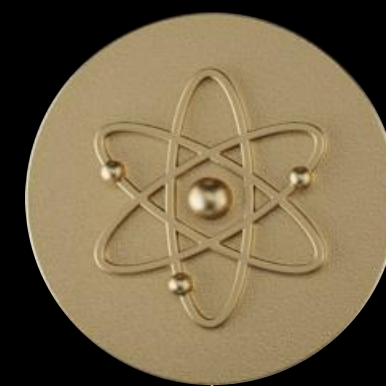


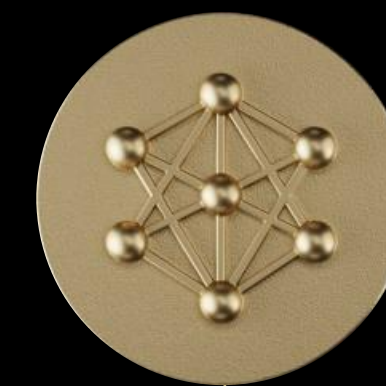


DPU Computing: Addressing HPC/AI Performance Bottlenecks

UCF 2022



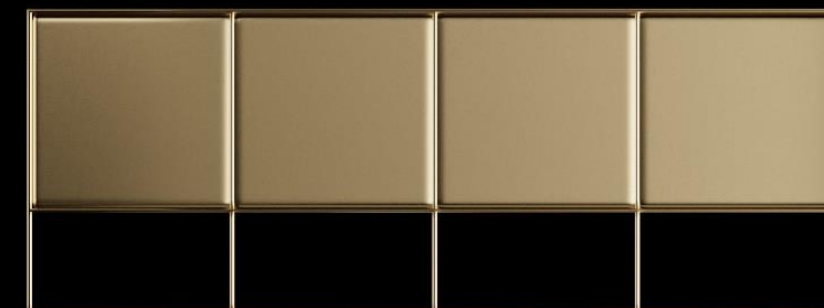
HPC



AI

ACCELERATION LIBRARIES

GPU



CPU

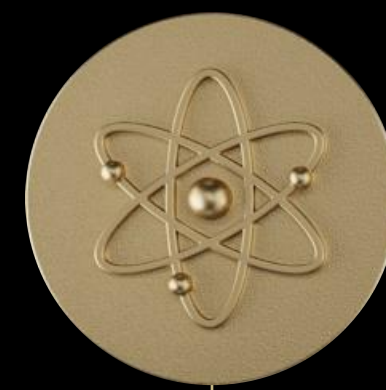


Infrastructure Workload

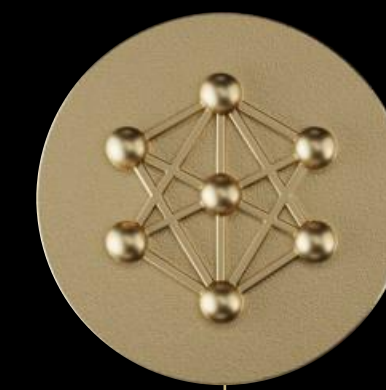
- Communications
- Storage
- Security and Isolation

NIC



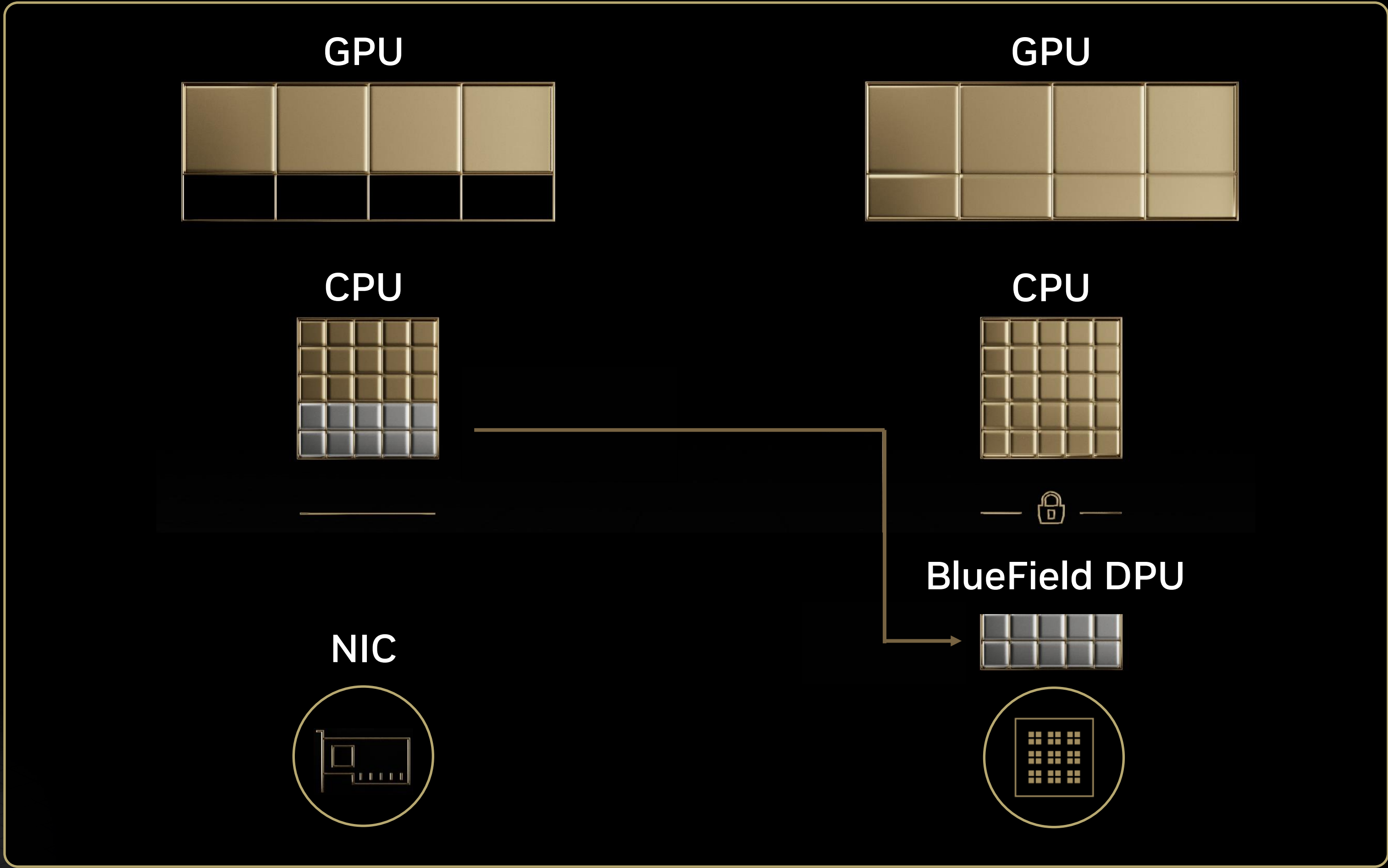


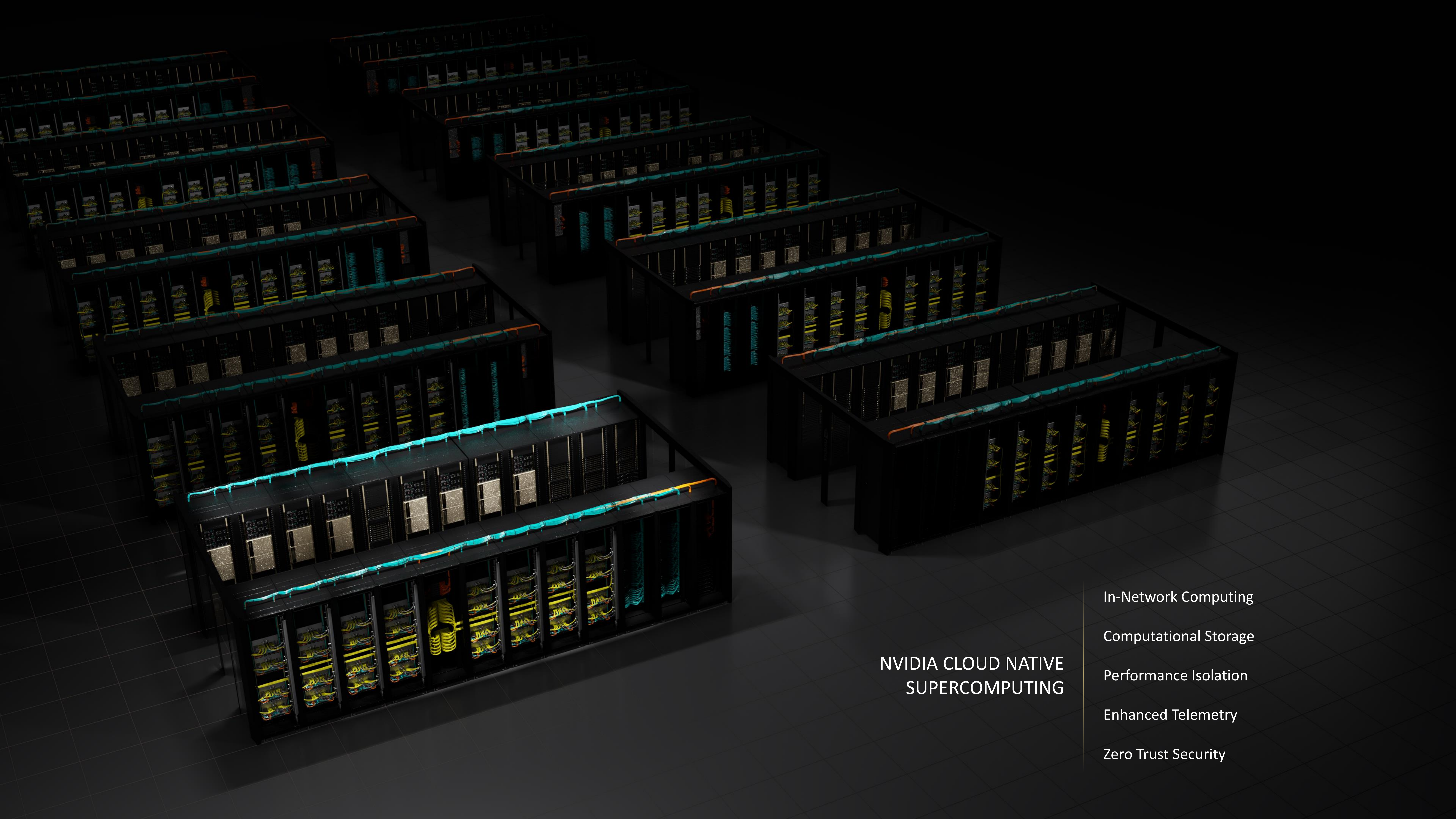
HPC



AI

ACCELERATION LIBRARIES





NVIDIA CLOUD NATIVE
SUPERCOMPUTING

In-Network Computing

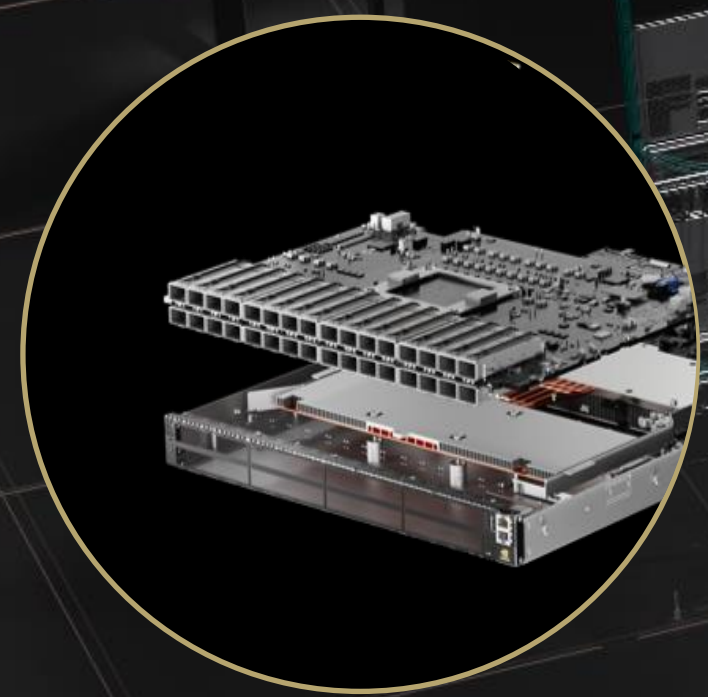
Computational Storage

Performance Isolation

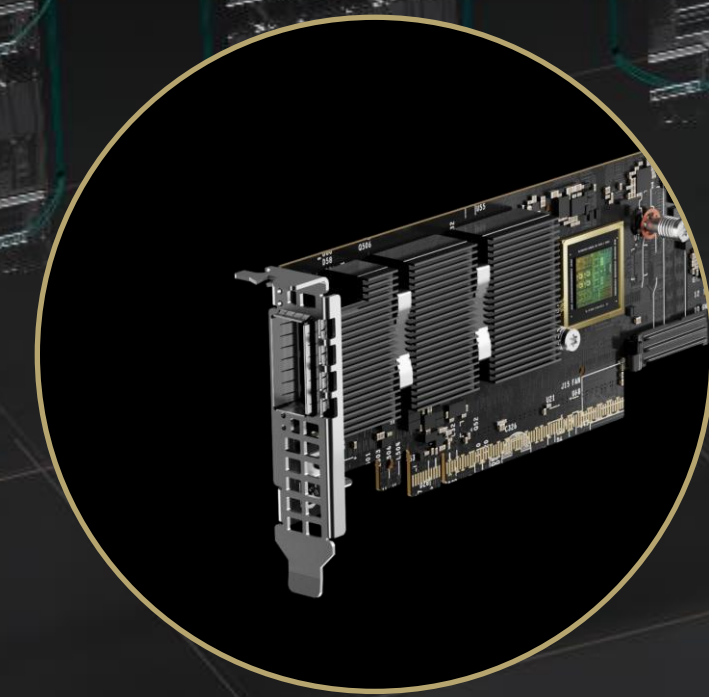
Enhanced Telemetry

Zero Trust Security

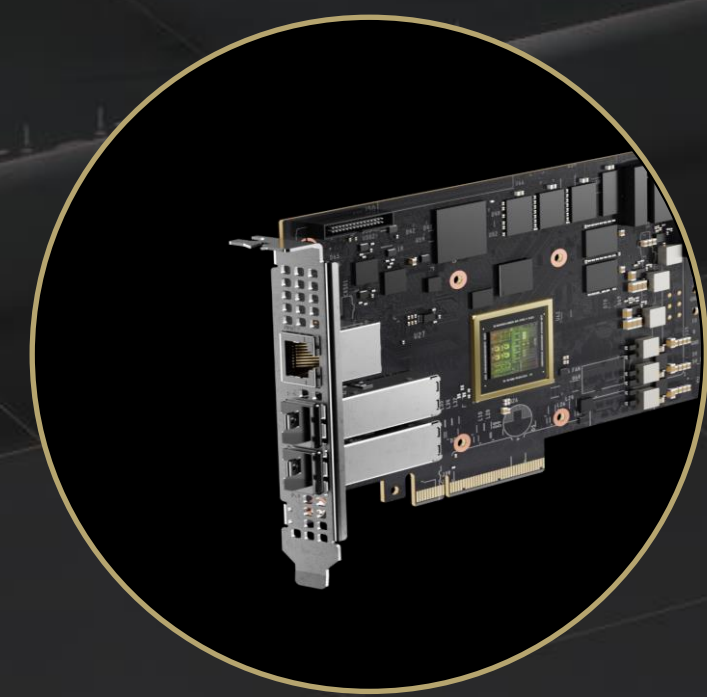
Cloud Native Supercomputing With NVIDIA Quantum-2 IB Platform



QUANTUM-2 INFINIBAND SWITCH
Cloud Native Supercomputing Platform
SHARP In-Network Computing
Higher Scalability



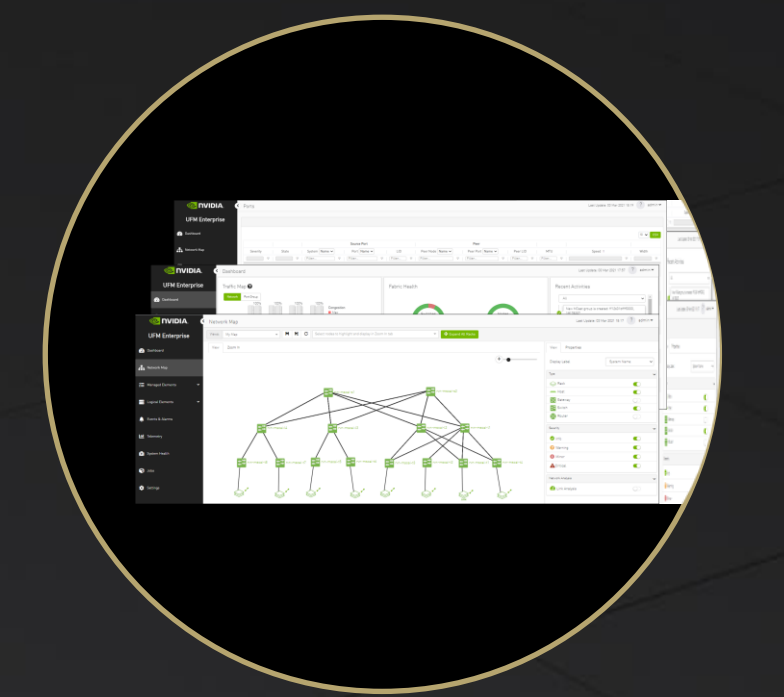
CONNECTX-7 SMARTNIC
Intelligent Offloads
Precision Timing
Software Defined Networking



BLUEFIELD-3/-X DPU
Intelligent Offloads
Precision Timing
Software Defined Networking



SKYWAY GATEWAY
InfiniBand to Ethernet
Low Latency
Load Balancing



UFM
Monitoring, Management, Orchestration
Predictive Maintenance
Anomaly Detection

BlueField Data Processing Unit

Data Center on a Chip

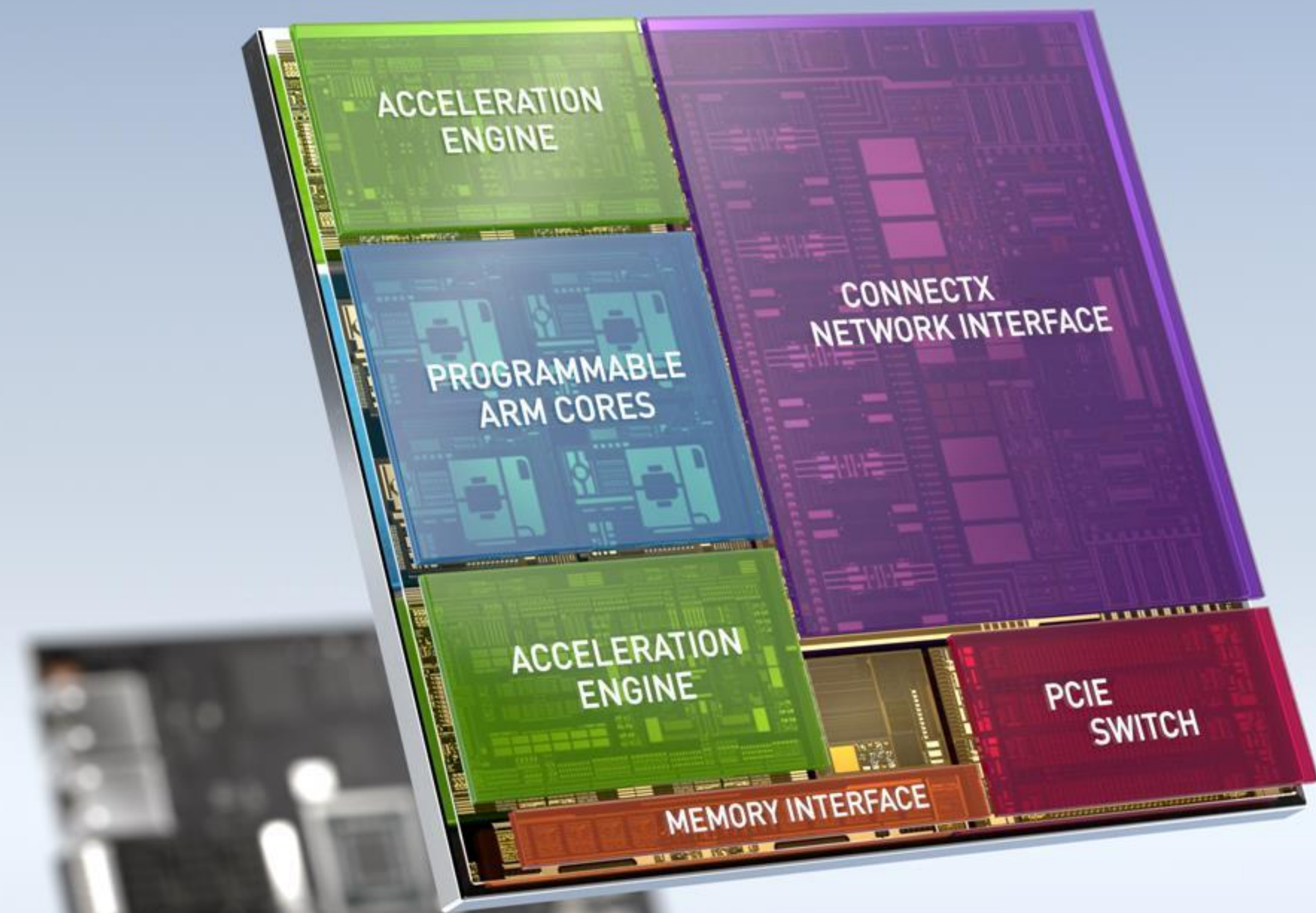
16 Arm 64-Bit Cores

16 Core / 256 Threads Datapath Accelerator

ConnectX InfiniBand / Ethernet

DDR memory interface

PCIe switch



BlueField Data Processing Unit

Data Center on a Chip

16 Arm 64-Bit Cores

16 Core / 256 Threads Datapath Accelerator

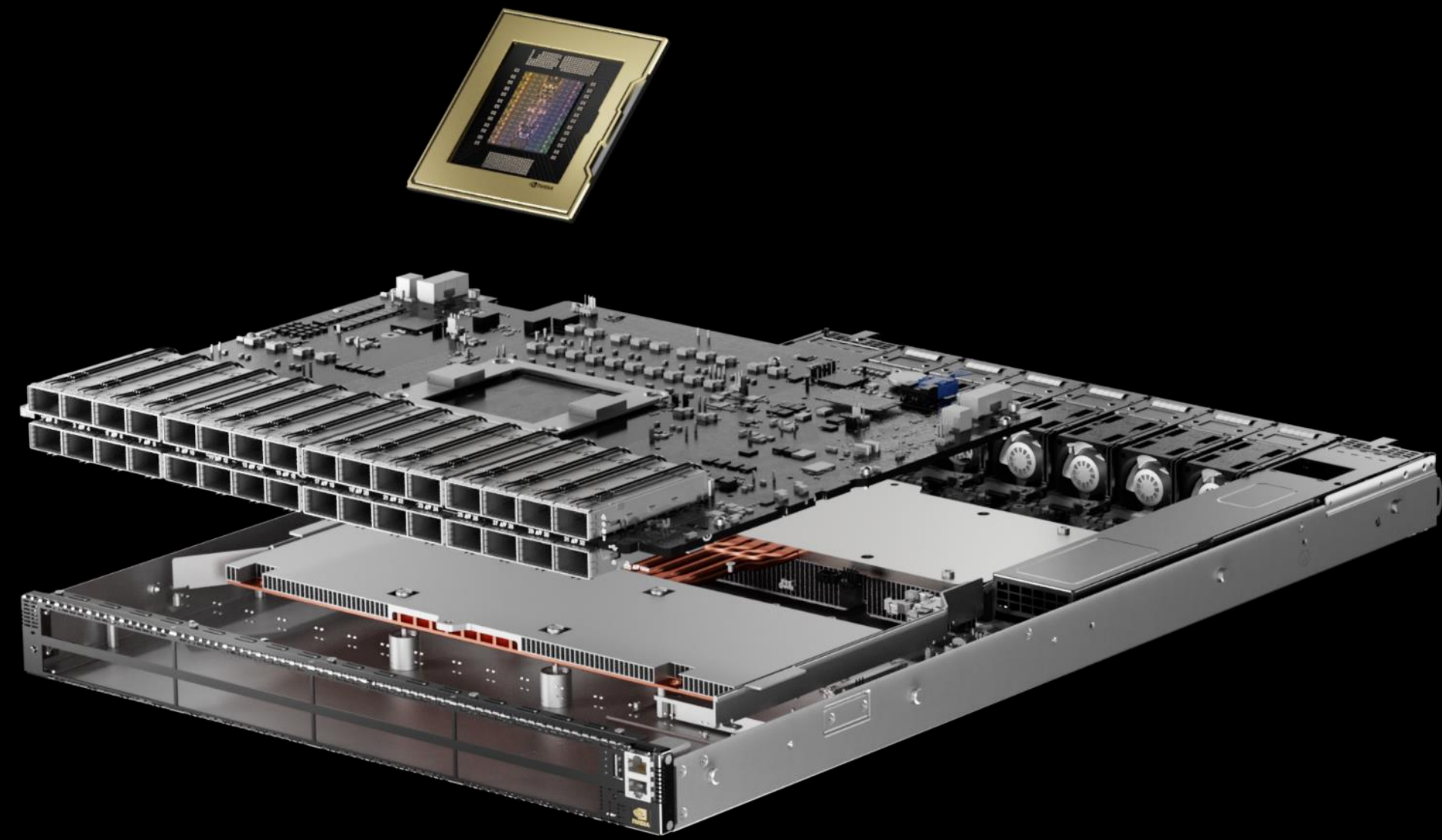
ConnectX InfiniBand / Ethernet

DDR memory interface

PCIe switch

	BlueField-2	BlueField-3
Network Bandwidth	200Gb/s	400Gb/s
RDMA msg rate	215Mpps	370Mpps
Compute	SPECINT2K17: 9.8	SPECINT2K17: 42
Memory Bandwidth	17GB/s	80GB/s

NVIDIA Quantum-2 400G 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



BLUEFIELD-3 INFINIBAND

16 Arm 64-Bit Cores

16 Core / 256 Threads Datapath Accelerator

Full Transport Offload and Telemetry

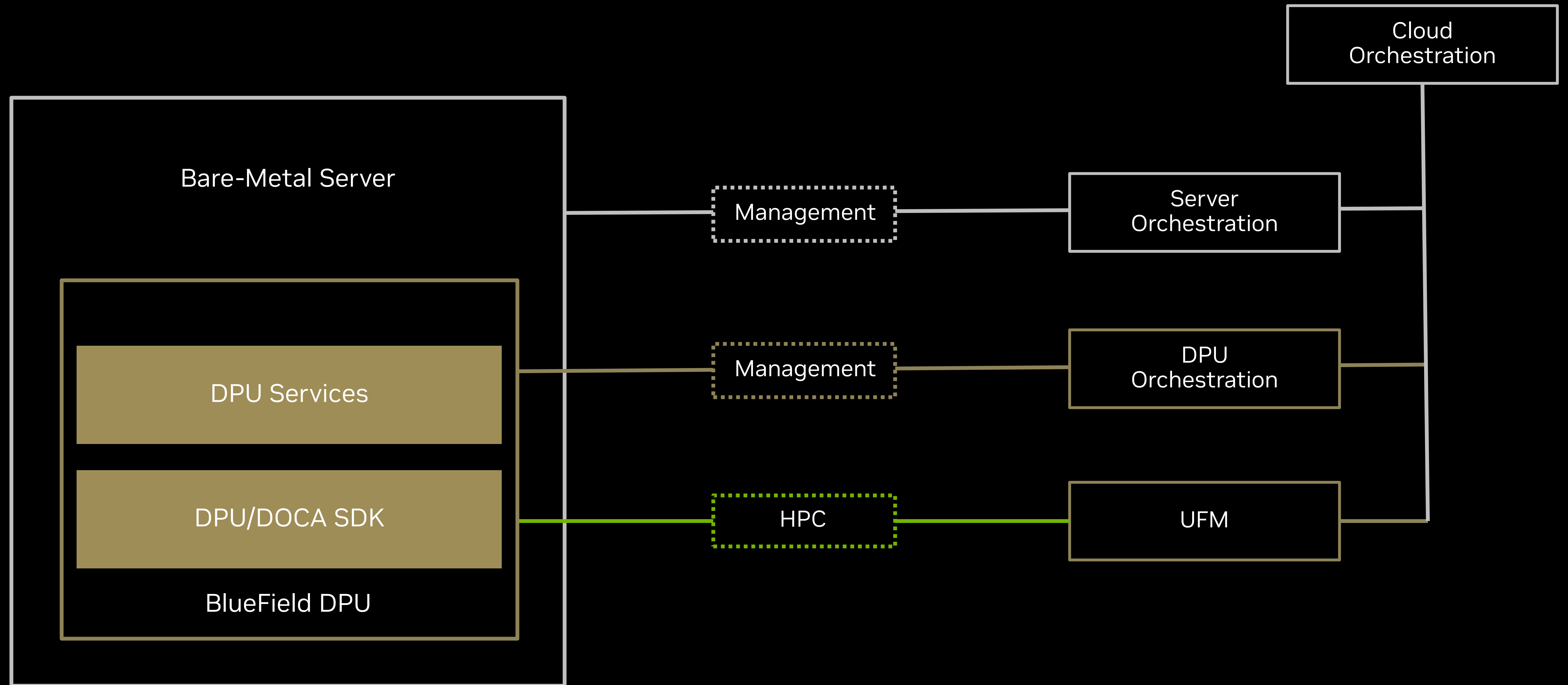
Hardware-Based RDMA / GPUDirect

MPI and NCCL Accelerations

Computational Storage

Security Engines

Delivering Cloud Native Supercomputing

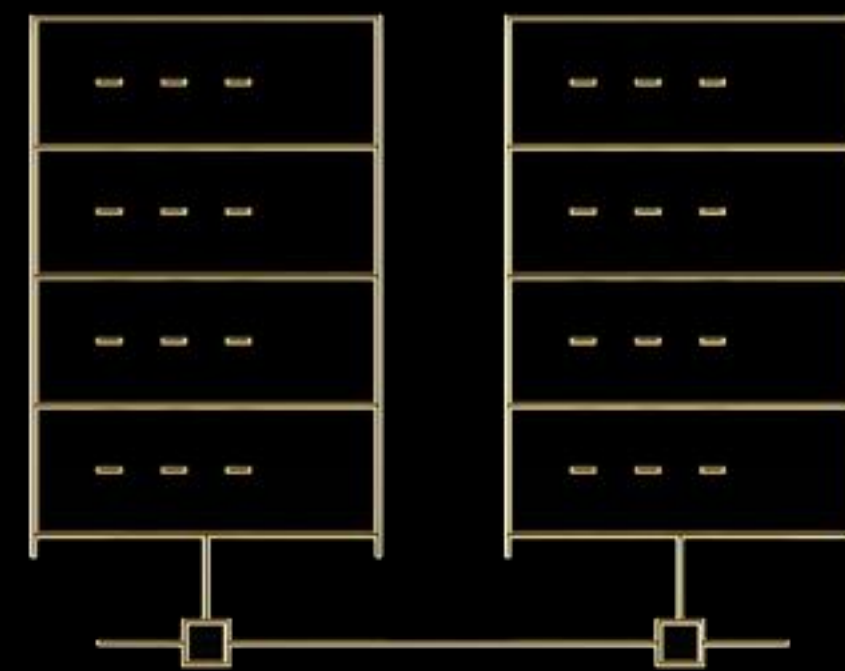


HPC Performance Bottlenecks

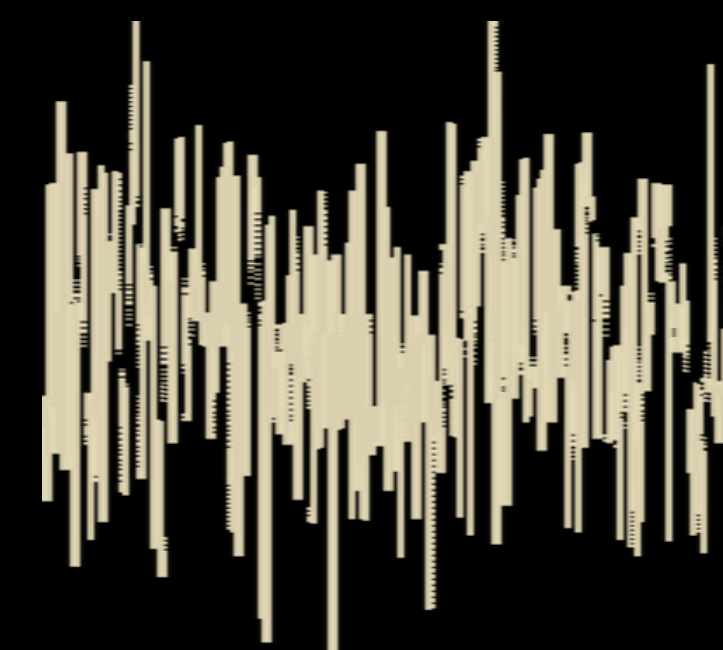
Synchronous Progress



Load Imbalanced



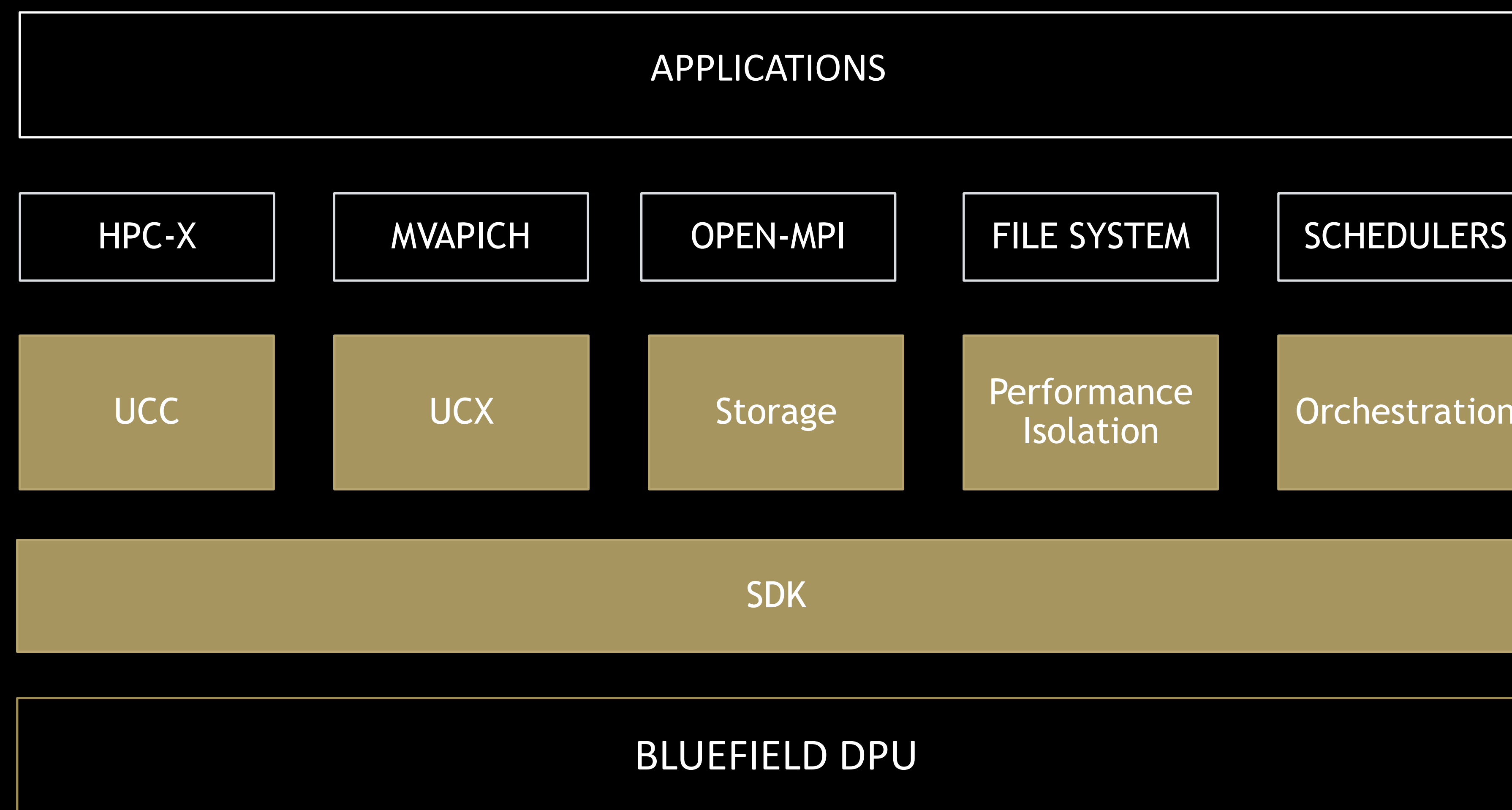
Infrastructure Jitter



Performance Interference

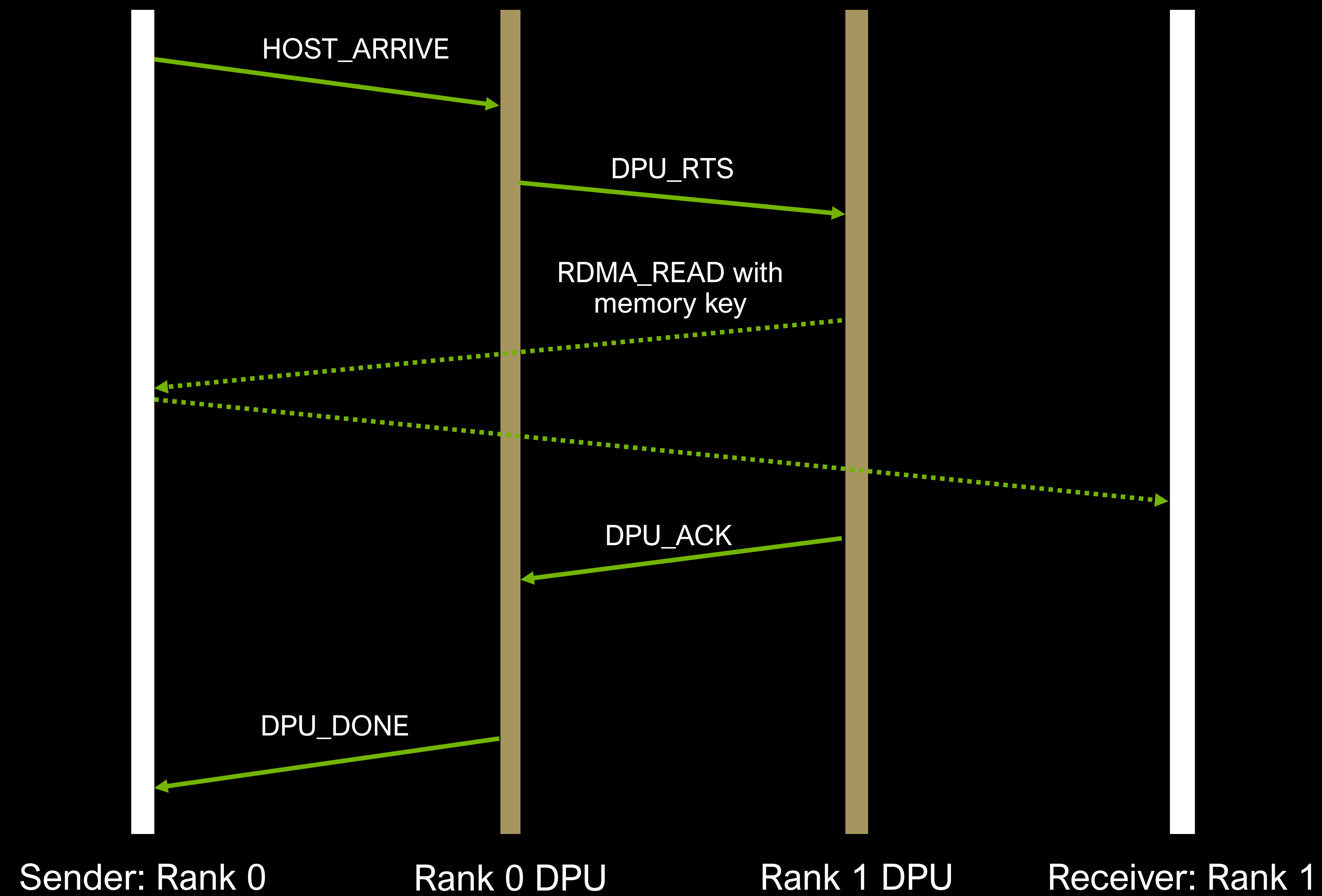


Accelerating HPC Applications with DPU/DOCA Services



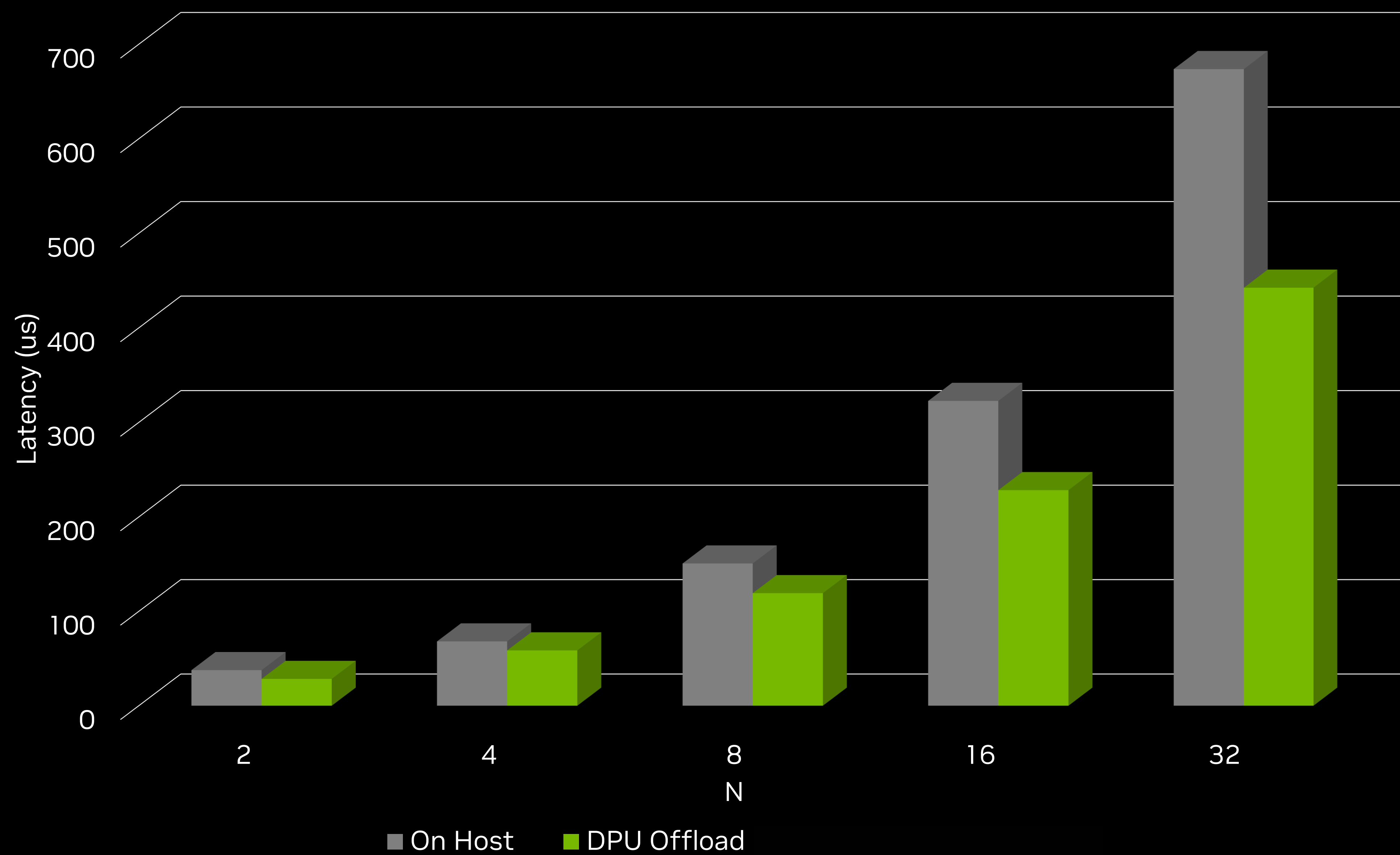
Offloading and Accelerating Data Exchange Example

An Element of Collective Algorithm

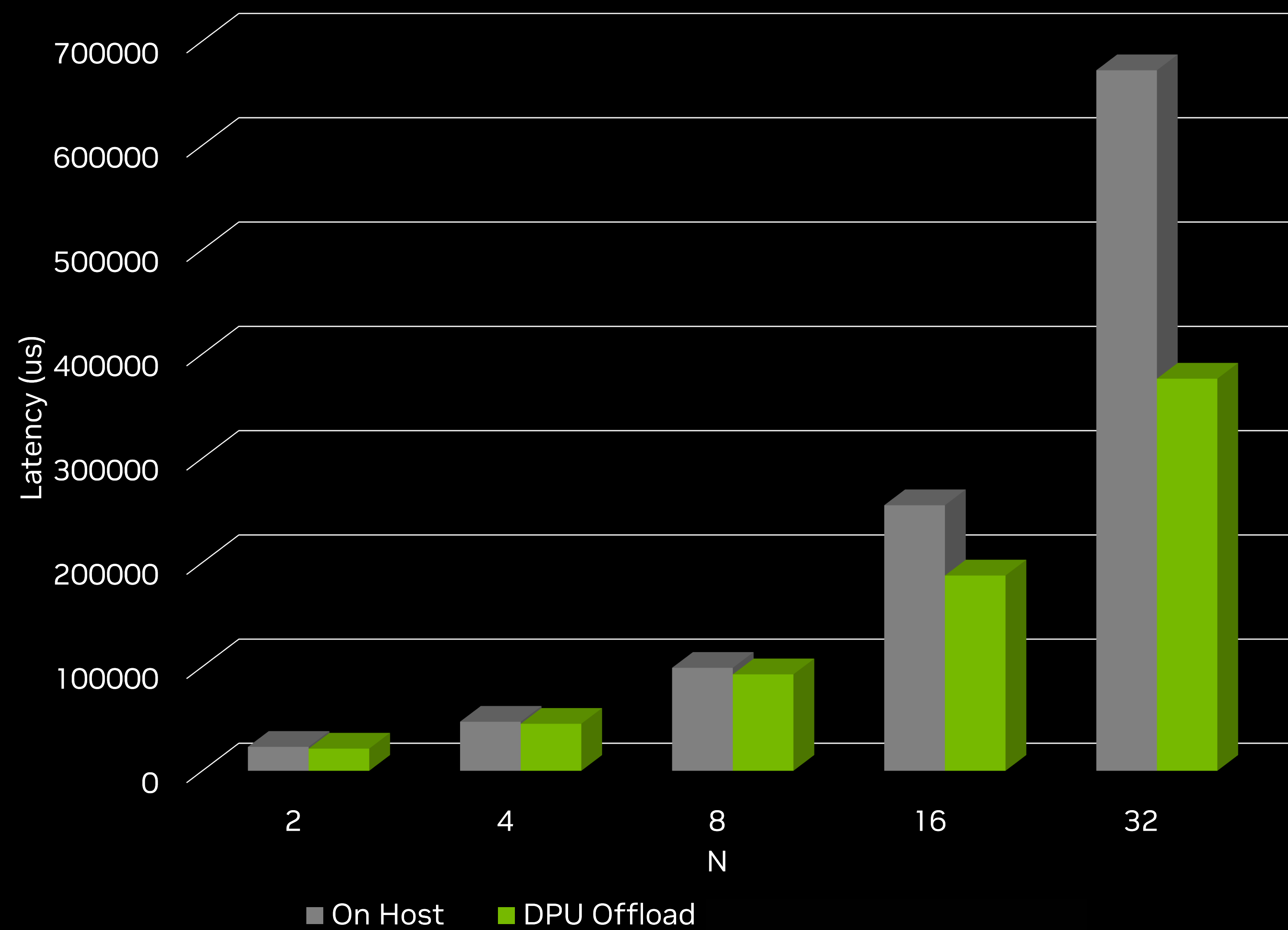


Alltoallv Latency

OSU Alltoallv 1 PPN, Size = 128 KB

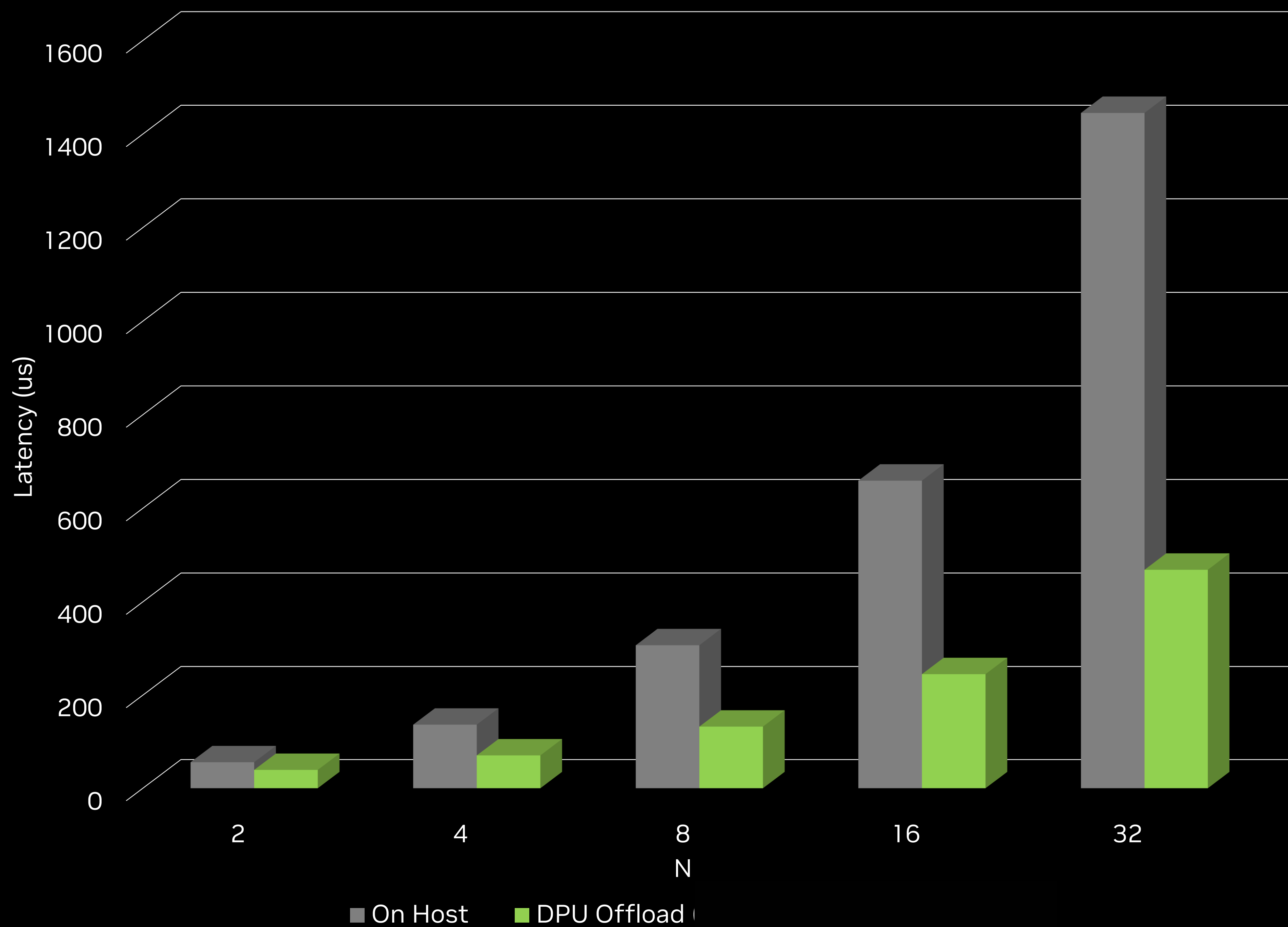


OSU Alltoallv 32 (full) PPN, Size = 128 KB

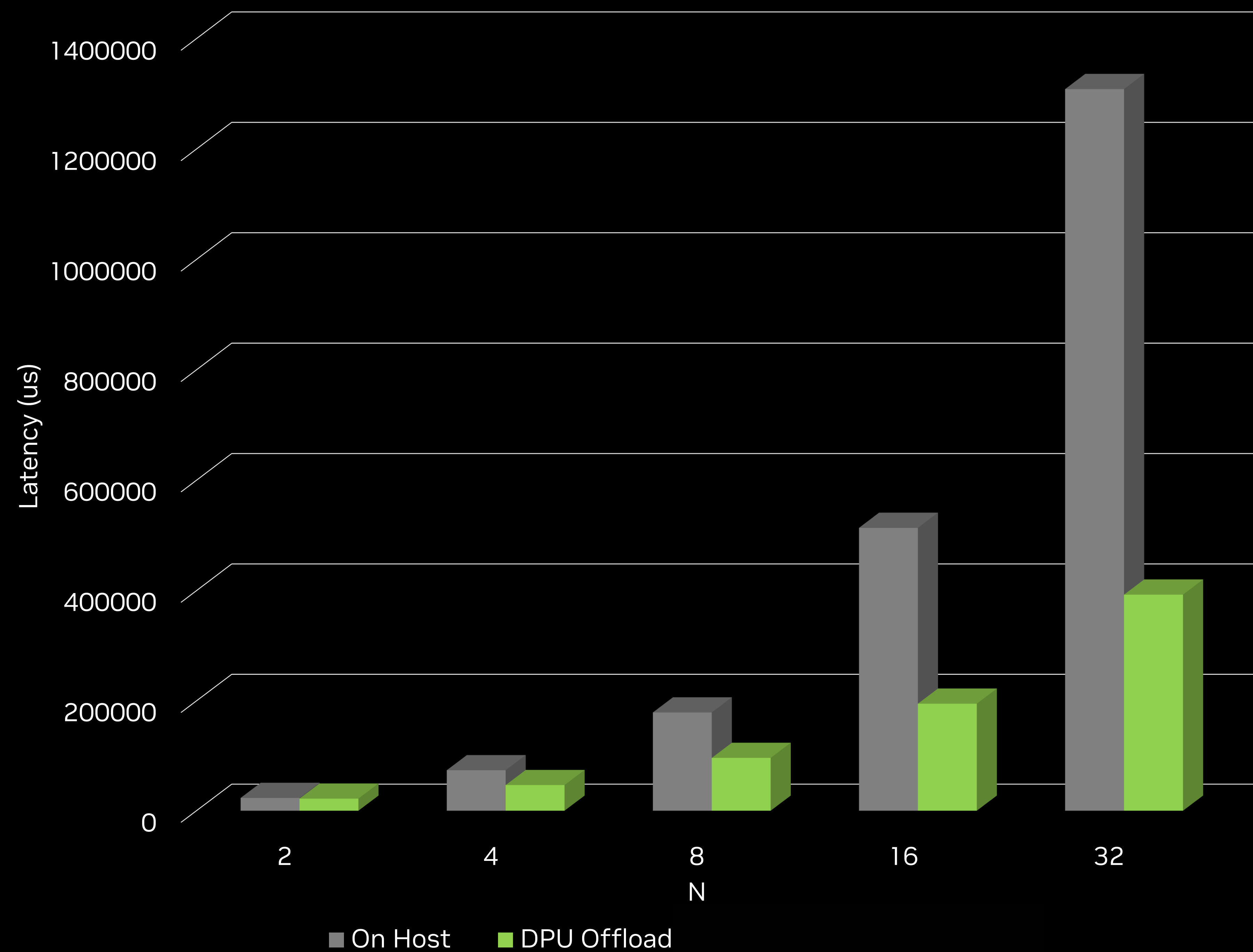


iAlltoallv latency

OSU ialltoallv 1 PPN, Size = 128 KB

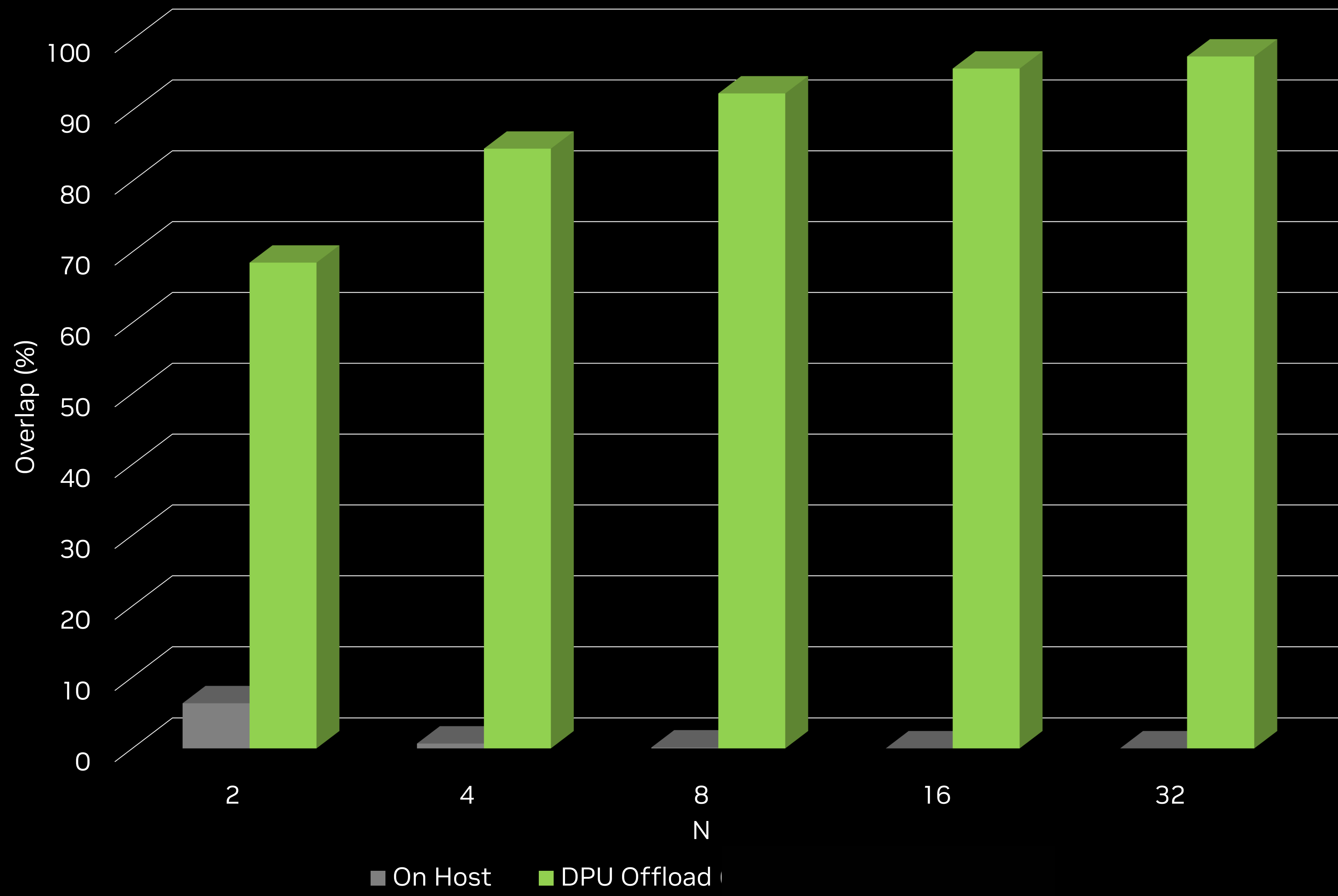


OSU ialltoallv 32 (full) PPN, Size = 128 KB

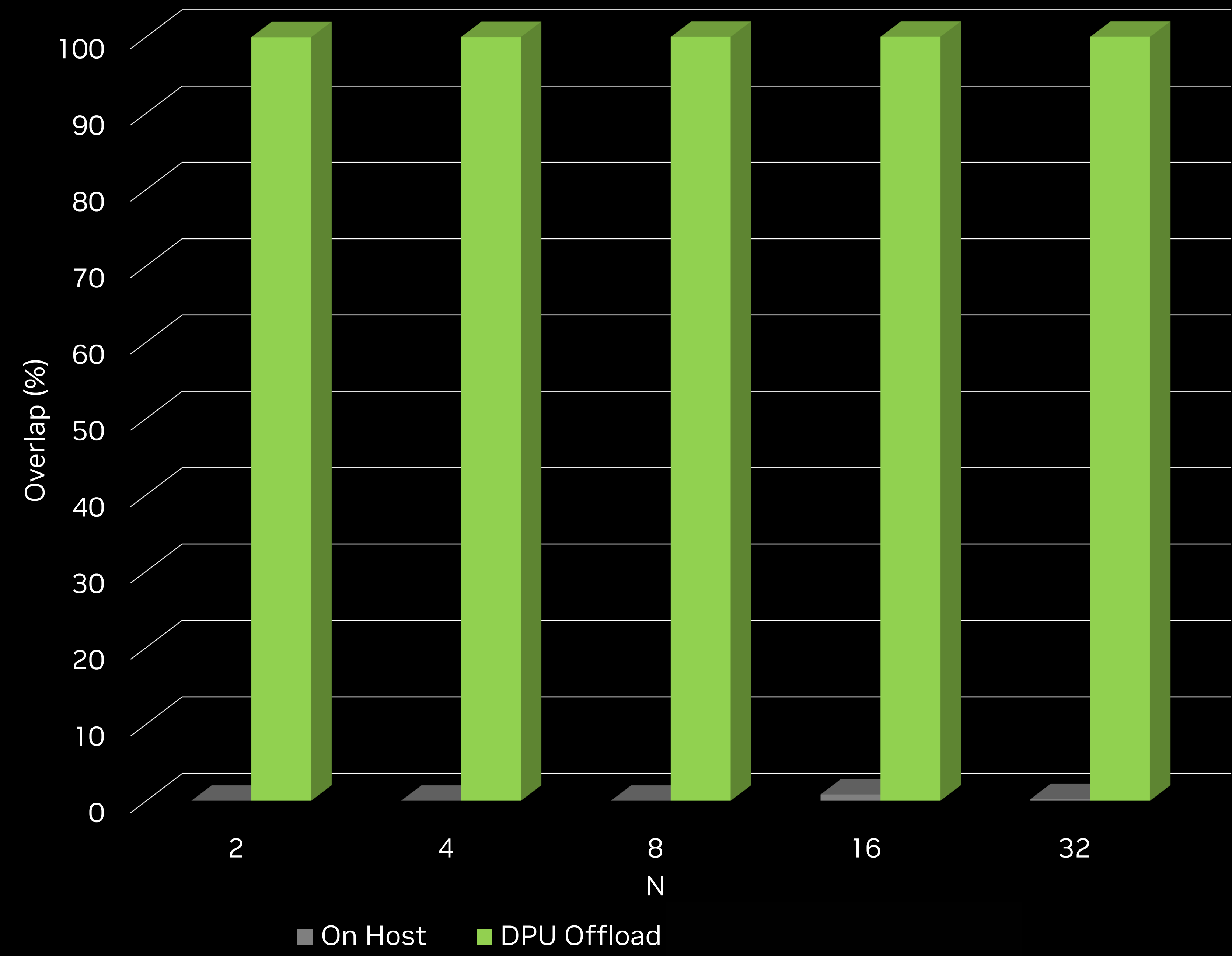


iAlltoallv compute/communication overlap

OSU ialltoallv 1 PPN, Size = 128 KB



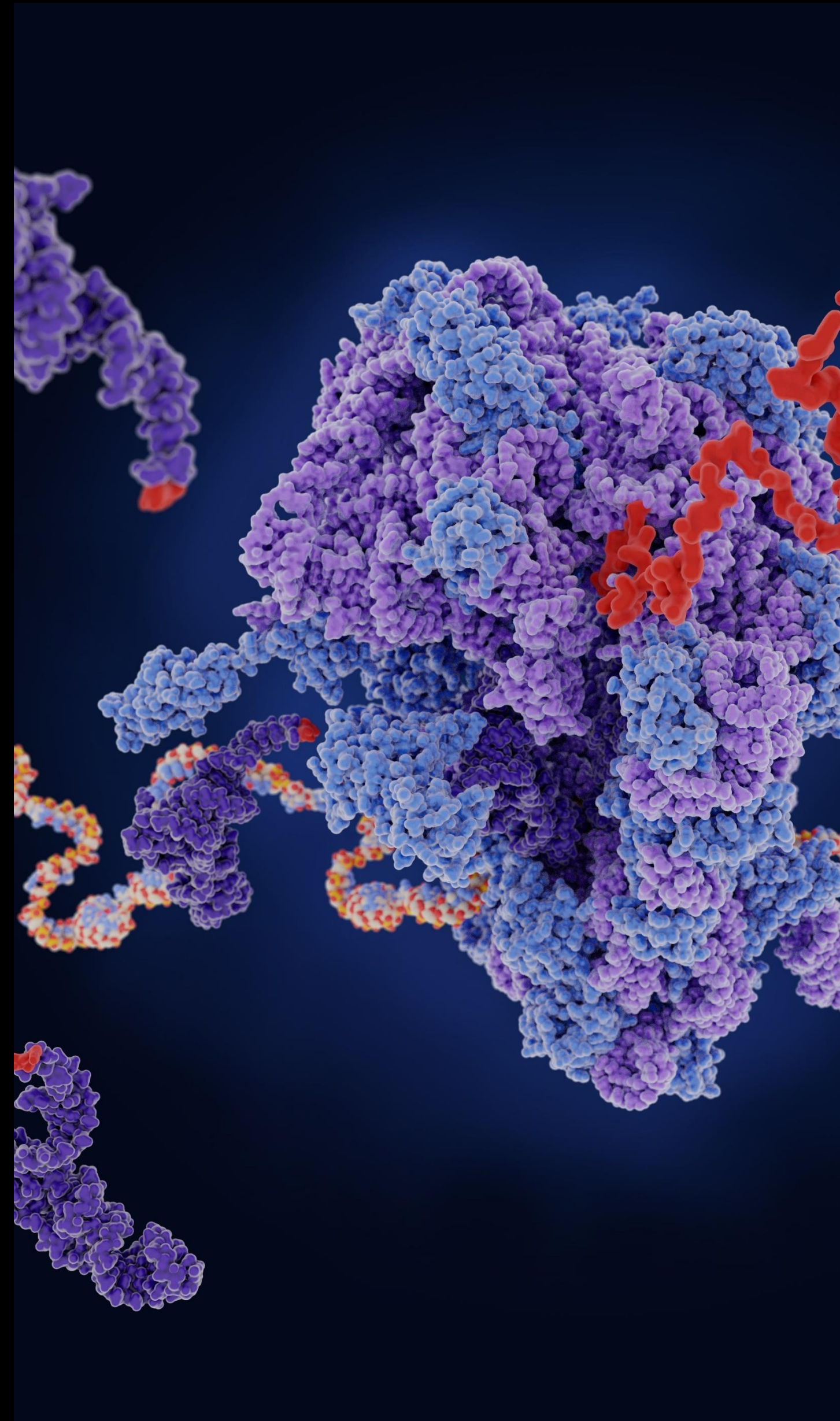
OSU ialltoallv 32 (full) PPN, Size = 128 KB



HIGHER APPLICATION PERFORMANCE

With BlueField DPU and Quantum In-Network Computing

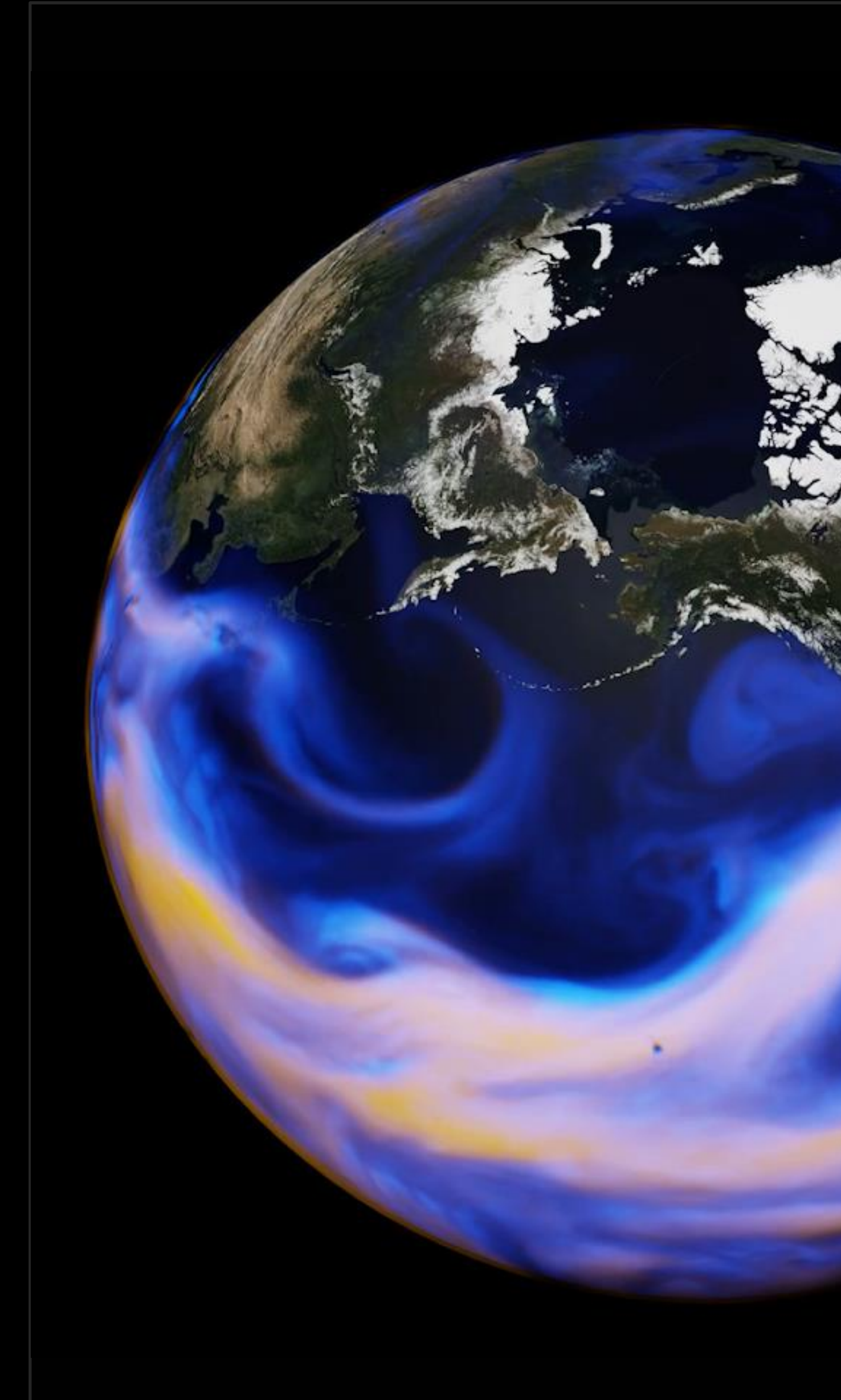
MOLECULAR DYNAMICS



MATHEMATICAL MODELING



WEATHER FORECASTING



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

HIGHER APPLICATION PERFORMANCE

With BlueField DPU and Quantum In-Network Computing

MOLECULAR DYNAMICS

MATHEMATICAL MODELING

WEATHER FORECASTING



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

