Interconnect Your Future

Paving the Road to Exascale

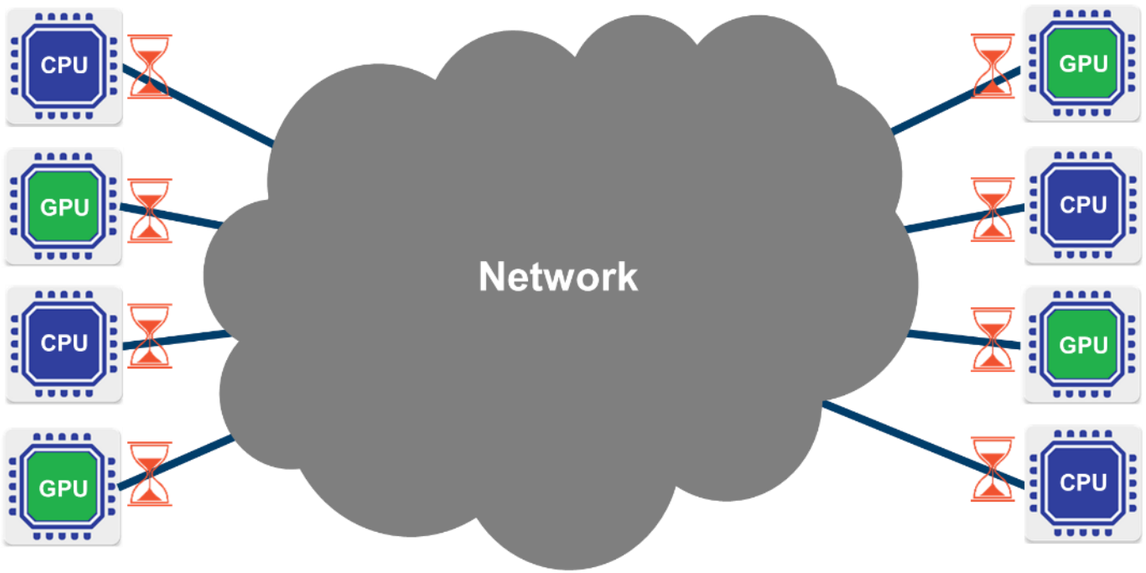
August 2017

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™

Exponential Data Growth – The Need for Intelligent and Faster Interconnect

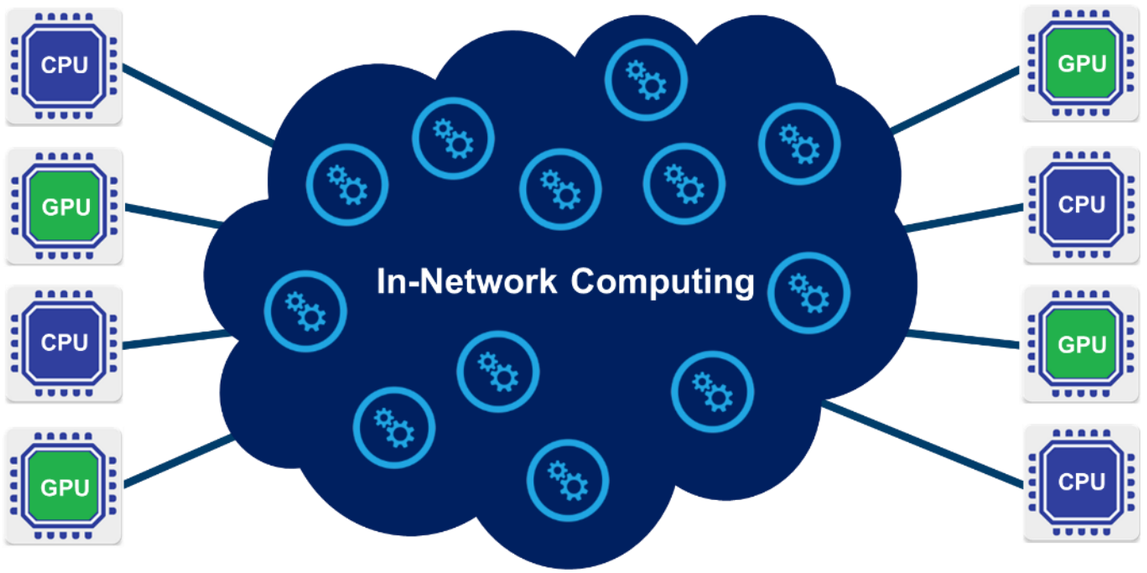
CPU-Centric (Onload)



Must Wait for the Data

Creates Performance Bottlenecks

Data-Centric (Offload)



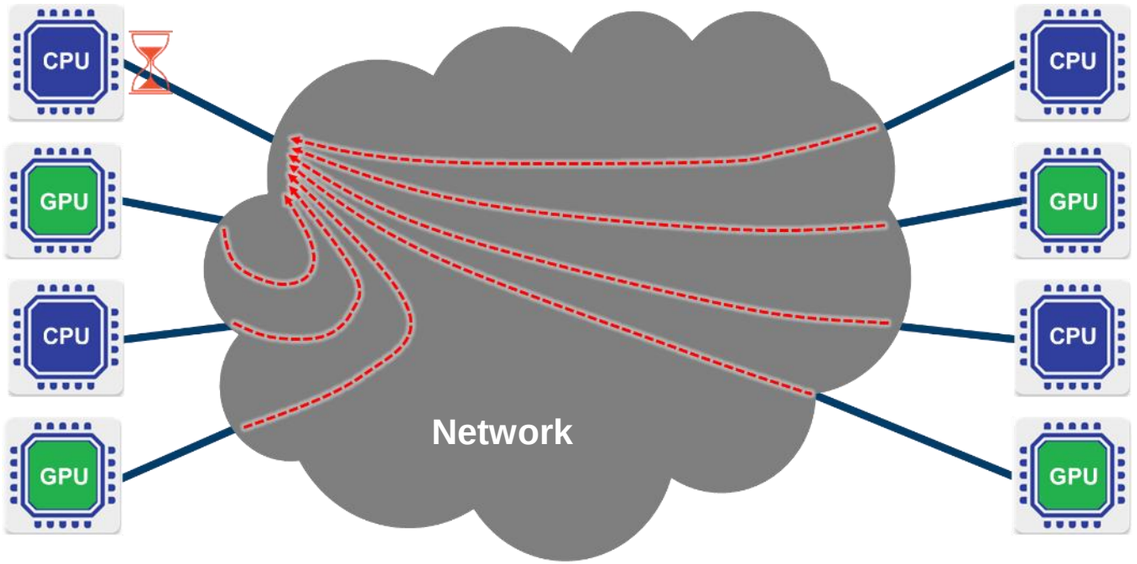
Analyze Data as it Moves!

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

© 2017 Mellanox Technologies

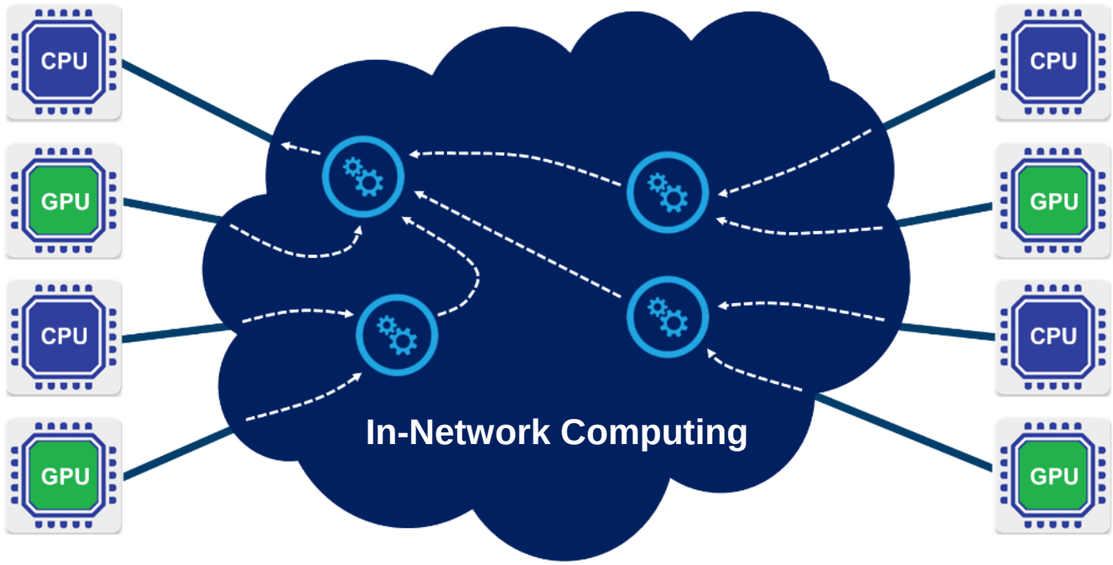
Data Centric Architecture to Overcome Latency Bottlenecks

CPU-Centric (Onload)



HPC / Machine Learning
Communications Latencies of 30-40us

Data-Centric (Offload)

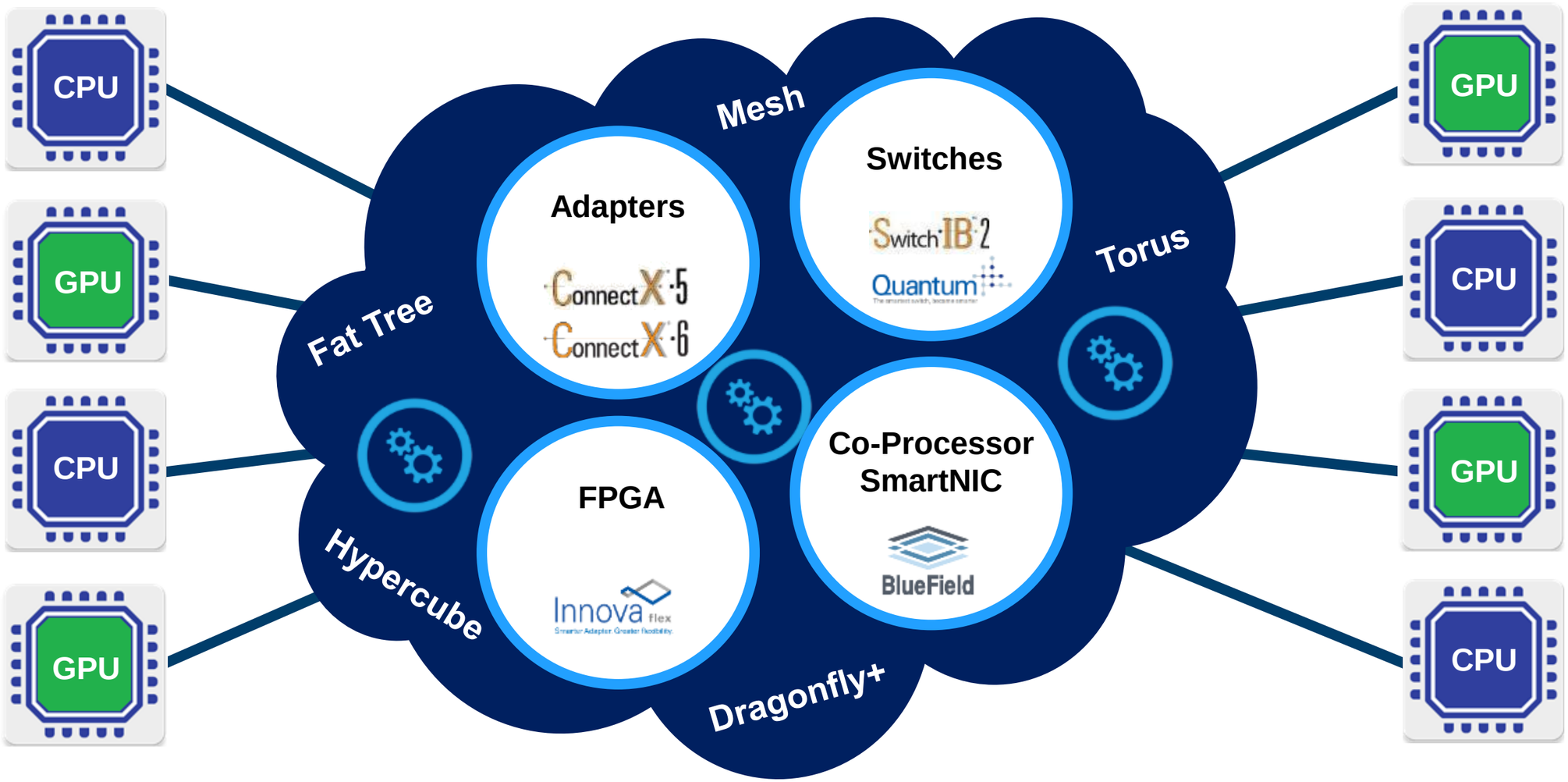


HPC / Machine Learning
Communications Latencies of 3-4us

Intelligent Interconnect Paves the Road to Exascale Performance

© 2017 Mellanox Technologies

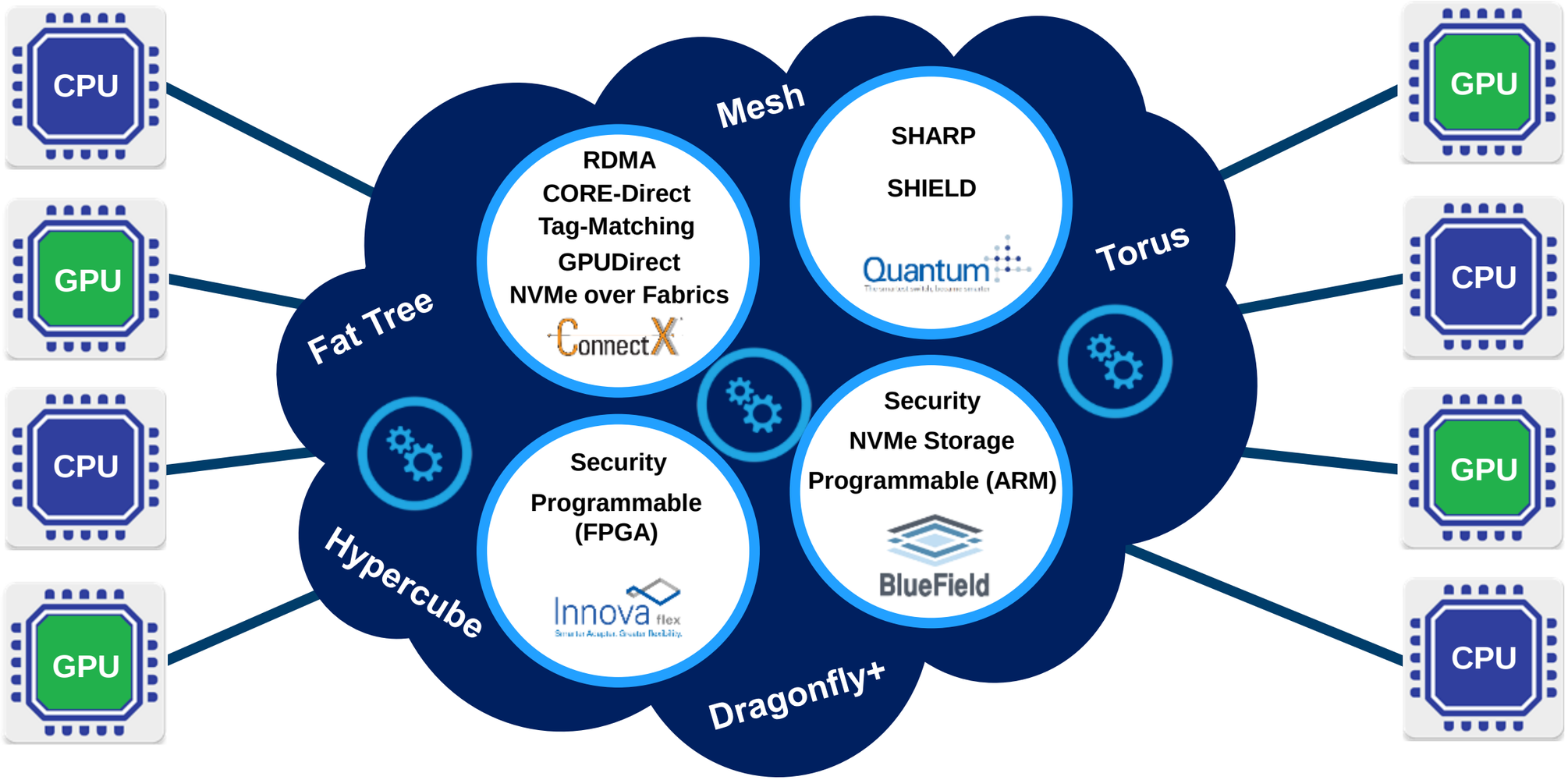
In-Network Computing to Enable Data-Centric Data Centers



In-Network Computing Key for Highest Return on Investment

© 2017 Mellanox Technologies

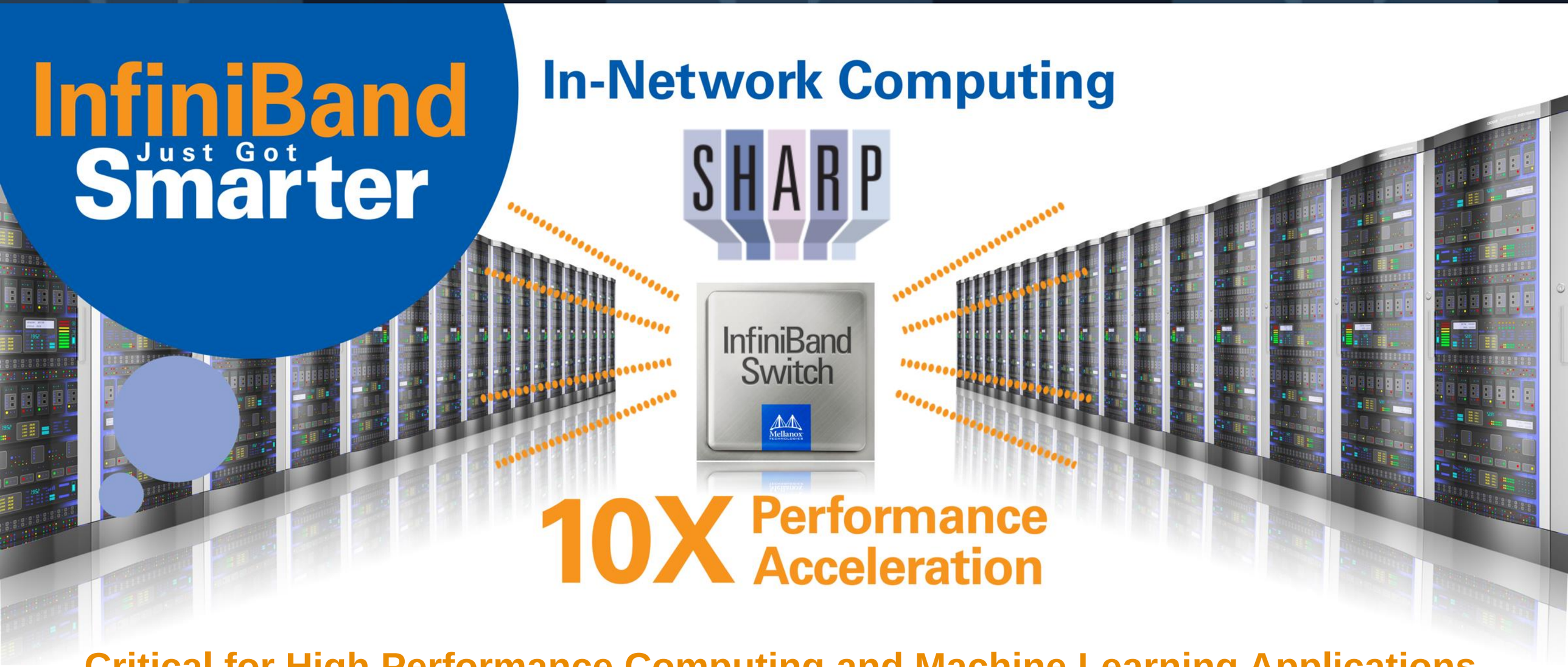
In-Network Computing to Enable Data-Centric Data Centers



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

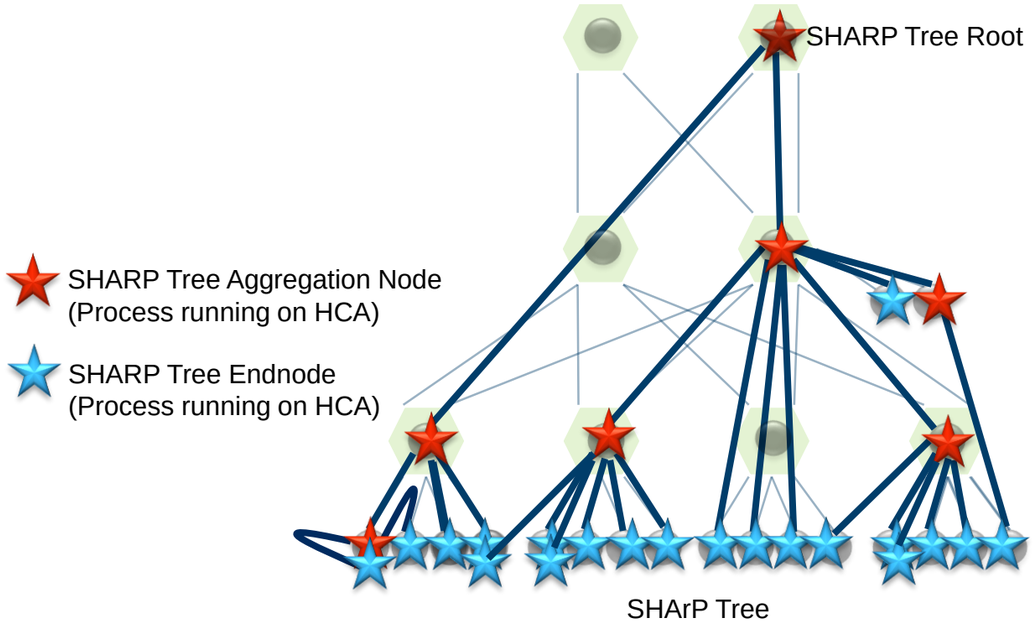
© 2017 Mellanox Technologies

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

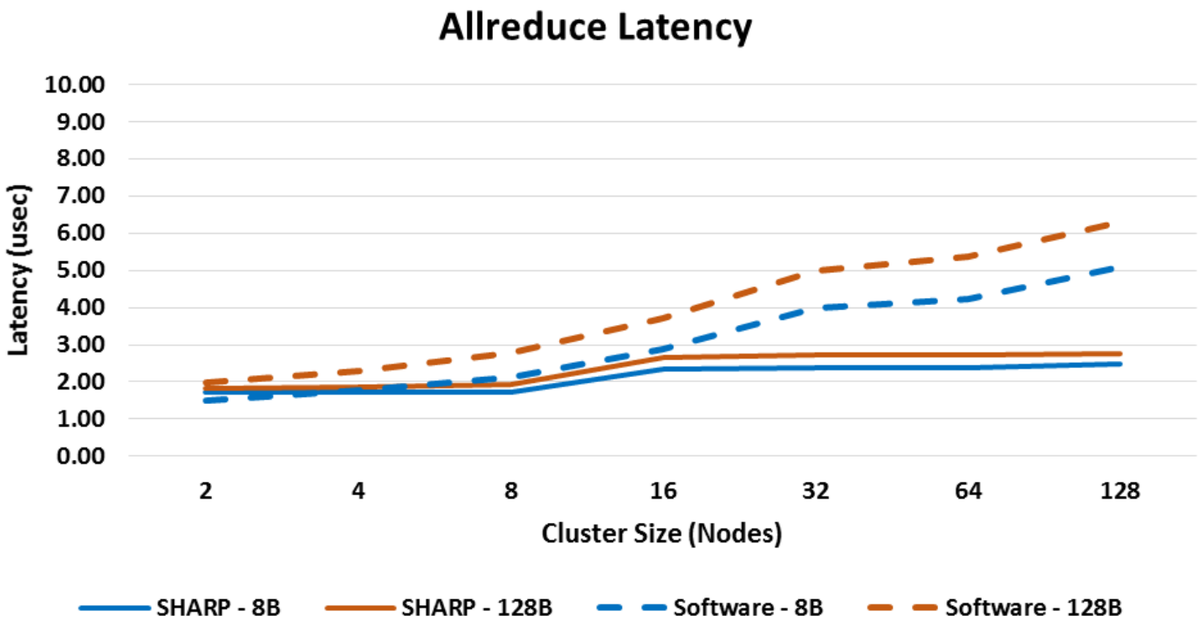
SHARP

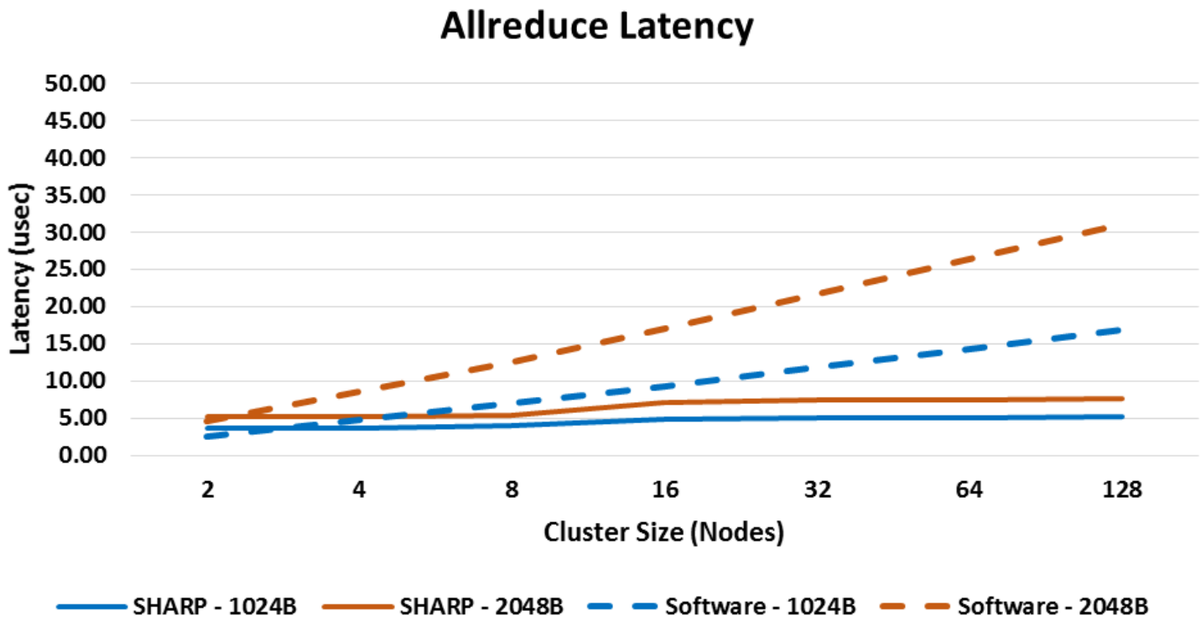
Scalable Hierarchical Aggregation and Reduction Protocol



© 2017 Mellanox Technologies

SHARP Allreduce Performance Advantages





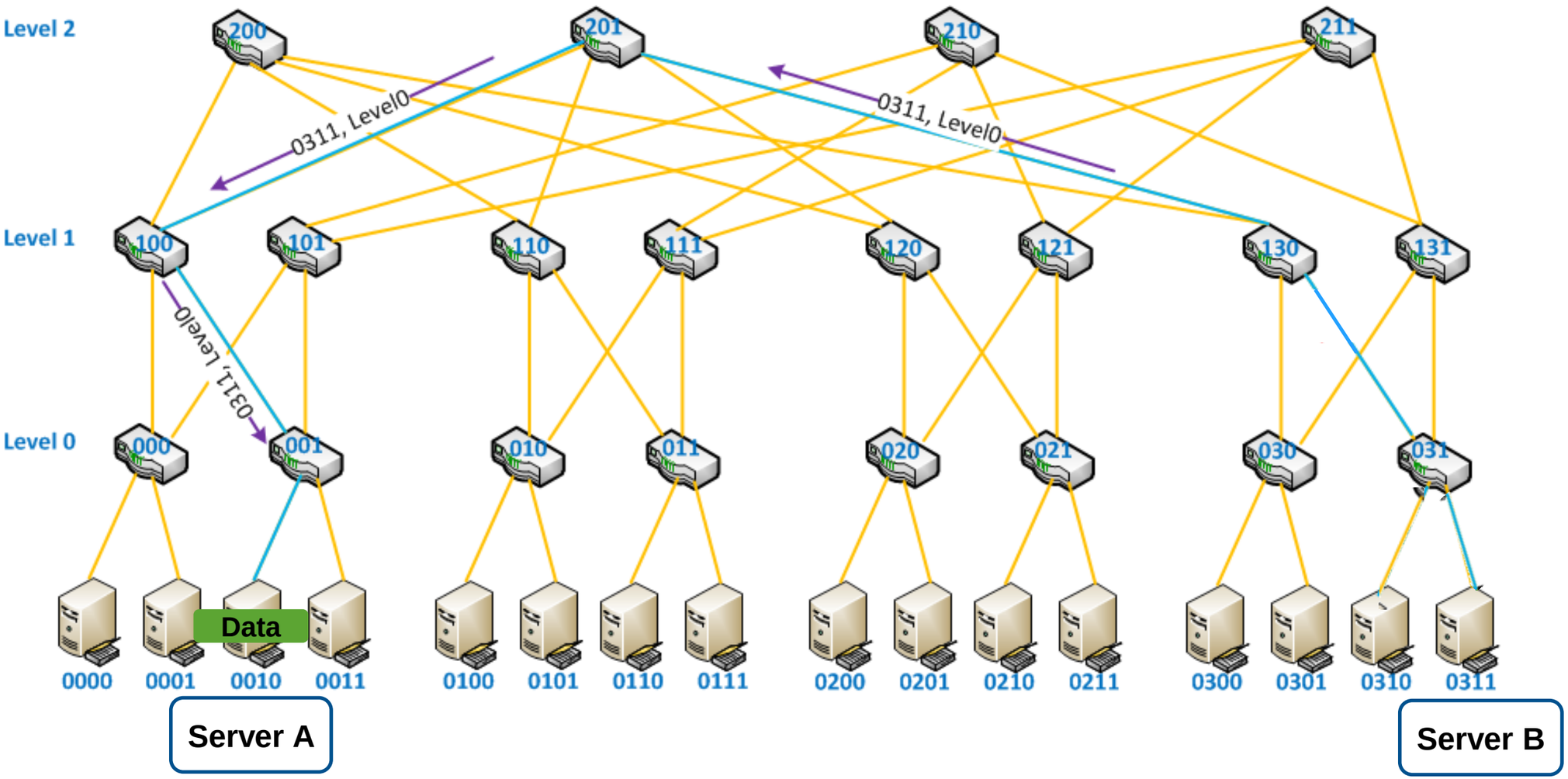
SHARP enables 75% Reduction in Latency
Providing Scalable Flat Latency

© 2017 Mellanox Technologies

SHIELD – Self Healing Interconnect Technology

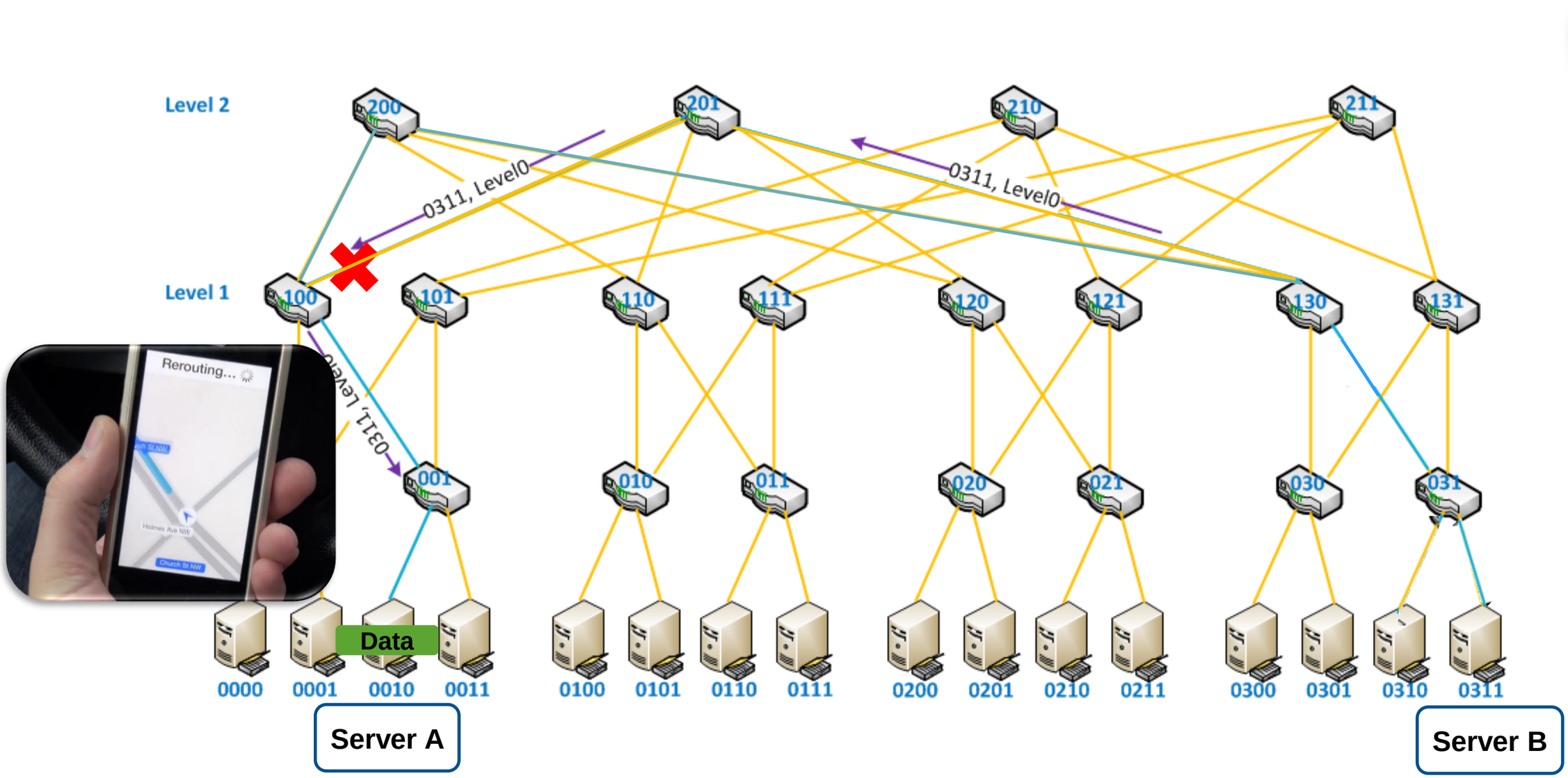
© 2017 Mellanox Technologies

Consider a Flow From A to B



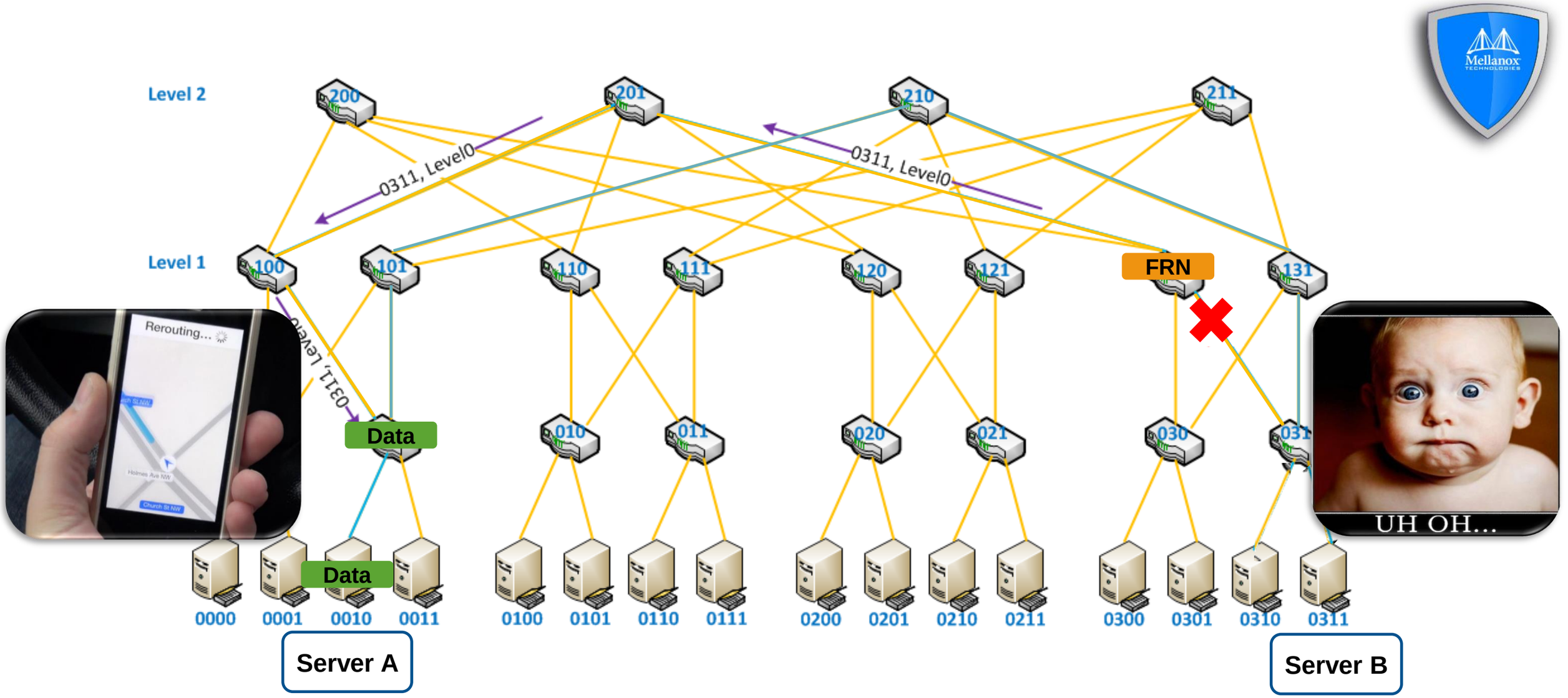
© 2017 Mellanox Technologies

The Simple Case: Local Fix



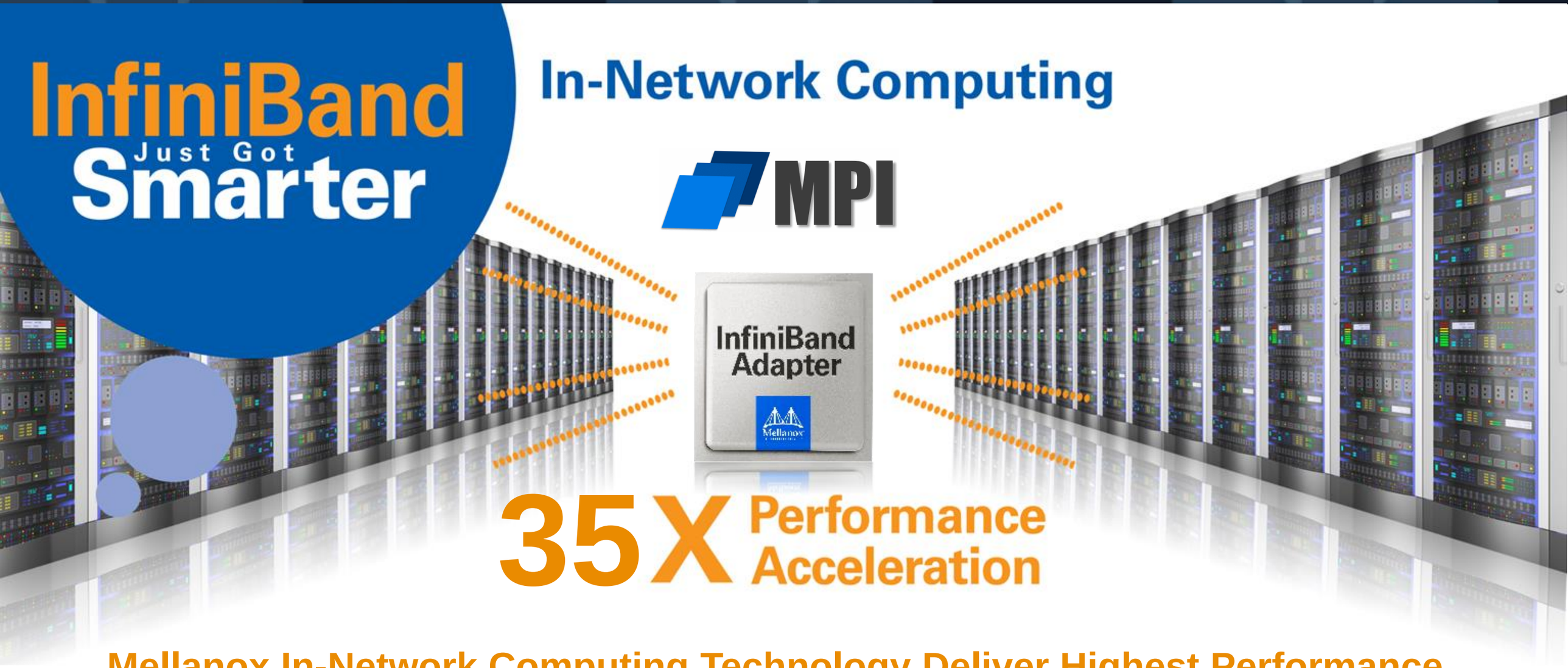
© 2017 Mellanox Technologies

The Remote Case: Using FRN's (Fault Recovery Notifications)



© 2017 Mellanox Technologies

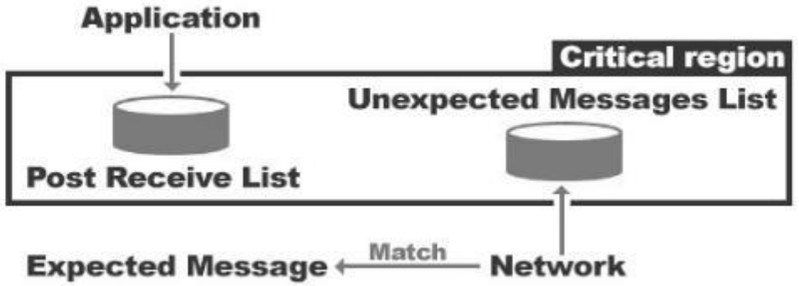
MPI Tag-Matching Hardware Engines

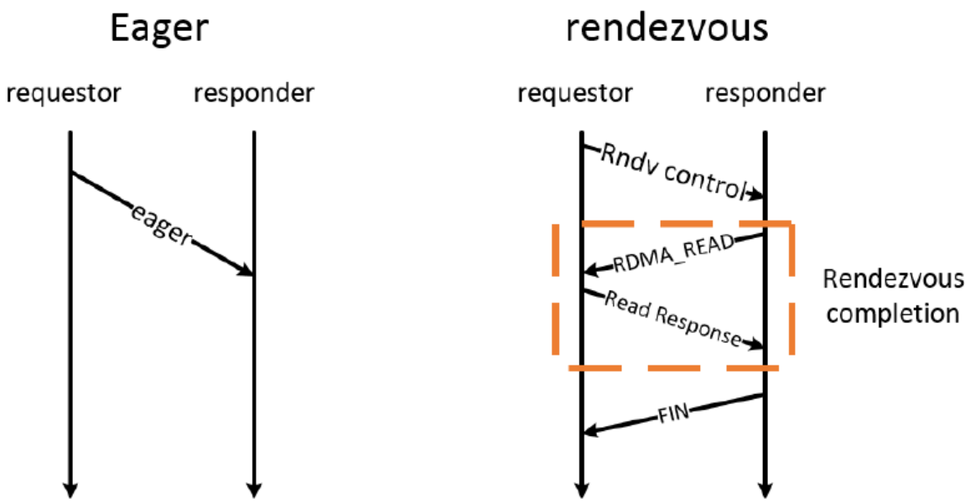


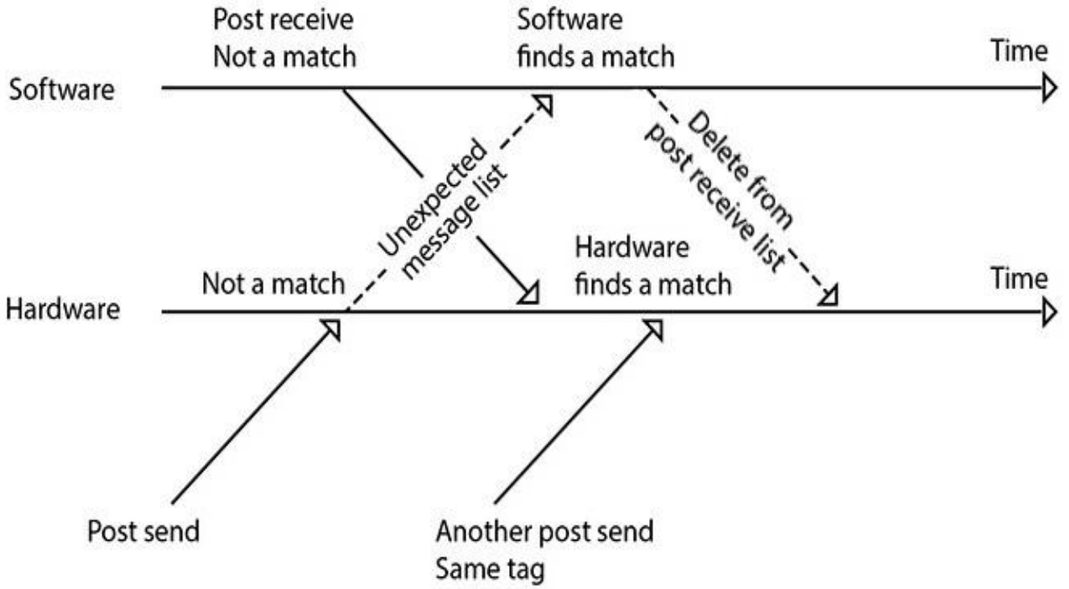
Mellanox In-Network Computing Technology Deliver Highest Performance

© 2017 Mellanox Technologies

MPI Tag-Matching Hardware Engines

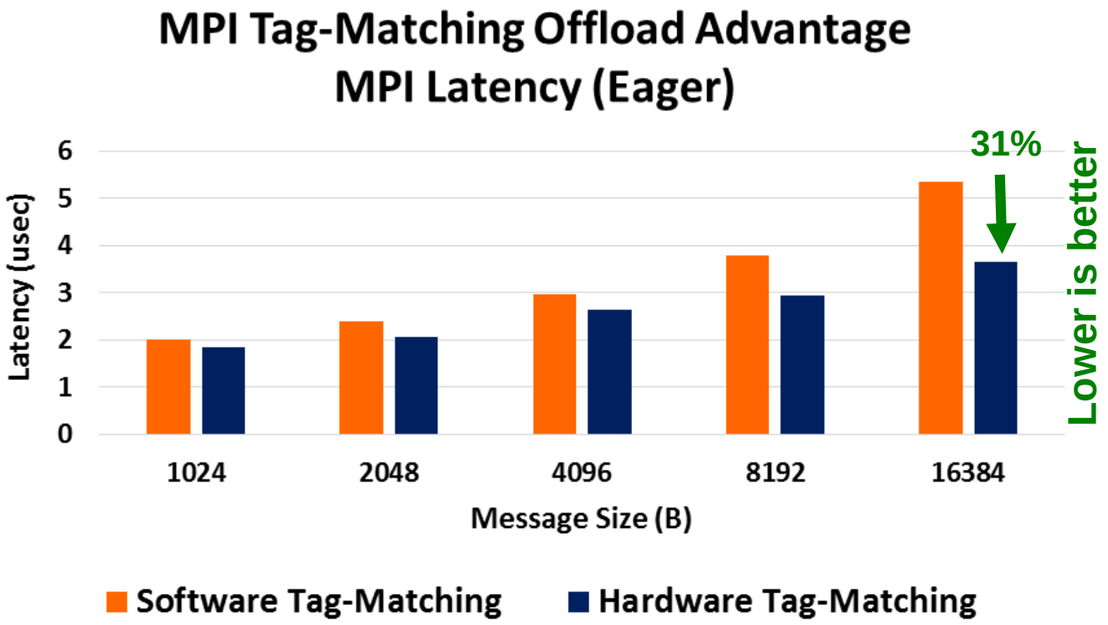


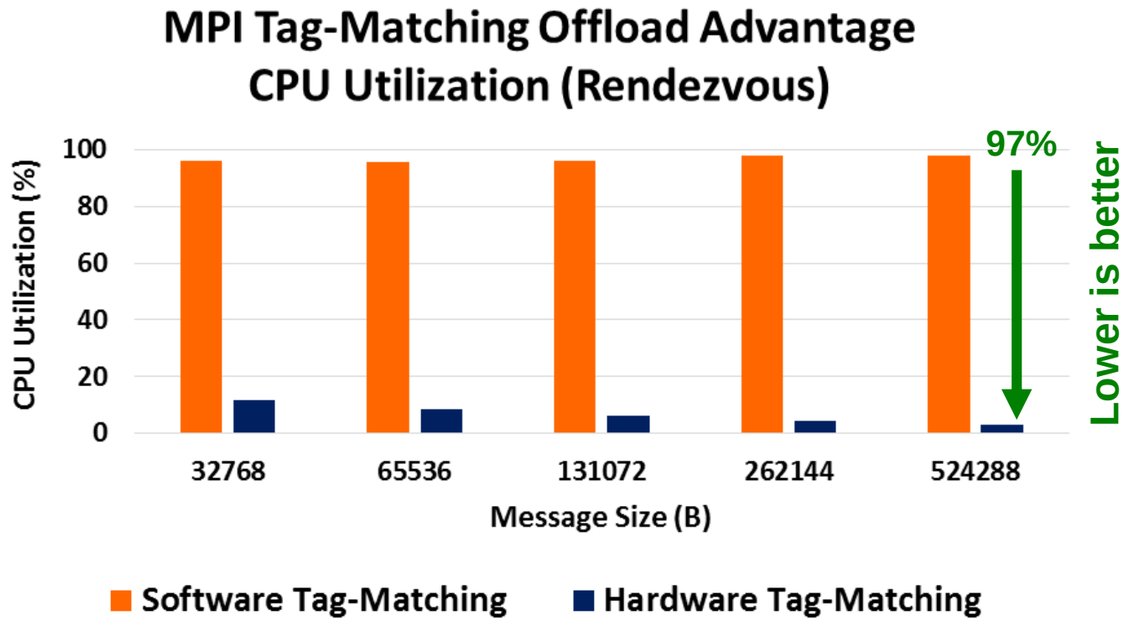




© 2017 Mellanox Technologies

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

Mellanox Roadmap

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

Mellanox to Connect Future #1 HPC Systems (Coral)

OAK RIDGE National Laboratory

“Summit” System

U.S. DEPARTMENT OF ENERGY

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

InfiniBand Delivers Best Return on Investment

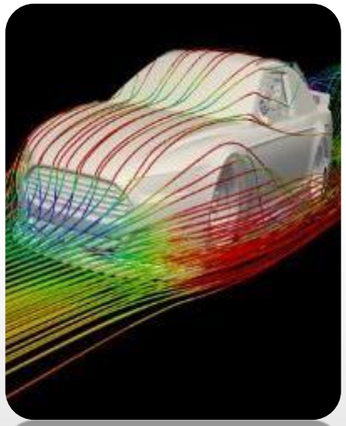
30%-250% Higher Return on Investment

Up to 50% Saving on Capital and Operation Expenses

Highest Applications Performance, Scalability and Productivity

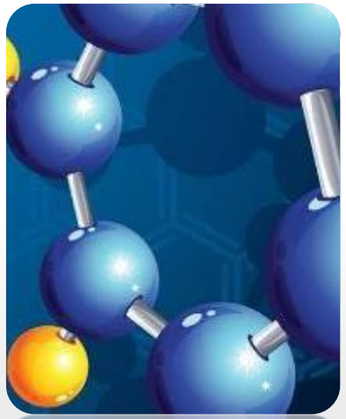
Weather

1.3X Better



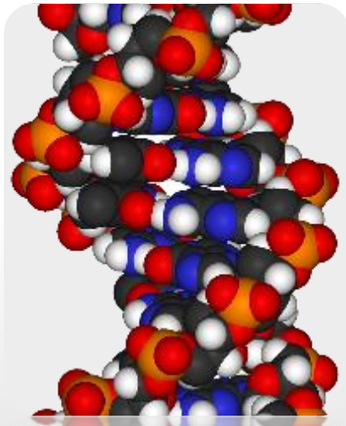
Automotive

2X Better



Chemistry

1.4X Better



Molecular Dynamics

2.5X Better

Genomics

1.3X Better

© 2017 Mellanox Technologies

Thank You

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™