



NVIDIA Cloud Native Supercomputing Platform

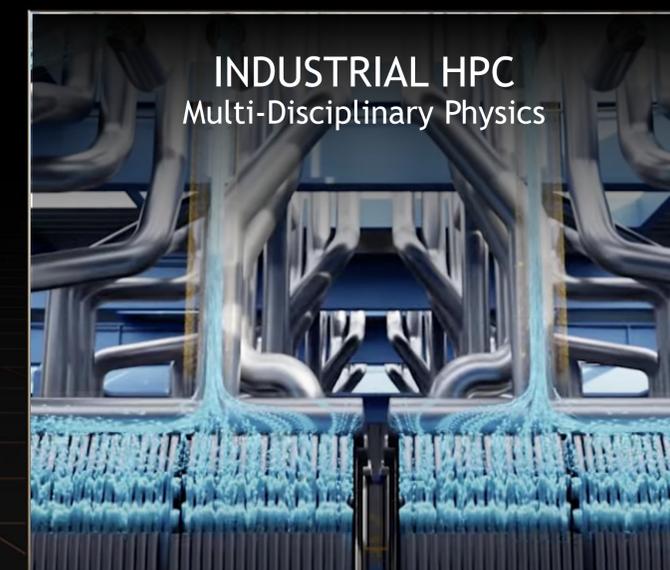
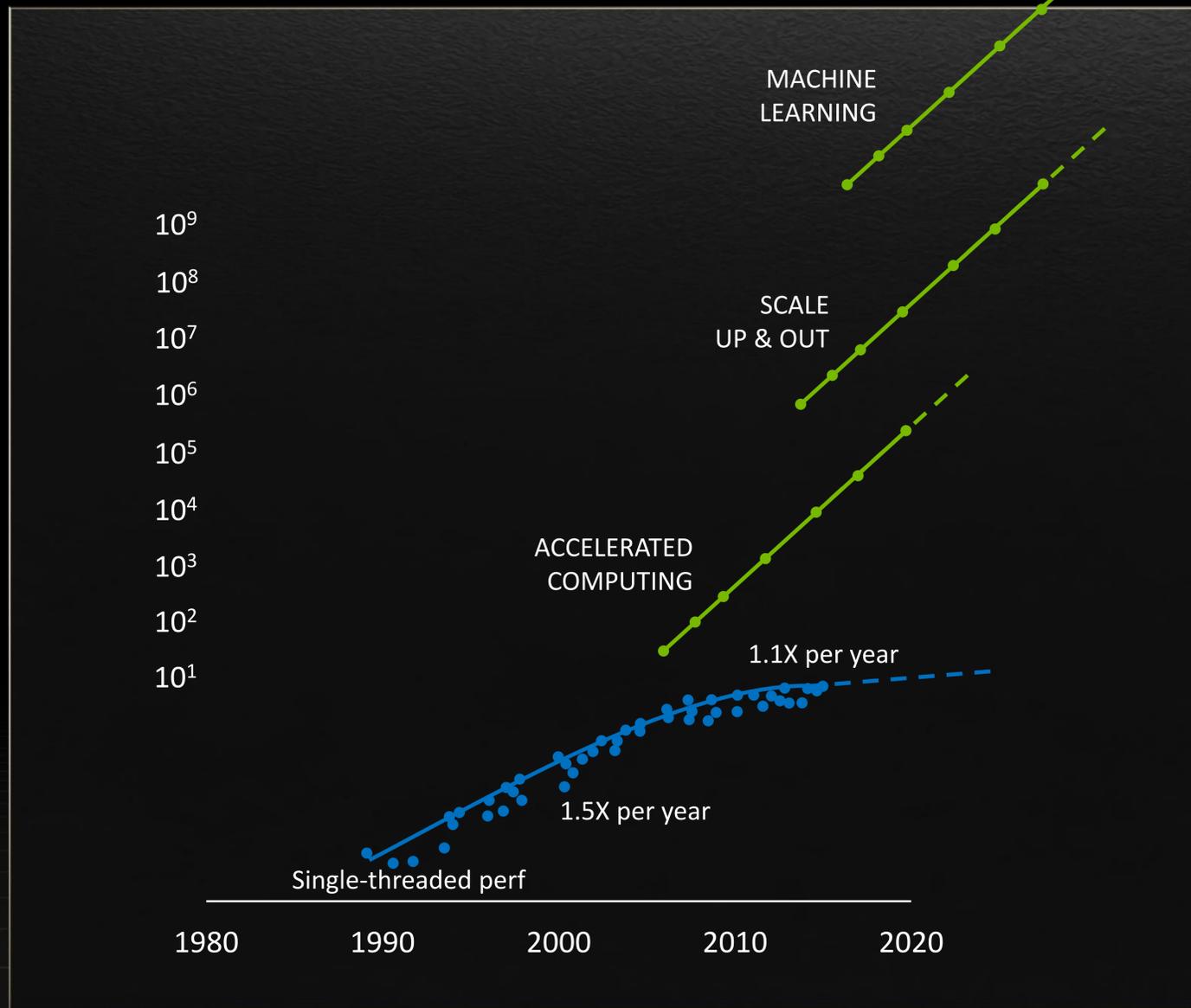
ご紹介

HPC/AI NETWORKING PRODUCT MARKETING DIRECTOR

MASAKI IWATANI

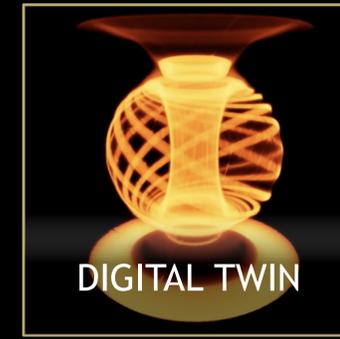
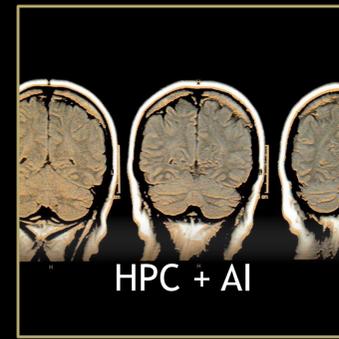
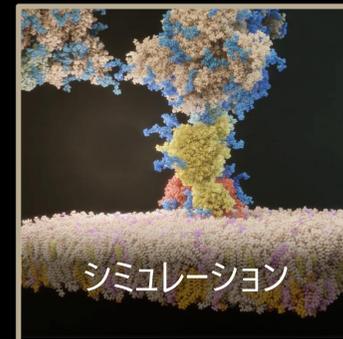
JUNE 2022

データおよびモデルサイズの爆発的増加



シミュレーション+AI：イノベーションとディスカバリーのための100万倍のスピードアップ

多種多様なアプリケーションを様々な場所を実施



AI Supercomputer



Cloud



Edge

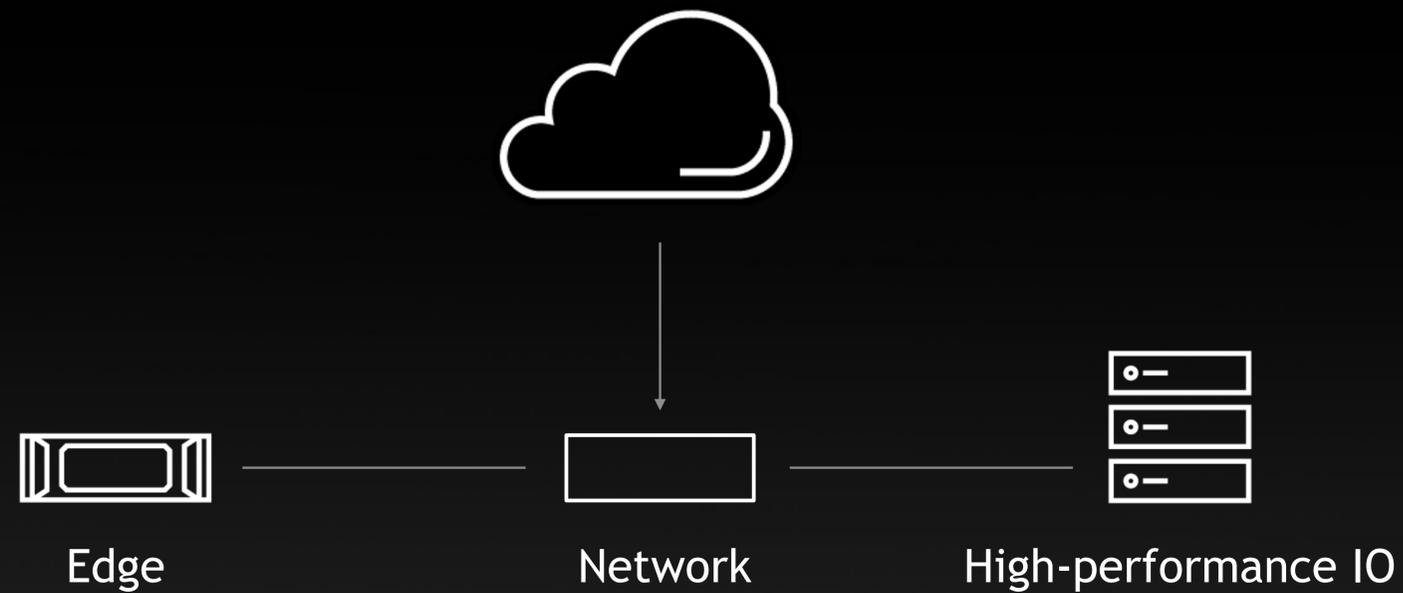
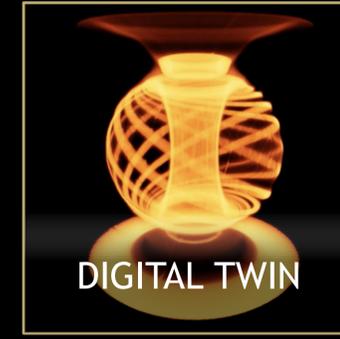
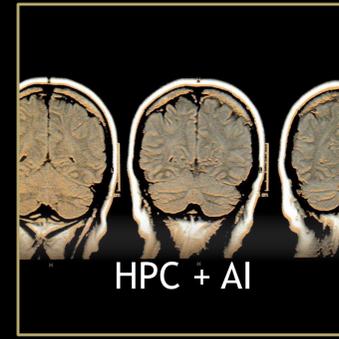
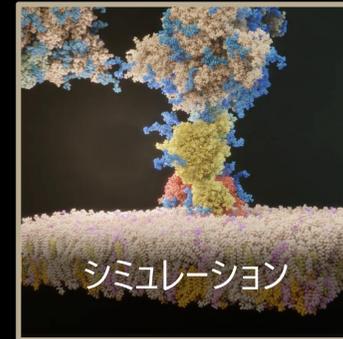


Network



High-performance IO

CLOUD NATIVE SUPERCOMPUTING



CLOUD NATIVE SUPERCOMPUTINGが実現出来るもの

NVIDIA CLOUD NATIVE
SUPERCOMPUTING

In-Network Computing

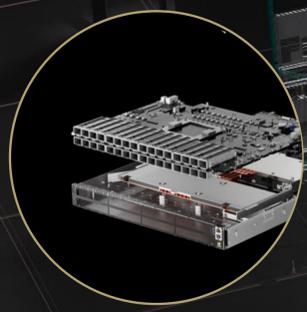
Computational Storage

Performance Isolation

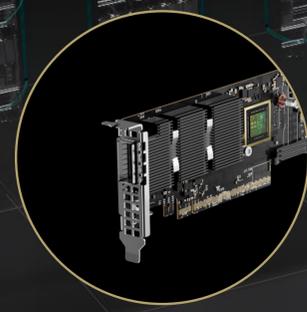
Enhanced Telemetry

Zero Trust Security

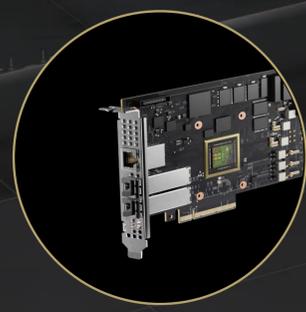
CLOUD NATIVE SUPERCOMPUTING 基盤であるNVIDIA QUANTUM-2 IB PLATFORM



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



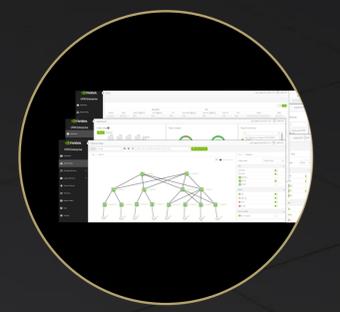
CONNECTX-7 SMARTNIC
Intelligent Offloads
Precision Timing
Software Defined Networking



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



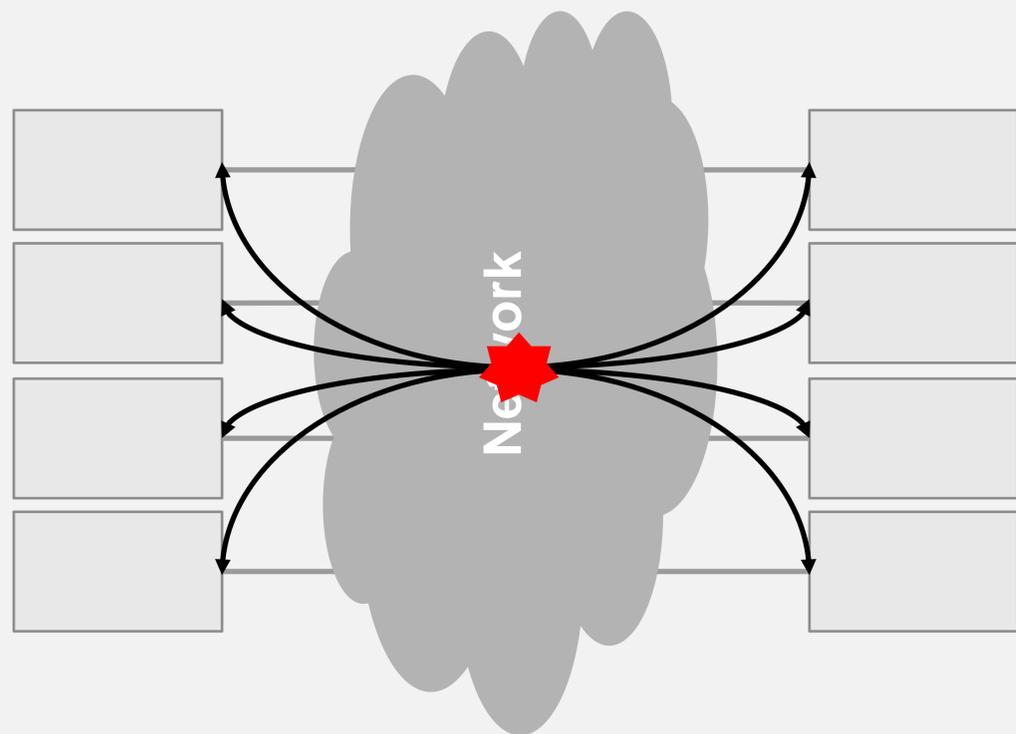
SKYWAY GATEWAY
InfiniBand to Ethernet
Low Latency
Load Balancing



UFM
Monitoring, Management, Orchestration
Predictive Maintenance
Anomaly Detection

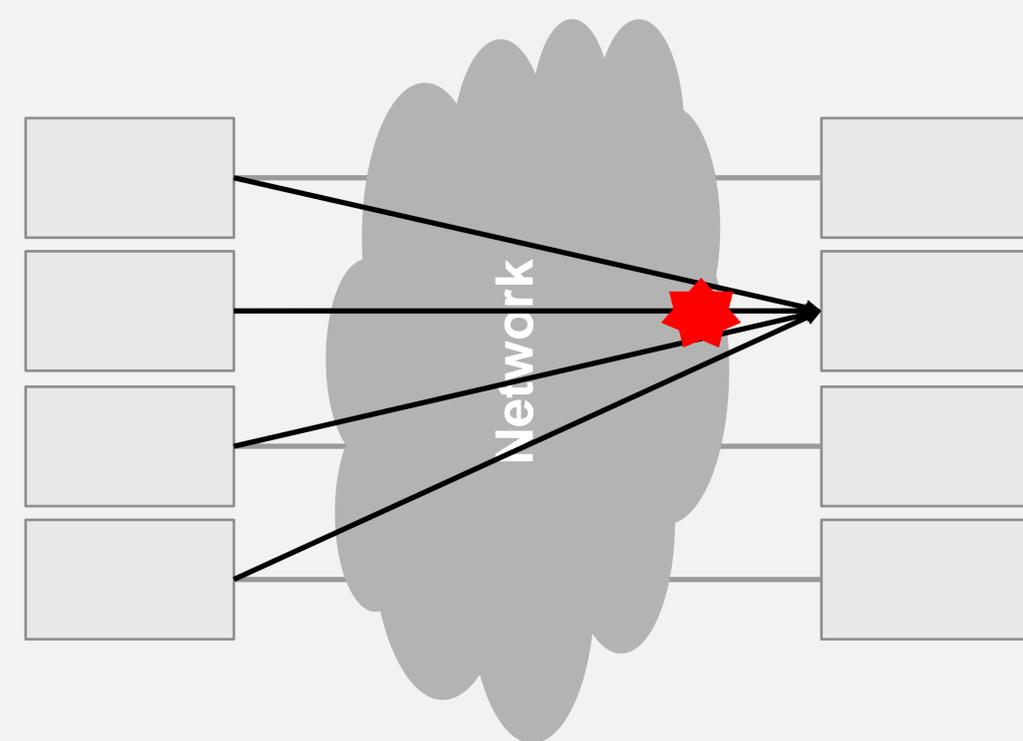
ネットワークの2つの輻輳による通信遅延

In-network 輻輳



Solution: Adaptive Routing

In-cast 輻輳

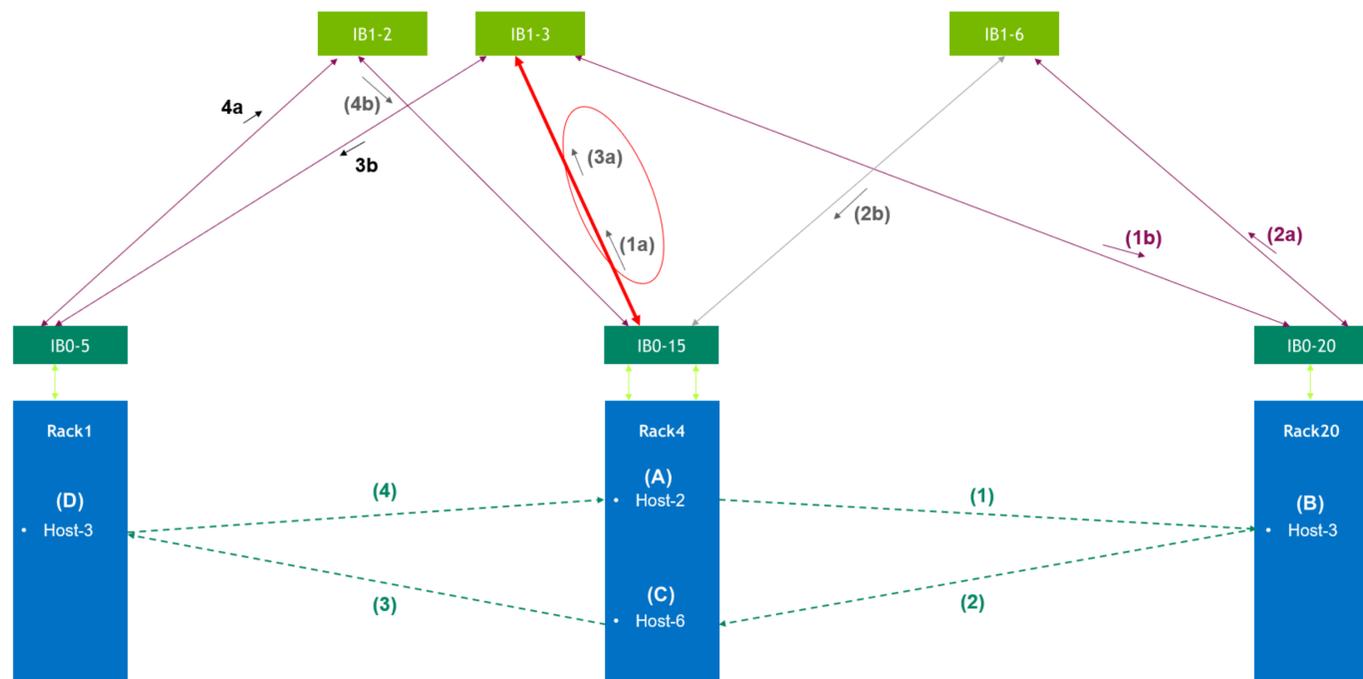


Solution: Congestion Control

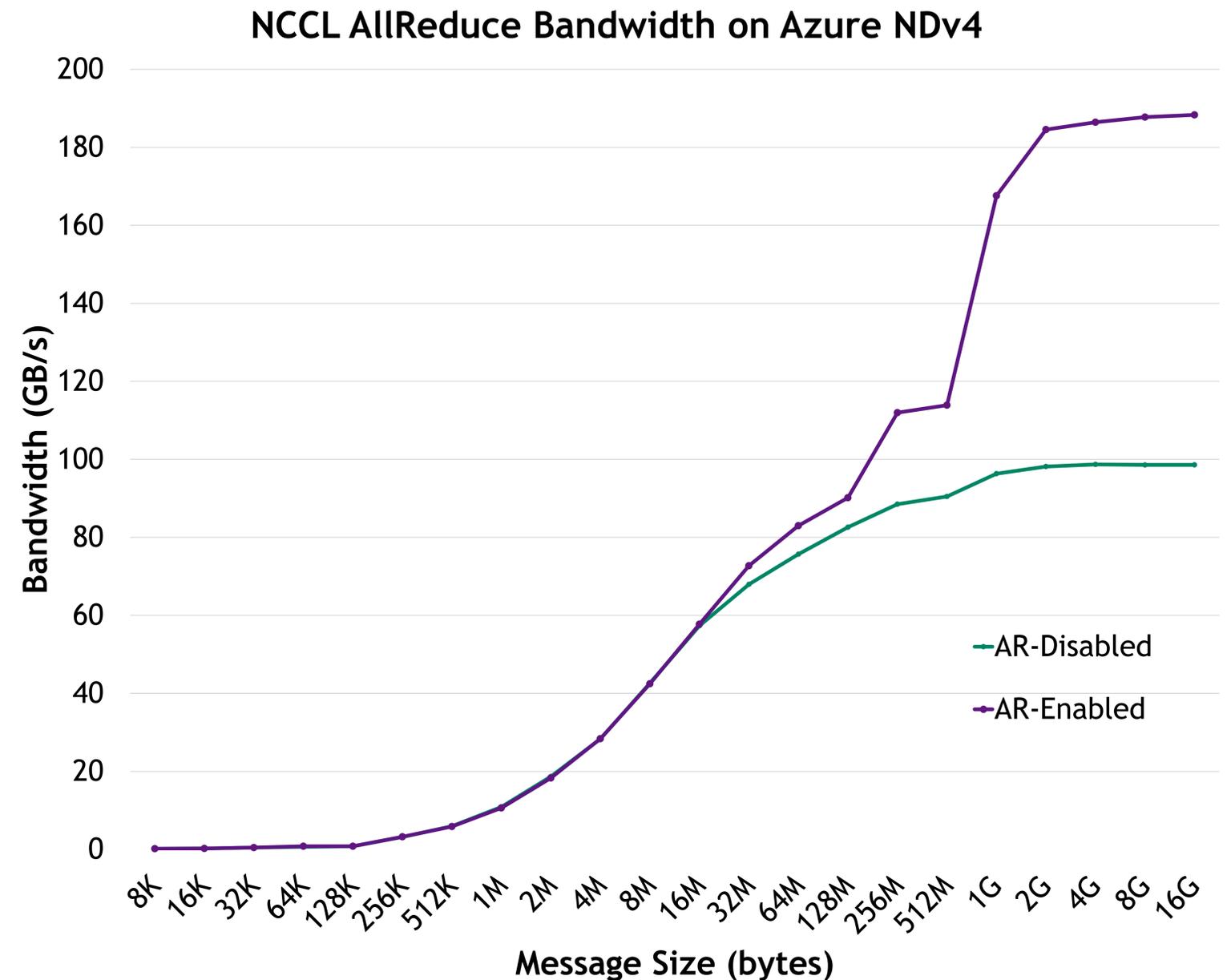
AZURE HPC パフォーマンス、ネットワーク利用率の最大化

InfiniBand Adaptive Routing

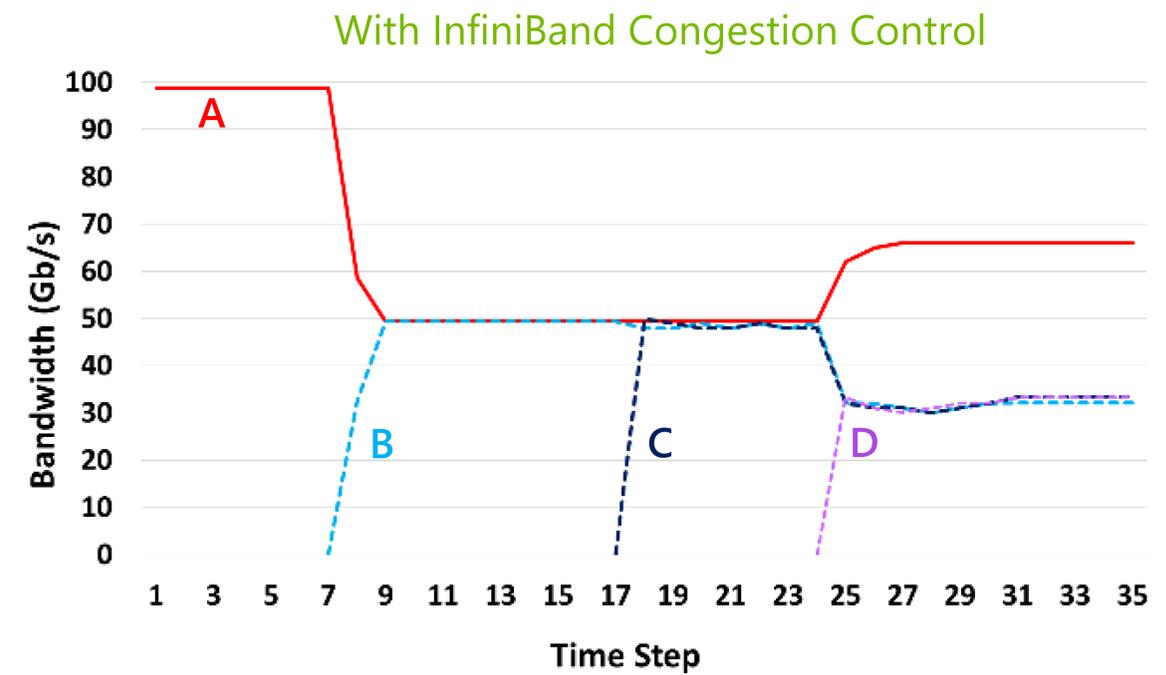
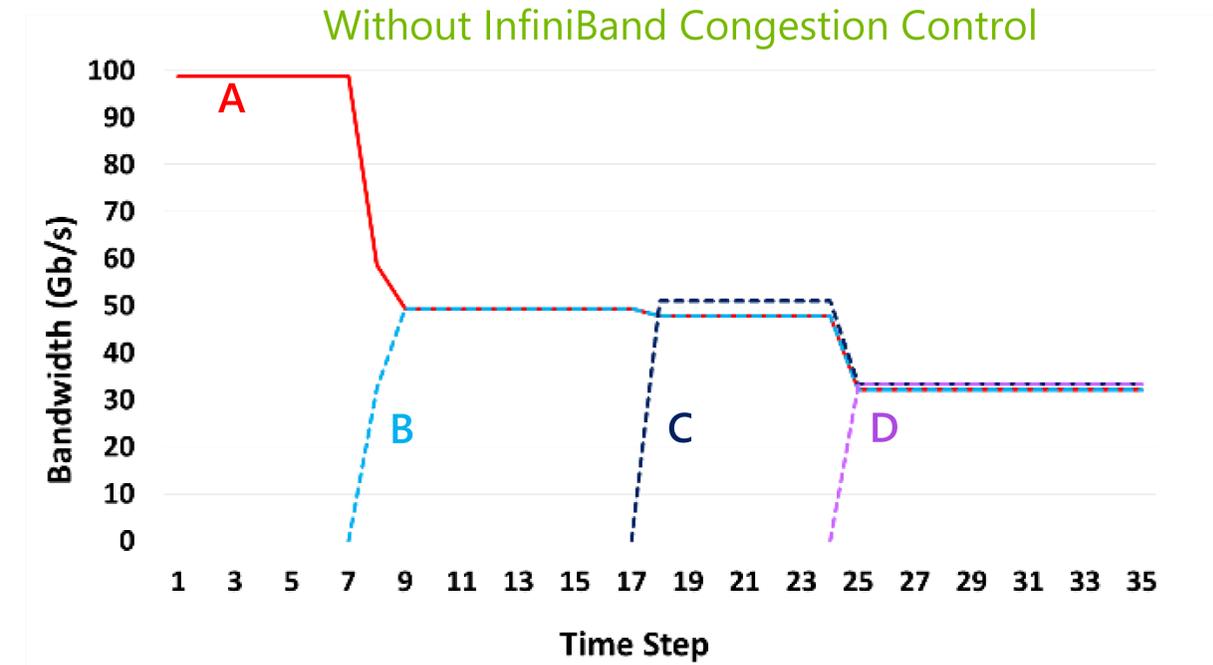
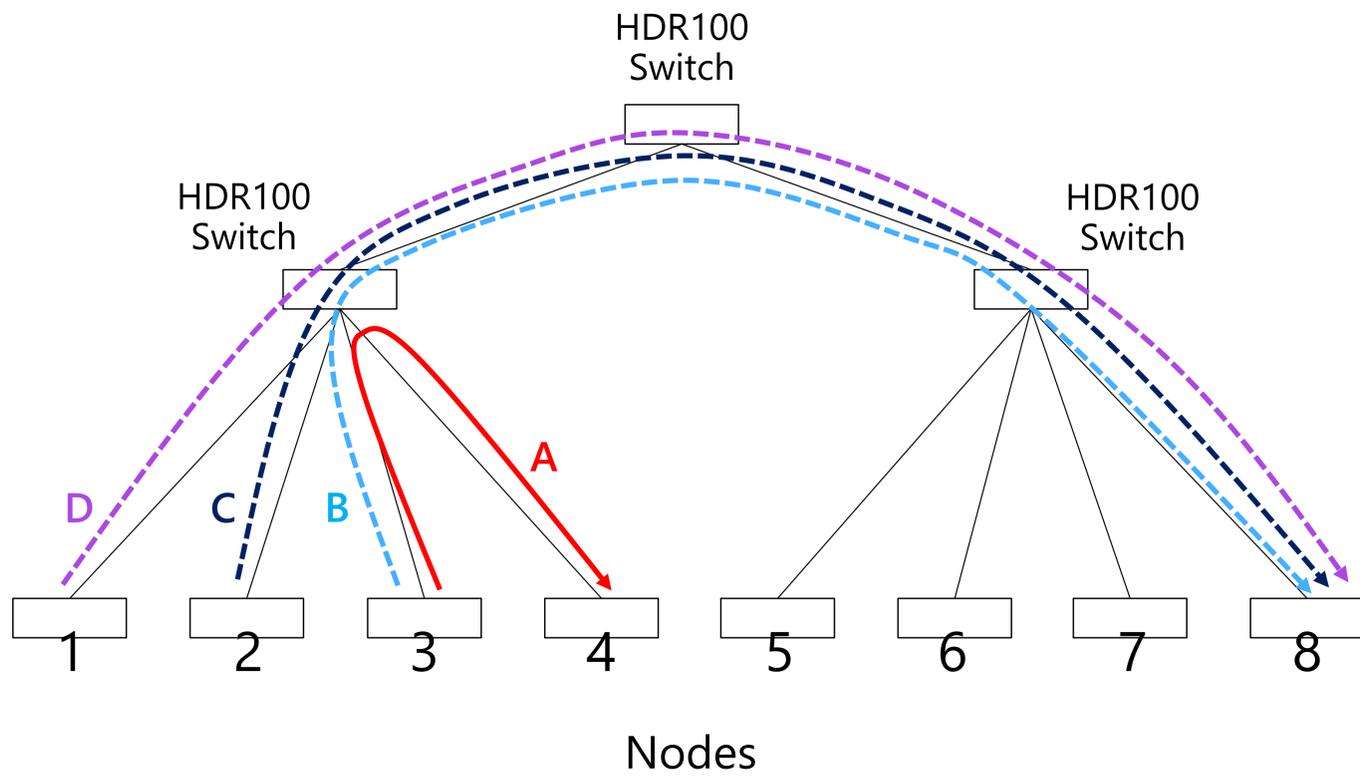
- Azure HPC InfiniBand Networks は
 - ノンブロッキング
 - オーバーサブスクリプション発生しない
- スタティックルーティングは、ネットワークの利用を制限する場合がある
 - 異なるネットワークトラフィックが同じネットワーク経路を共有することがある
 - Adaptive Routingはこれを防止し、シームレスなパフォーマンスを提供



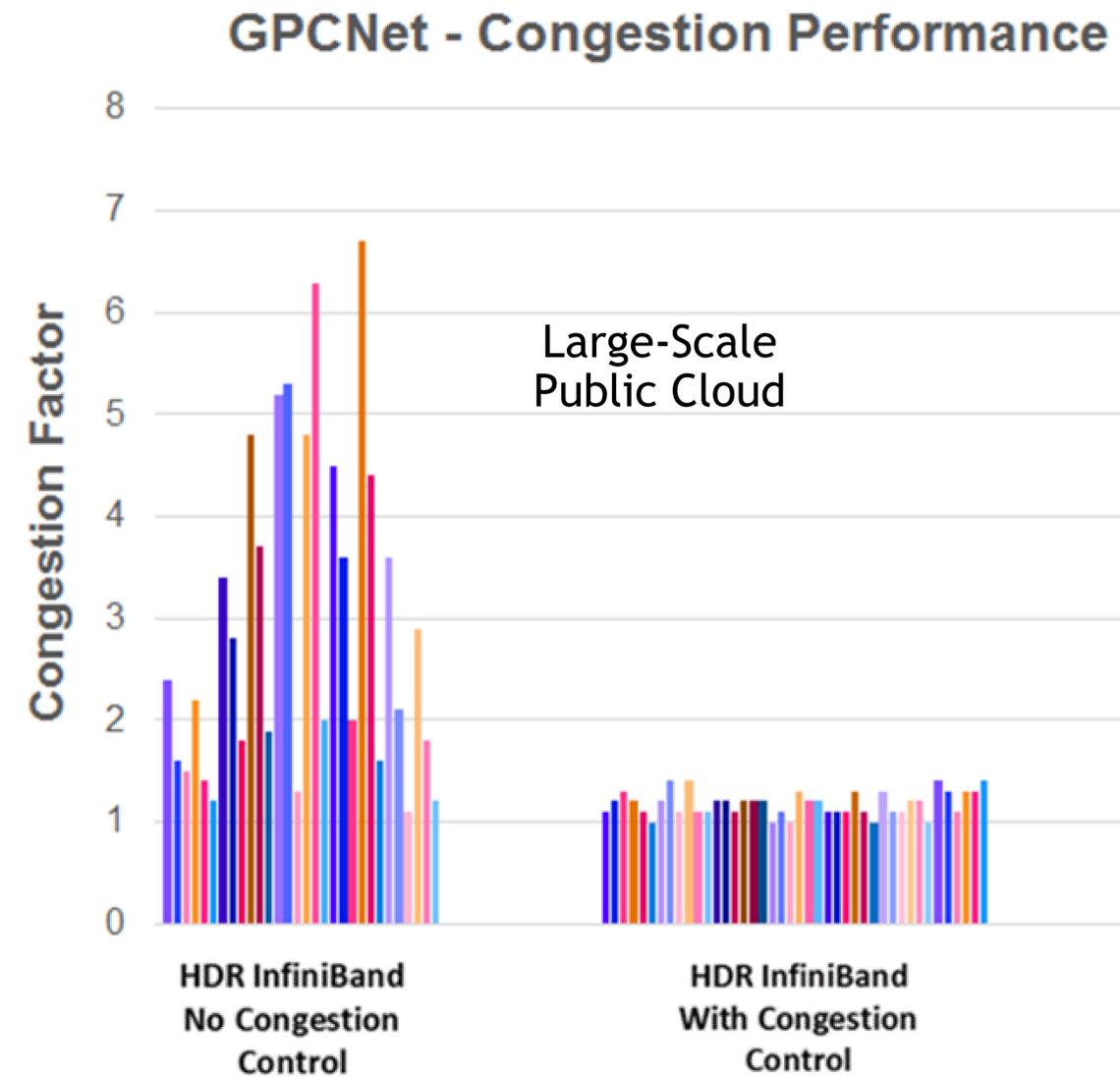
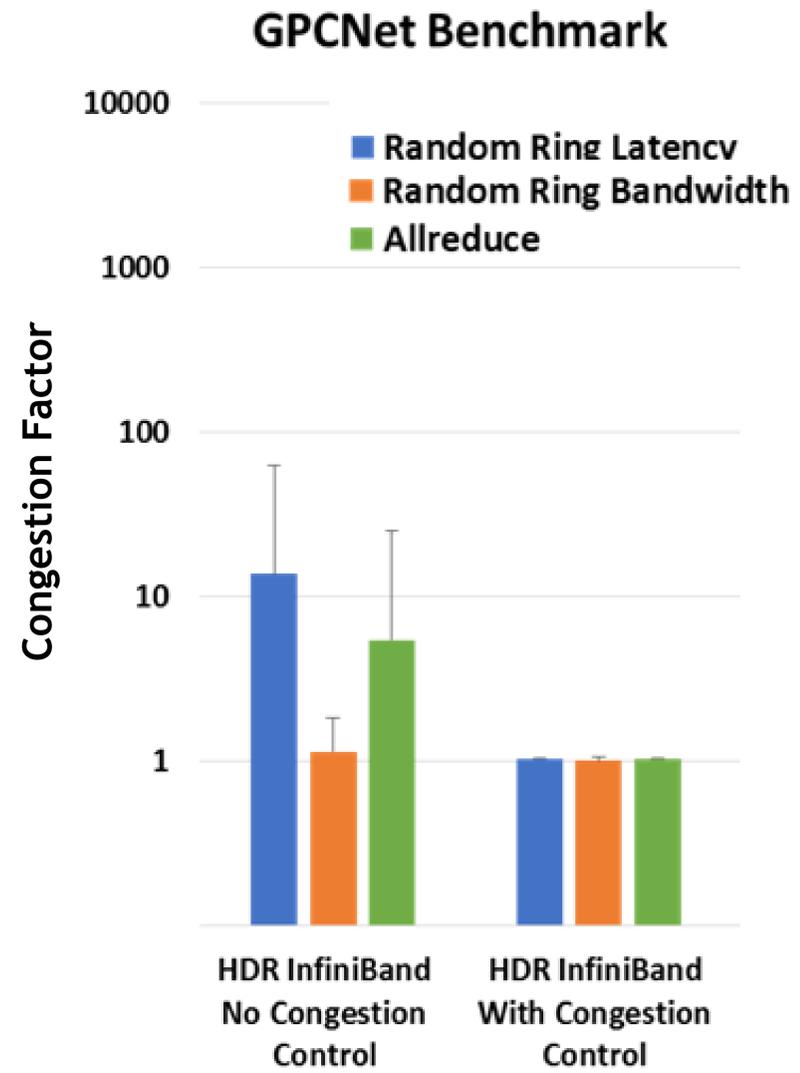
NCCL AllReduceジョブ中の通信経路。スタティックルーティングがリンクオーバーサブスクリプションを引き起こす



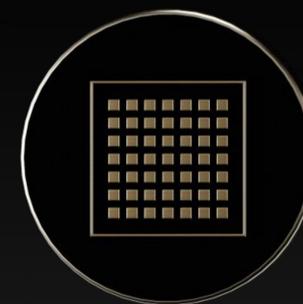
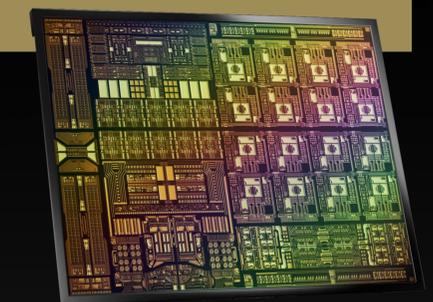
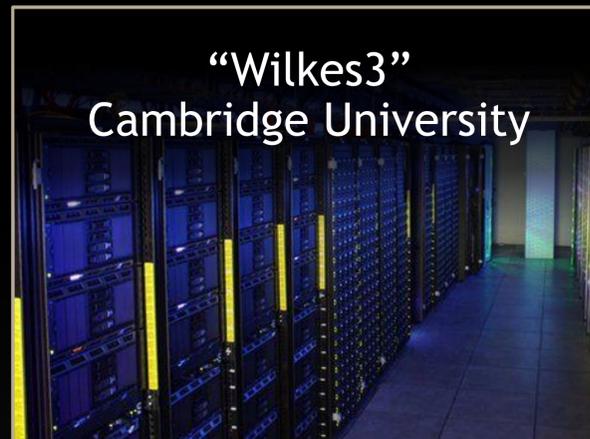
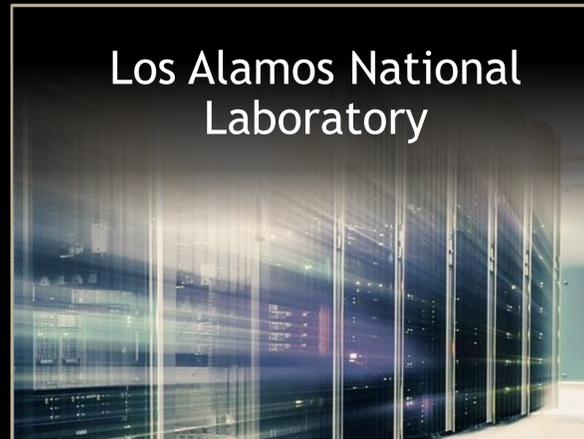
HDR INFINIBAND CONGESTION CONTROL



INFINIBAND CONGESTION CONTROL



クラウドネイティブ・スーパーコンピューティングを実現する



PROGRAMMABLE
COMPUTE

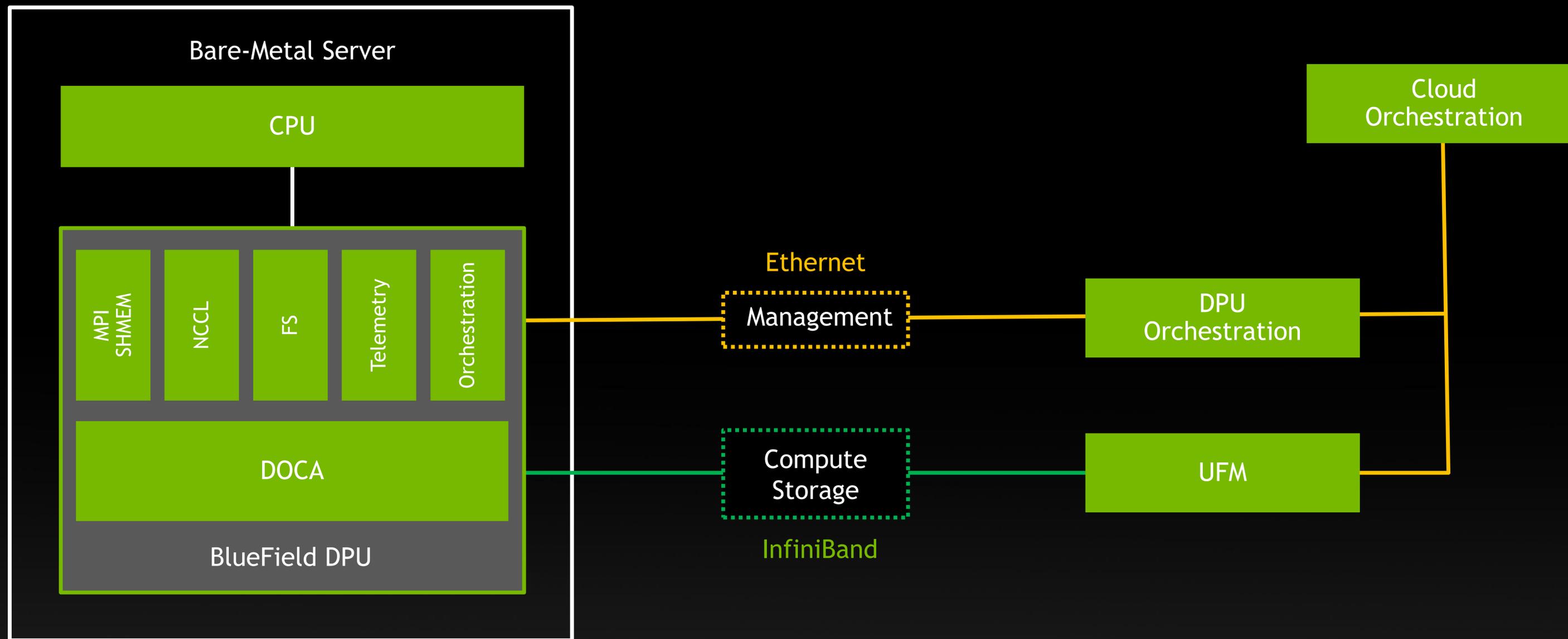


COMPUTATIONAL
STORAGE

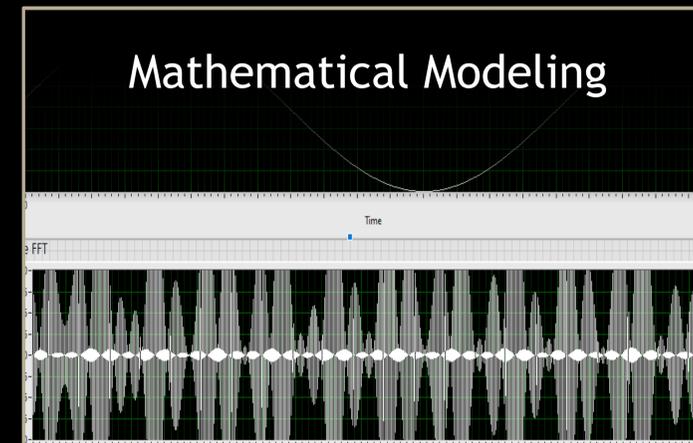
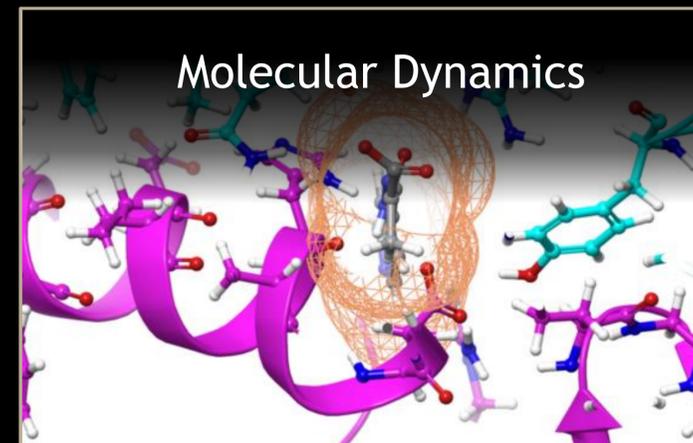
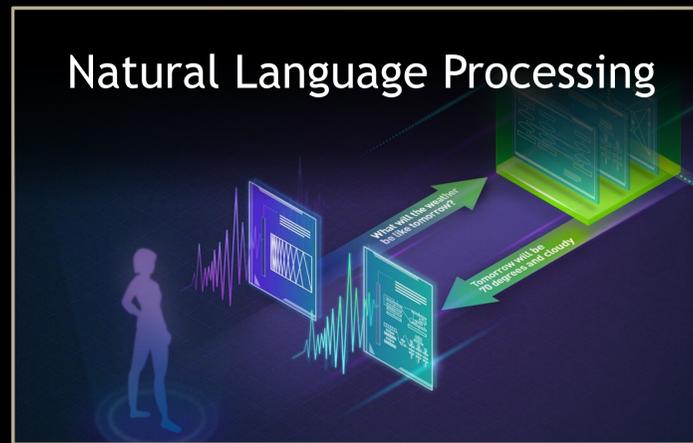


ZERO-TRUST
SECURITY

クラウドネイティブ・スーパーコンピューティングを実現する



CLOUD-NATIVE SUPERCOMPUTING DRIVING APPLICATION PERFORMANCE GAINS



Barcelona
Supercomputing Center

Durham
University

Georgia Institute
of Technology

Japan Meteorological
Agency

Los Alamos
National Laboratory

Ohio State
University

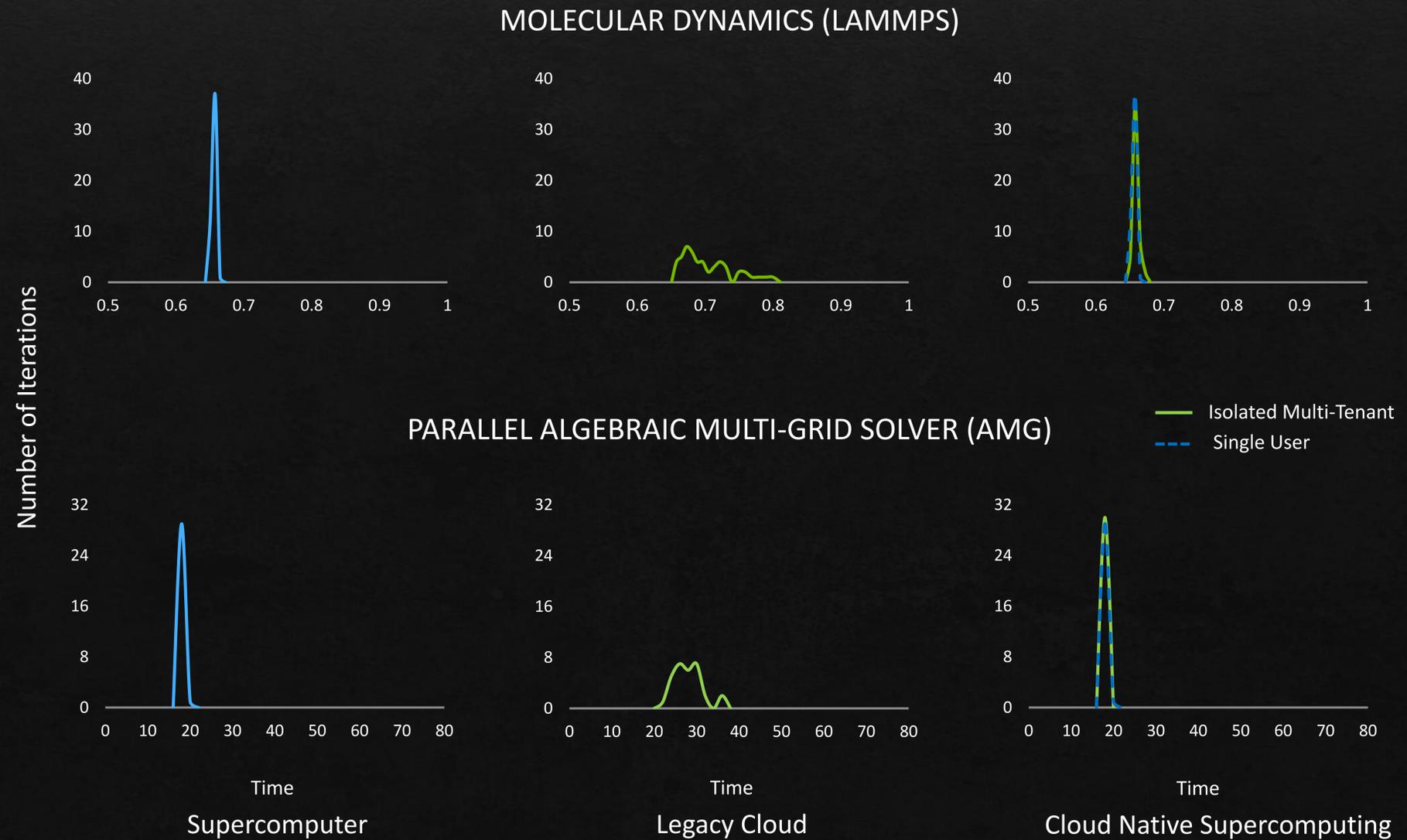
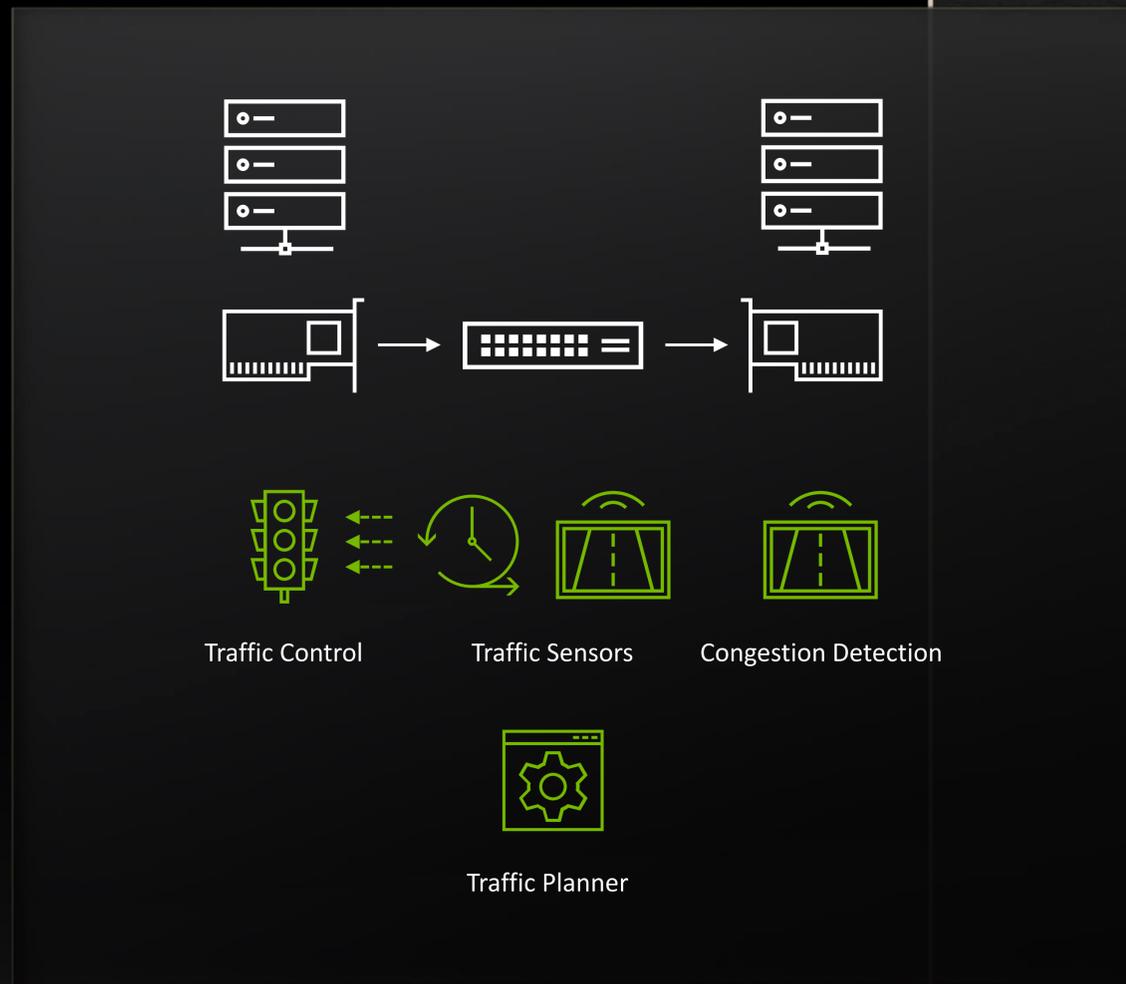
Sandia National
Laboratory

Technical University
Munich

University College
London

1.2X TO 1.4X HIGHER APPLICATION PERFORMANCE
With BlueField DPU and Quantum In-Network Computing

クラウドネイティブのスーパーコンピューティングでパフォーマンスアイソレーションを実現

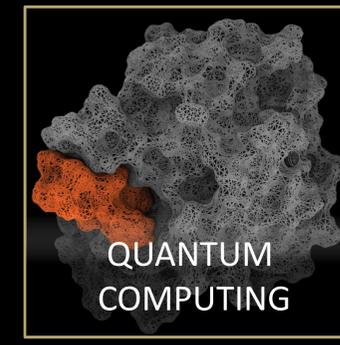
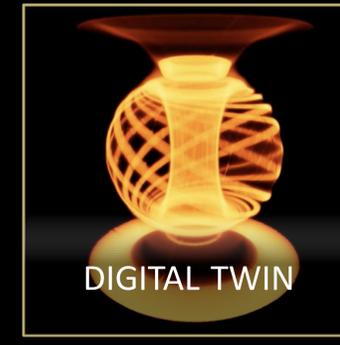
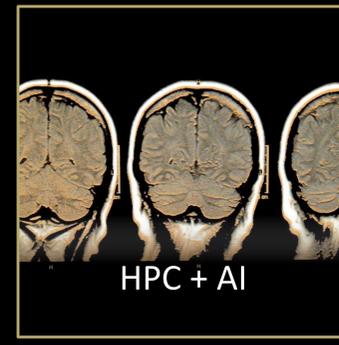
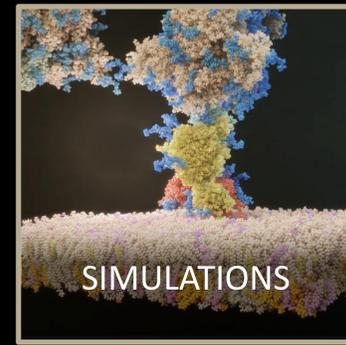


Bluefield DPU

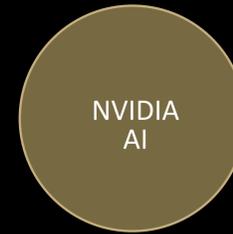
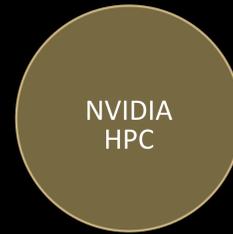


Quantum InfiniBand

APPLICATIONS



PLATFORM



SYSTEM SOFTWARE



HARDWARE

