In-Network Computing

Paving the Road to Exascale

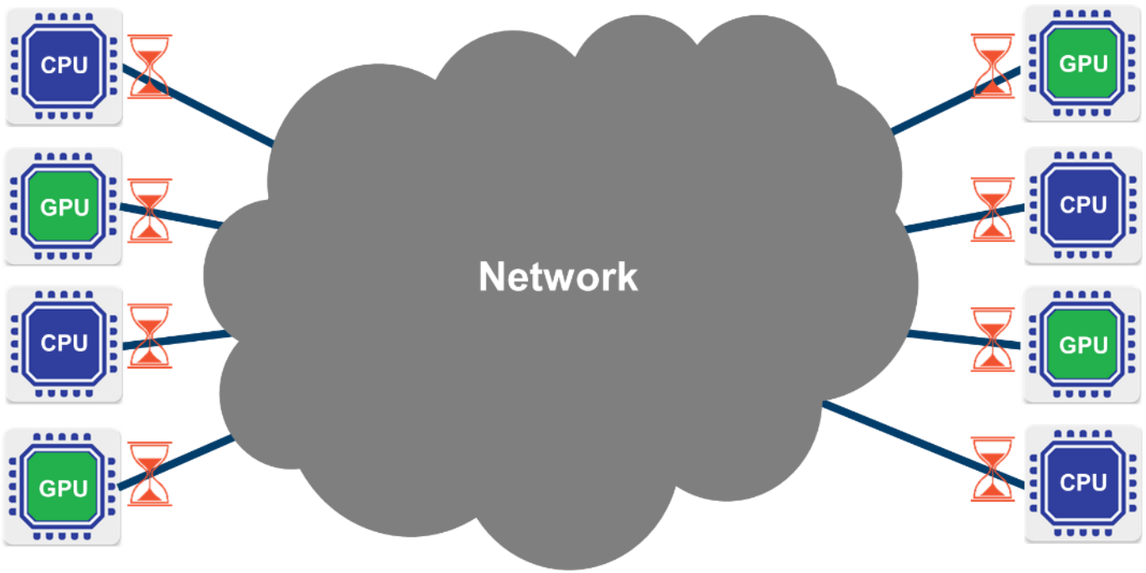
5th Annual MVAPICH User Group (MUG) Meeting, August 2017

Mellanox TECHNOLOGIES

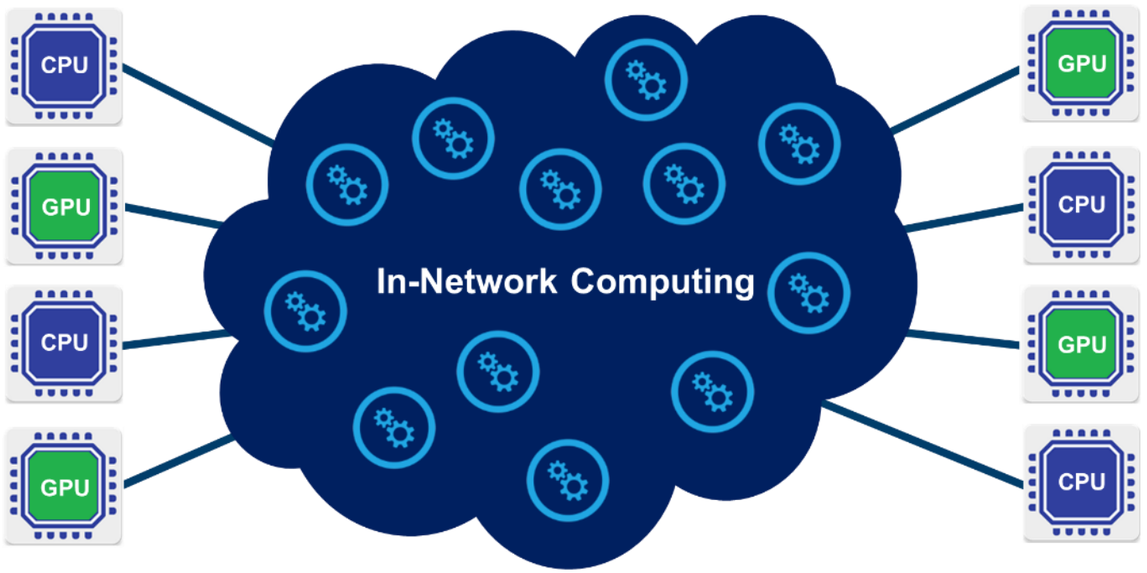
Connect. Accelerate. Outperform.™

Exponential Data Growth – The Need for Intelligent and Faster Interconnect

CPU-Centric (Onload)



Data-Centric (Offload)



Must Wait for the Data

Creates Performance Bottlenecks

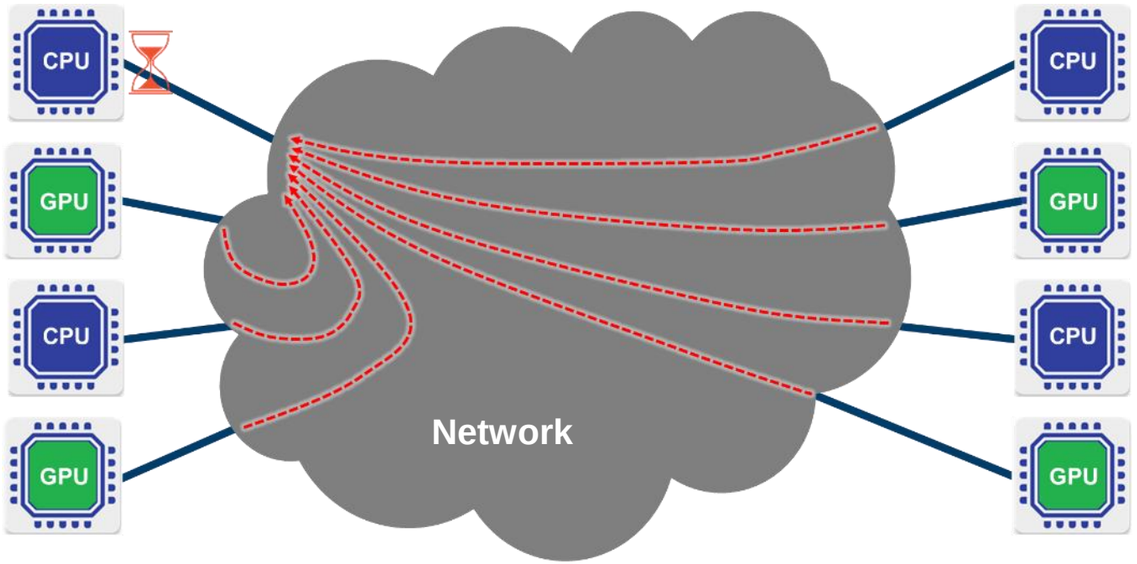
Analyze Data as it Moves!

Faster Data Speeds and In-Network Computing Enable Higher Performance and Scale

© 2017 Mellanox Technologies

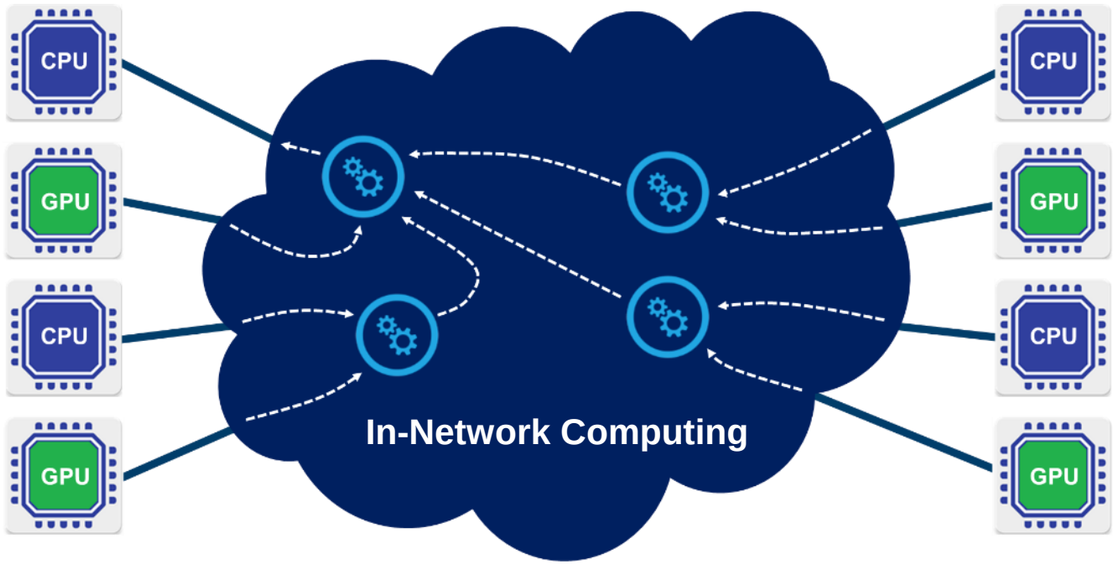
Data Centric Architecture to Overcome Latency Bottlenecks

CPU-Centric (Onload)



HPC / Machine Learning
Communications Latencies of 30-40us

Data-Centric (Offload)

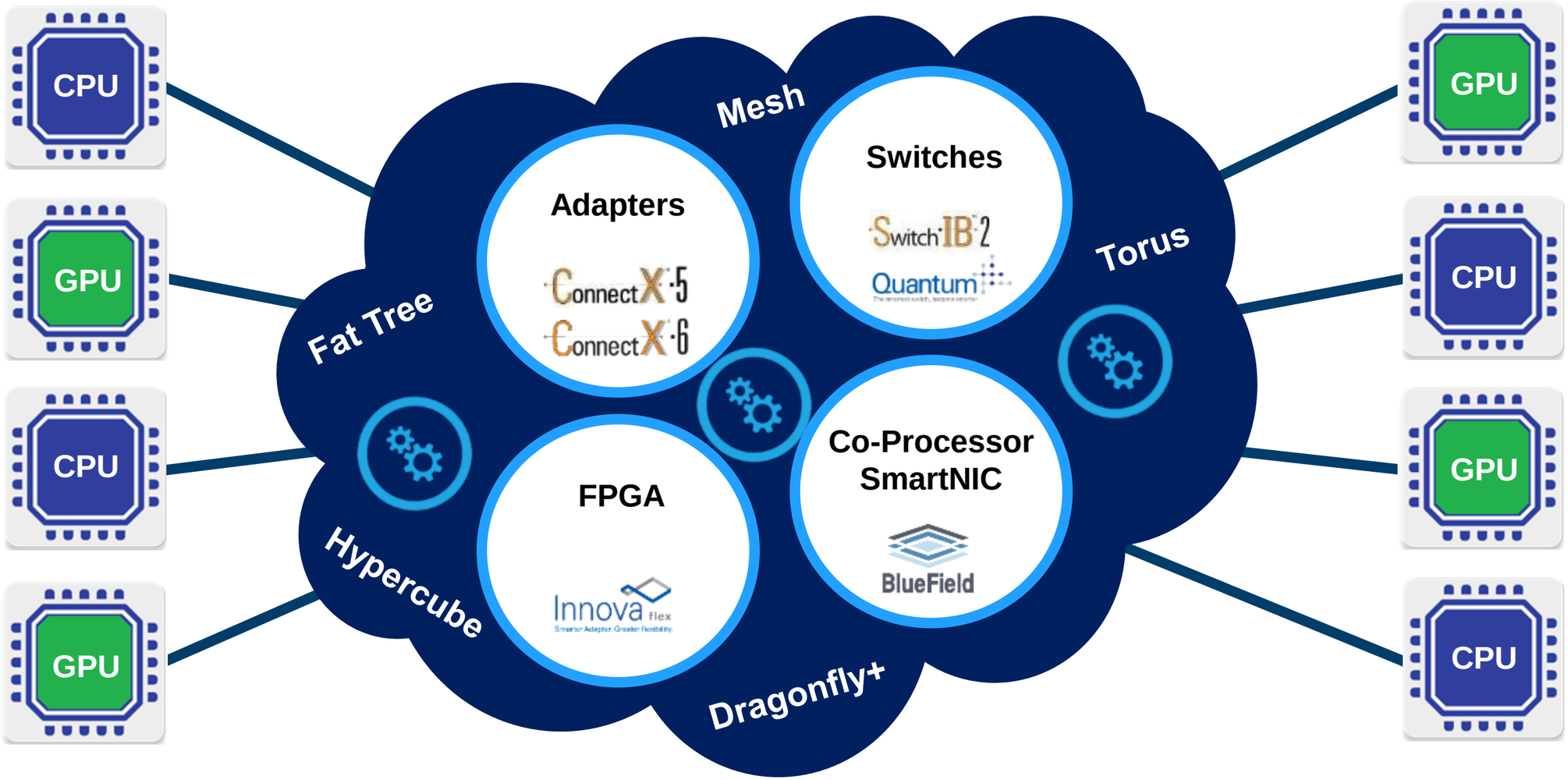


HPC / Machine Learning
Communications Latencies of 3-4us

Intelligent Interconnect Paves the Road to Exascale Performance

© 2017 Mellanox Technologies

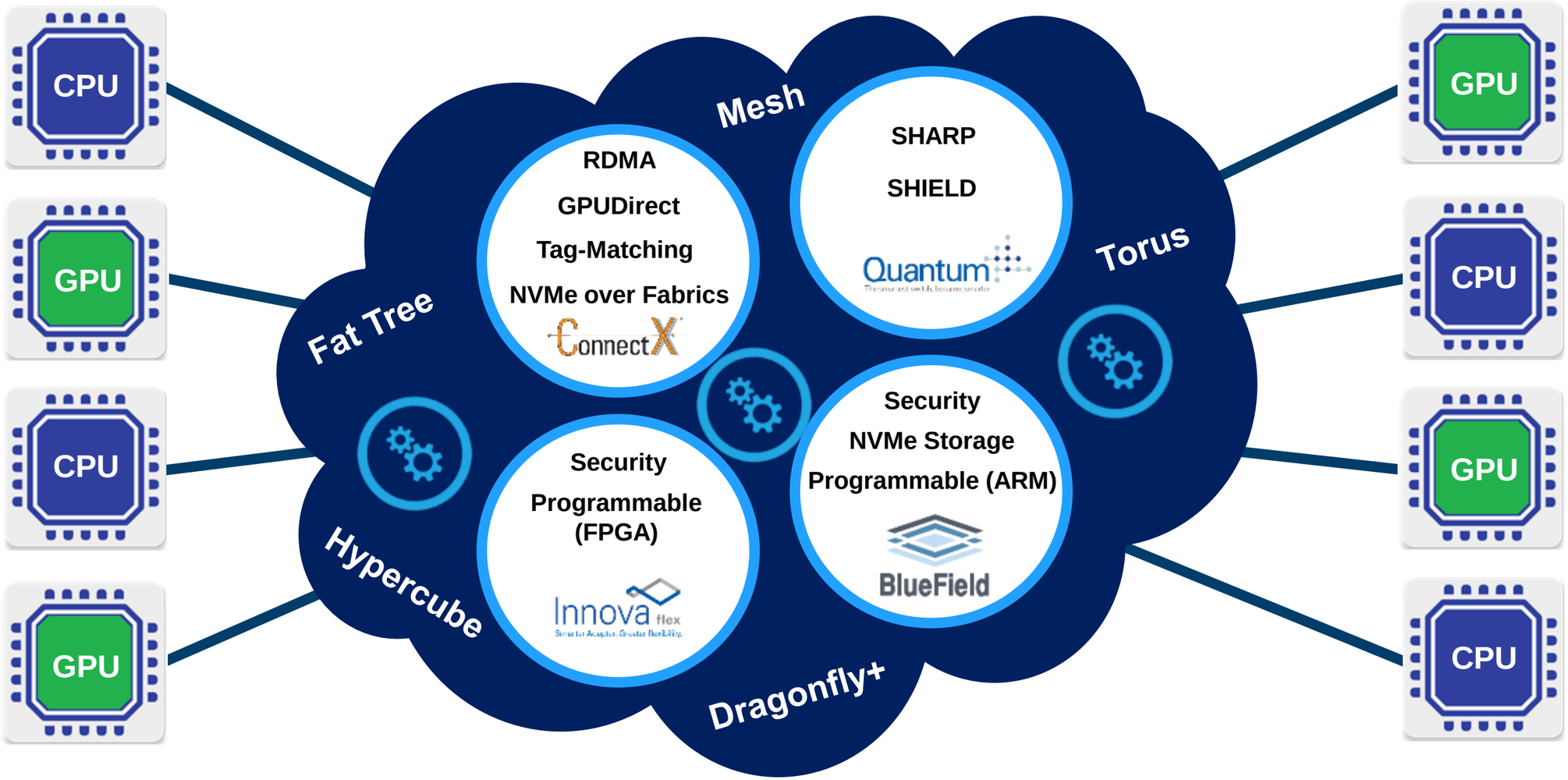
In-Network Computing to Enable Data-Centric Data Centers



In-Network Computing Key for Highest Return on Investment

© 2017 Mellanox Technologies

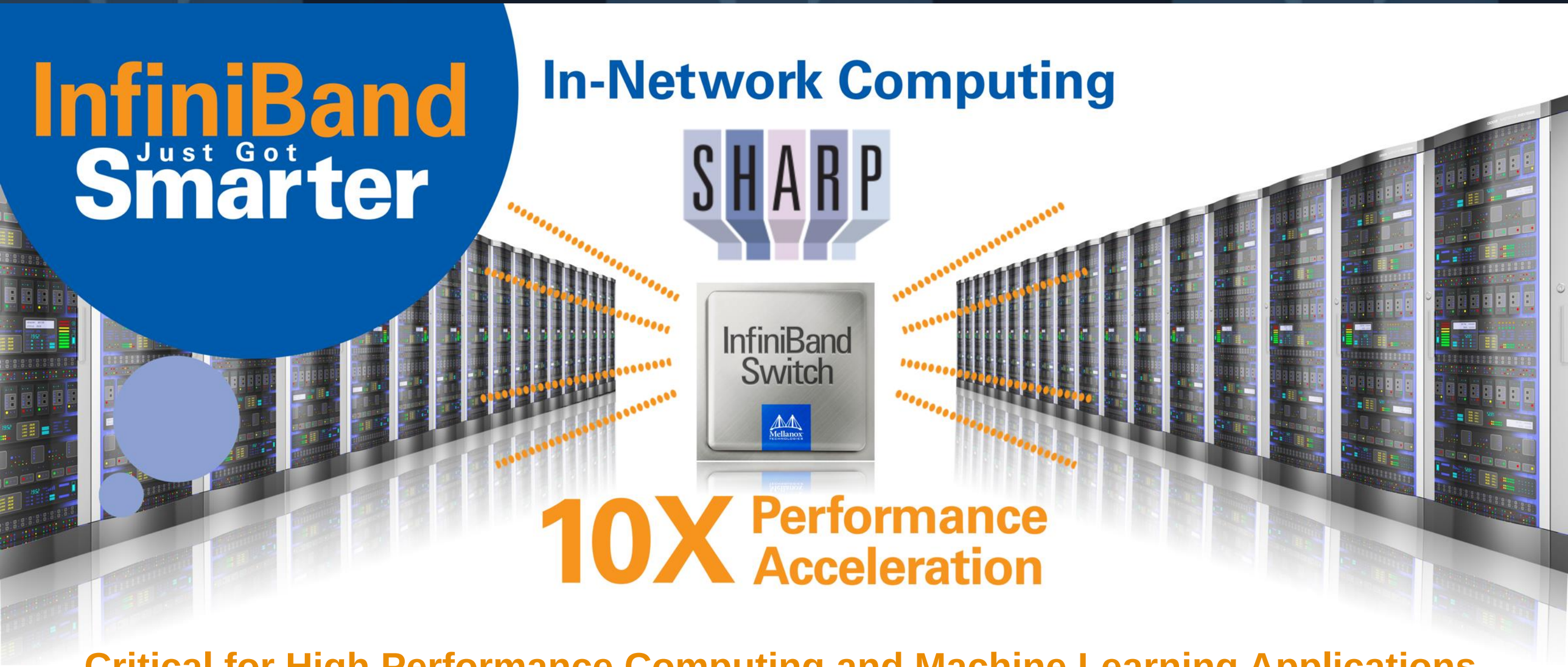
In-Network Computing to Enable Data-Centric Data Centers



In-Network Computing Key for Highest Return on Investment

© 2017 Mellanox Technologies

In-Network Computing Advantages with SHARP Technology



Critical for High Performance Computing and Machine Learning Applications

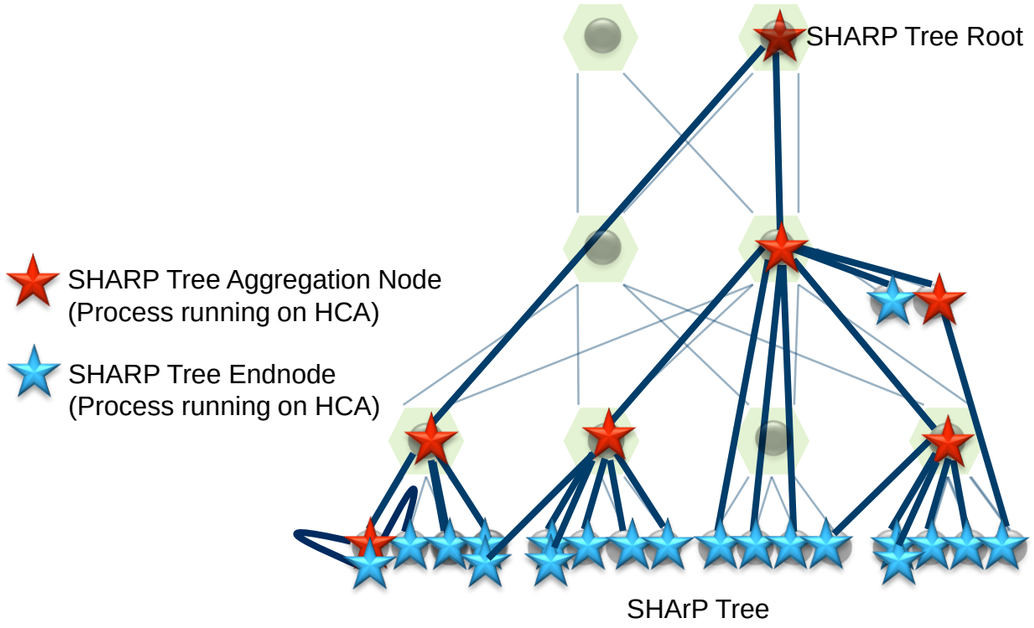
© 2017 Mellanox Technologies

Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)

- Reliable Scalable General Purpose Primitive
  - In-network Tree based aggregation mechanism
  - Large number of groups
  - Multiple simultaneous outstanding operations
- Applicable to Multiple Use-cases
  - HPC Applications using MPI / SHMEM
  - 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, 16/32/64/128 bits

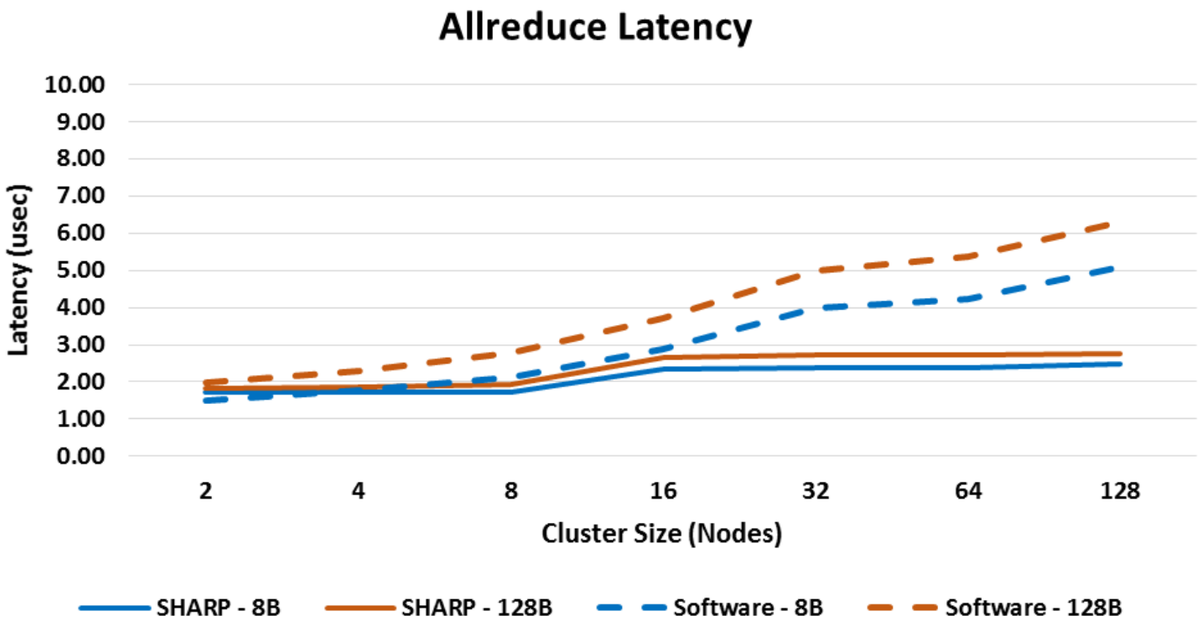
SHARP

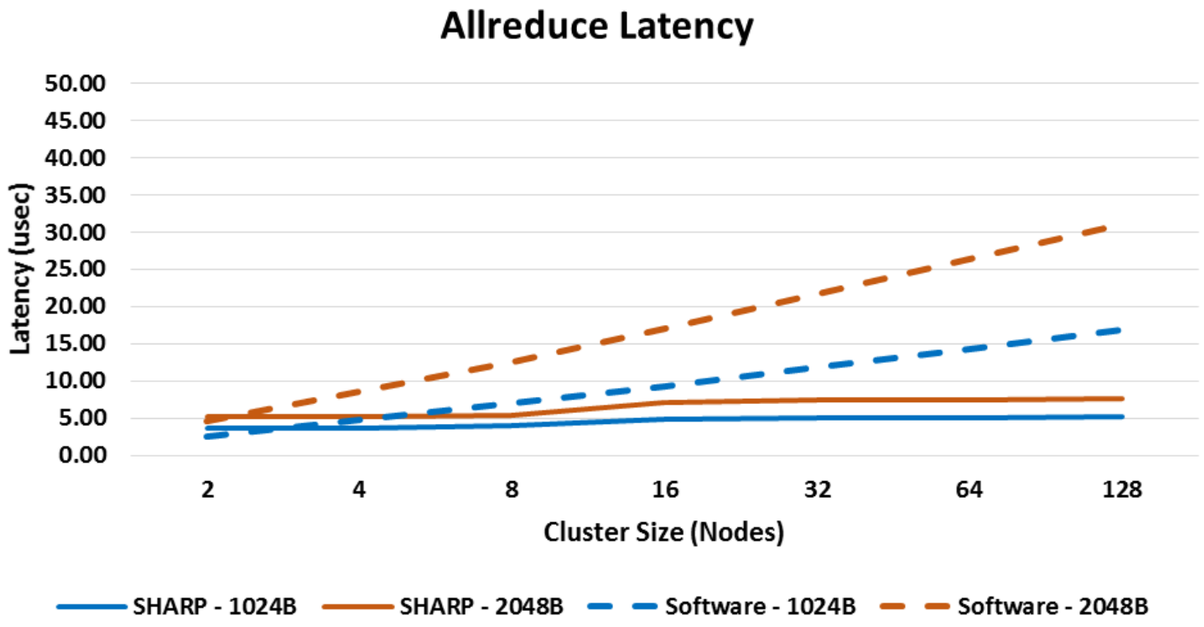
Scalable Hierarchical Aggregation and Reduction Protocol



© 2017 Mellanox Technologies

SHARP Allreduce Performance Advantages





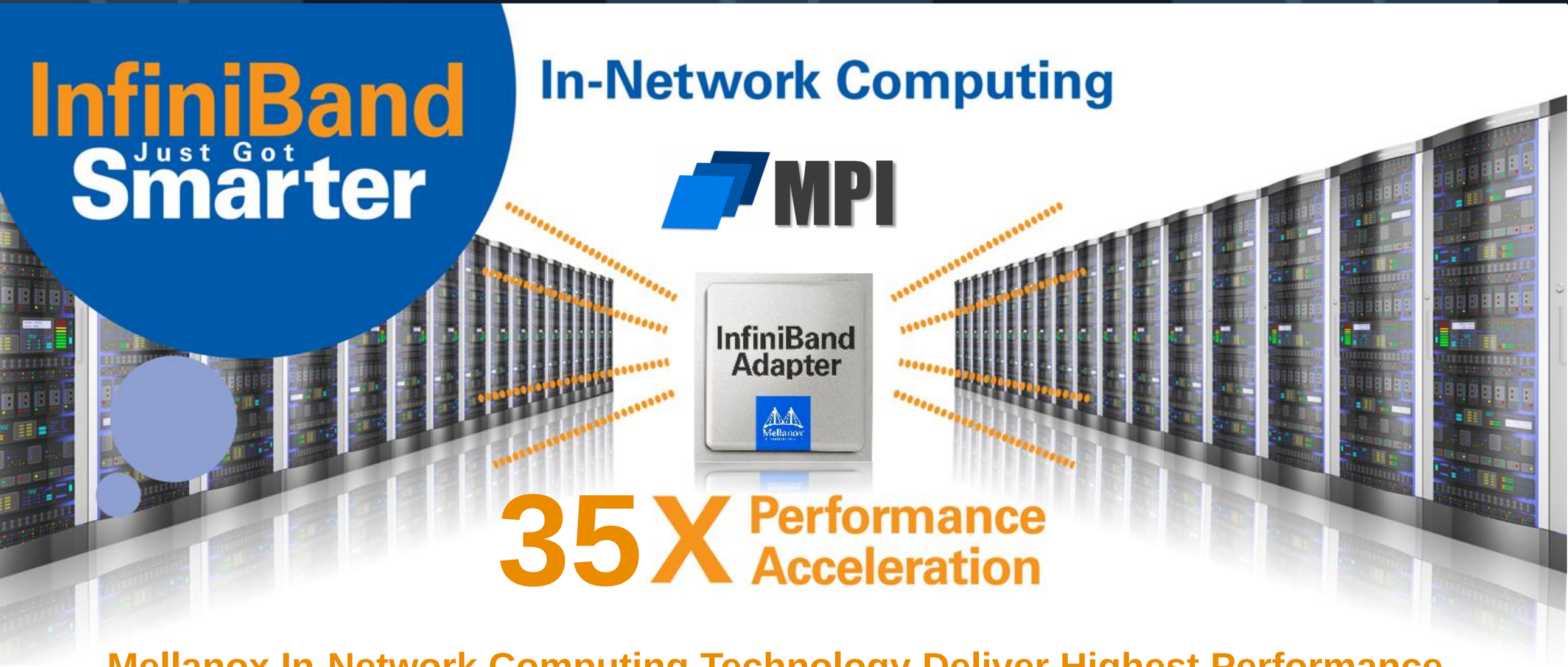
SHARP enables 75% Reduction in Latency
Providing Scalable Flat Latency

© 2017 Mellanox Technologies

SHIELD – Self Healing Interconnect Technology

© 2017 Mellanox Technologies

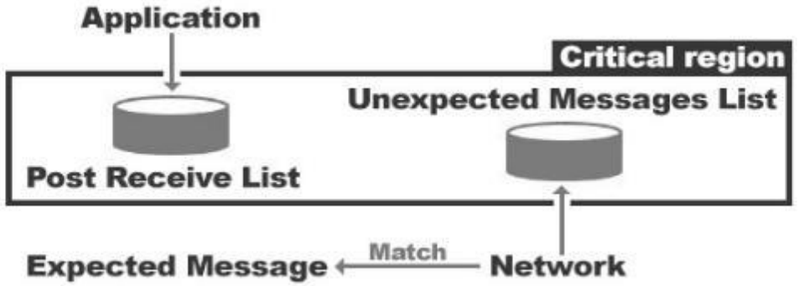
MPI Tag-Matching Hardware Engines

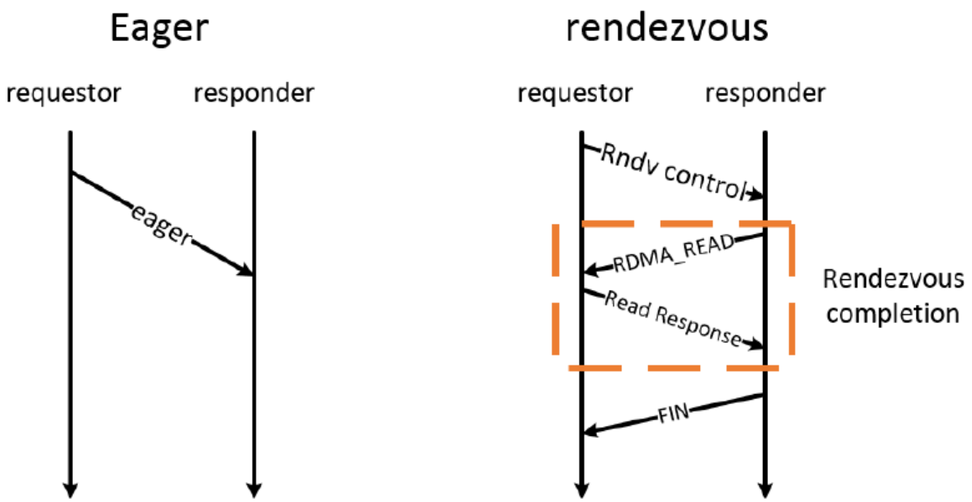


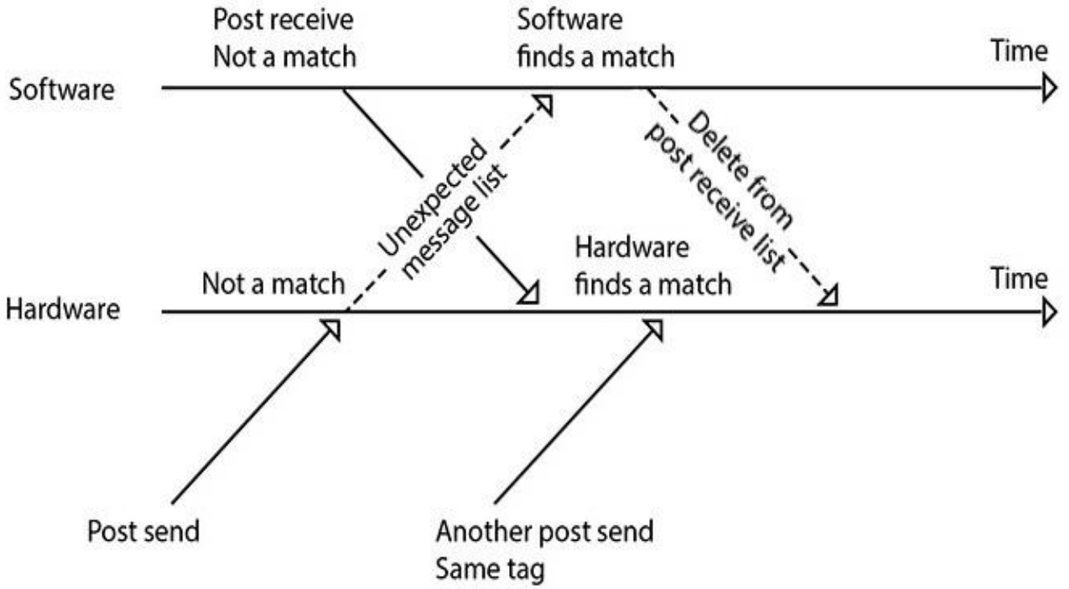
Mellanox In-Network Computing Technology Deliver Highest Performance

© 2017 Mellanox Technologies

MPI Tag-Matching Hardware Engines

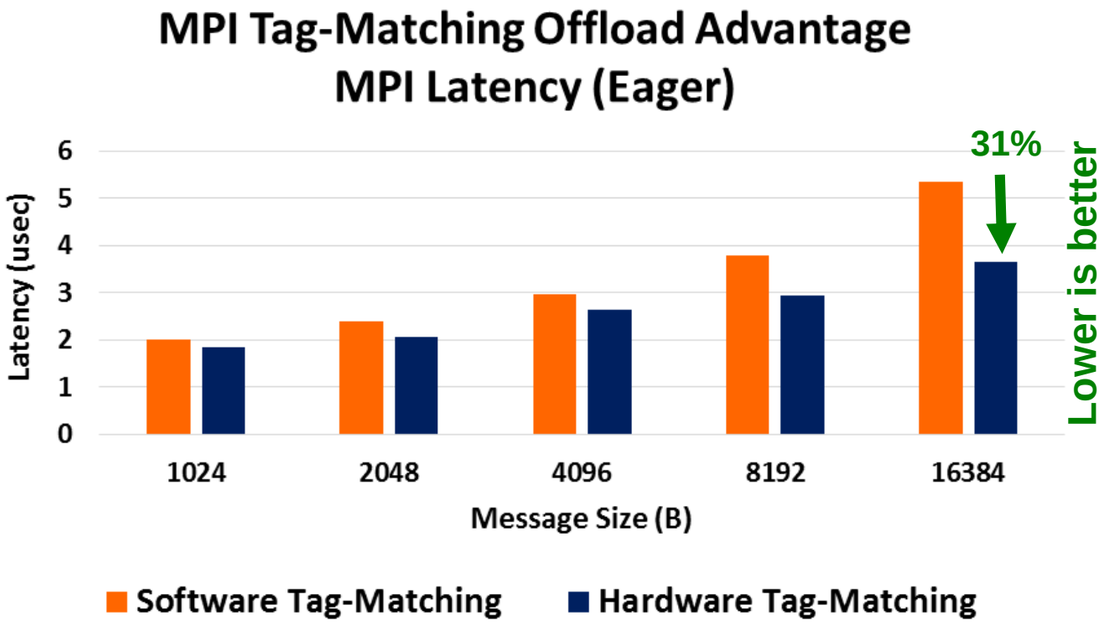


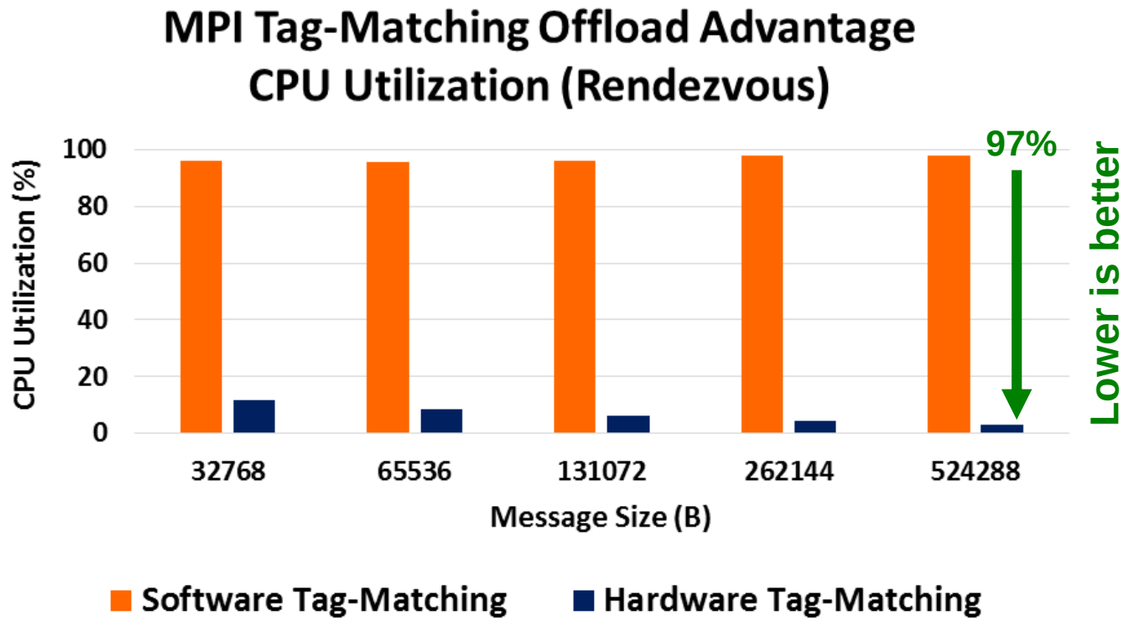




© 2017 Mellanox Technologies

MPI Tag-Matching Offload Advantages





- 31% lower latency and 97% lower CPU utilization for MPI operations
- Performance comparisons based on ConnectX-5

Mellanox In-Network Computing Technology Deliver Highest Performance

© 2017 Mellanox Technologies

Multi-Host Socket Direct Technology

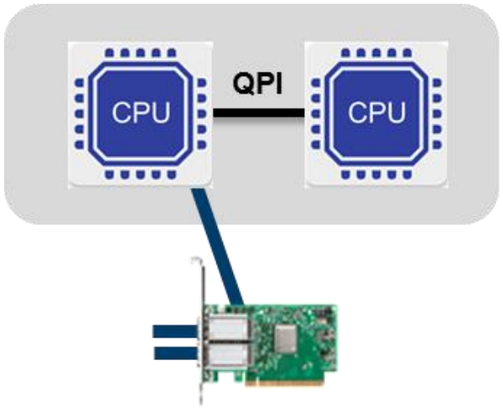
Innovative Solution to Dual-Socket Servers

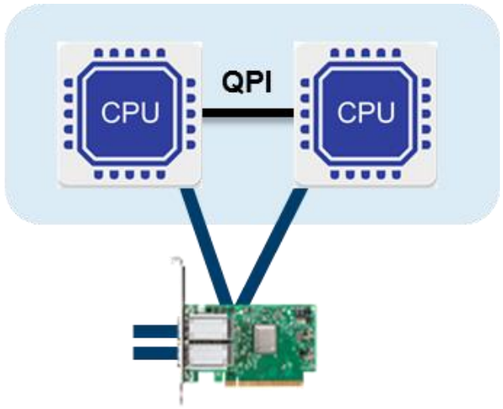
Multi-Host Socket-Direct Adapters Increase Server Return on Investment

30%-60% Better CPU Utilization

50%-80% Lower Data Latency

15%- 28% Better Data Throughout





ConnectX®-5

Available for All Servers (x86, Power, ARM, etc.)

Highest Applications Performance, Scalability and Productivity

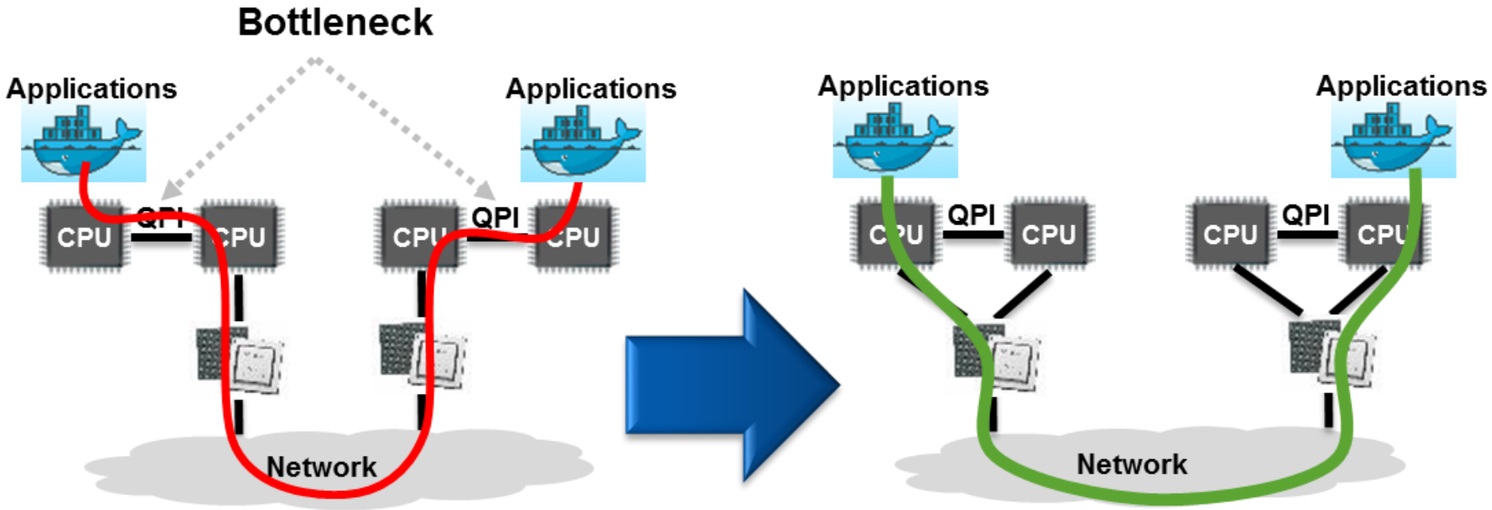
© 2017 Mellanox Technologies

Multi-Host Socket-Direct Adapters Increase Server Return on Investment

30%-60% Better CPU Utilization

50%-80% Lower Data Latency

15%- 28% Better Data Throughout



ConnectX®-5

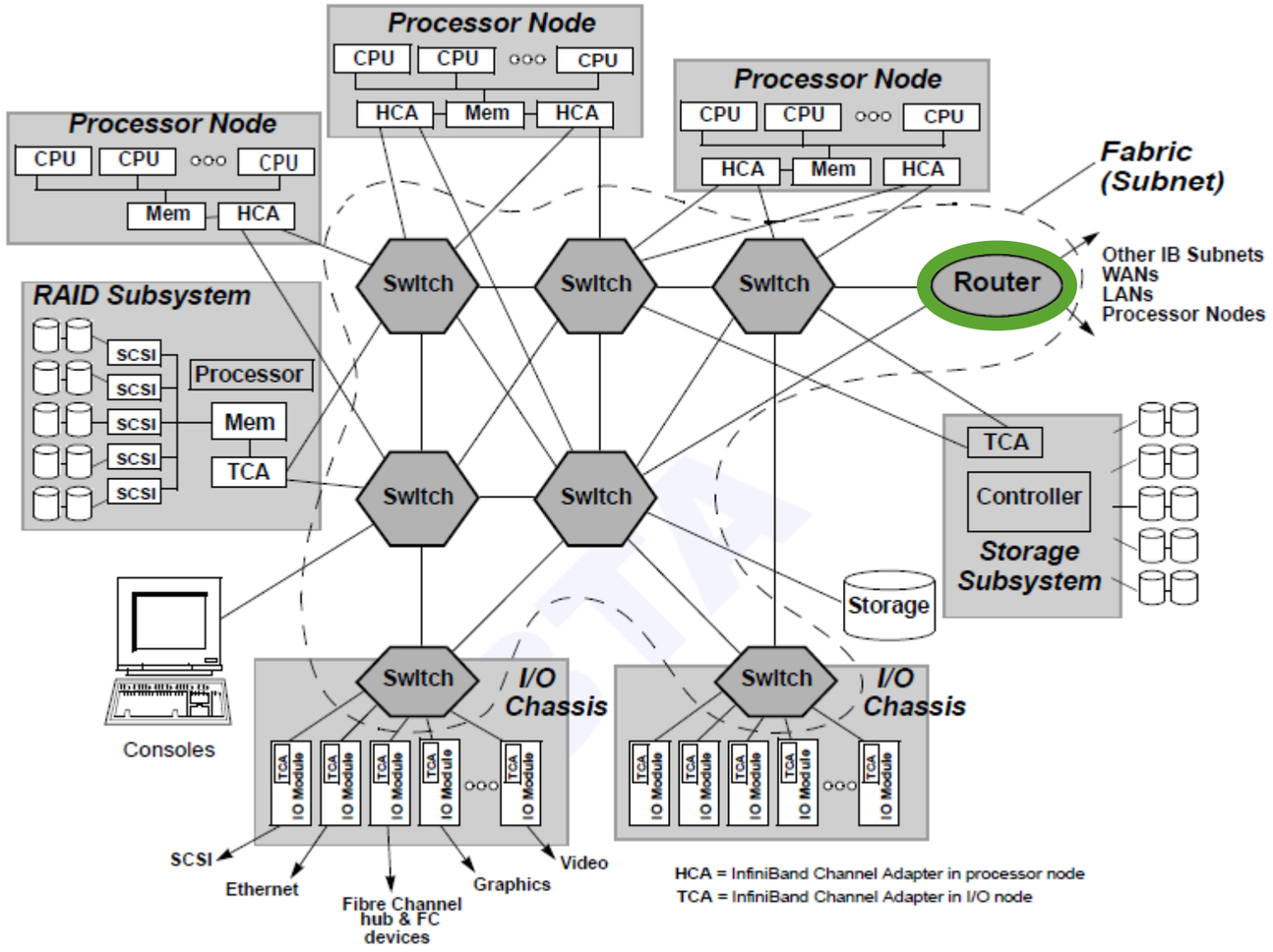
Available for All Servers (x86, Power, ARM, etc.)

Highest Applications Performance, Scalability and Productivity

© 2017 Mellanox Technologies

InfiniBand Router

InfiniBand Architecture



Native InfiniBand Connectivity Between Different InfiniBand Subnets (Each Subnet can Include 40K nodes)

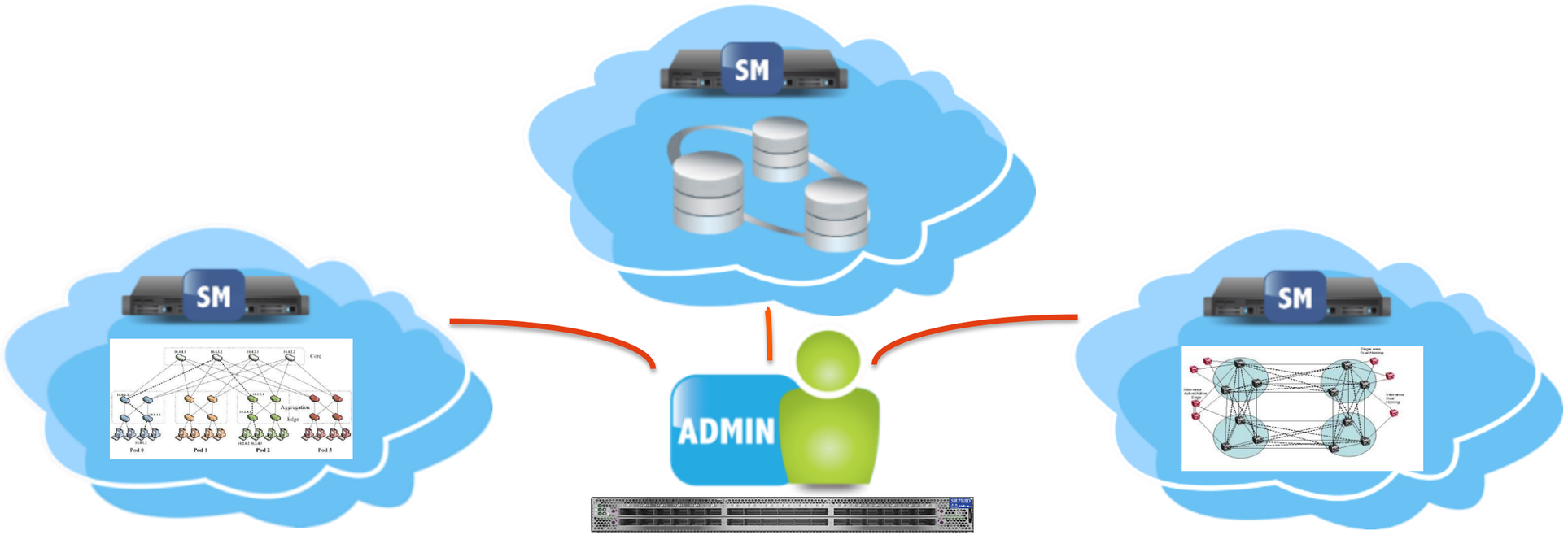
Isolation Between Different InfiniBand Networks (Each Network can be Managed Separately)

Native InfiniBand Connectivity Between Different Network Topologies (Fat-Tree, Torus, Dragonfly, etc.)

© 2017 Mellanox Technologies

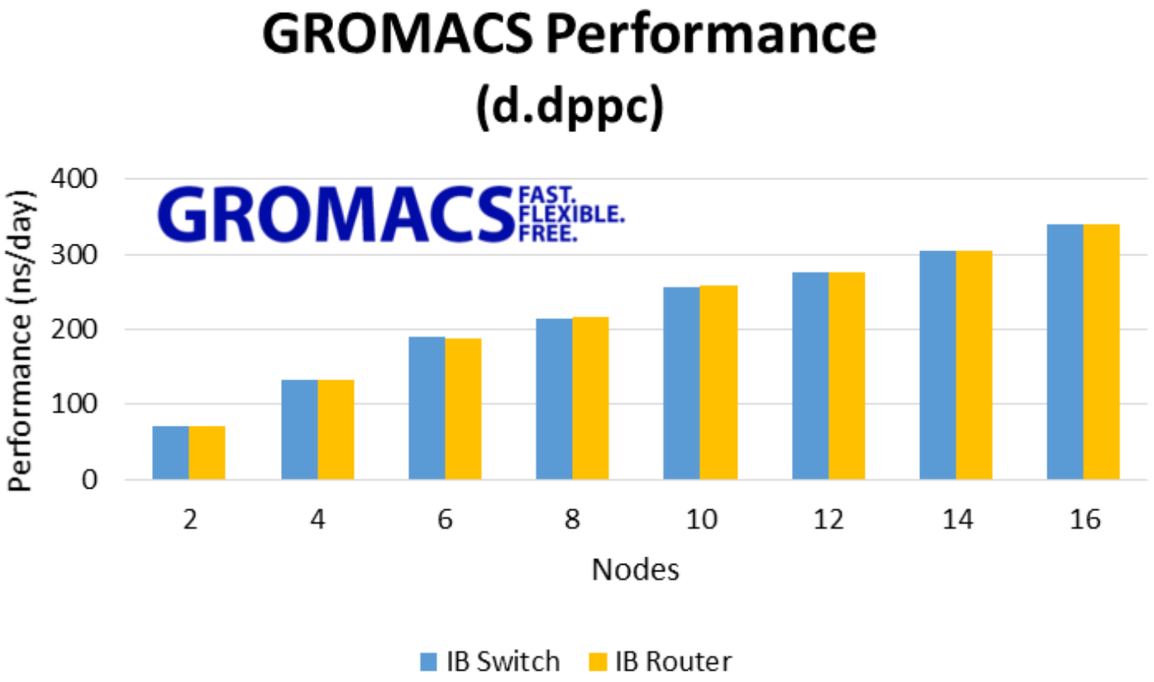
InfiniBand Isolation Enables Great Flexibility

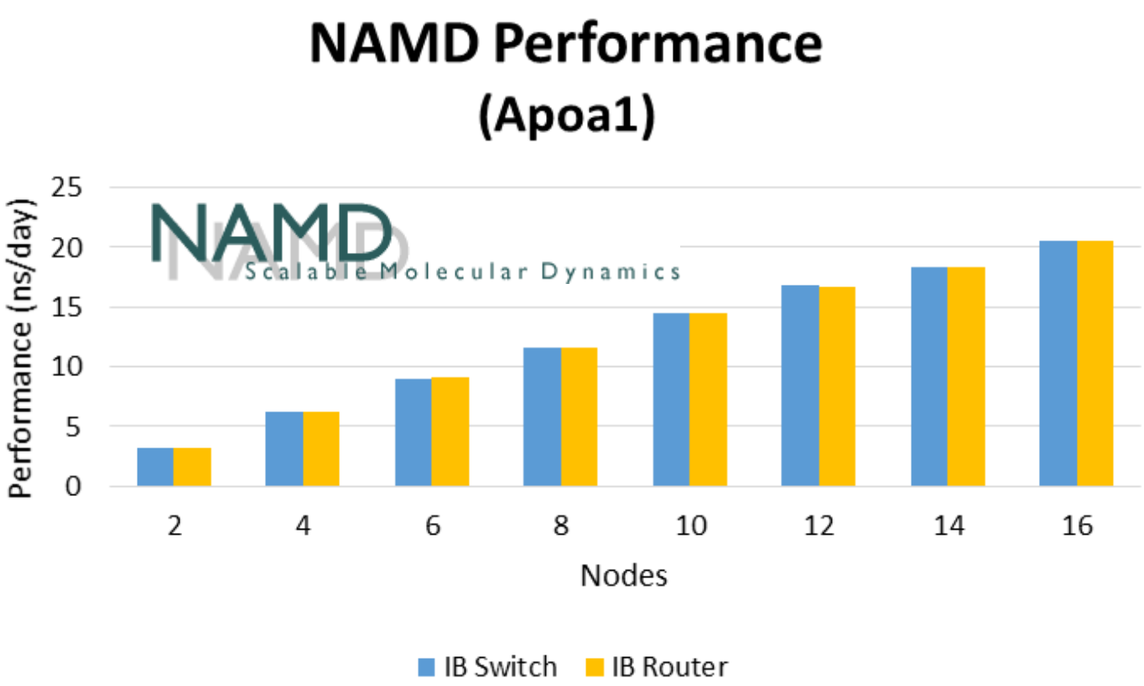
- Enable separation and fault resilience between multiple InfiniBand subnets
- Enable sharing a common storage network by multiple compute infrastructure
- Connect different topologies used by the different subnets



© 2017 Mellanox Technologies

Application Performance





Native InfiniBand Router – Native InfiniBand Performance

© 2017 Mellanox Technologies

Mellanox Solutions

Future Proof Your Data Center

Mellanox to Connect Future #1 HPC Systems (Coral)

OAK RIDGE National Laboratory

“Summit” System

U.S. DEPARTMENT OF ENERGY

Lawrence Livermore National Laboratory

“Sierra” System

Paving the Path to Exascale

© 2017 Mellanox Technologies

Highest-Performance 100Gb/s Interconnect Solutions

Adapters — ConnectX®-5

100Gb/s Adapter, 0.6us latency
175-200 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 3.2Tb/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

© 2017 Mellanox Technologies

Highest-Performance 200Gb/s Interconnect Solutions

Adapters

ConnectX-6

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

Switch

Quantum
The smartest switch, became smarter

40 HDR (200Gb/s) InfiniBand Ports
80 HDR100 InfiniBand Ports
Throughput of 16Tb/s, <90ns Latency

Switch

Spectrum-2

16 400GbE, 32 200GbE, 128 25/50GbE Ports
(10 / 25 / 40 / 50 / 100 / 200 GbE)
Throughput of 6.4Tb/s

Interconnect

LinkX

Transceivers
Active Optical and Copper Cables
(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)

VCSELs, Silicon Photonics and Copper

Software

HPC-X

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

© 2017 Mellanox Technologies

Thank You

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™