand with Arm Cores
PCIe Gen3 and Gen4, DDR4
AI Application Accelerators
In-Network Computing



Quantum Switch

40-ports HDR 200G InfiniBand
80-ports 100G HDR100
In-Network Computing
1U system: 40 x 200G, 80 x 100G
Modular: 800 x 200G, 1600 x 100G



LinkX Cables

Copper Cables
Active Copper Cables
Optical Transceivers

INFINIBAND DATA CENTER APPLIANCES



METROX-2
Extending InfiniBand to
40km Reach



SKYWAY
InfiniBand to Ethernet Gateway
8 x 100G / 200G ports InfiniBand
8 x 100G / 200G ports Ethernet



UFM CYBER-AI
Cyber Intelligence and Analytics
Management and Monitoring

UFM CYBER-AI

Management, Monitoring, Orchestration, Cyber Intelligence and Analytics

Network setup, connectivity validation and secure cable management

Automated network discovery and network provisioning

Network telemetry and traffic monitoring, congestion discovery

Performance, health and fault monitoring

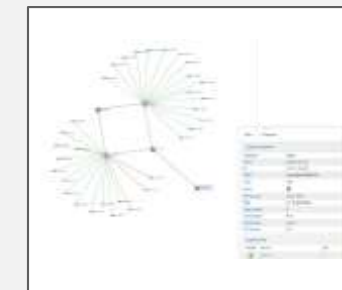
Centralized management for global software updates and configuration

Job scheduler provisioning, network provisioning

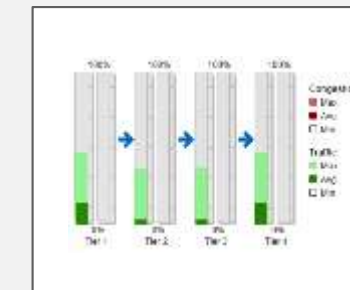
Detects performance degradations, anomalies and usage changes

Provides alerts of abnormal system and application behavior

Provides alerts for potential system failures



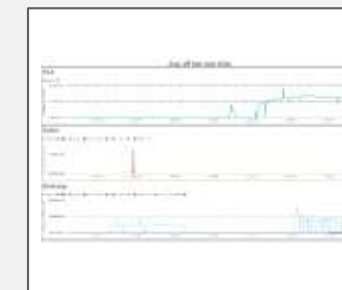
Network Validation



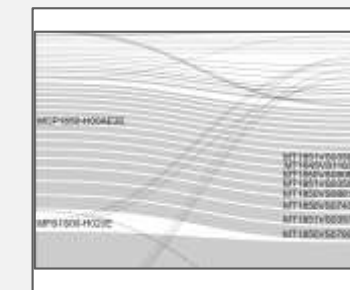
Congestion Mapping



Prediction Dashboard



Real-Time Analysis



Performance Monitoring



Secure Cable Management

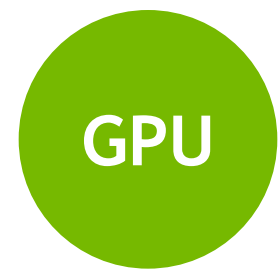




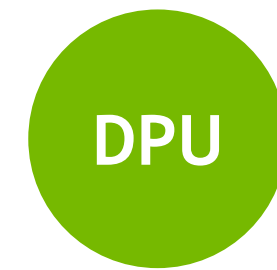
DATA PROCESSING UNITS (DPU)

REINVENTING THE DATA CENTER

The Data Center is the New Unit of Computing



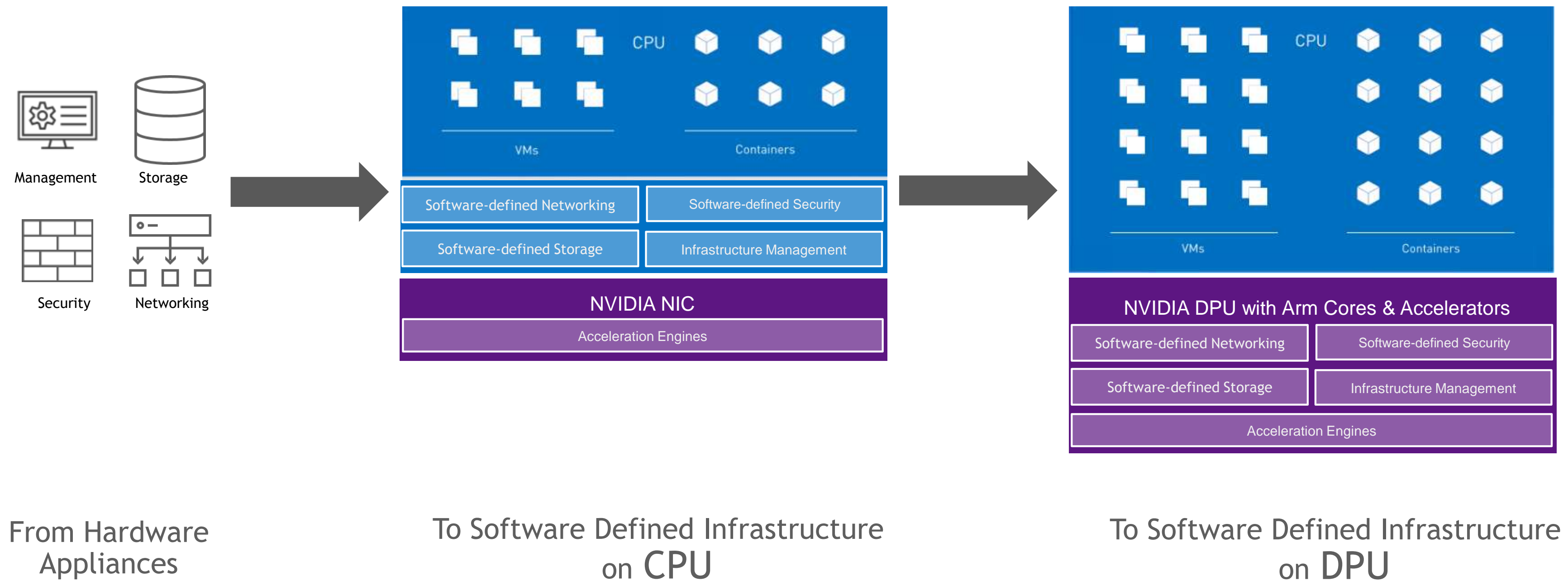
Accelerate Computing
AI & Machine Learning



Software defined,
Hardware-accelerated

INTRODUCING THE DATA PROCESSING UNIT

Software Defined Data Center Infrastructure-on-a-Chip



BLUEFIELD-2 DATA PROCESSING UNIT

Data Center Infrastructure-on-a-Chip

6.9B Transistors

8 64-bit Arm CPUs Cores

Dual 16-way VLIW Engine

100 Gbps IPsec

100 Gbps Video Streaming

5M NVMe IOPs



INTRODUCING NVIDIA DOCA

Data Center Infrastructure-on-a-Chip Architecture

Multiple OS

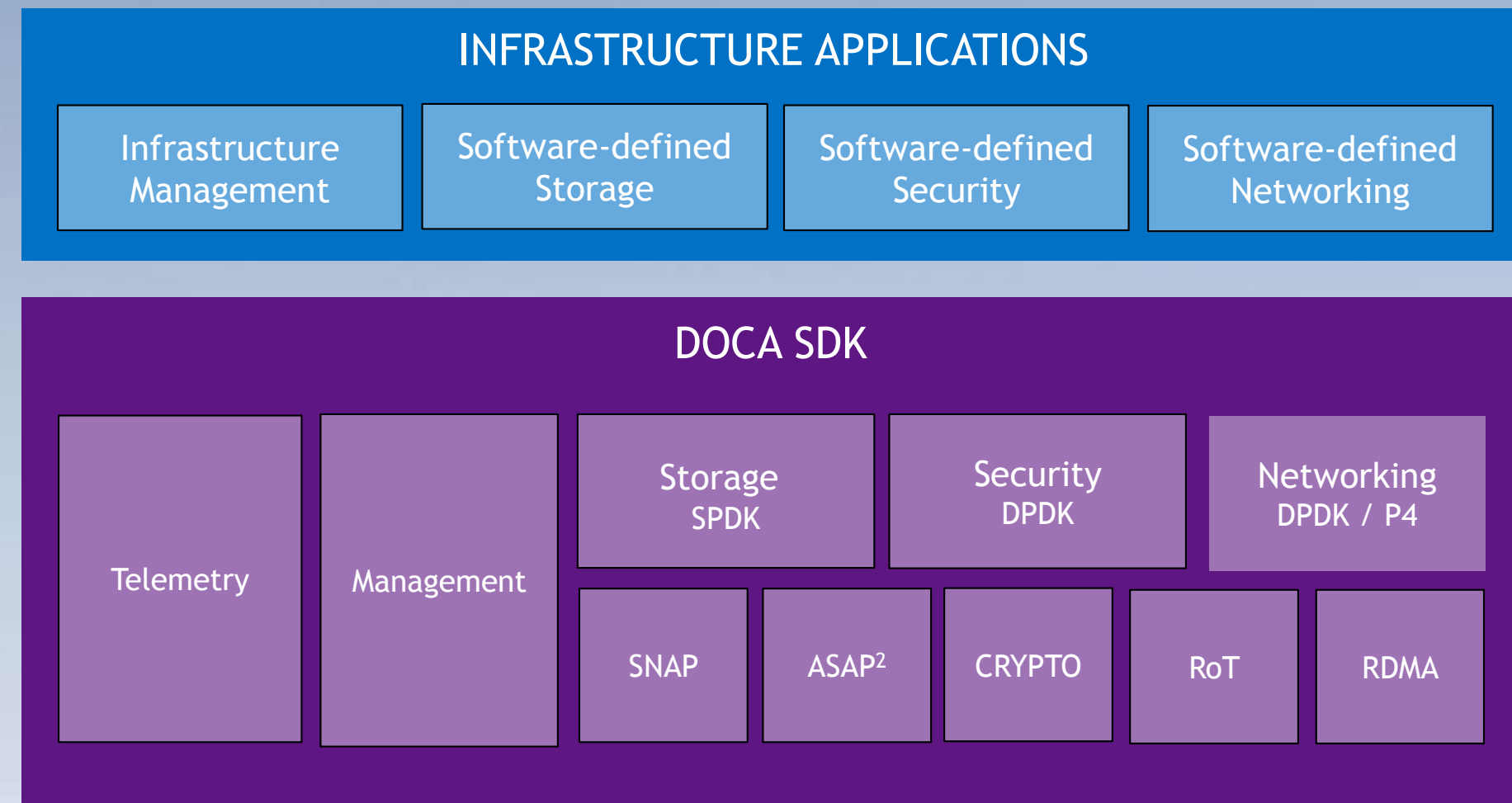
All Programmable Resources

Full Application on DPU

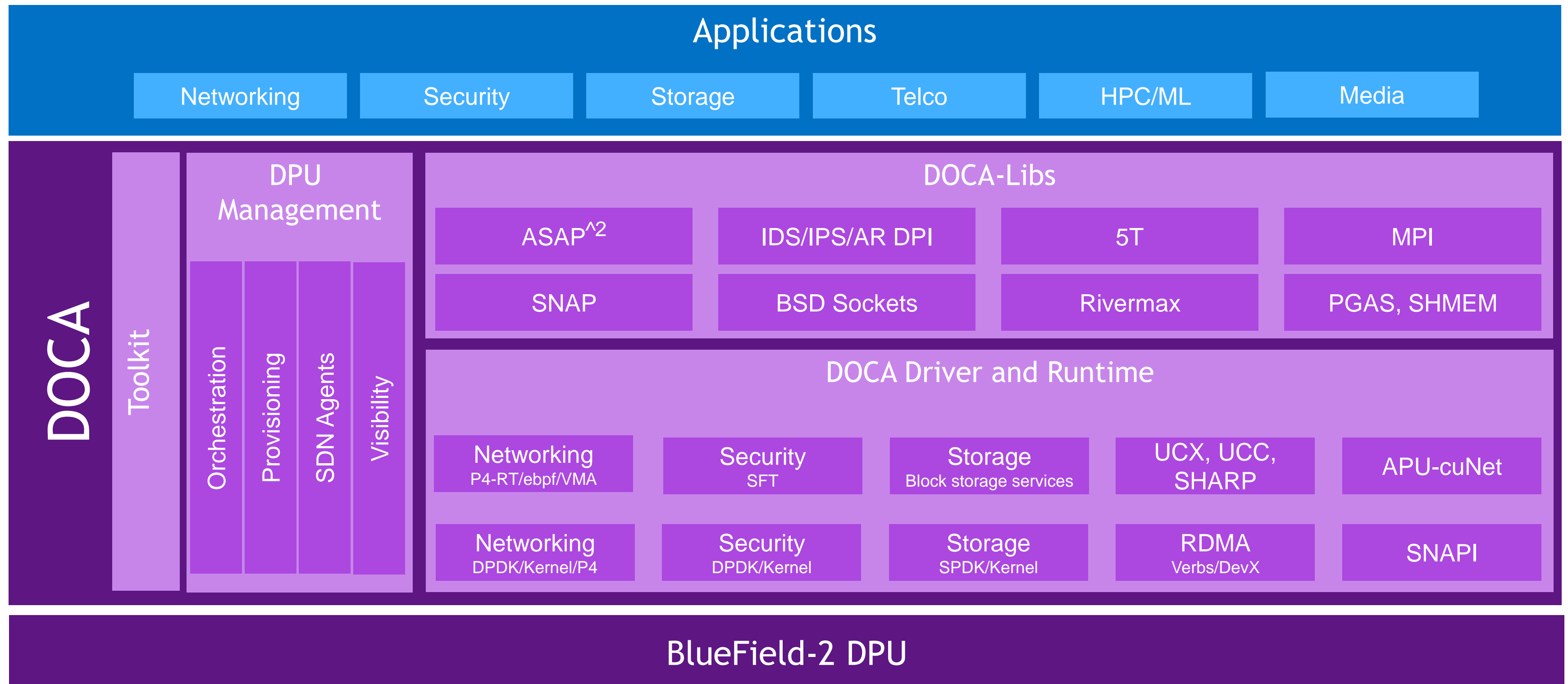
Open Source, Open APIs

Backward and Forward Compatibility

<https://developer.nvidia.com/networking/doca>



DOCA STACK



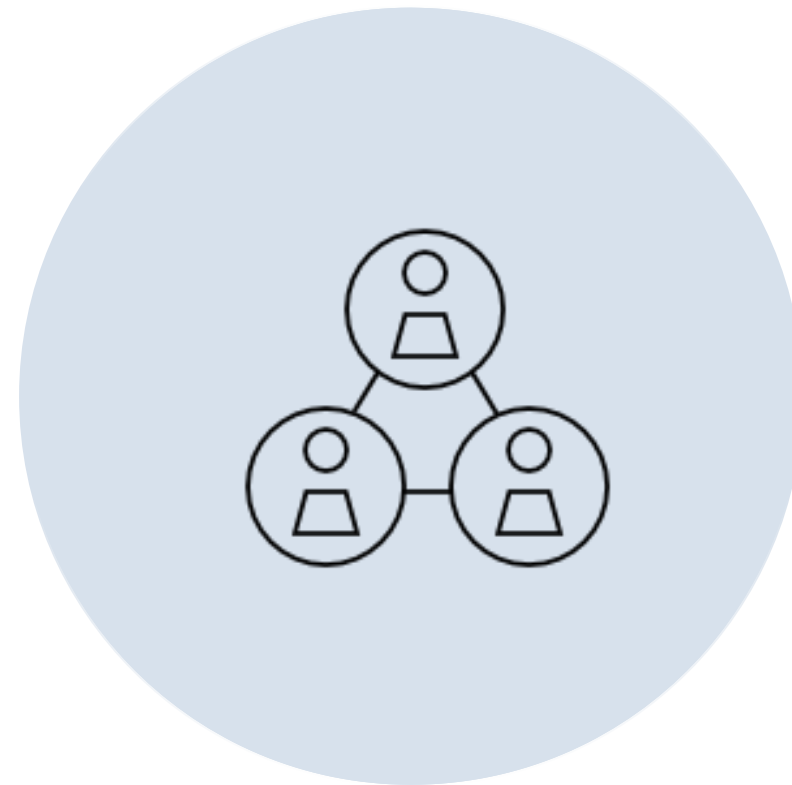


**INFINIBAND DPU
CLOUD-NATIVE SUPERCOMPUTING**

CLOUD NATIVE SUPERCOMPUTER



BARE-METAL PERFORMANCE



MULTI TENANCY

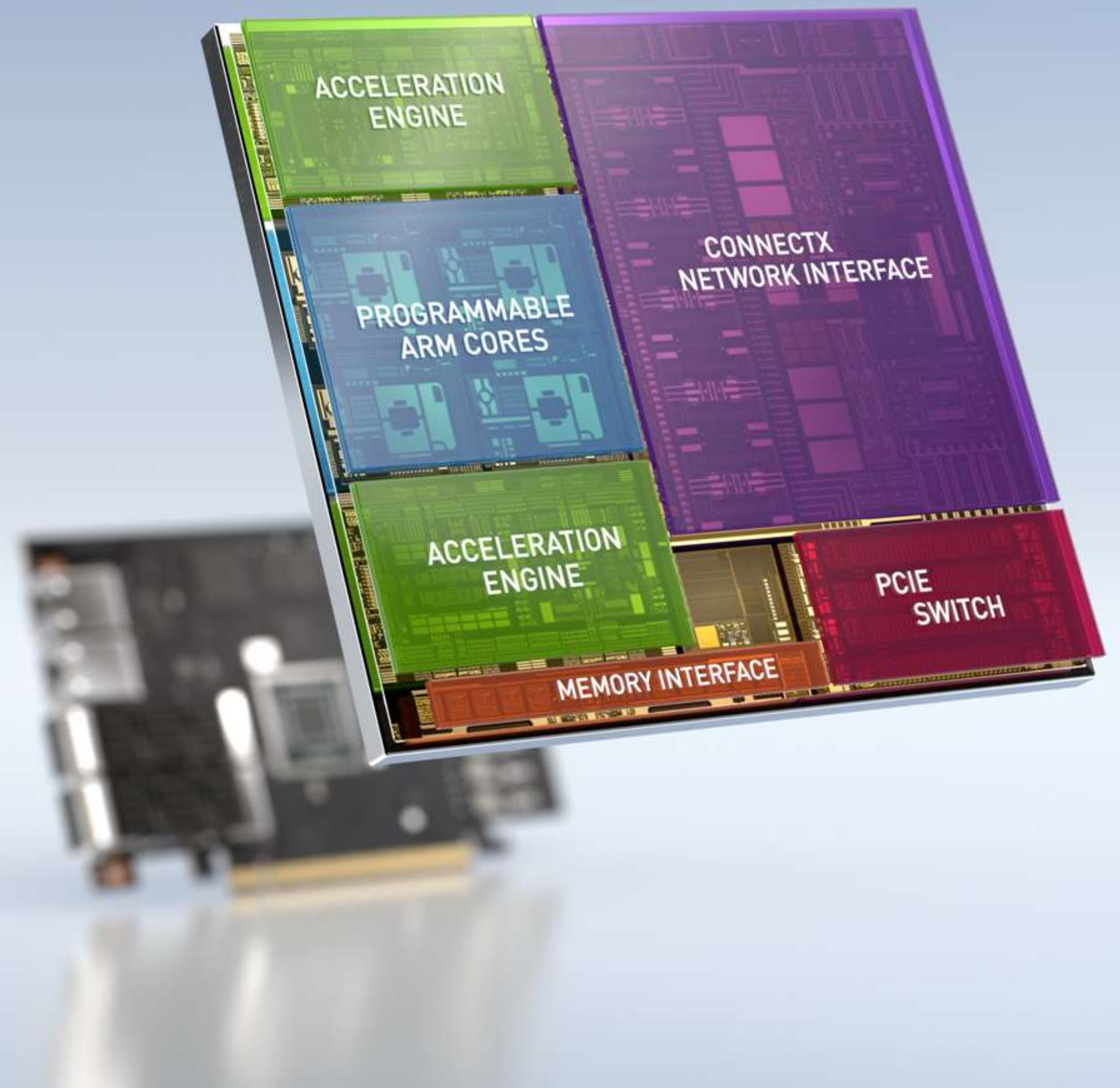


SECURITY



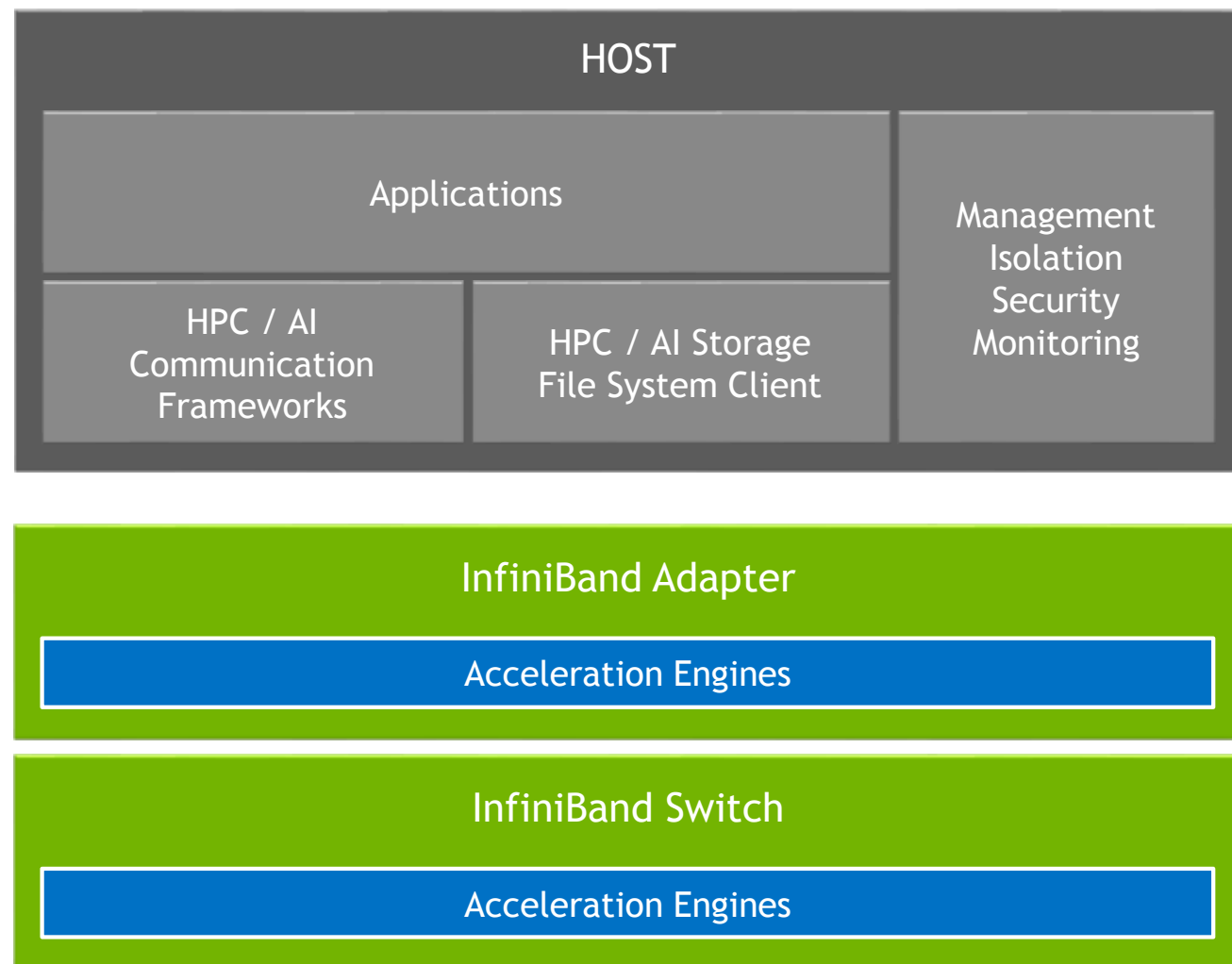
CONFIGURABLE SERVICES

BLUEFIELD DPU - THE CLOUD-NATIVE SUPERCOMPUTING INFRASTRUCTURE PLATFORM

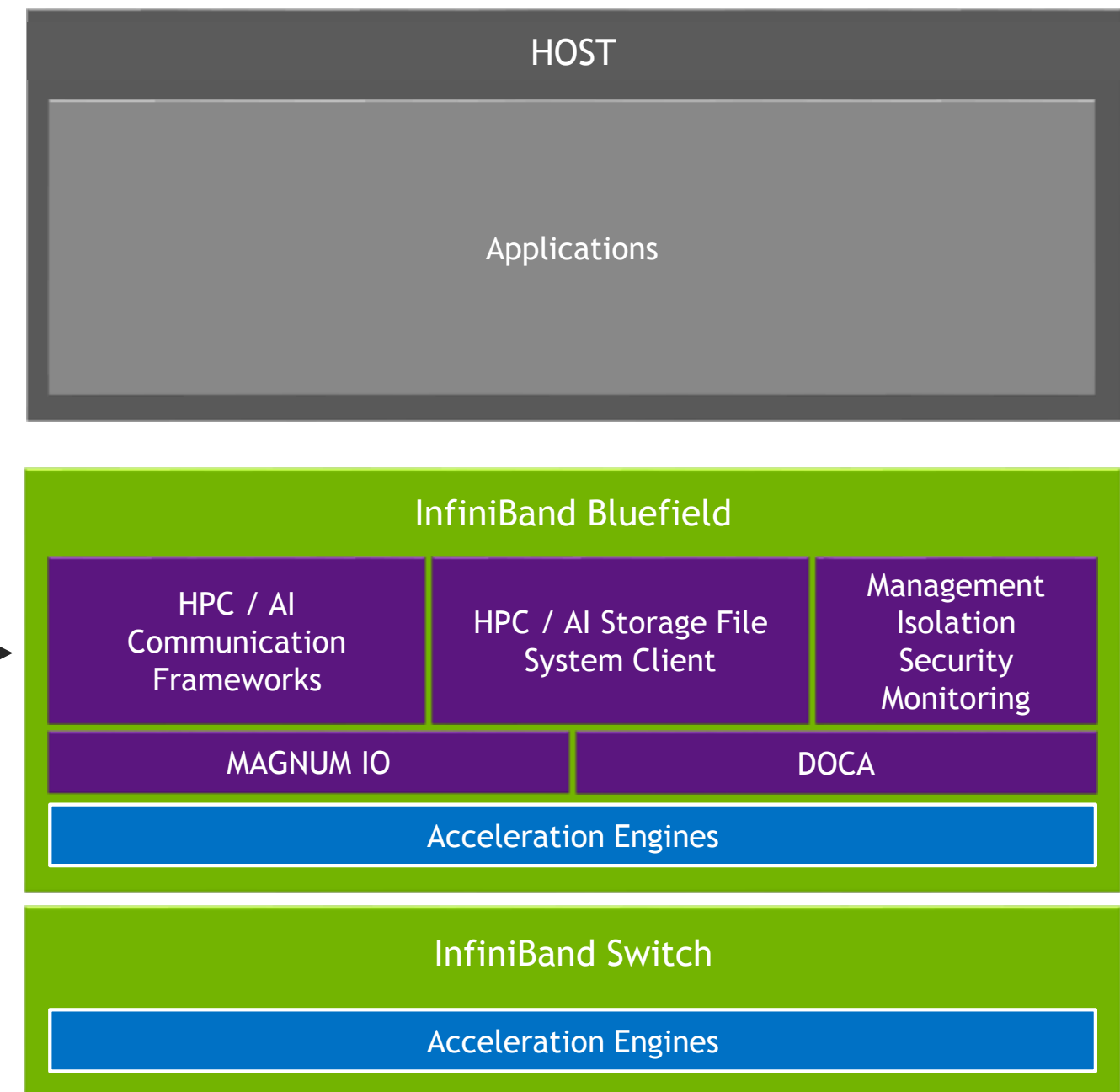


BLUEFIELD DPU - THE CLOUD NATIVE SUPERCOMPUTING INFRASTRUCTURE PLATFORM

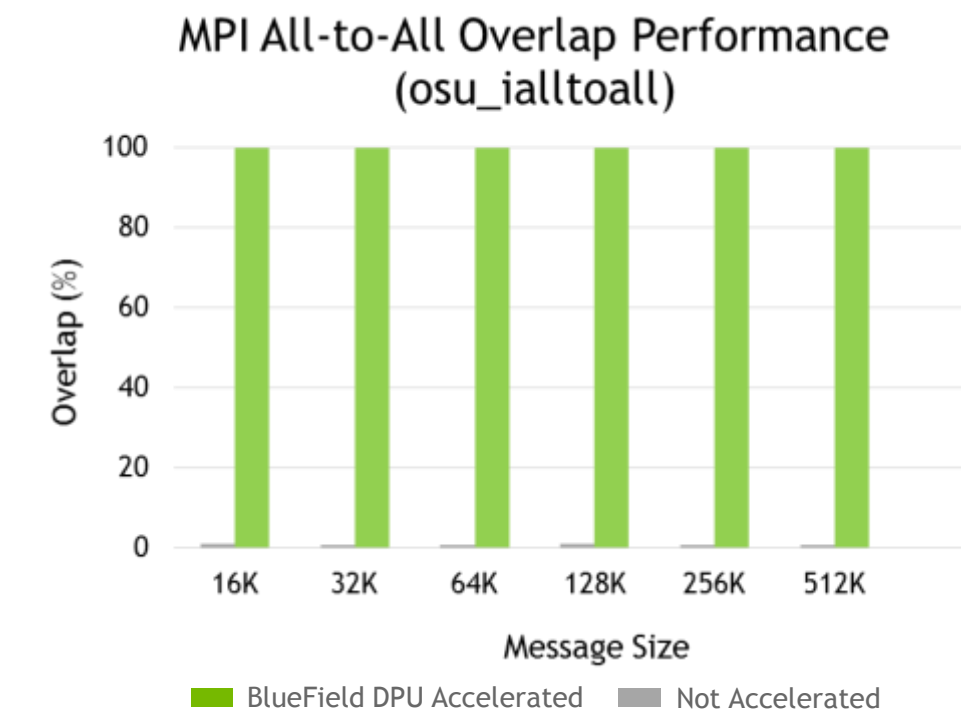
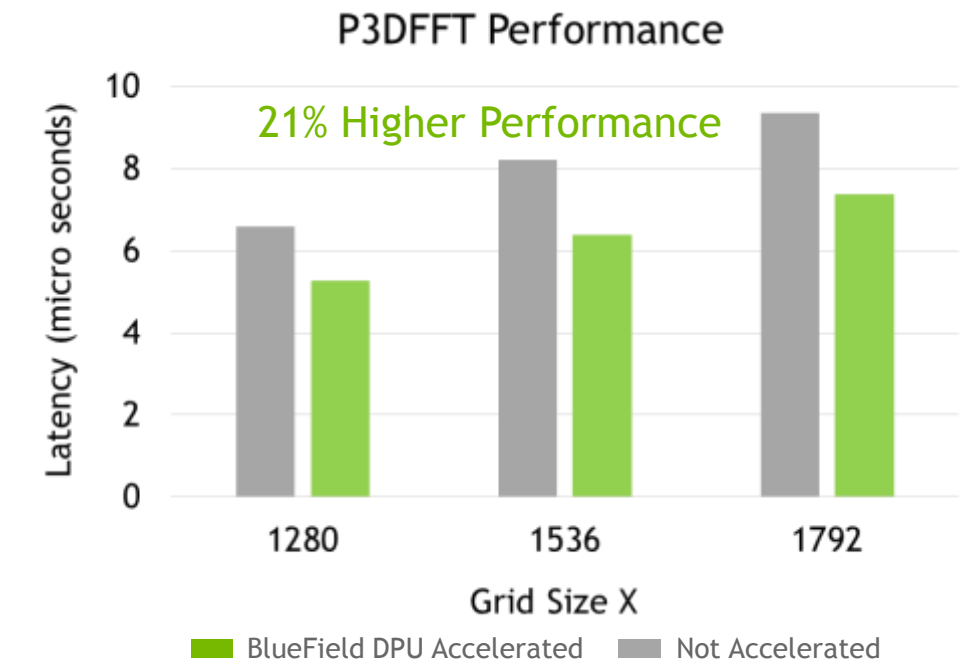
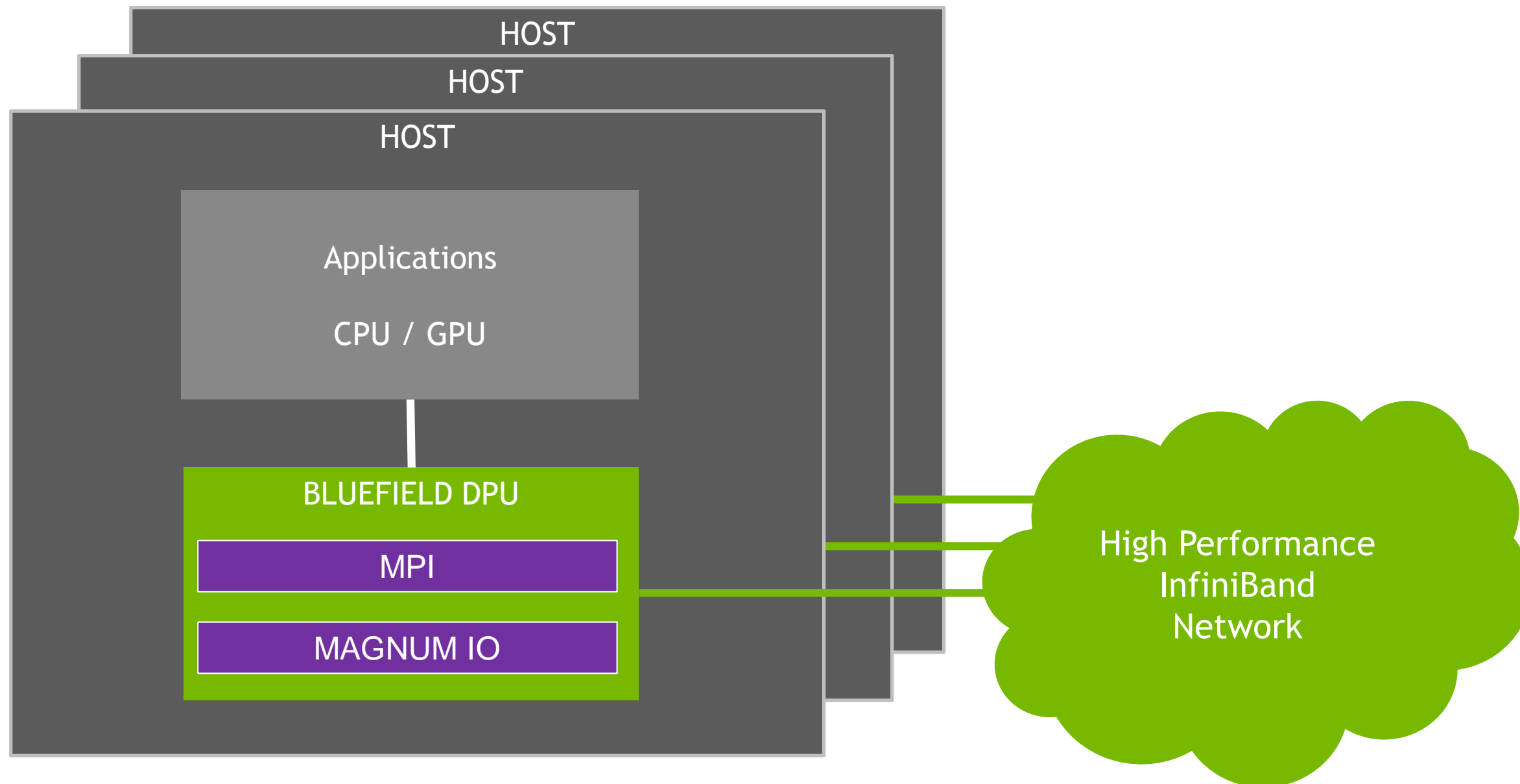
TRADITIONAL SUPERCOMPUTING



CLOUD NATIVE SUPERCOMPUTING



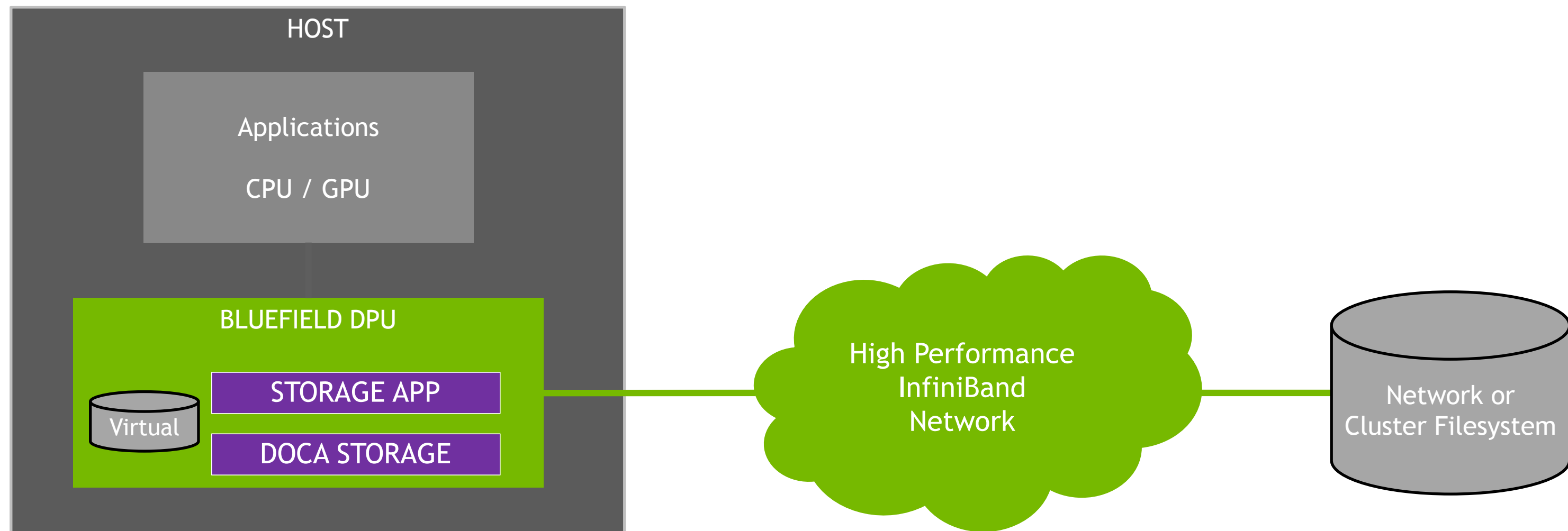
BLUEFIELD DPU - HPC AND AI COMMUNICATION FRAMEWORKS OFFLOAD



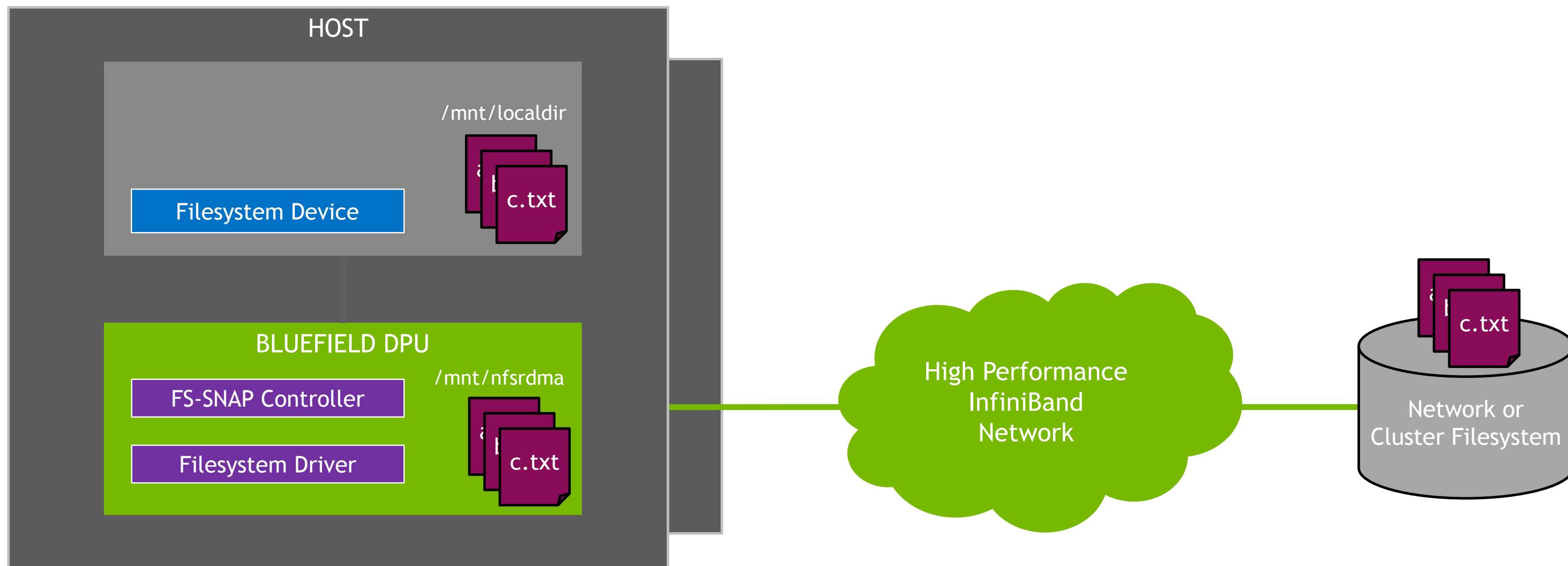
Courtesy of Ohio State University MVA PICH team and X-ScaleSolutions

Eight servers, Dual Socket Intel® Xeon® 16-core CPUs E5-2697A V4 @ 2.60 GHz (32 processes per node), NVIDIA BlueField-2 HDR100 DPUs and ConnectX-6 HDR100 adapters, NVIDIA Mellanox HDR Quantum Switch QM7800 40-Port 200Gb/s HDR InfiniBand, 256GB DDR4 2400MHz RDIMMs memory and 1TB 7.2K RPM SATA 2.5" hard drive per node.

BLUEFIELD DPU - HPC STORAGE FRAMEWORKS VIRTUALIZATION AND OFFLOAD



BLUEFIELD DPU - HPC STORAGE FRAMEWORKS VIRTUALIZATION AND OFFLOAD





**SCALABLE HIERARCHICAL
AGGREGATION AND
REDUCTION PROTOCOL
(SHARP)**

SCALABLE HIERARCHICAL AGGREGATION AND REDUCTION PROTOCOL (SHARP)

In-network Tree based aggregation mechanism

Multiple simultaneous outstanding operations

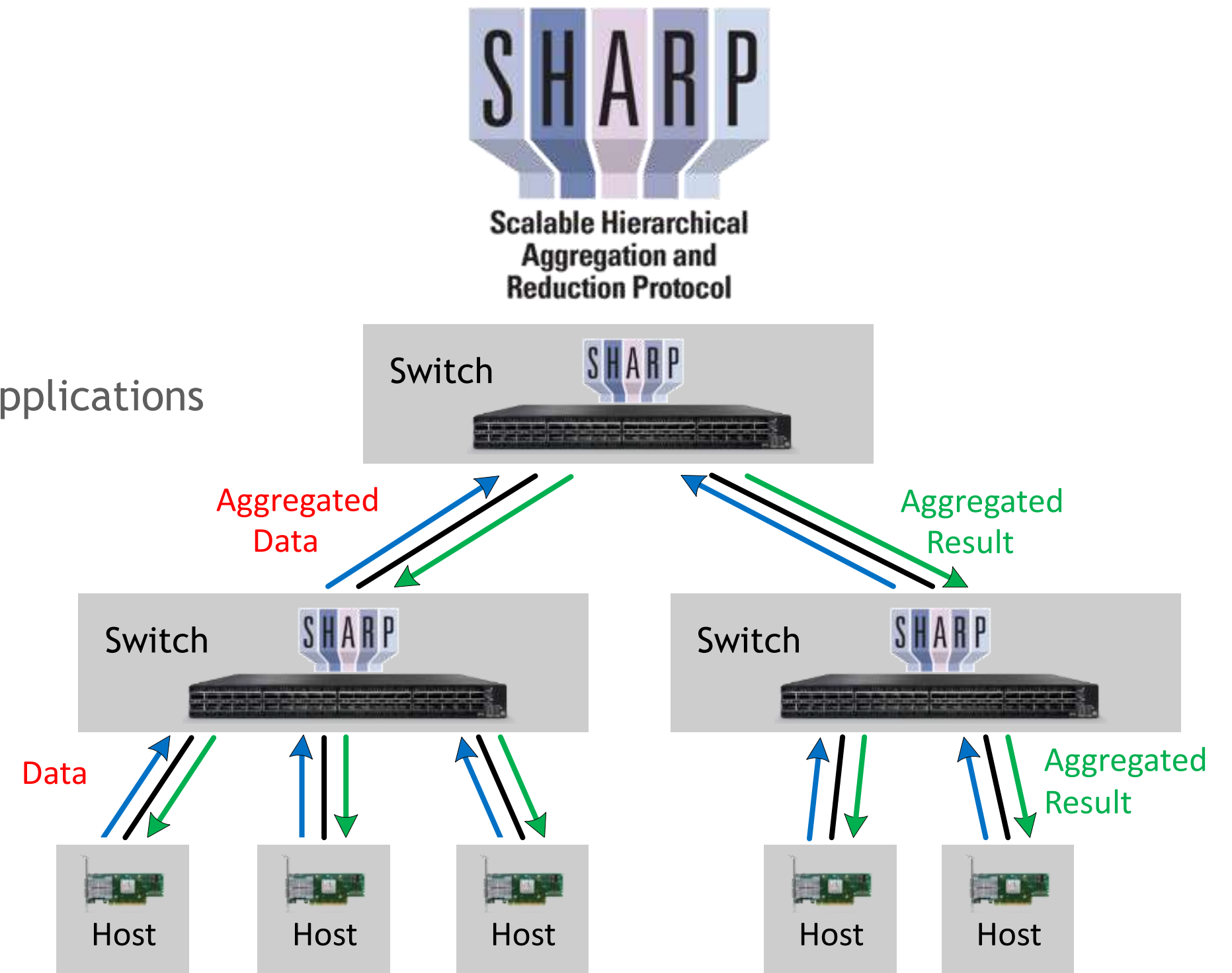
For HPC (MPI / SHMEM) and Distributed Machine Learning applications

Scalable High Performance Collective Offload

Barrier, Reduce, All-Reduce, Broadcast and more

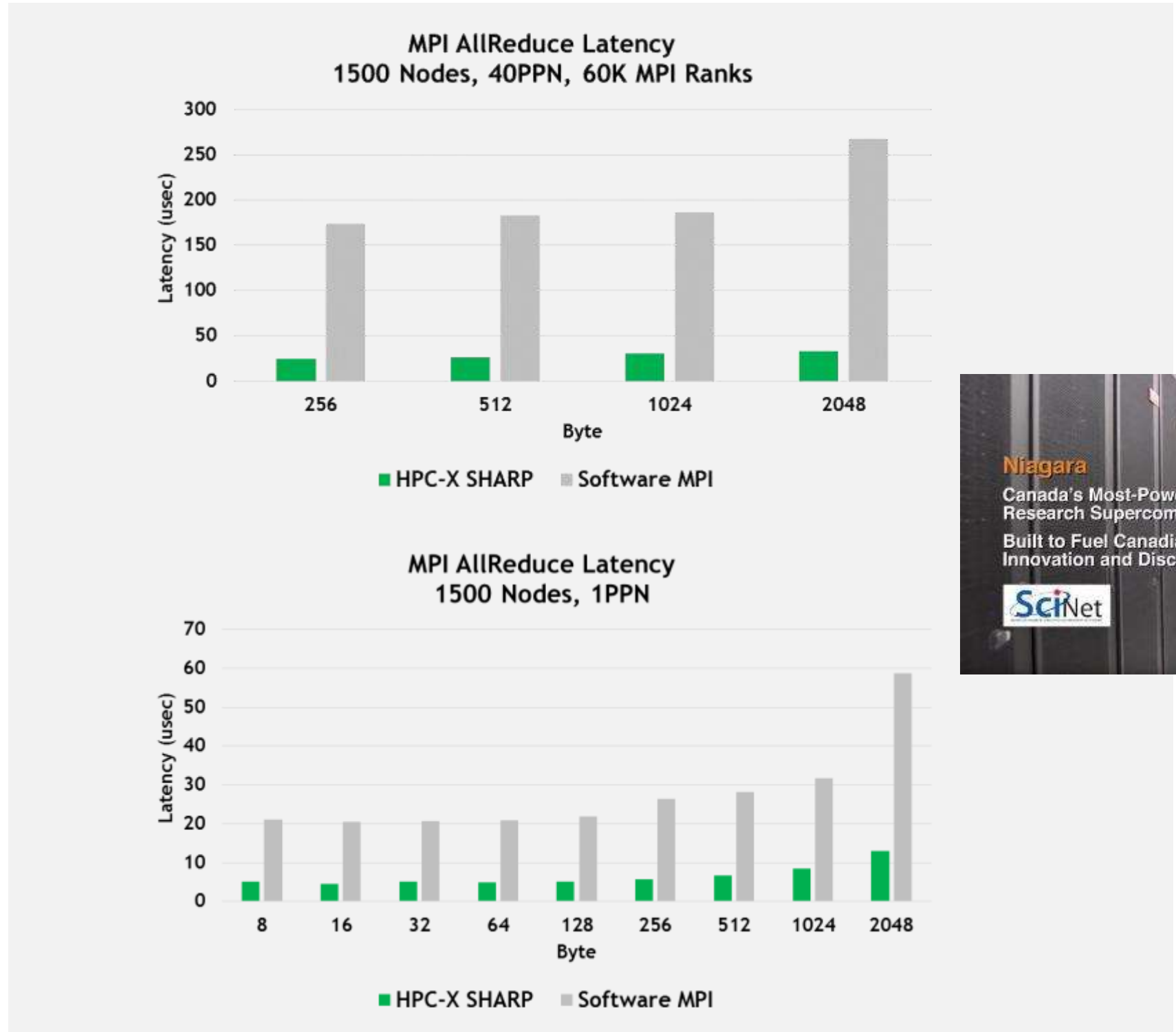
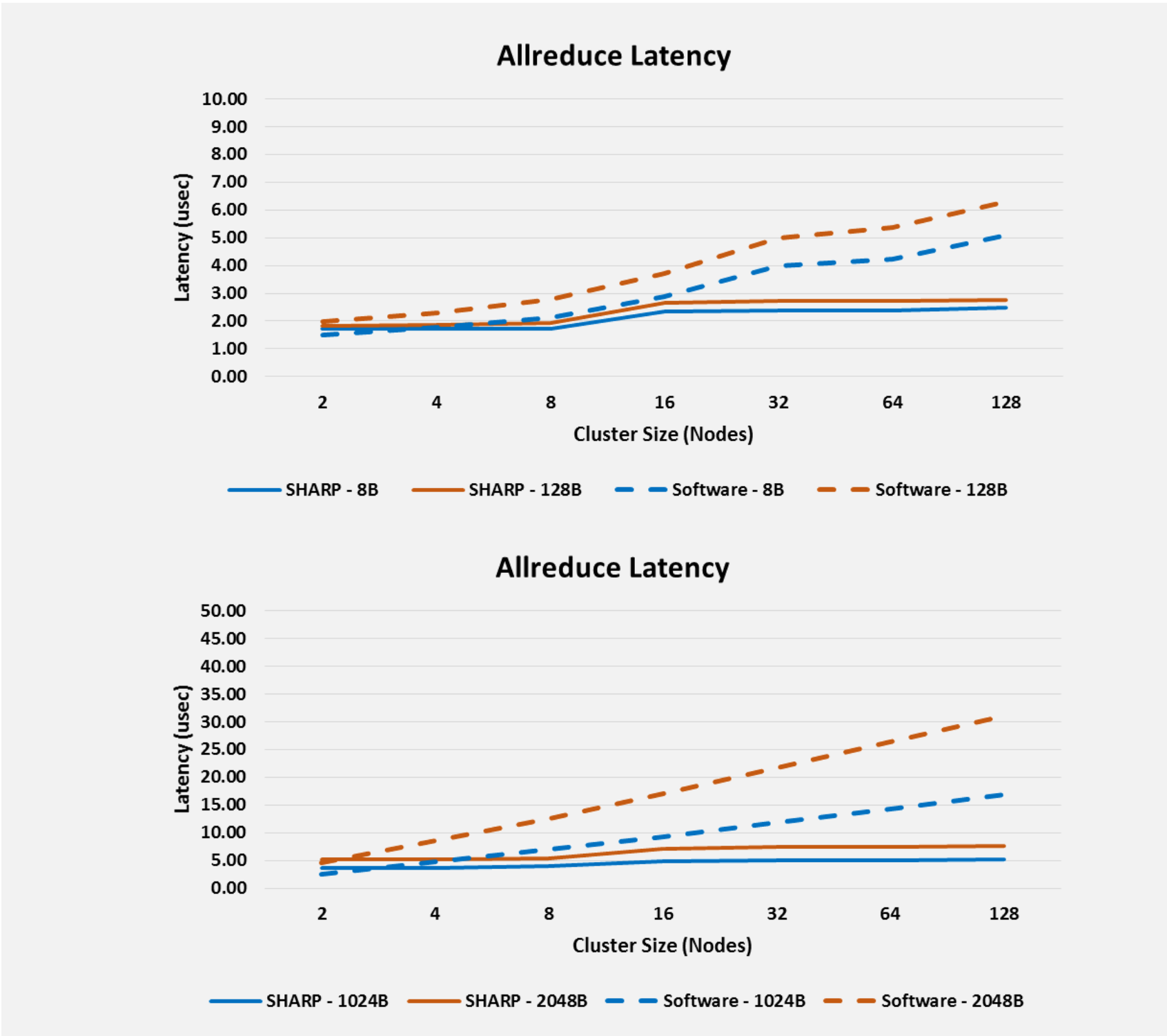
Sum, Min, Max, Min-loc, max-loc, OR, XOR, AND

Integer and Floating-Point, 16/32/64 bits



SHARP ALLREDUCE PERFORMANCE ADVANTAGES

Providing Flat Latency, 7X Higher Performance



SHARP ACCELERATES AI PERFORMANCE

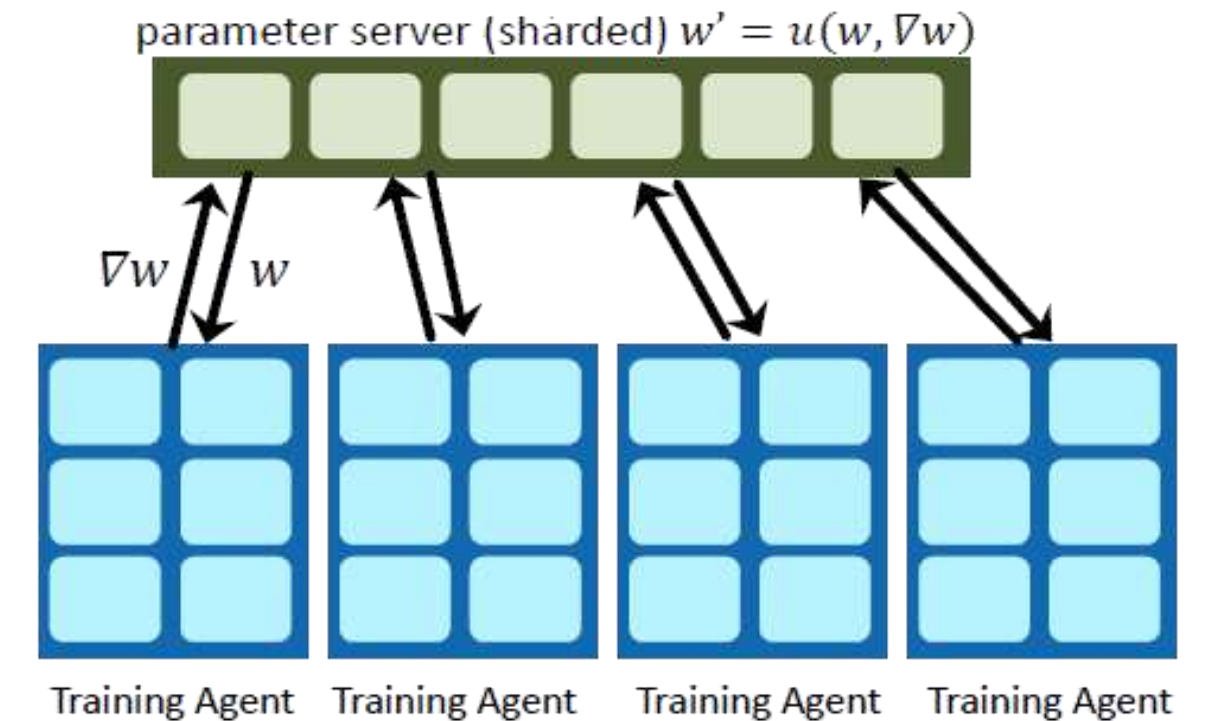
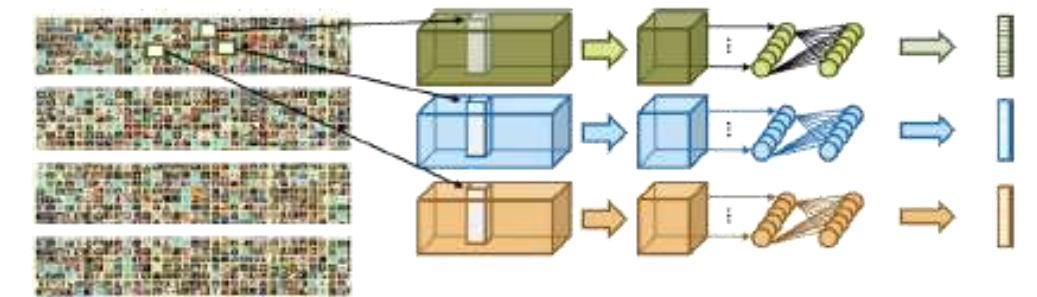
The CPU in a parameter server becomes the bottleneck



**Scalable Hierarchical
Aggregation and
Reduction Protocol**

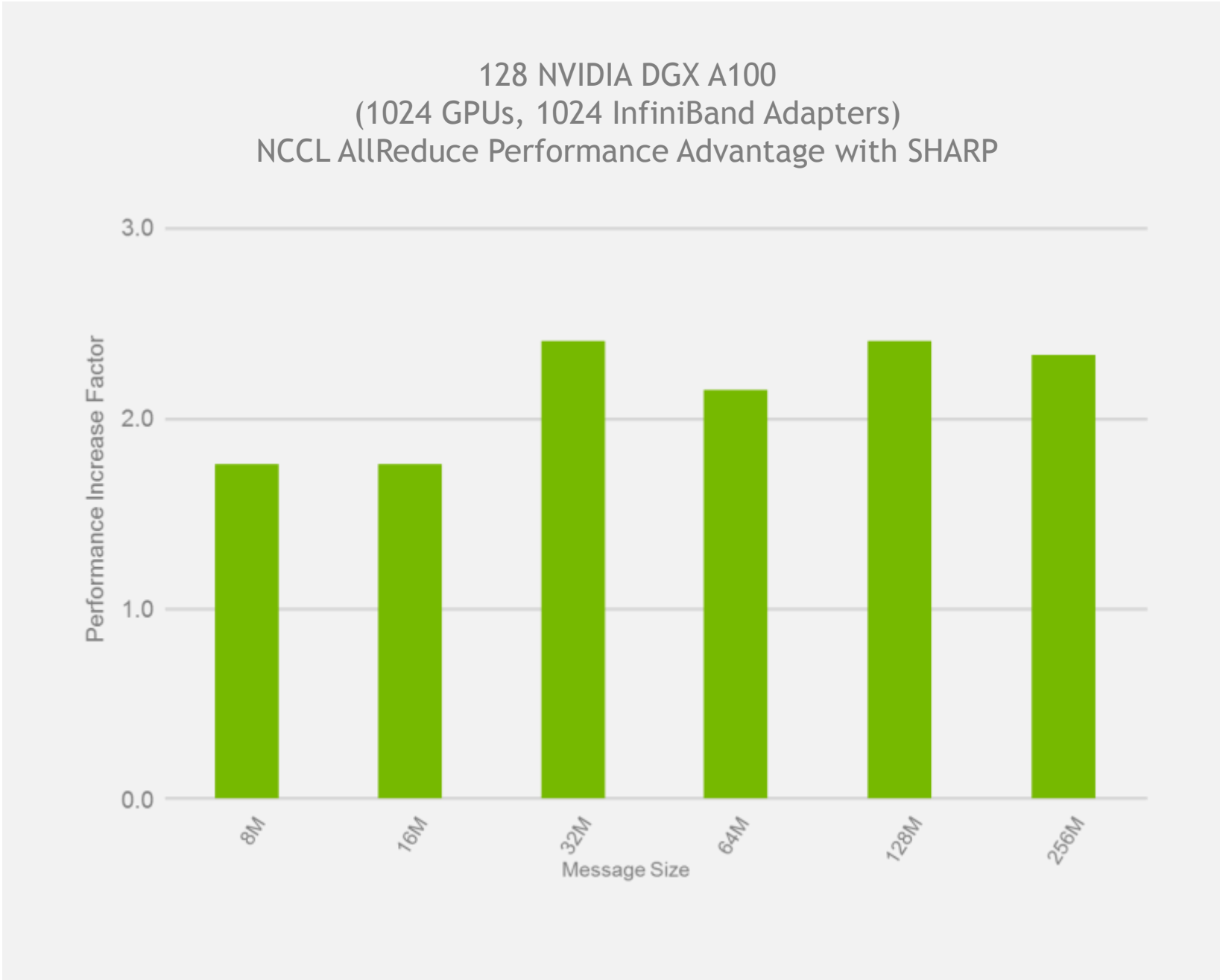
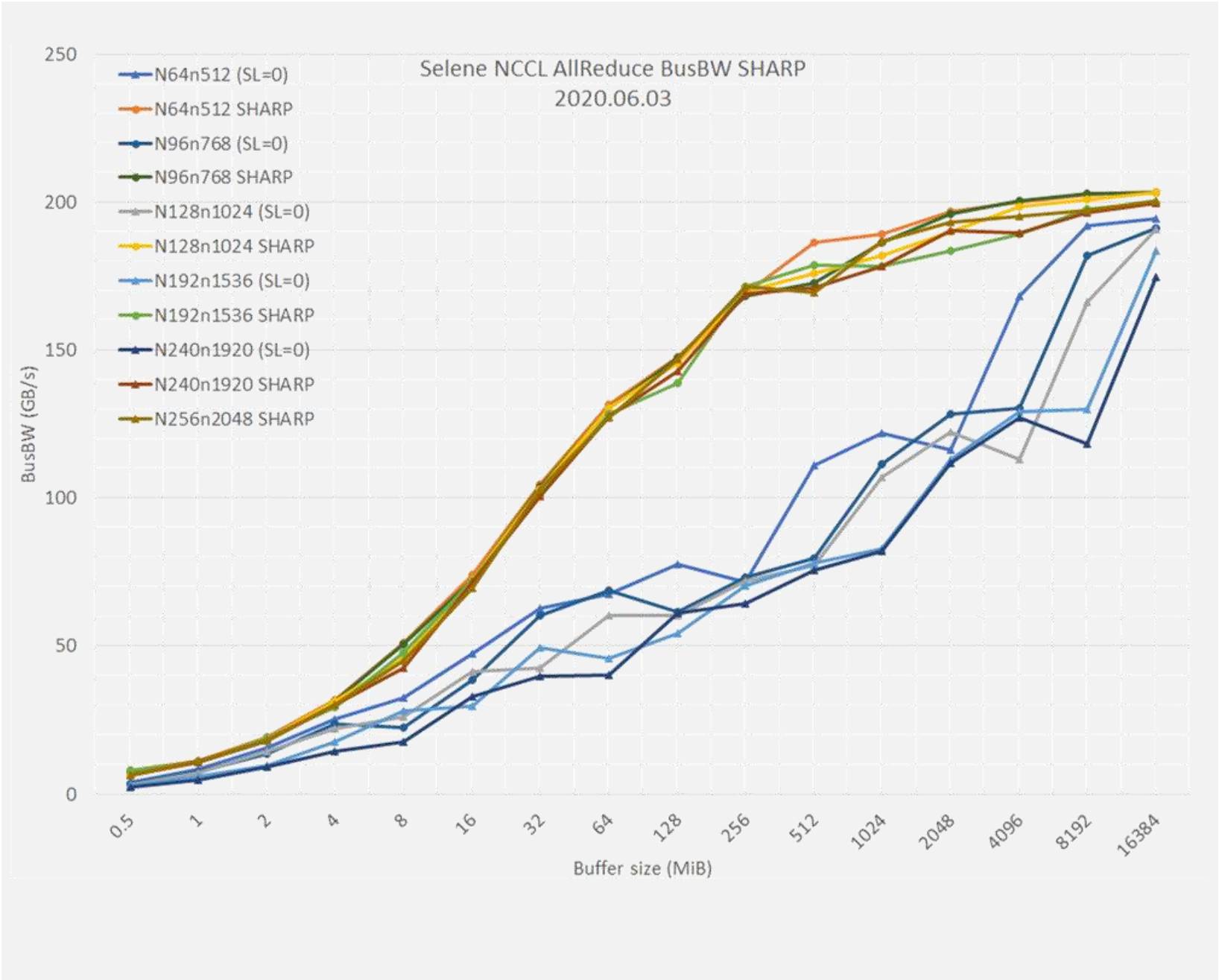


Performs the Gradient Averaging
Replaces all physical parameter servers
Accelerate AI Performance



INFINIBAND SHARP AI PERFORMANCE ADVANTAGE

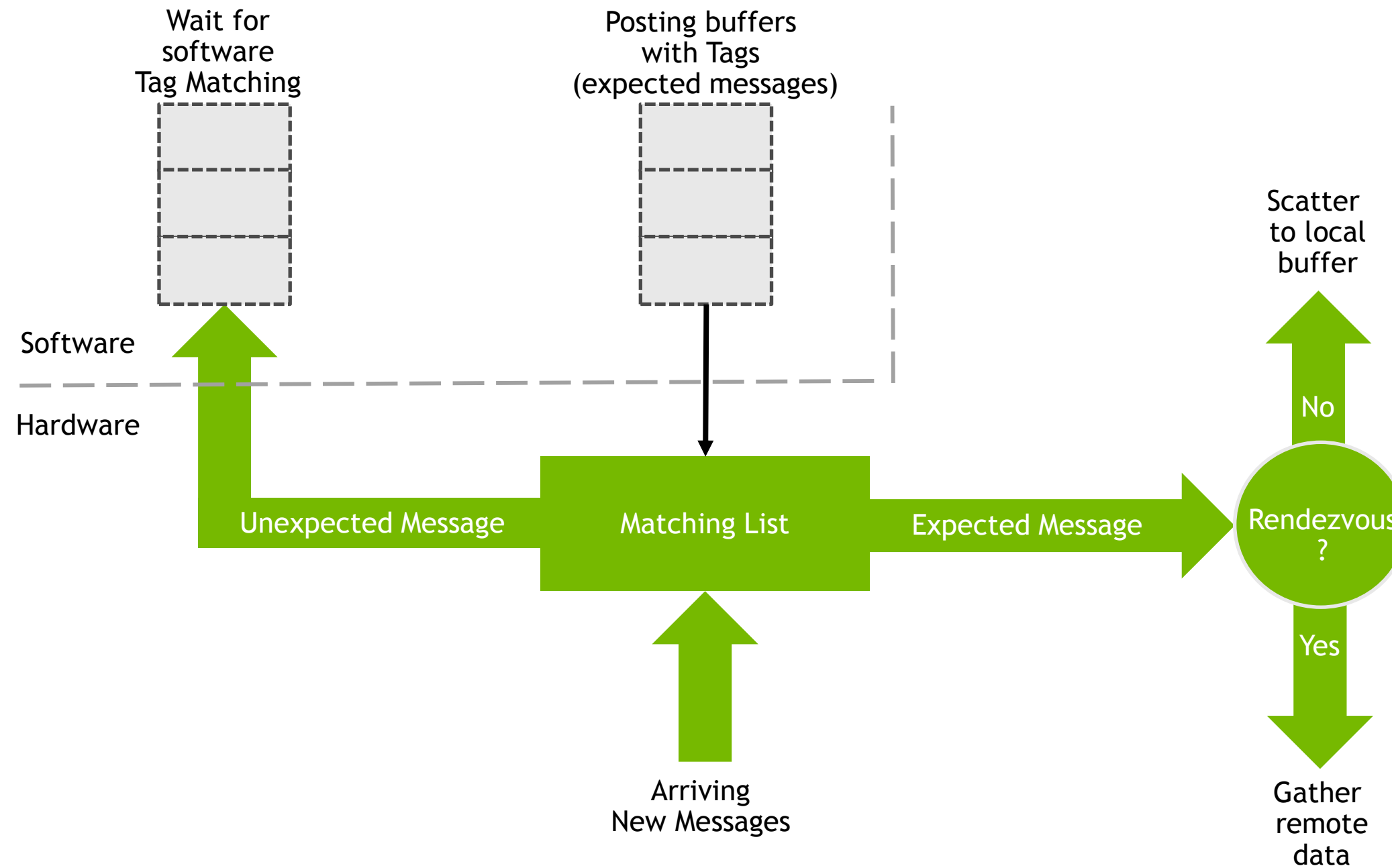
2.5X Higher Performance





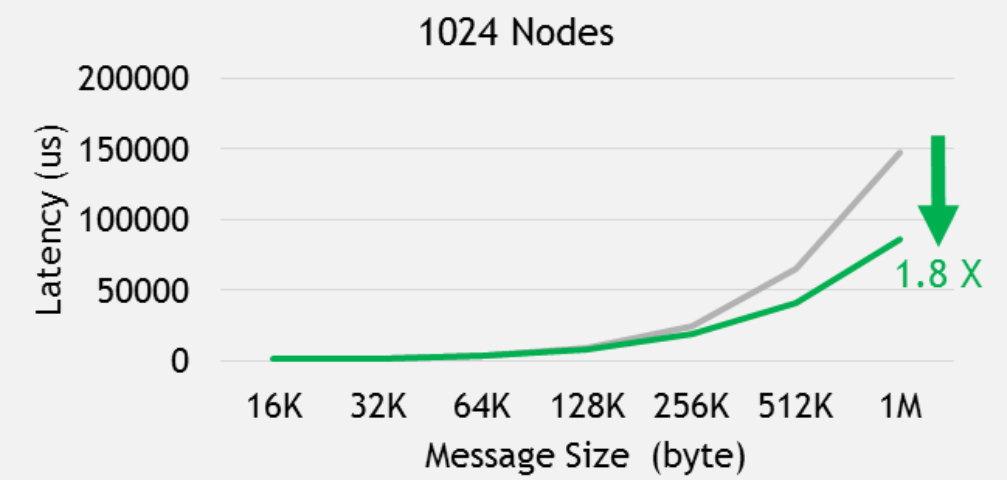
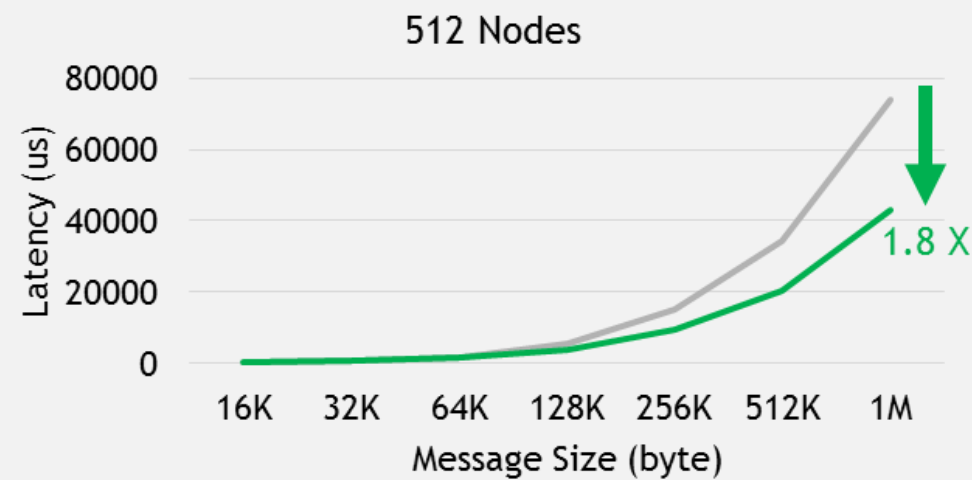
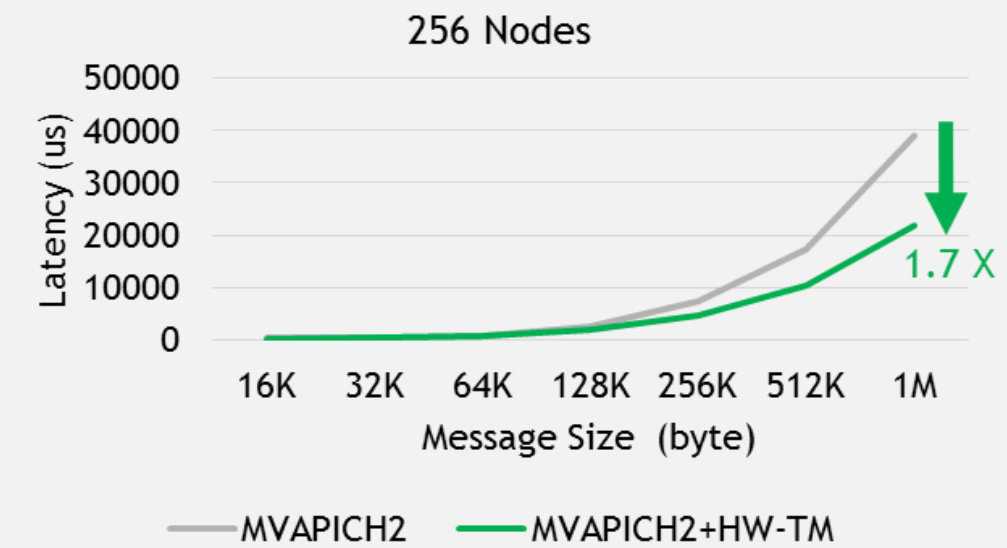
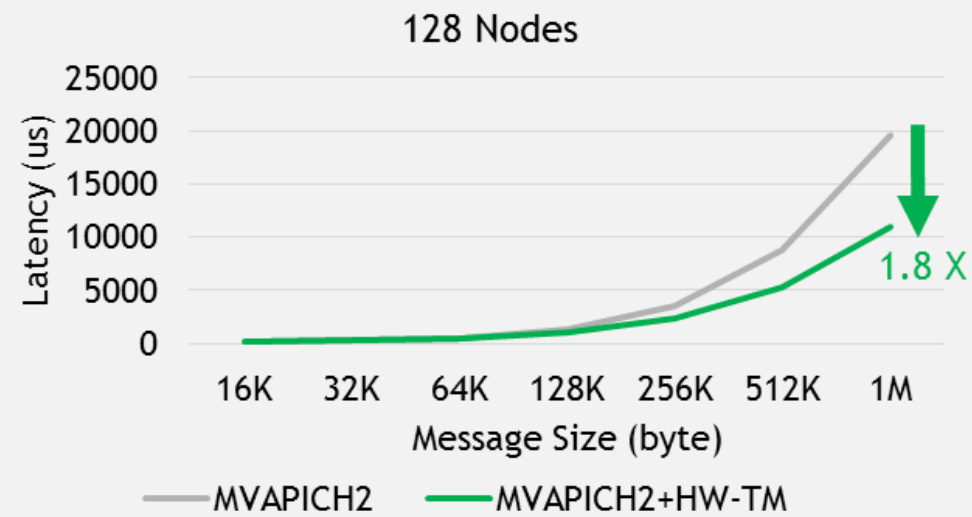
MPI TAG MATCHING HARDWARE ENGINE

INFINIBAND MPI TAG MATCHING HARDWARE ENGINE



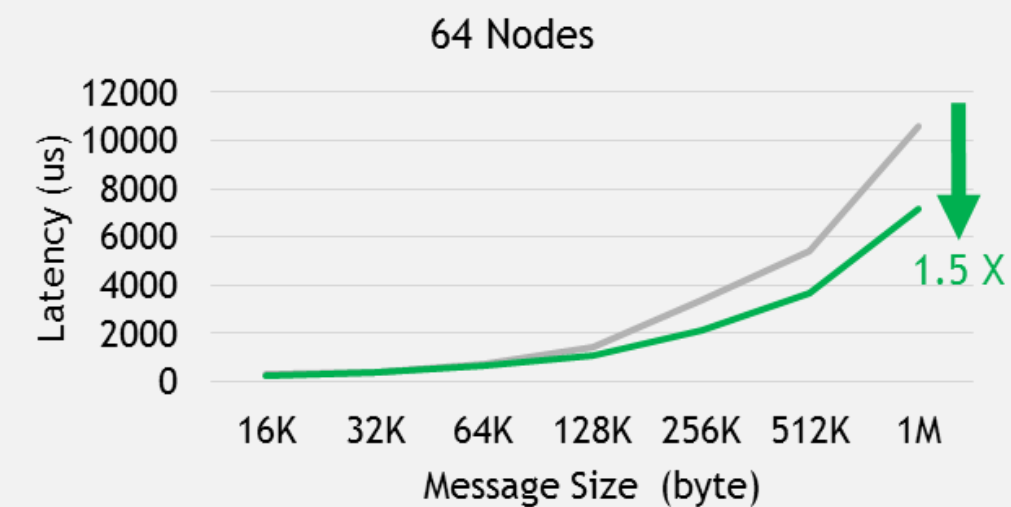
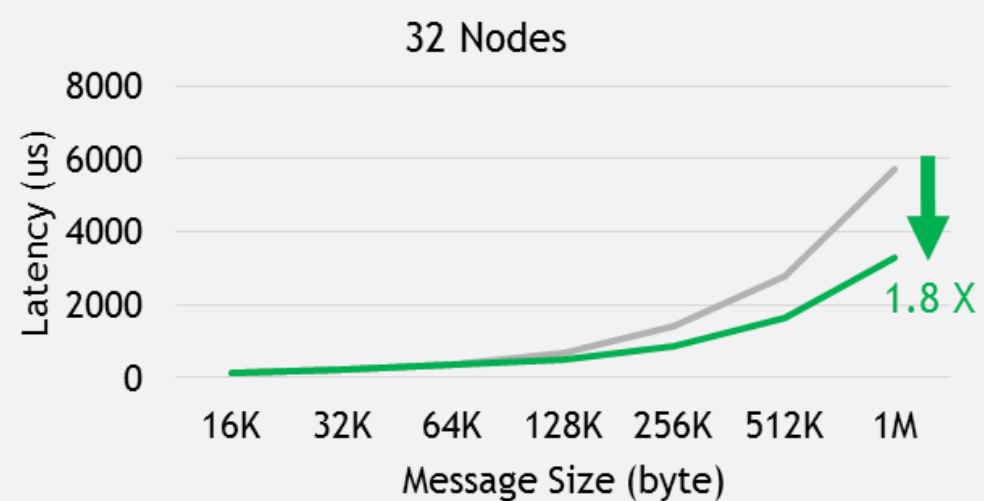
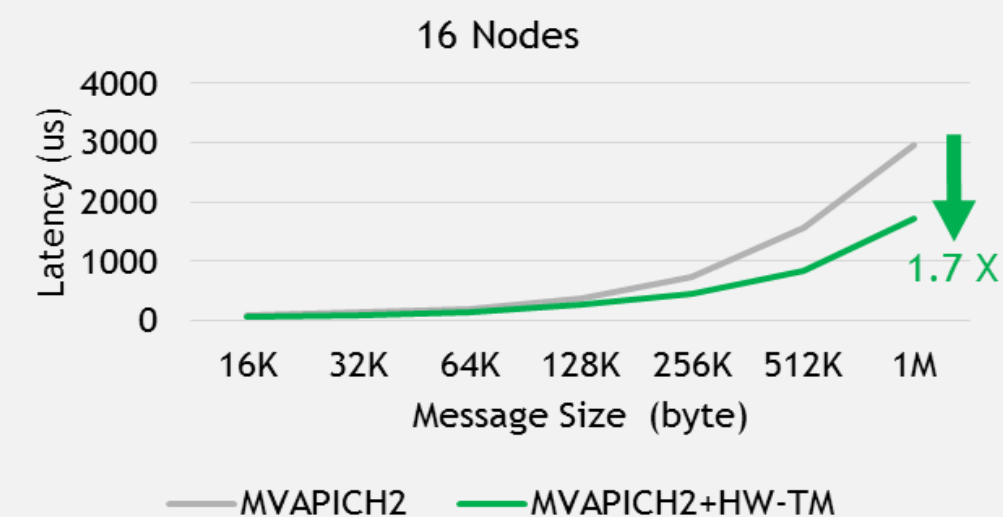
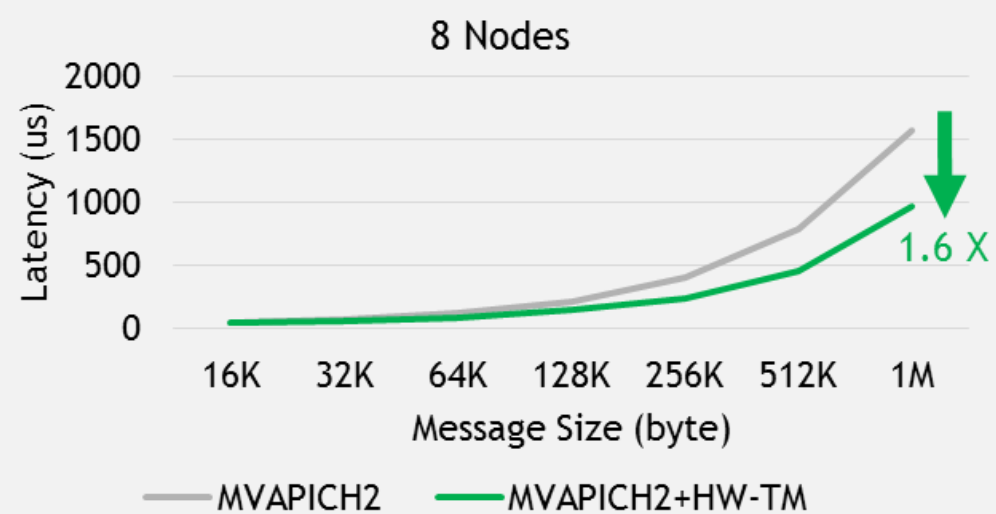
HARDWARE TAG MATCHING PERFORMANCE ADVANTAGES

1.8X Higher MPI_Isscatterv Performance on TACC Frontera



HARDWARE TAG MATCHING PERFORMANCE ADVANTAGES

1.8X higher MPI_lalltoall Performance on TACC Frontera



HARDWARE TAG MATCHING PERFORMANCE ADVANTAGES

Nearly 100% Compute - Communication Overlap

