Interconnect Your Future

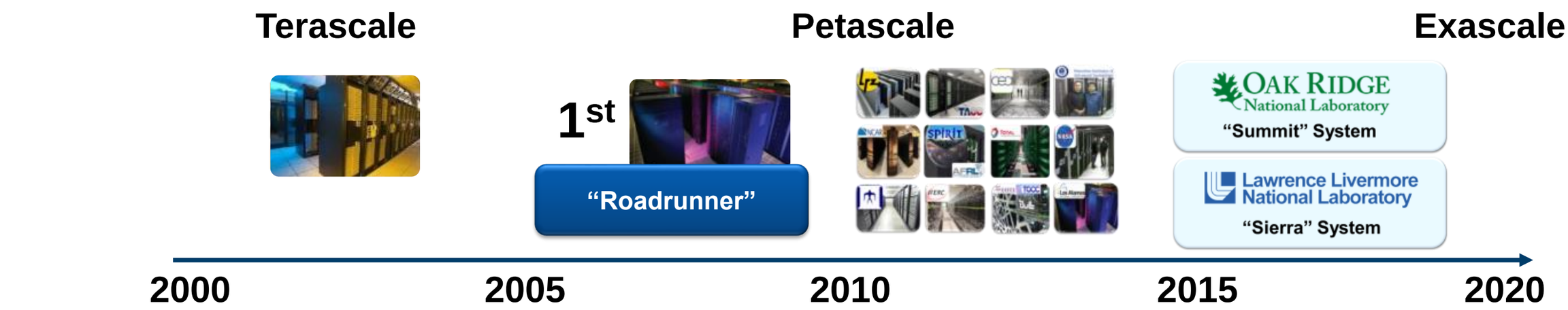
Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™

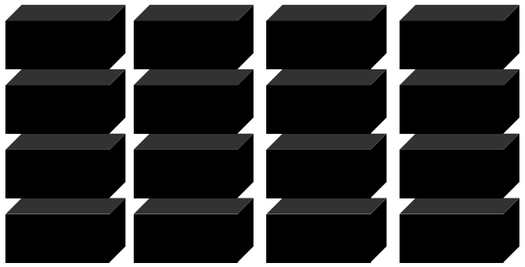
September 2016

The Ever Growing Demand for Higher Performance

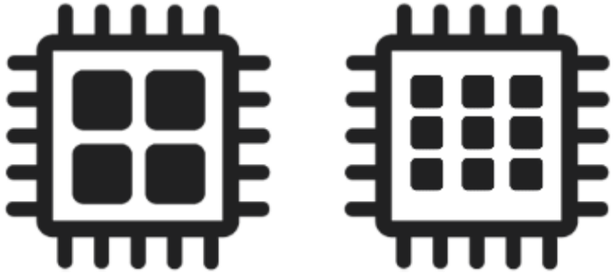
Performance Development



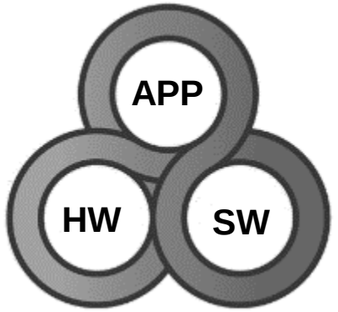
The Interconnect is the Enabling Technology



SMP to Clusters



Single-Core to Many-Core



Application

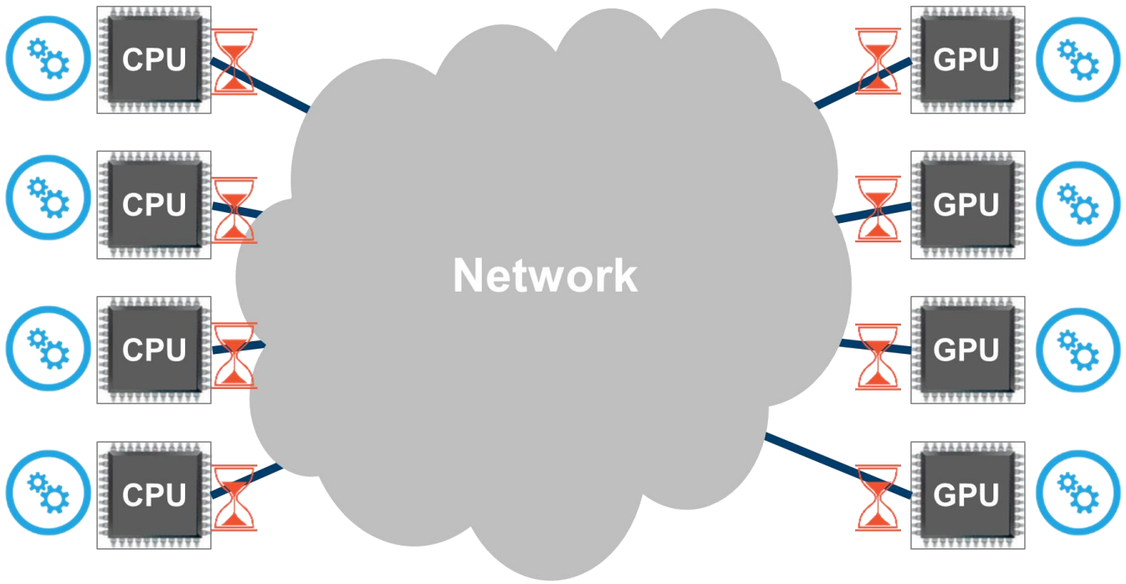
Software

Hardware

Co-Design

The Intelligent Interconnect to Enable Exascale Performance

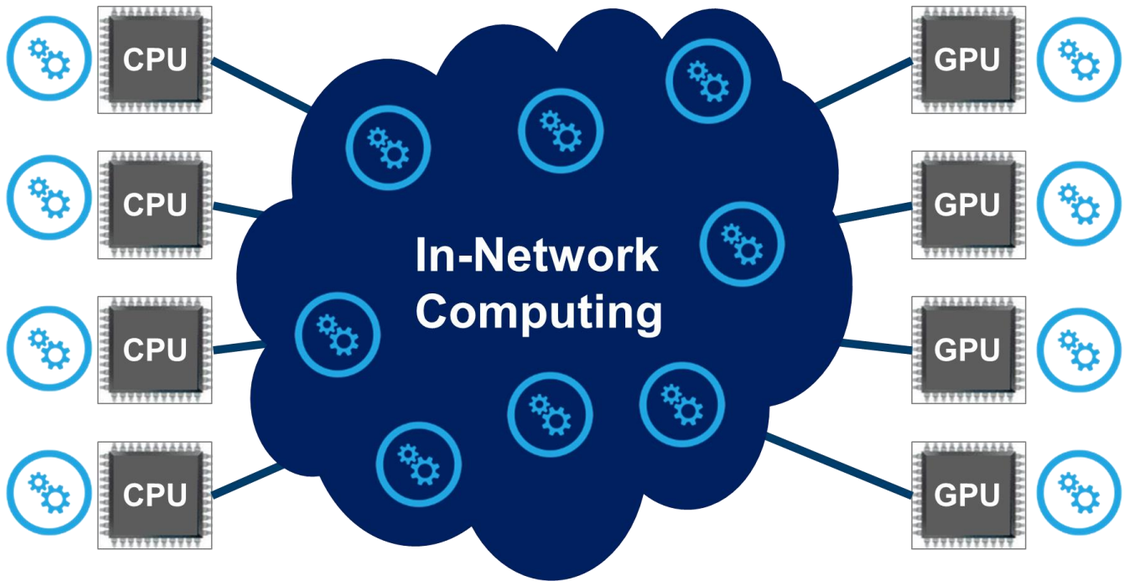
CPU-Centric



Limited to Main CPU Usage
Results in Performance Limitation

Must Wait for the Data
Creates Performance Bottlenecks

Co-Design



Creating Synergies
Enables Higher Performance and Scale

Work on The Data as it Moves
Enables Performance and Scale

ISC'16: Introducing ConnectX-5 World's Smartest Adapter

State of the Smart

a new generation of co-processors emerges

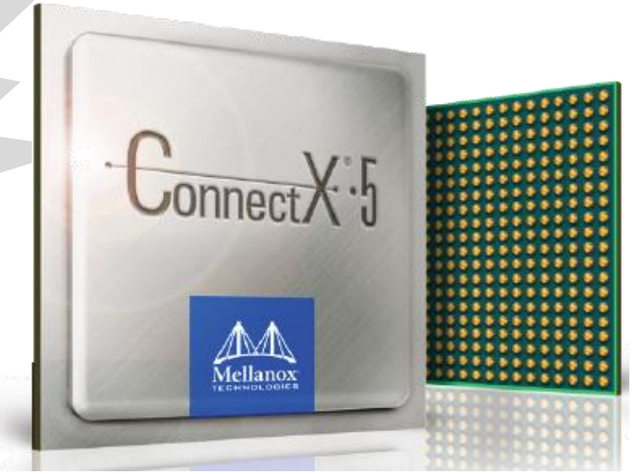
Mellanox Smart Interconnect

Switch IB 2

SHArP

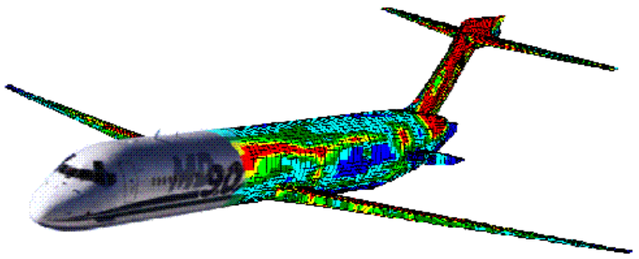
NEW!

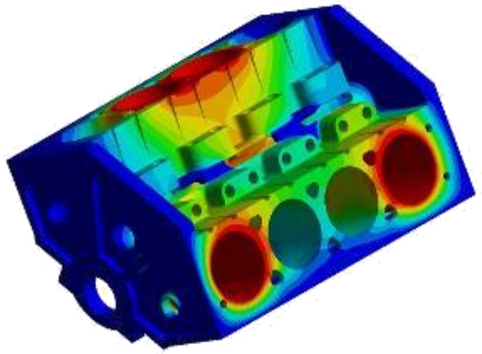
ConnectX-5

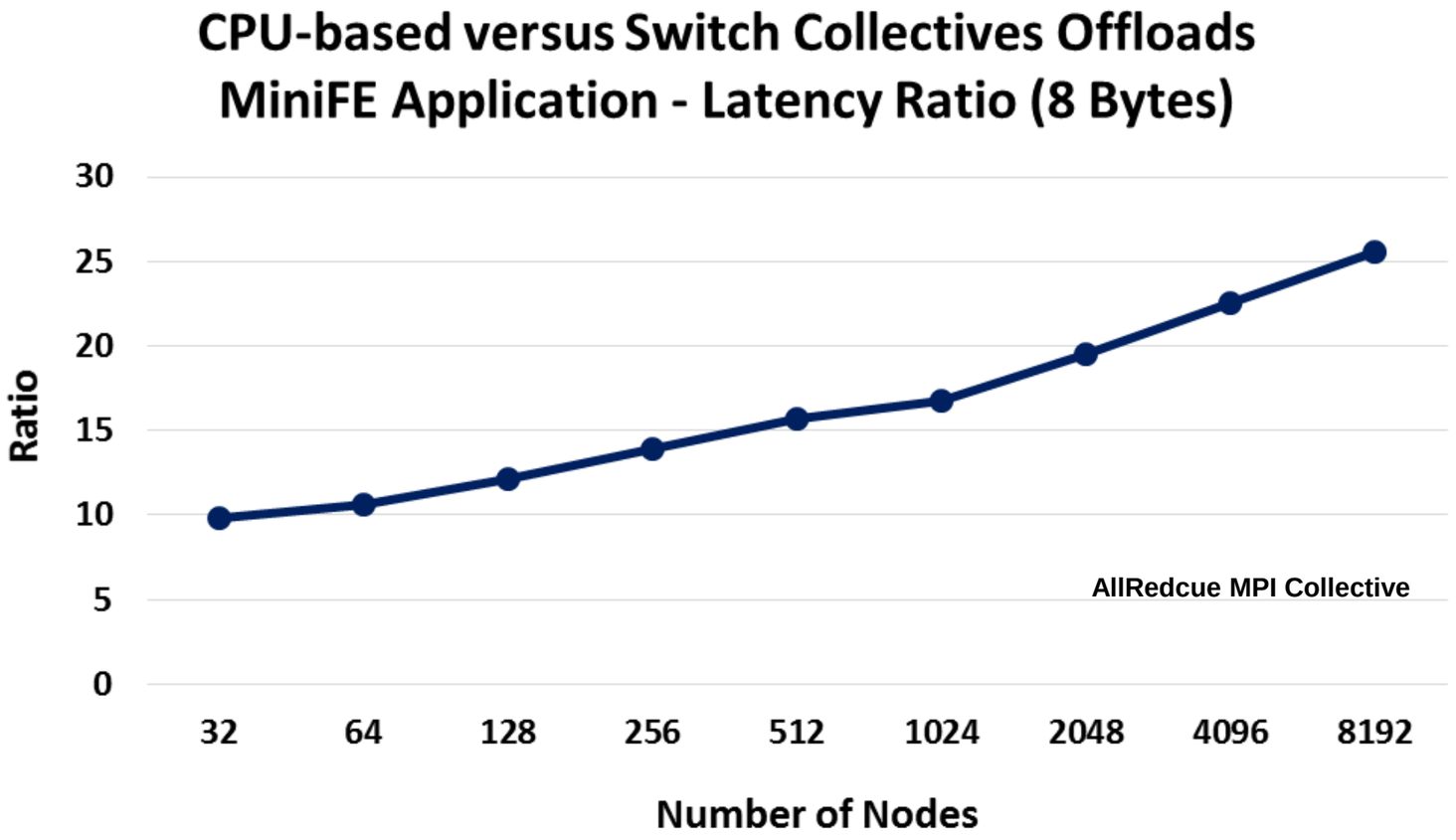


SHArP Performance Advantage

- MiniFE is a Finite Element mini-application
  - Implements kernels that represent implicit finite-element applications







10X to 25X Performance Improvement!

Highest-Performance 100Gb/s Interconnect Solutions

Adapters — ConnectX®-5

100Gb/s Adapter, 0.6us latency
200 million messages per second
(10 / 25 / 40 / 50 / 56 / 100Gb/s)

Switch — SwitchIB™-2

36 EDR (100Gb/s) Ports, <90ns Latency
Throughput of 7.2Tb/s
7.02 Billion msg/sec (195M msg/sec/port)

Switch — Spectrum™

32 100GbE Ports, 64 25/50GbE Ports
(10 / 25 / 40 / 50 / 100GbE)
Throughput of 6.4Tb/s

Interconnect — LinkX™

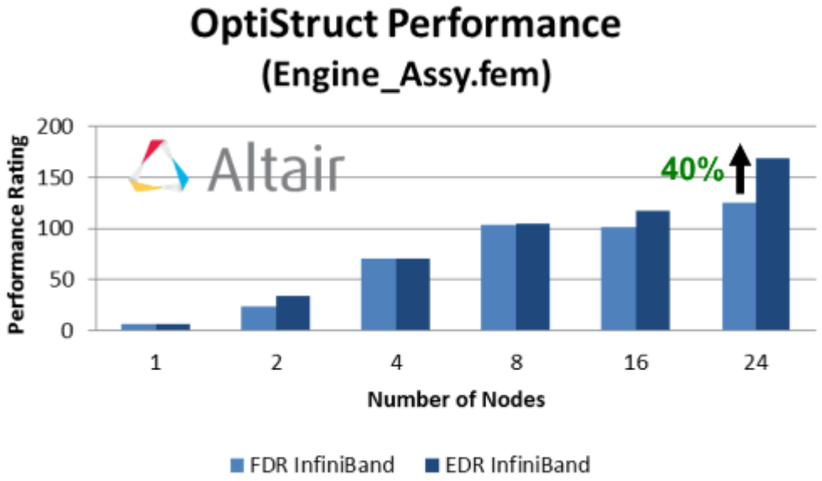
Transceivers
Active Optical and Copper Cables
(10 / 25 / 40 / 50 / 56 / 100Gb/s)

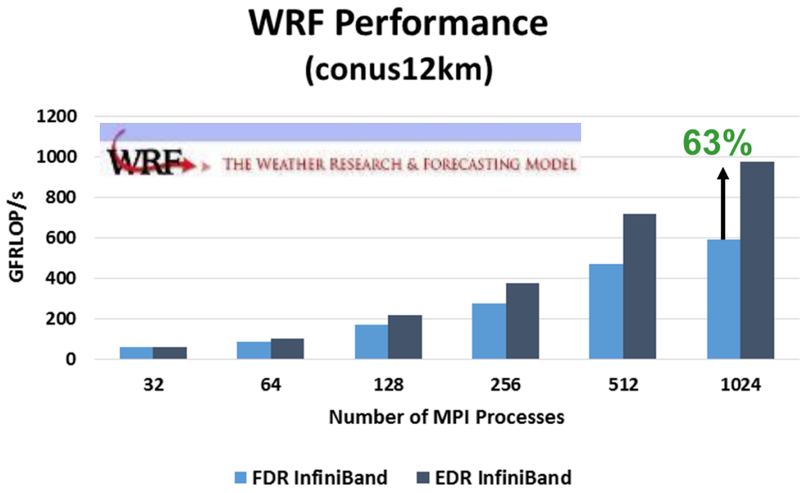
VCSELs, Silicon Photonics and Copper

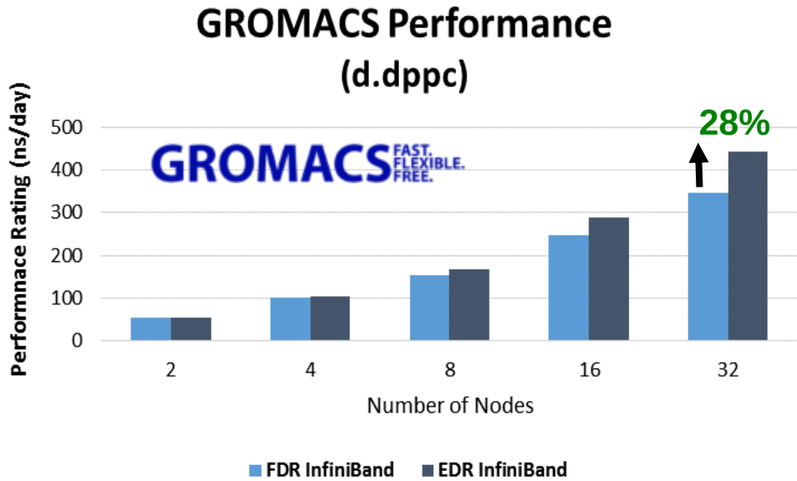
Software — HPC-X™

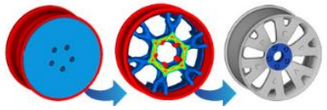
MPI, SHMEM/PGAS, UPC
For Commercial and Open Source Applications
Leverages Hardware Accelerations

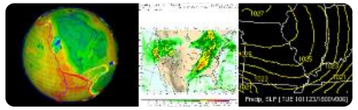
The Performance Advantage of EDR 100G InfiniBand (28-80%)

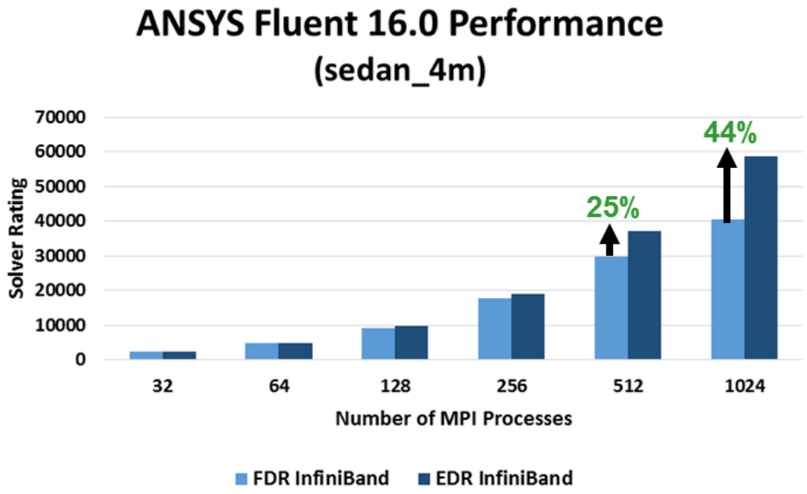


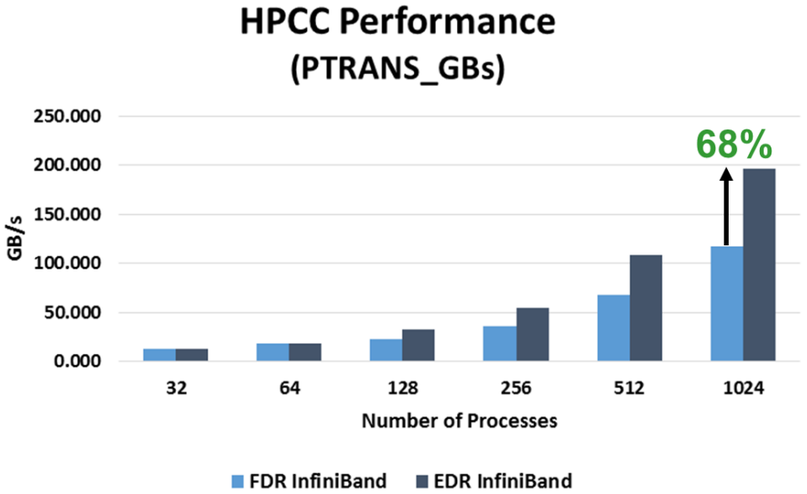


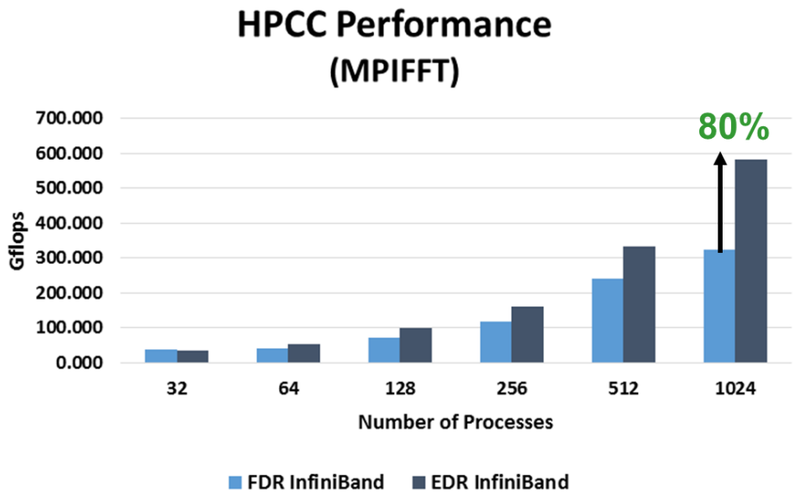












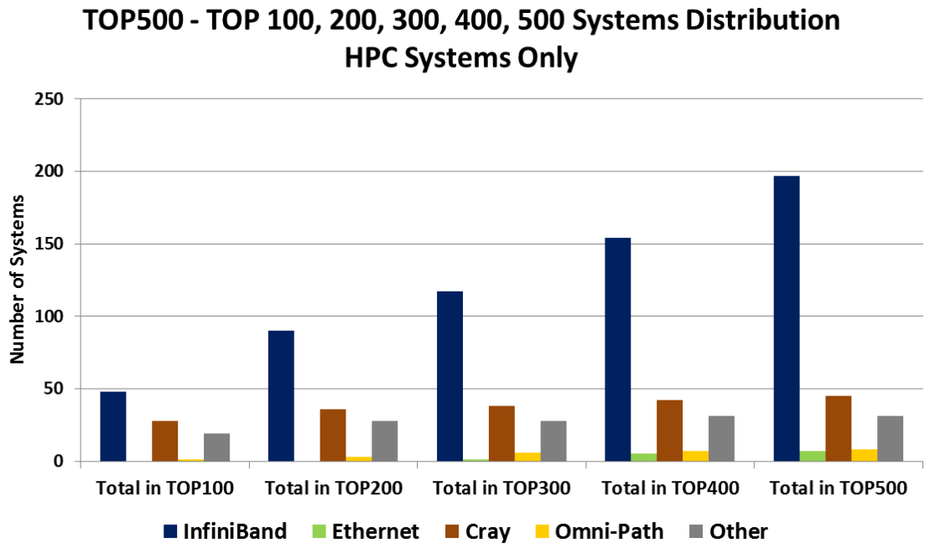
Mellanox Connects the World's Fastest Supercomputer

National Supercomputing Center in Wuxi, China

#1 on the TOP500 Supercomputing List

- 93 Petaflop performance, 3X higher versus #2 on the TOP500
- 40K nodes, 10 million cores, 256 cores per CPU
- Mellanox adapter and switch solutions

- The TOP500 list has evolved, includes HPC & Cloud / Web2.0 Hyperscale systems
- Mellanox connects 41.2% of overall TOP500 systems
- Mellanox connects 70.4% of the TOP500 HPC platforms
- Mellanox connects 46 Petascale systems, Nearly 50% of the total Petascale systems



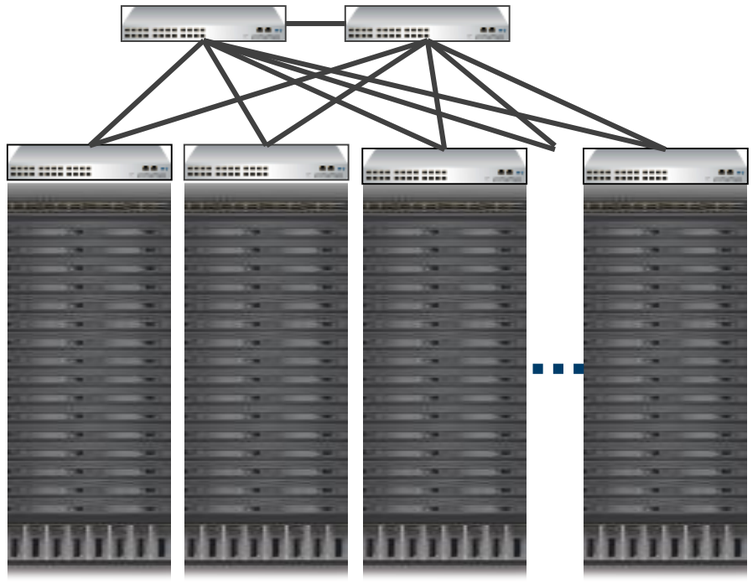
InfiniBand is the Interconnect of Choice for HPC Compute and Storage Infrastructures

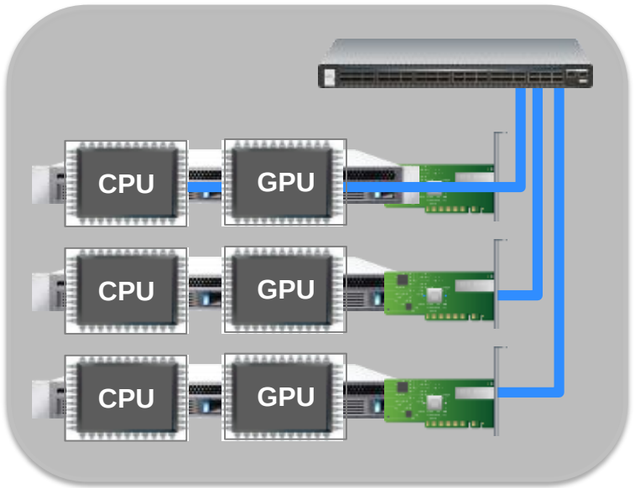
Machine Learning

GPUDirect Enables Efficient Training Platform for Deep Neural Network

(a) Face (b) Body (c) Cat

THE OHIO STATE UNIVERSITY





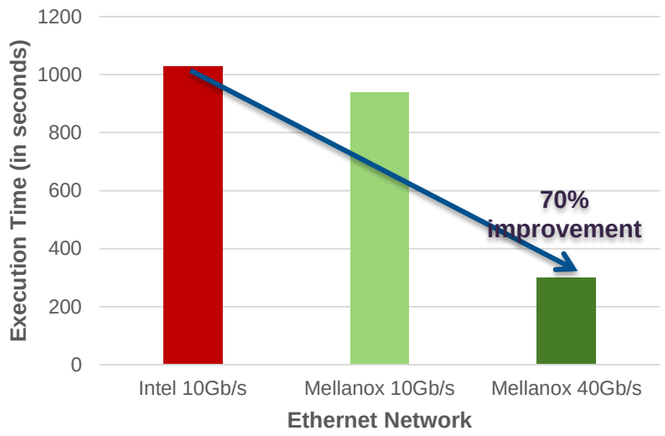
1K nodes (16K cores) for 1 week

3 Nodes with 3 GPUs for 3 days
Mellanox InfiniBand and GPU-Direct

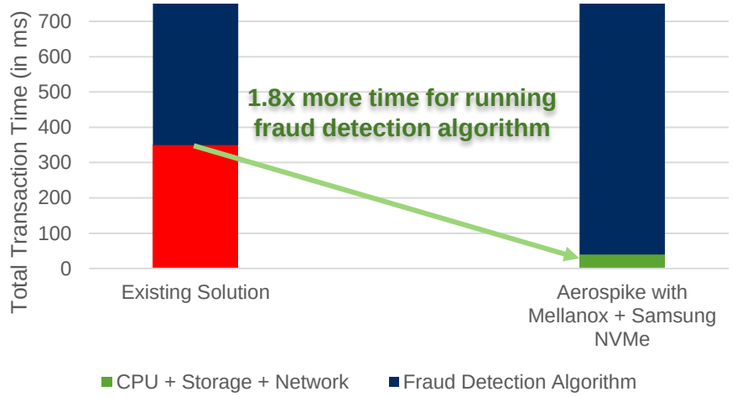
Deep Learning - Transforming Data to Intelligence

Enable Real-time Decision Making

TeraSort



Fraud Detection workload



Big Sur Machine Learning Platform

MAPR

Mellanox TECHNOLOGIES
Connect. Accelerate. Outperform.

SAMSUNG

Mellanox TECHNOLOGIES
Connect. Accelerate. Outperform.

AEROSPIKE

DELL

Mellanox TECHNOLOGIES
Connect. Accelerate. Outperform.

Innova flex

Cassandra

Couchbase

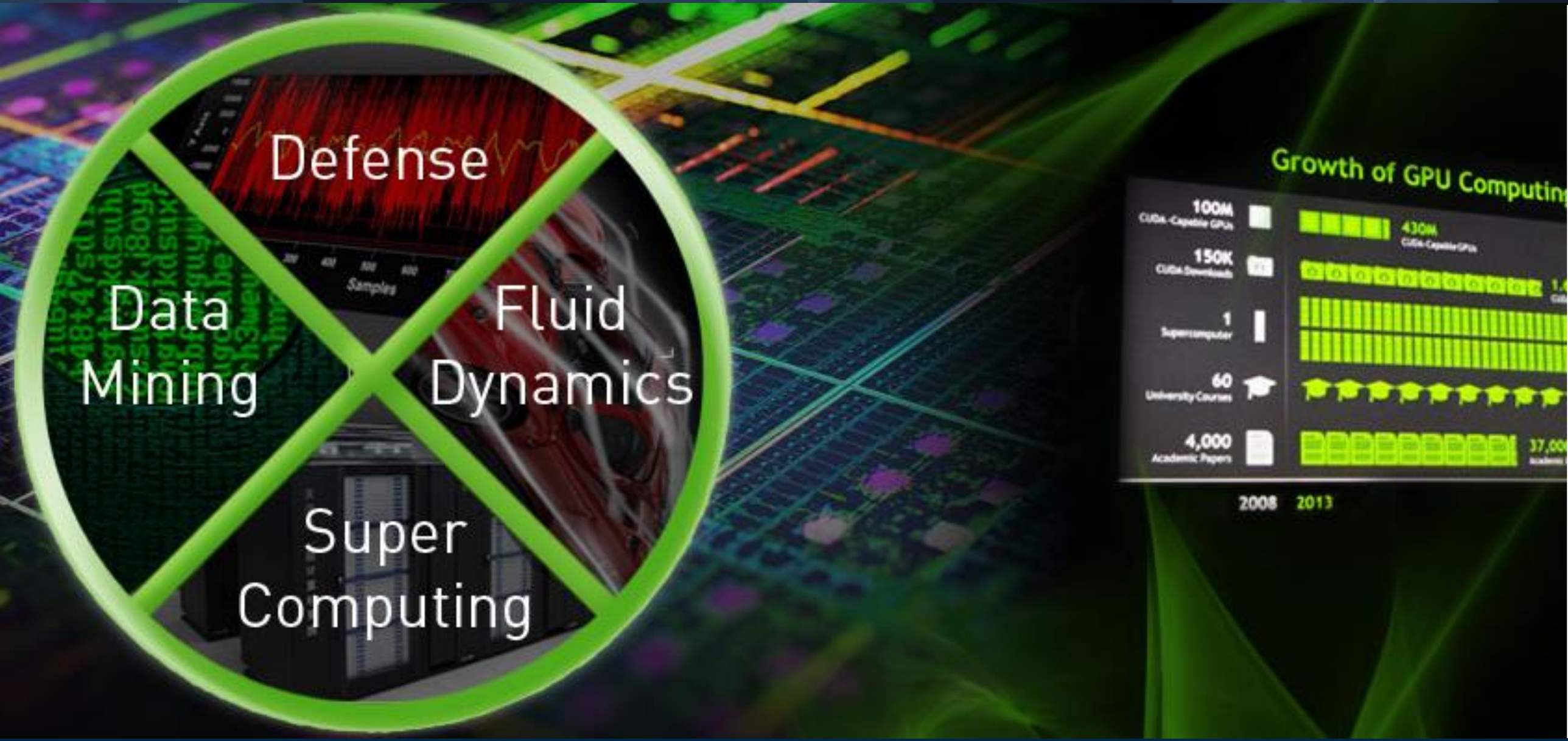
Hortonworks

MEMCACHED

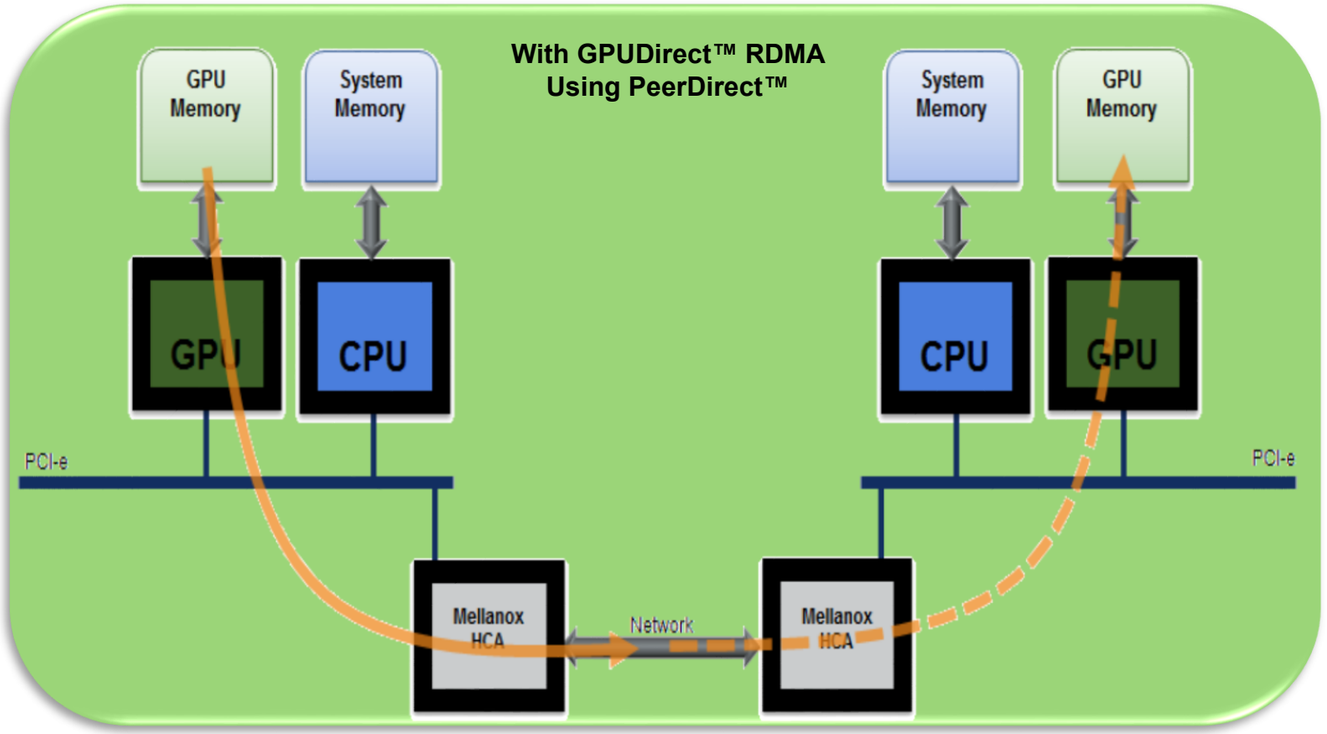
mongoDB

redis

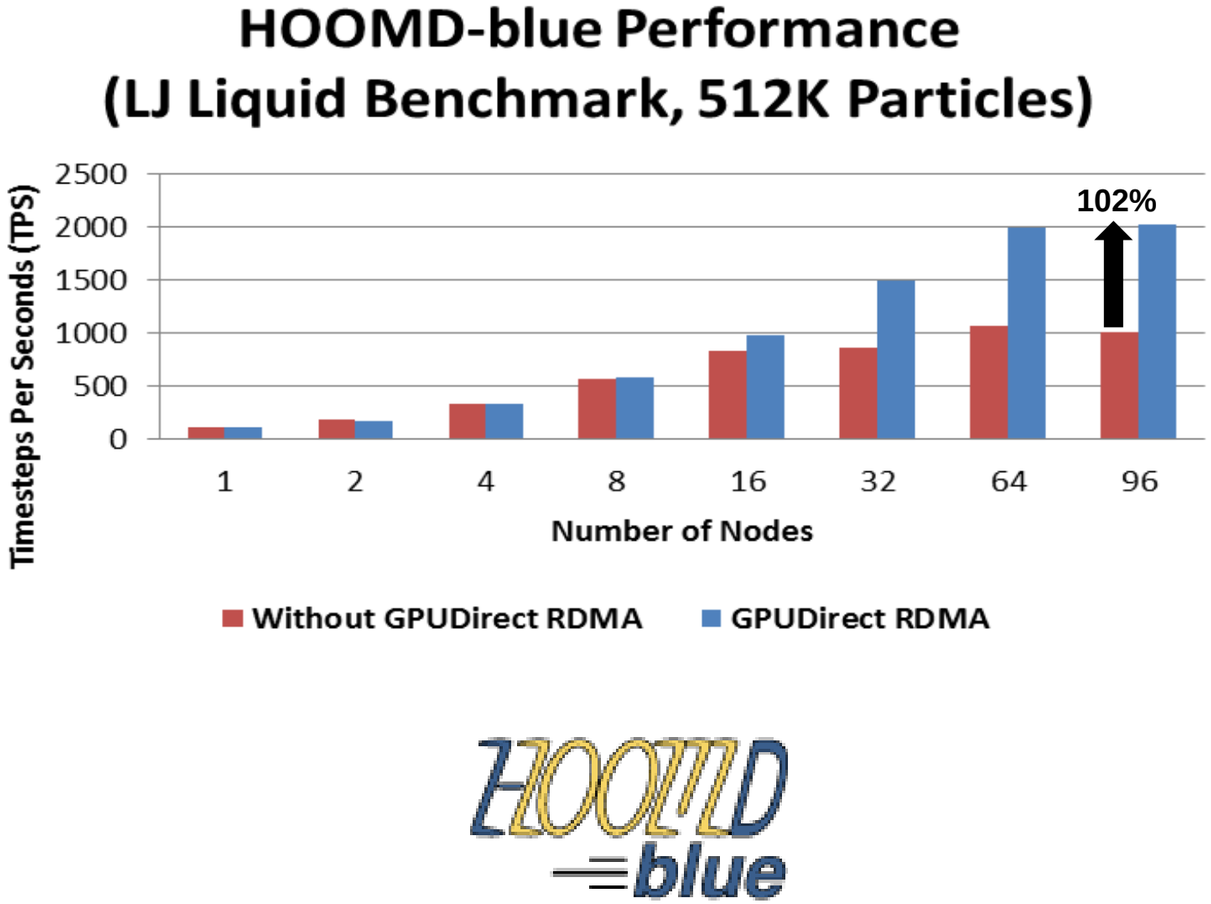
GPU Computing

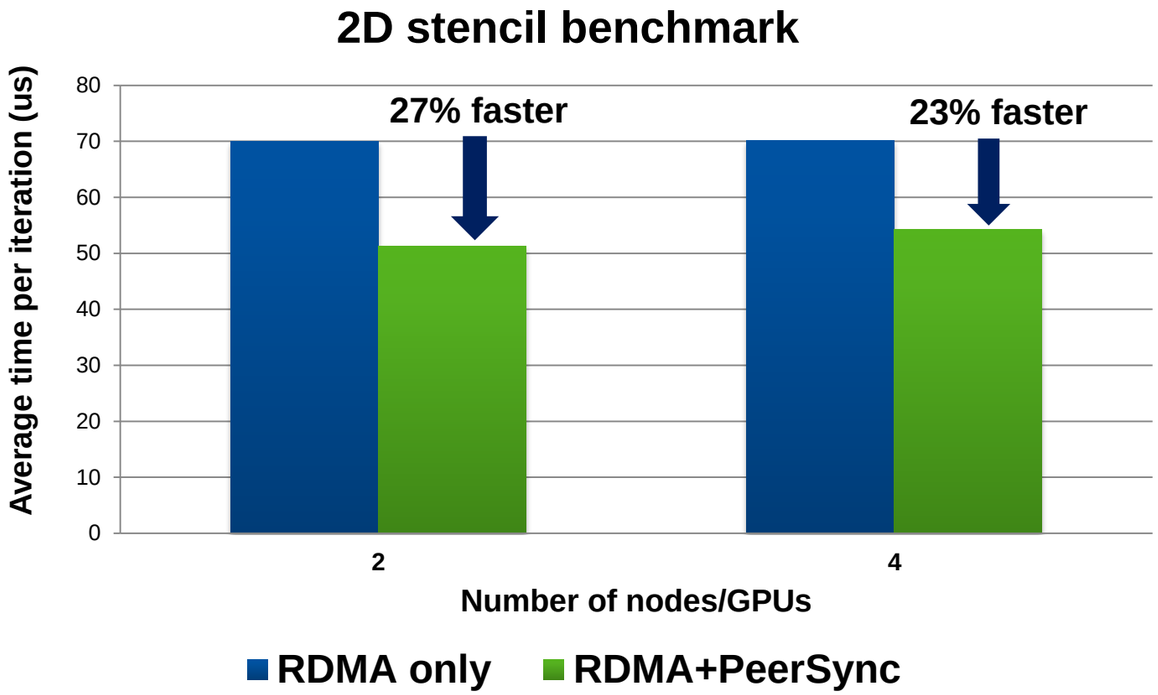


GPUDirect™ RDMA Ecosustem



Mellanox PeerDirect™ with NVIDIA GPUDirect RDMA

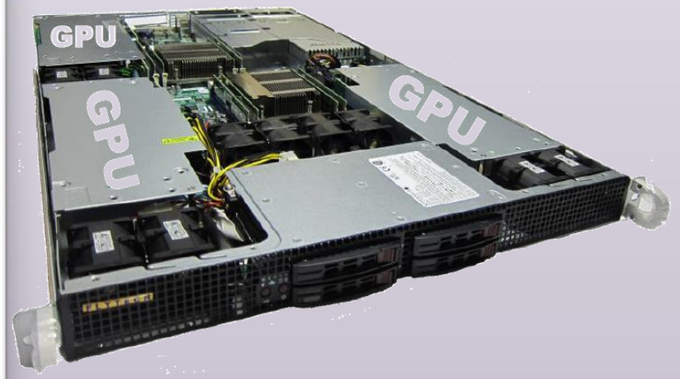


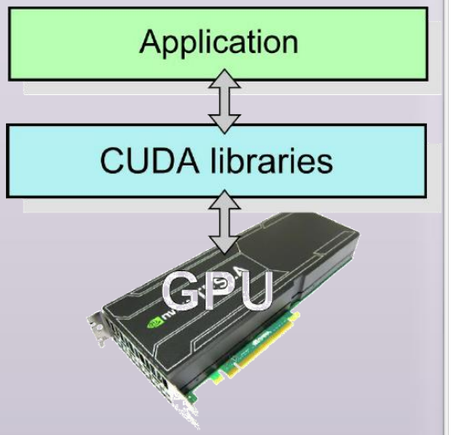


GPU – from Server to Service

Basic GPU computing

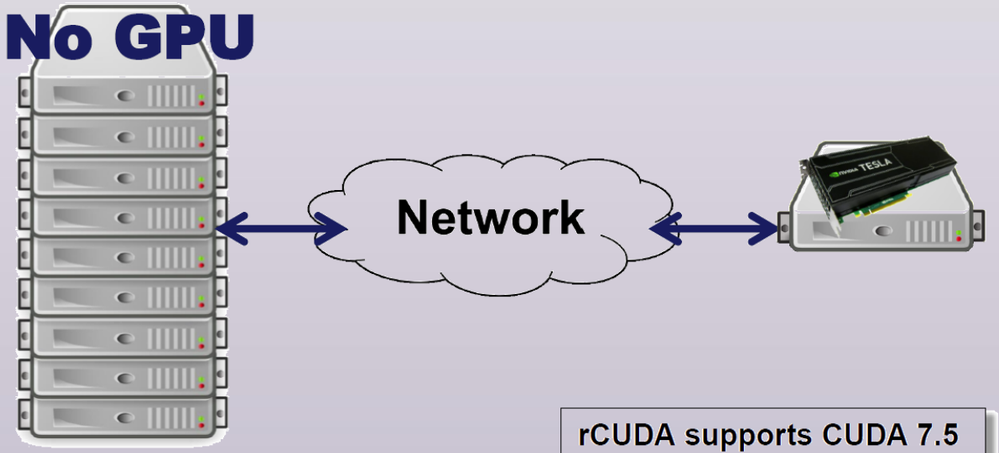
Basic behavior of CUDA





rCUDA – remote CUDA

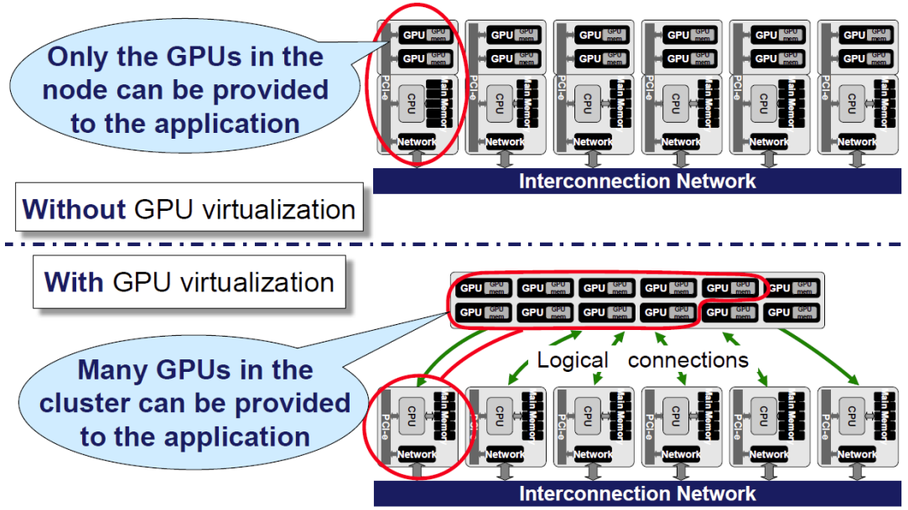
A **software** technology that enables **a more flexible use of GPUs** in computing facilities

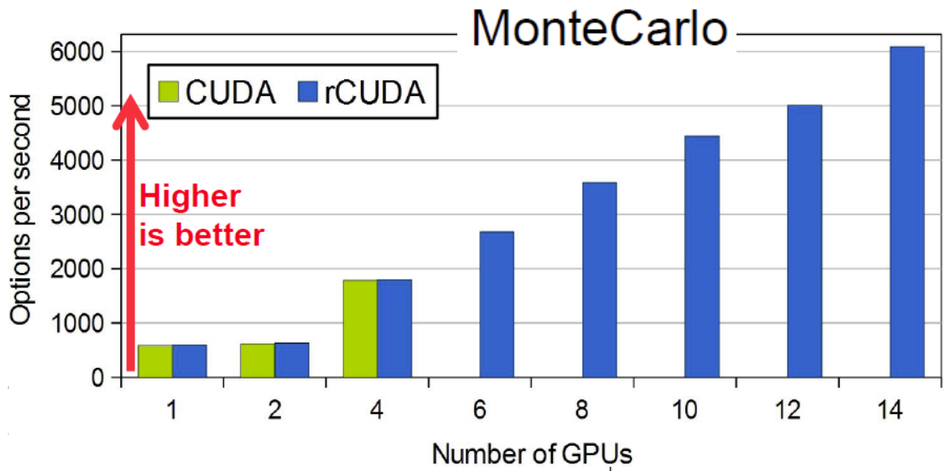


rCUDA supports CUDA 7.5

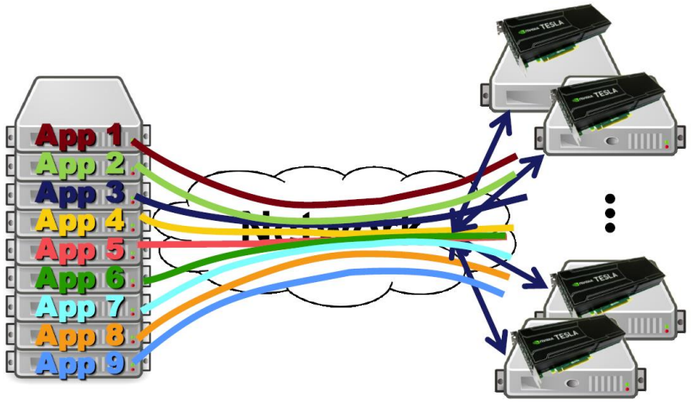
GPU Virtualization – Benefits

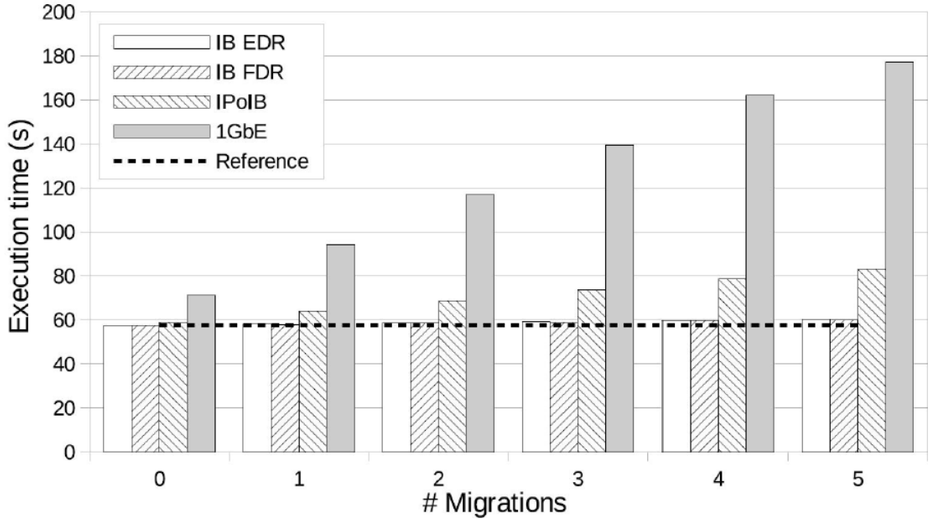
Flexible Resources Assignment





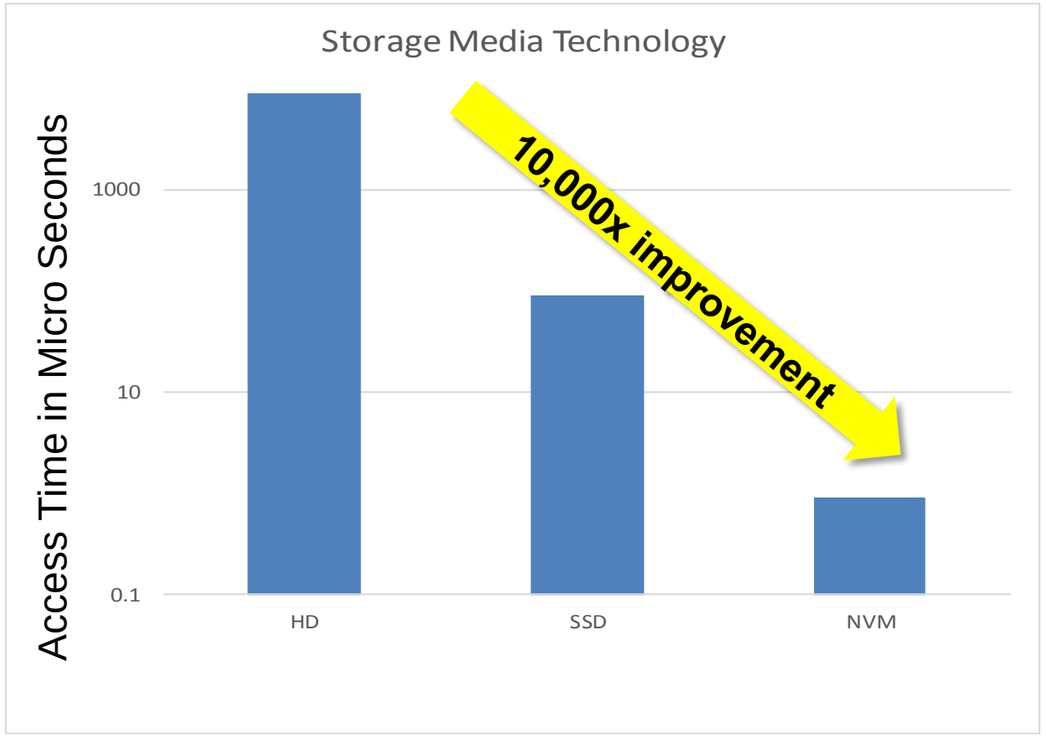
Resources' Consolidation and Management

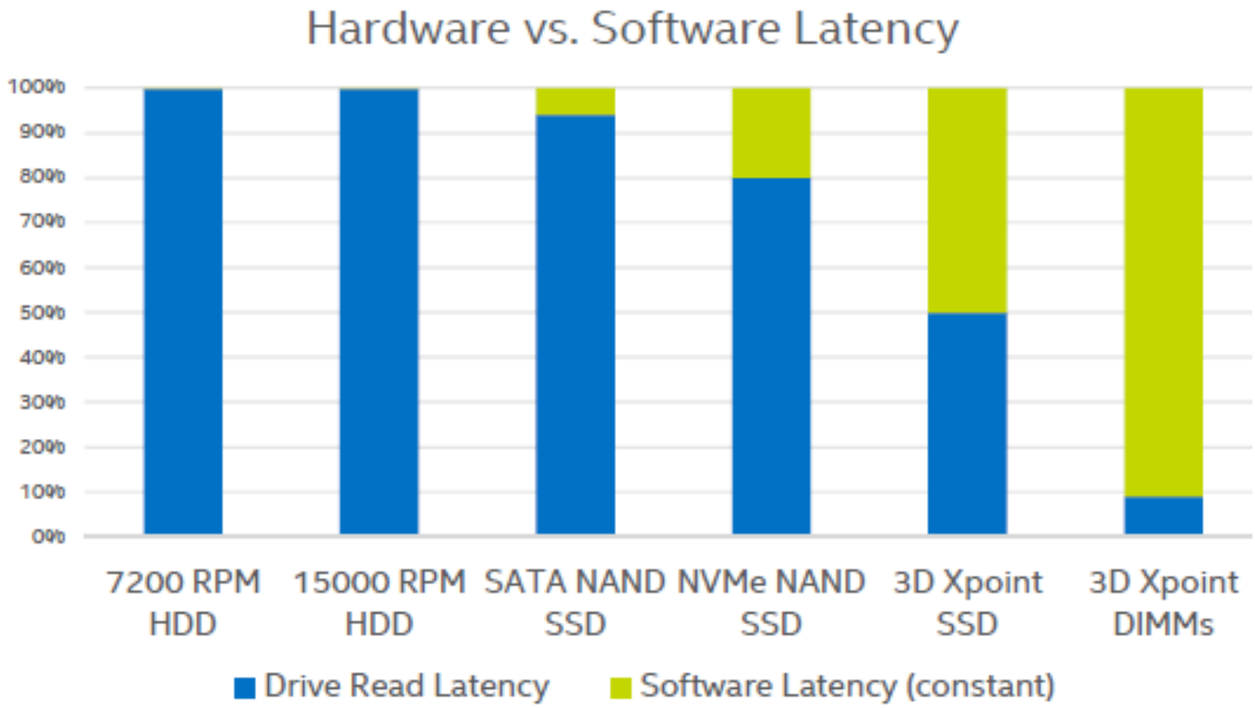




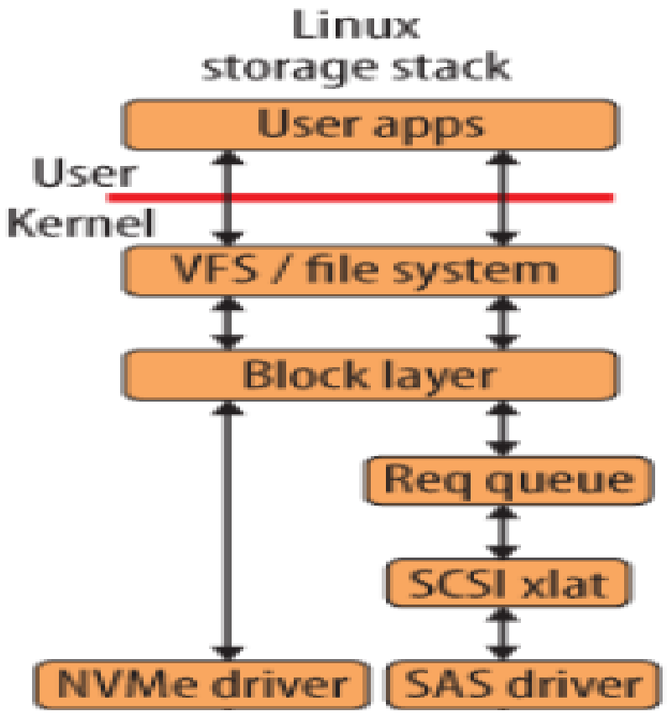
Storage and Data Access

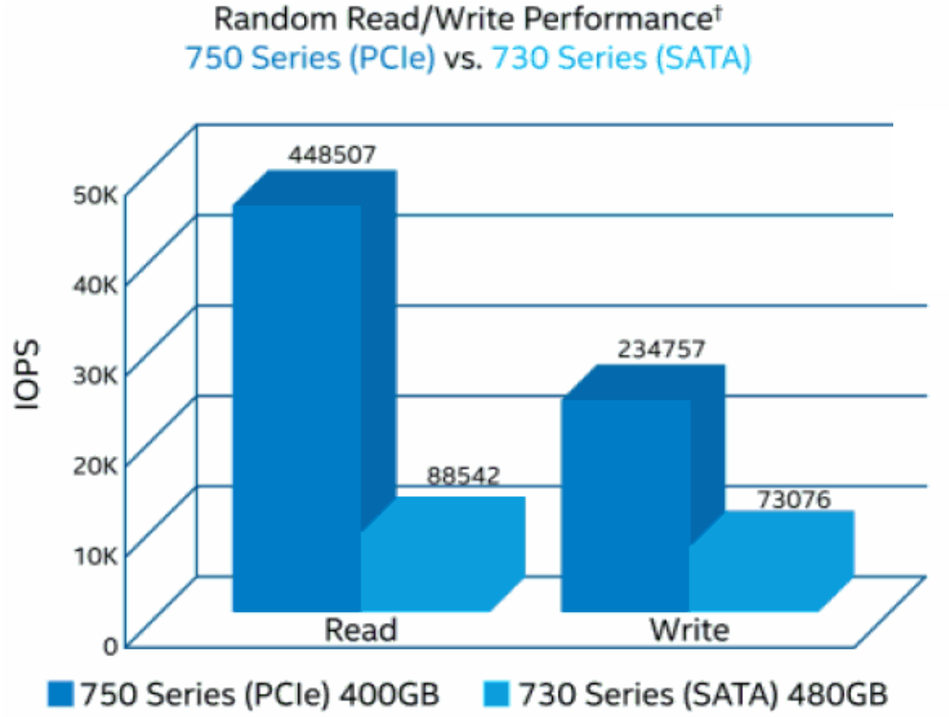
Storage Technology Evolution





NVMe Technology



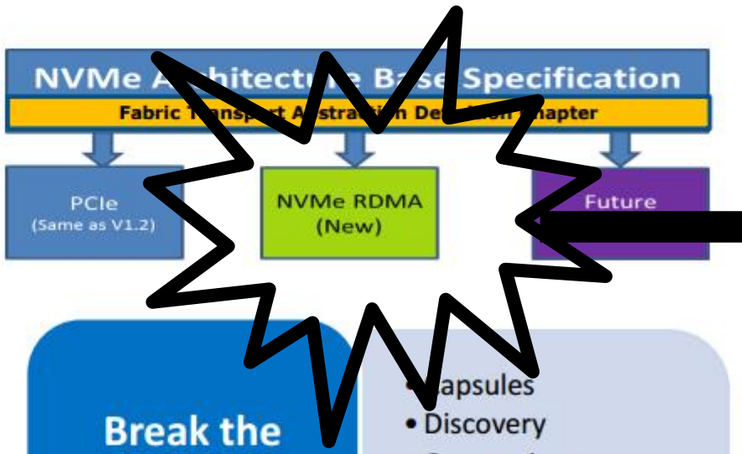


NVME over Fabrics – RDMA-based networking storage

Specification Strategy and Breakdown of Work

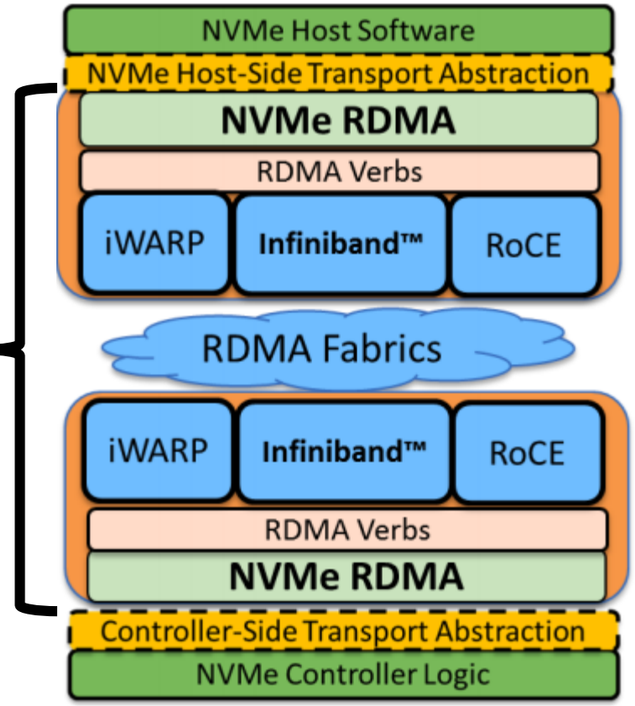
Do not create a standalone specification

- Initial goal is to minimize changes to existing specification
- Cleanly separate out the non-PCIe NVMe Transport layers through separate chapters/sections
  - Fabrics Core (concepts and RDMA binding)
  - Fabrics Base Differences (SGL changes, etc.)
- Long-term goal is to create a Transport agnostic base spec

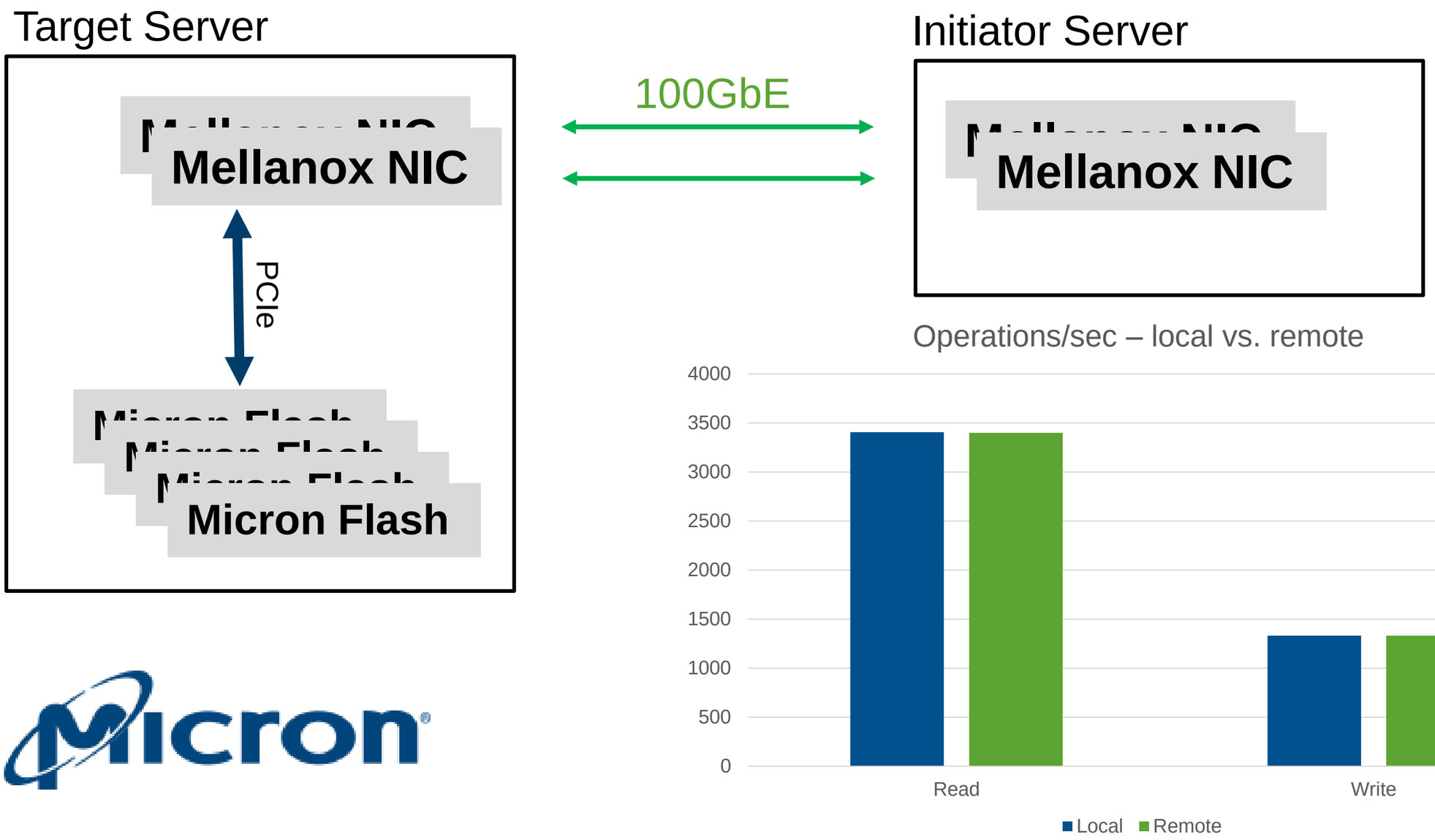


Break the work into functional sub-sections

- Capsules
- Discovery
- Connections
- Flow Control
- Naming
- Binding
- Error Handling



RDMA-based Remote NVME Access (NVME over Fabrics)



BlueField System-on-a-Chip (SoC) Solution

Storage

NVMe Flash Storage Arrays

Scale-Out Storage (NVMe over Fabric)



NFV

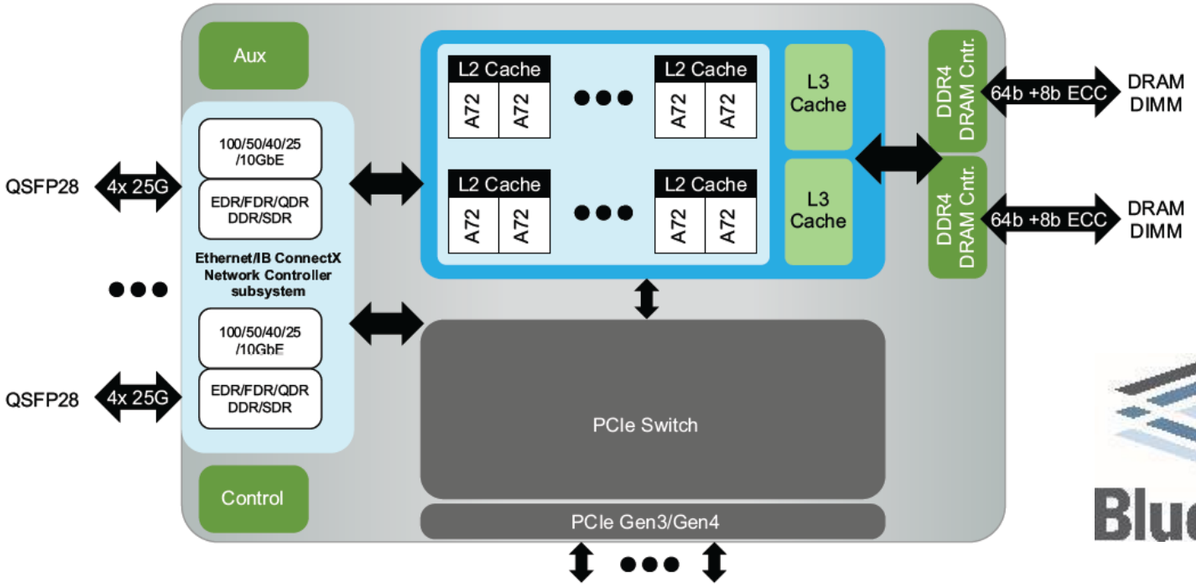
Accelerating & Virtualizing VNFs

Open vSwitch (OVS), SDN

Overlay networking offloads


- Integration of ConnectX5 + Multicore ARM
- State of the art capabilities
  - 10 / 25 / 40 / 50 / 100G Ethernet & InfiniBand
  - PCIe Gen3 / Gen4
  - Hardware acceleration offload
    - RDMA, RoCE, NVMeF, RAID
- Family of products
  - Range of ARM core counts and I/O ports/speeds
  - Price/Performance points

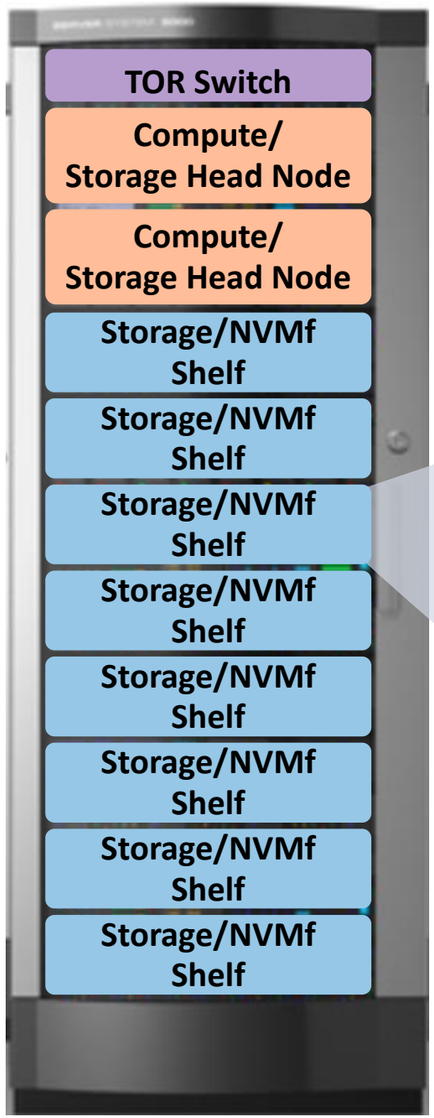
BlueField Architecture

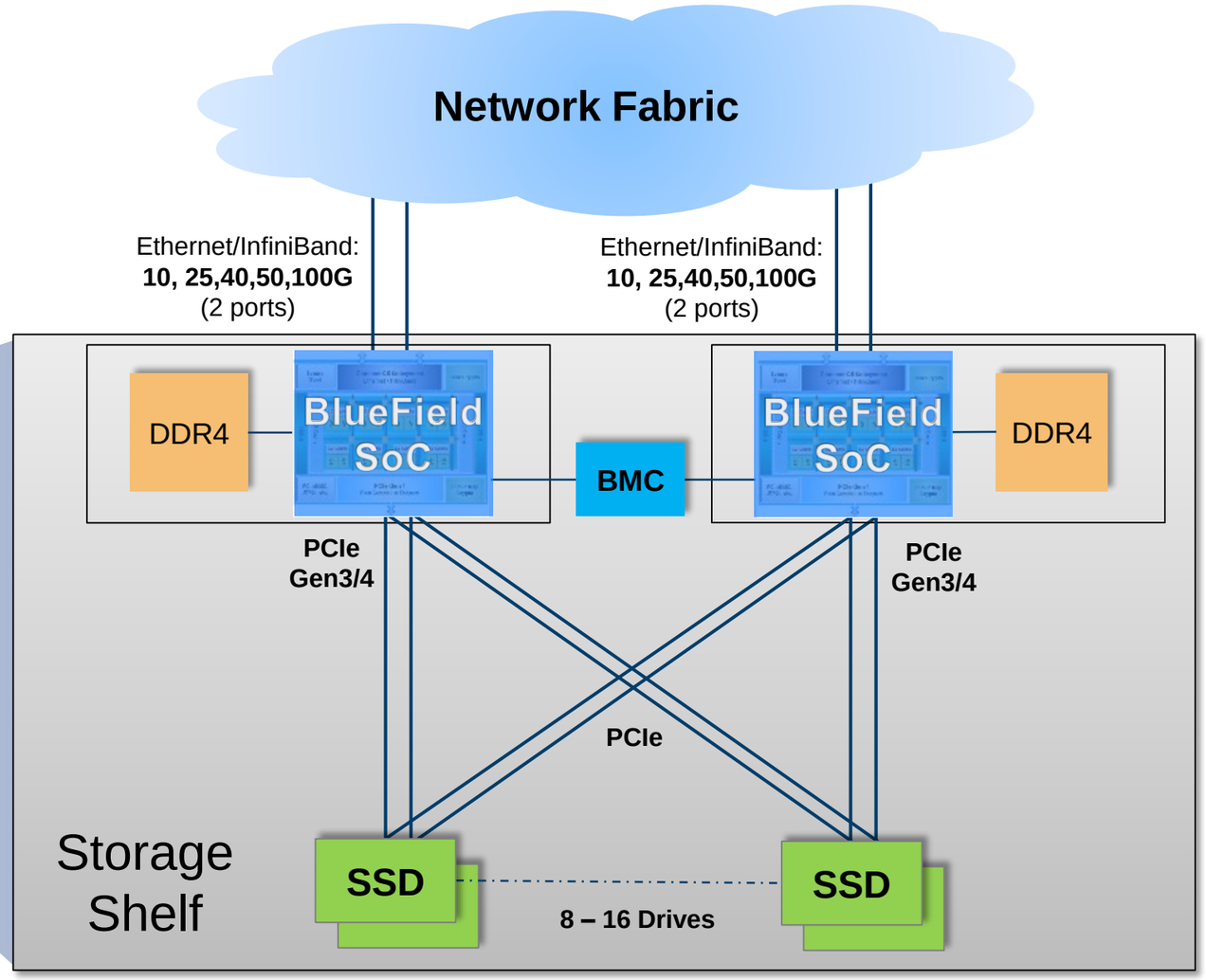


BlueField

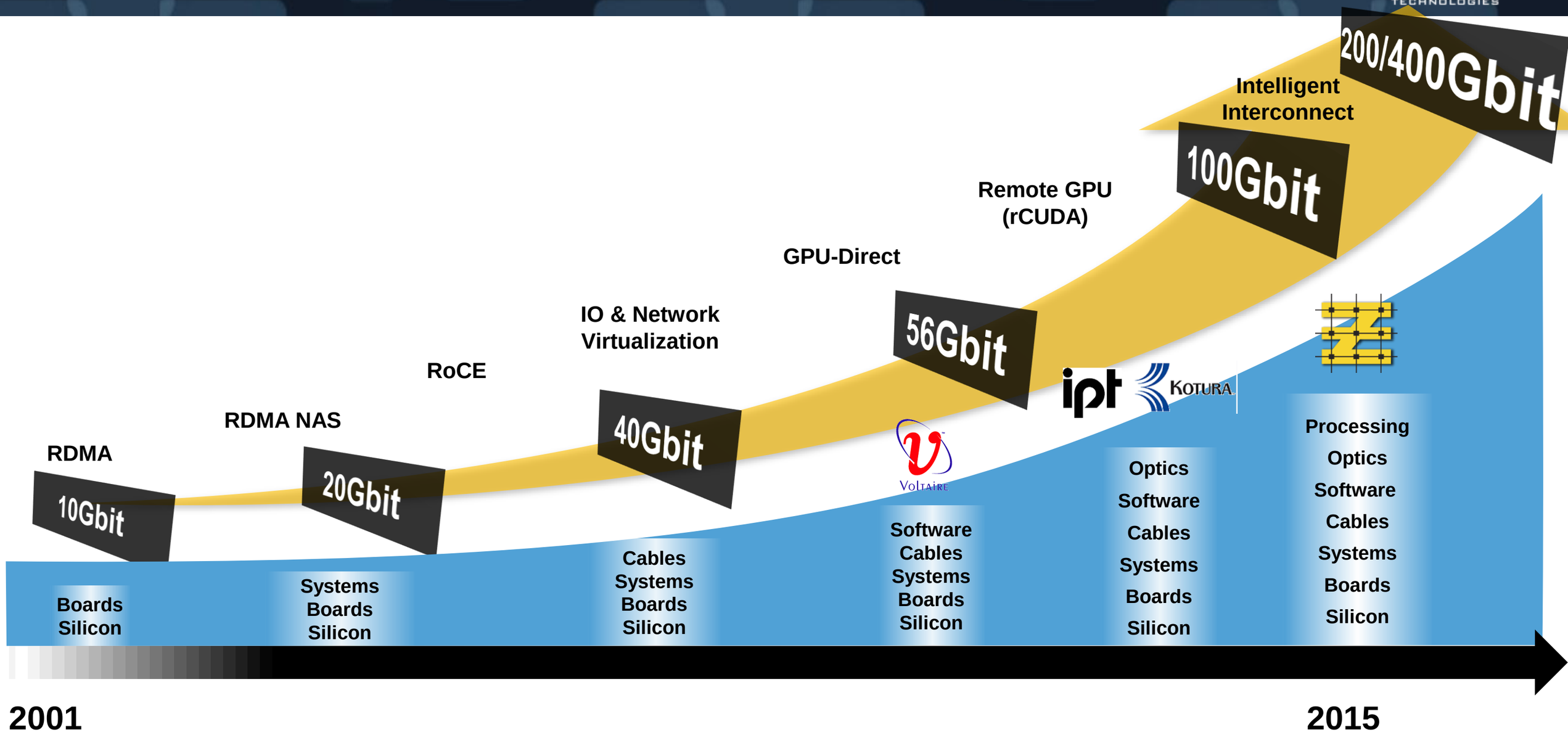
Scale-Out NVMe Storage System with BlueField

Rack view

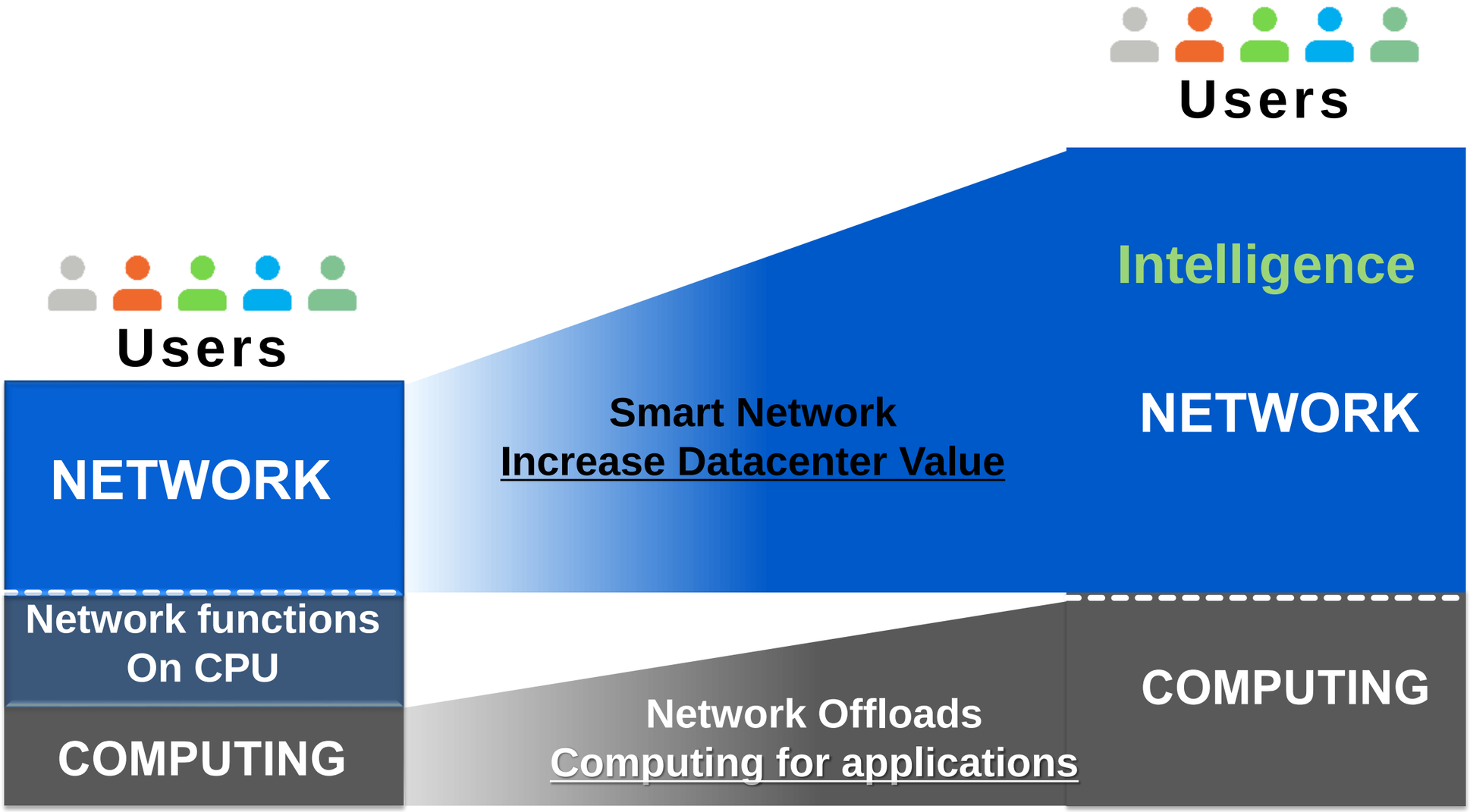




Technology Leadership



Strategy: Network-based Computing



Computing Evolution

Bare Metal Programming

Scale-out Services

Thank You

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™