HPC Interconnect Technology update

Paving the Road to Exascale – HPC User Forum

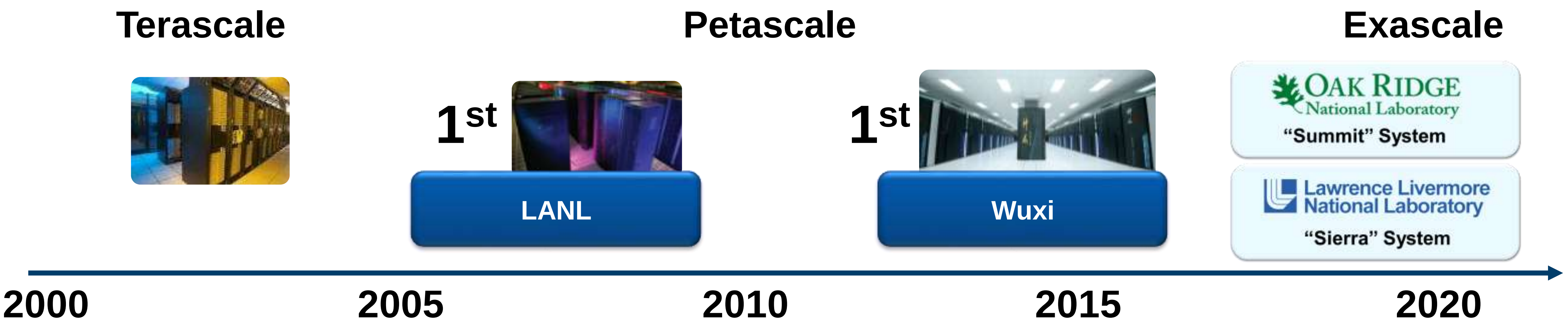
2017

Mellanox TECHNOLOGIES

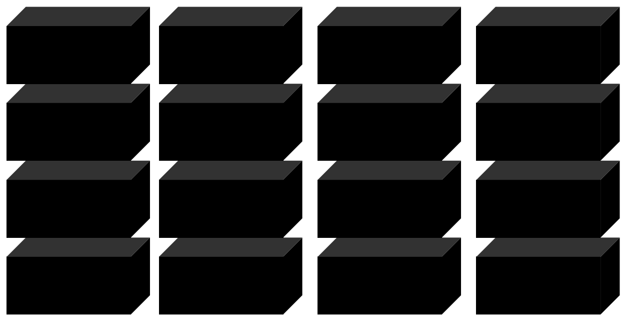
Connect. Accelerate. Outperform.™

The Ever-Growing Demand for Higher Performance

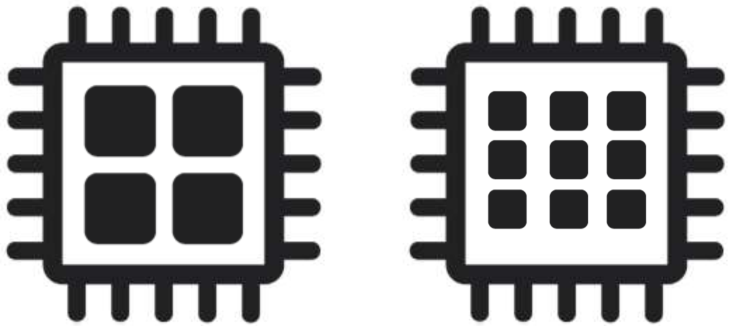
Performance Development



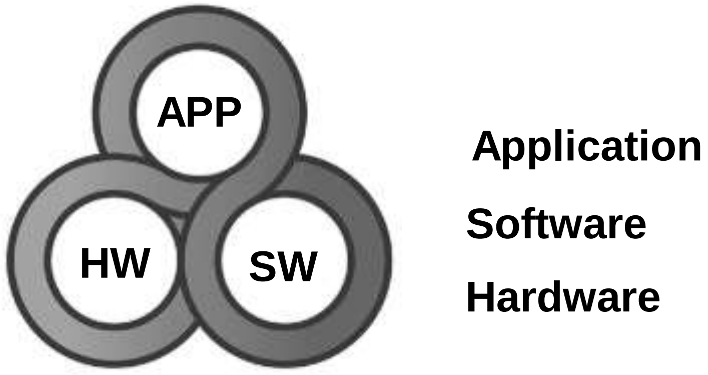
The Interconnect is the Enabling Technology



SMP to Clusters



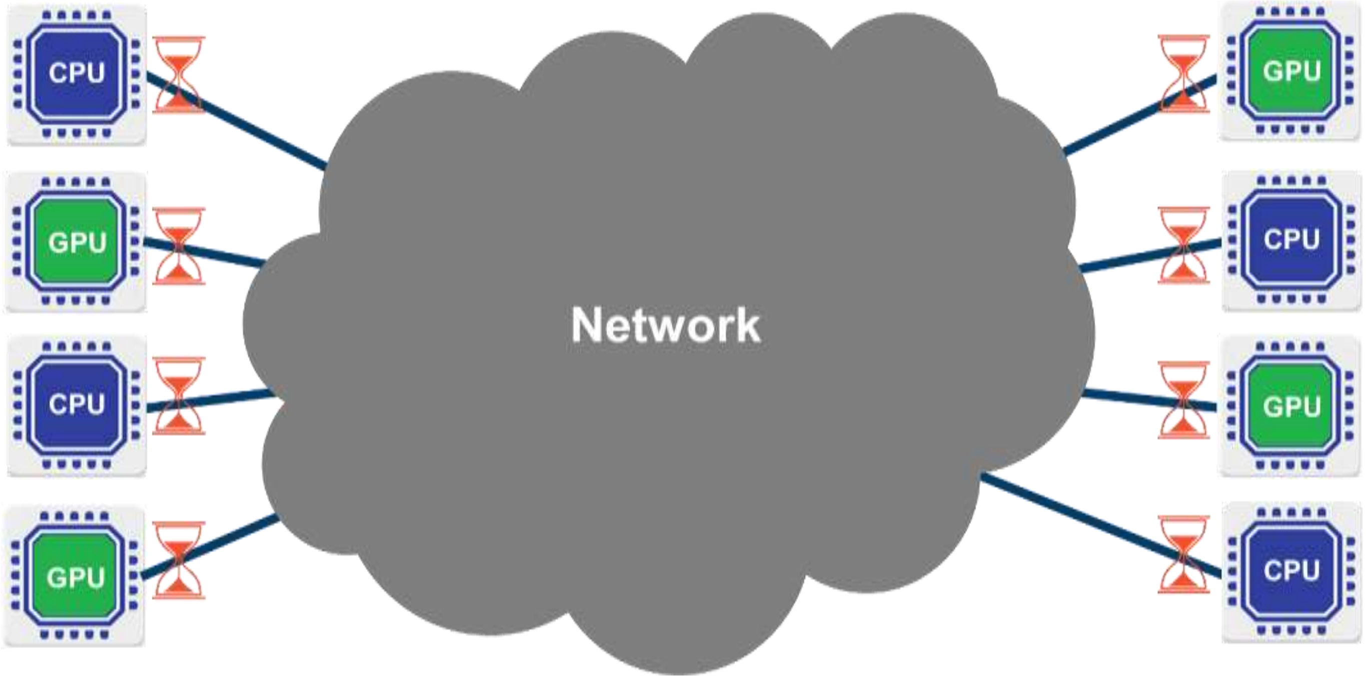
Single-Core to Many-Core



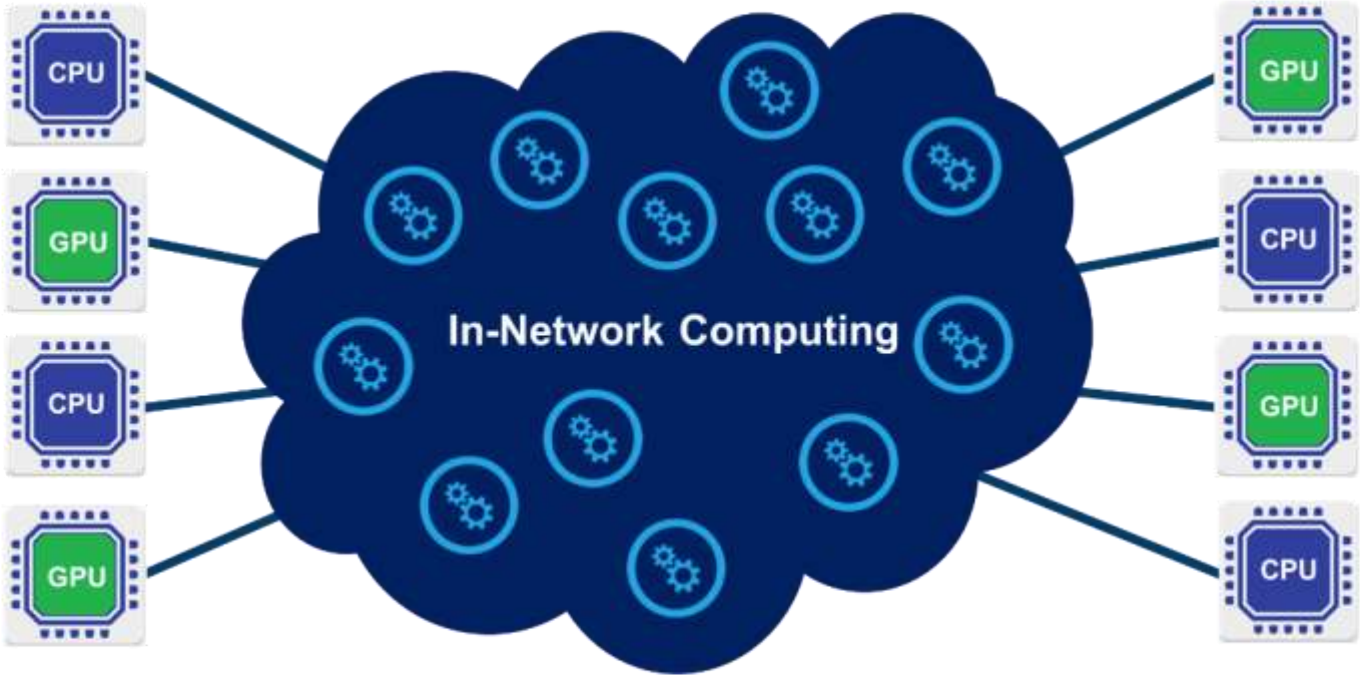
Co-Design

Exponential Data Growth – The Need for Intelligent and Faster Interconnect

CPU-Centric (Onload)



Data-Centric (Offload)



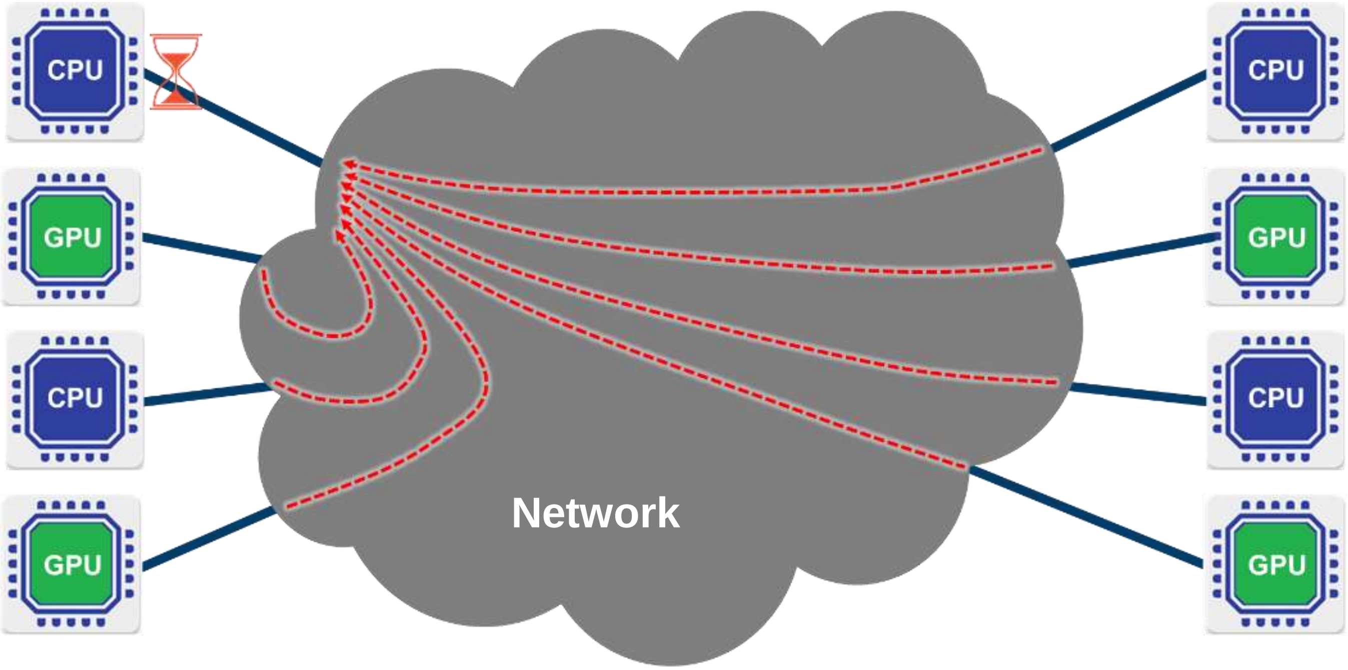
Must Wait for the Data
Creates Performance Bottlenecks

Analyze Data as it Moves!

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

Data Centric Architecture to Overcome Latency Bottlenecks

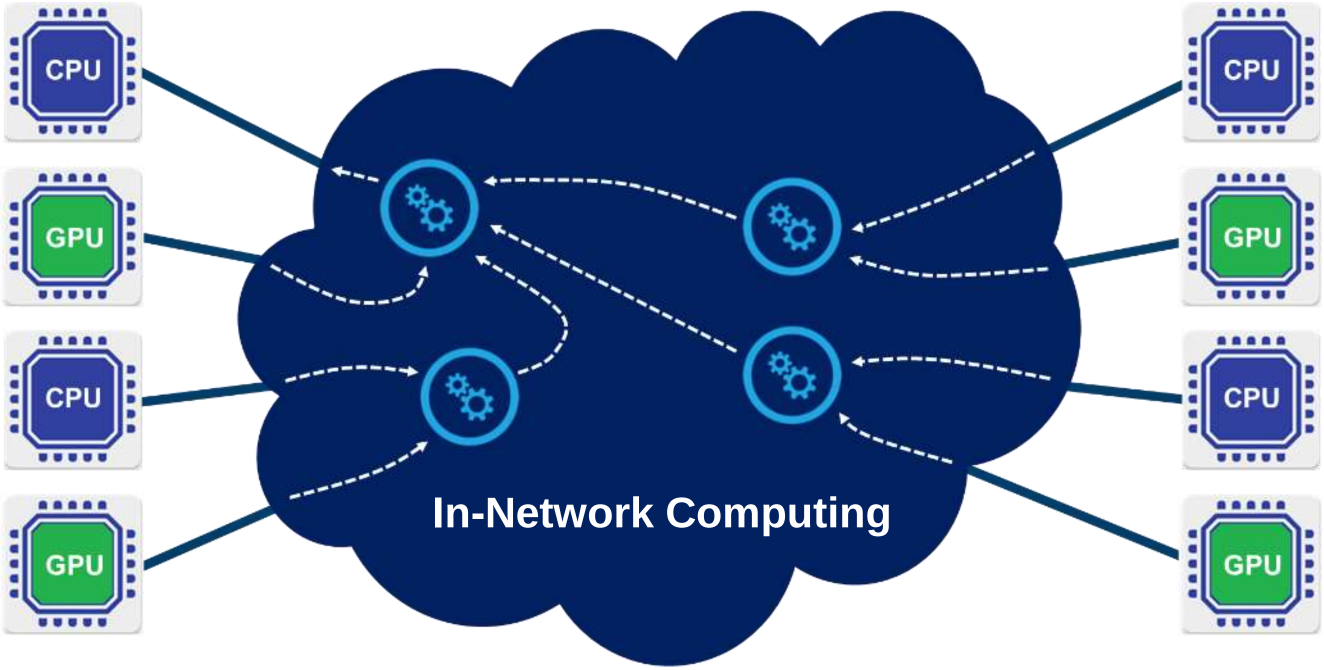
CPU-Centric (Onload)



HPC / Machine Learning

Communications Latencies of 30-40us

Data-Centric (Offload)

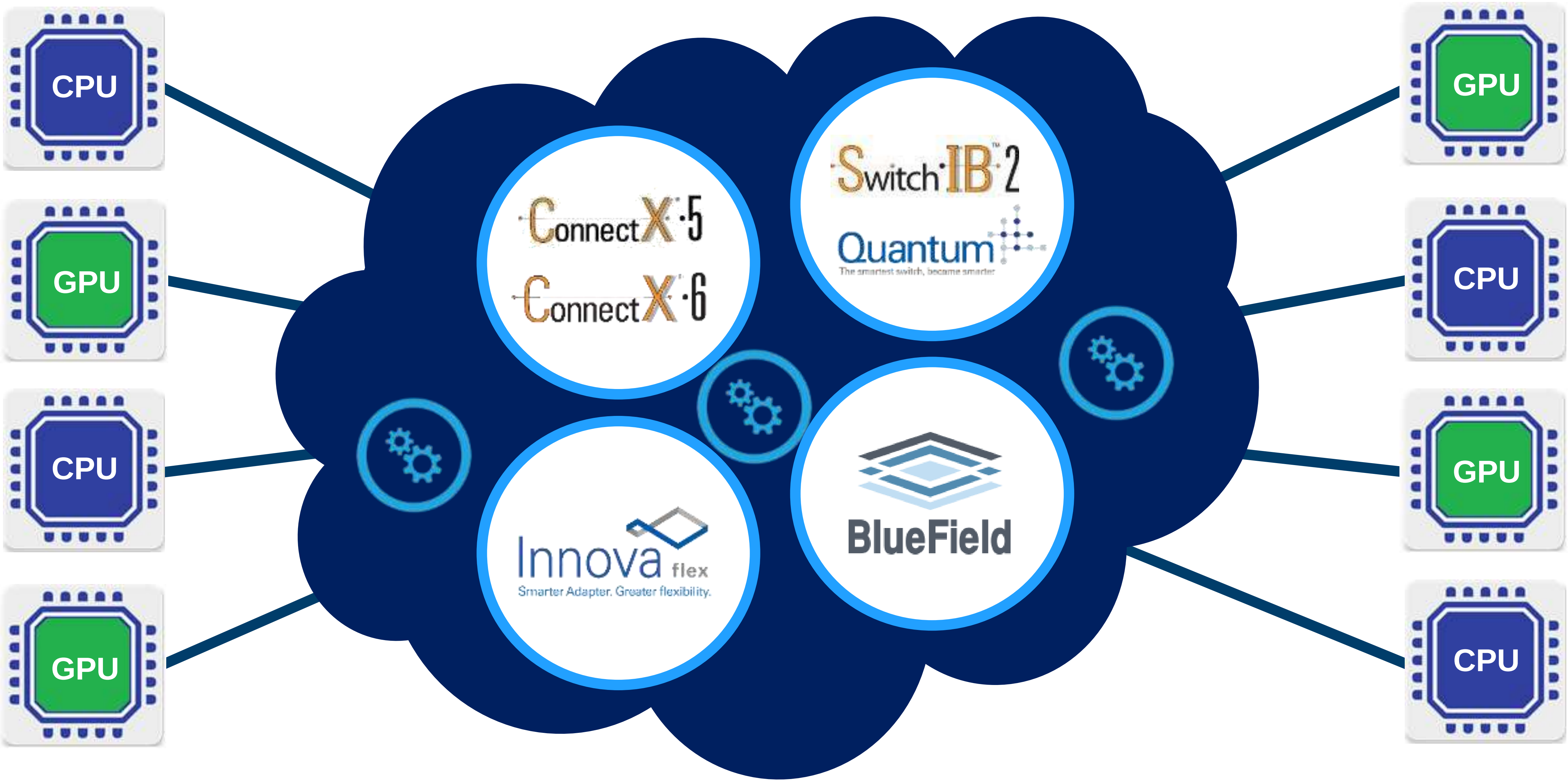


HPC / Machine Learning

Communications Latencies of 3-4us

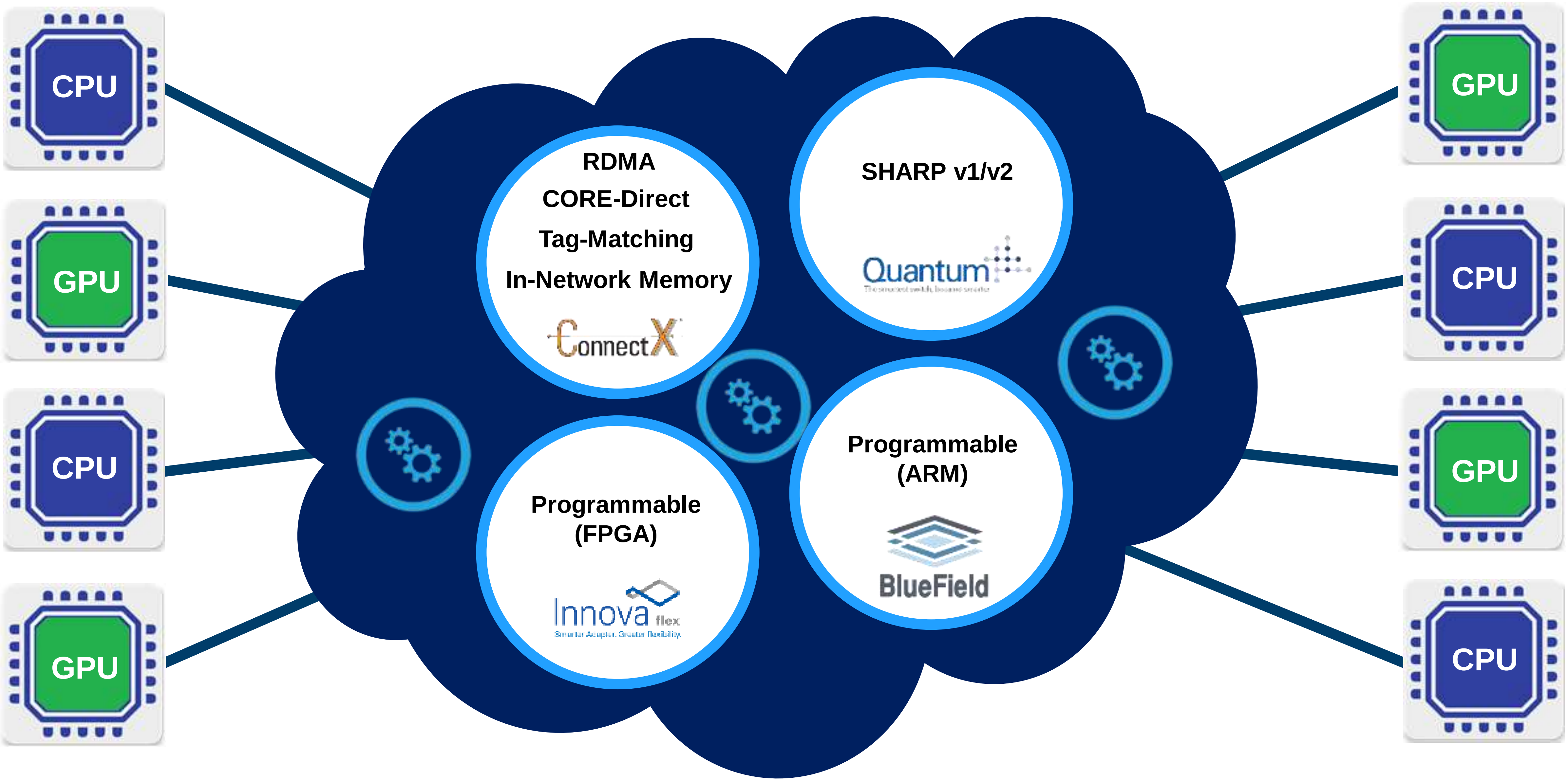
Intelligent Interconnect Paves the Road to Exascale Performance

In-Network Computing to Enable Data-Centric Data Center



In-Network Computing Key for Highest Return on Investment

In-Network Computing to Enable Data-Centric Data Center



In-Network Computing Key for Highest Return on Investment

In-Network Computing and Acceleration Engines

RDMA GPUDirect

Most Efficient Data Access and Data Movement for Compute and Storage platforms, SRIOV for HPC Clouds

200G with <1%CPU Utilization
10X Performance Improvement with GPUDirect

Collectives

CORE-Direct and SHARP Technologies Executes and Manages Data Aggregation and Reduction Algorithms

Accelerates MPI, PGAS/SHMEM and UPC Communication Performance, Accelerates Machine Learning Training Algorithms

Storage

NVMe over Fabrics Offloads, T10-DIF and Erasure Coding offloads

Efficient End-to-End Data Protection, Background Check-Pointing (burst-buffer) and More. Increase System Performance and CPU Availability

Network Transport

All Communications Managed and Operated by the Network Hardware; Adaptive Routing and Congestion Management, Dynamic Connected Transport (DCT)

Maximizes CPU Availability for Applications, increases Network Efficiency and Scalability

Tag Matching

MPI Tag-Matching Offload
MPI Rendezvous Protocol Offload

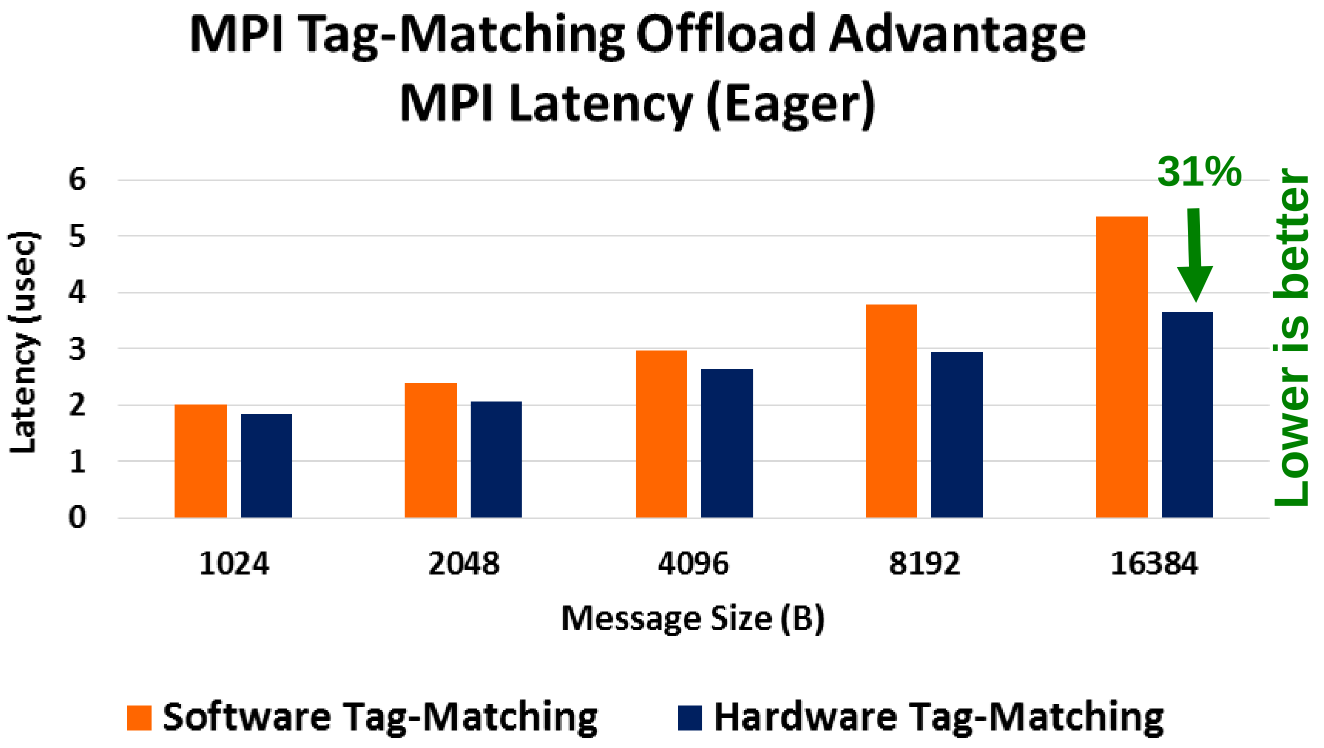
Accelerates MPI Application Performance

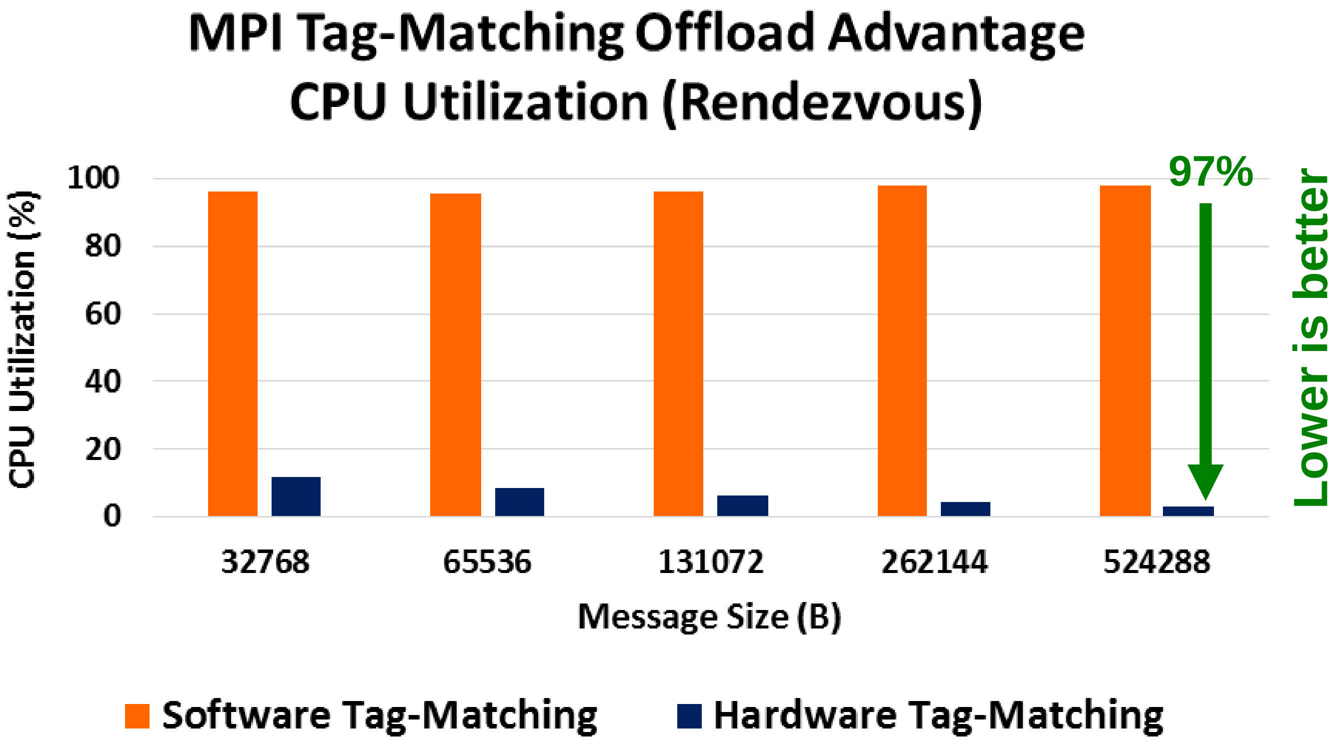
Security

Data Encryption / Decryption (IEEE XTS standard) and Key Management; Federal Information Processing Standards (FIPS) Compliant

Enhances Data Security Options, Enables Protection Between Users Sharing the Same Resources (Different Keys)

MPI Tag-Matching Offload Advantages



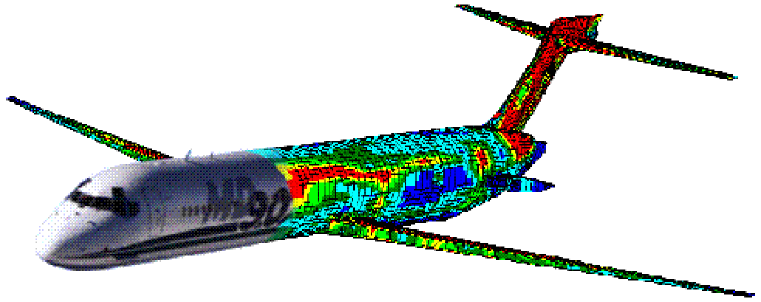


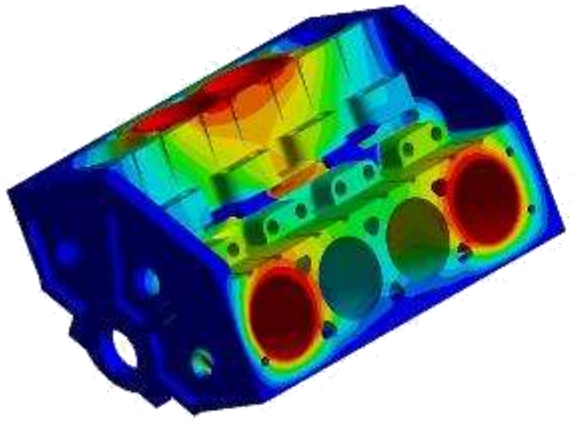
- 31% lower latency and 97% lower CPU utilization for MPI operations
- Performance comparisons based on ConnectX-5

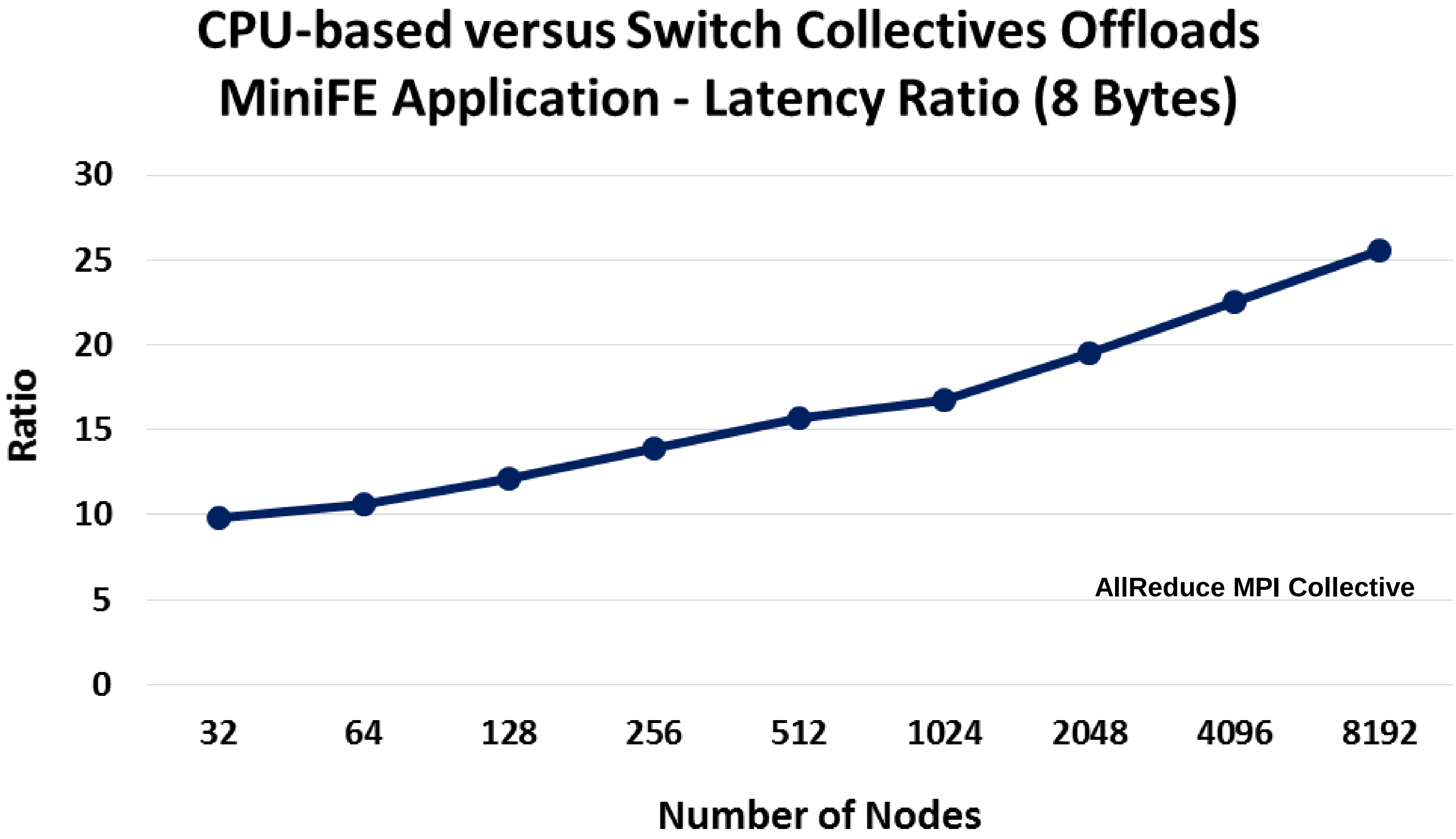
Mellanox In-Network Computing Technology Deliver Highest Performance

SHARP Performance Advantage

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



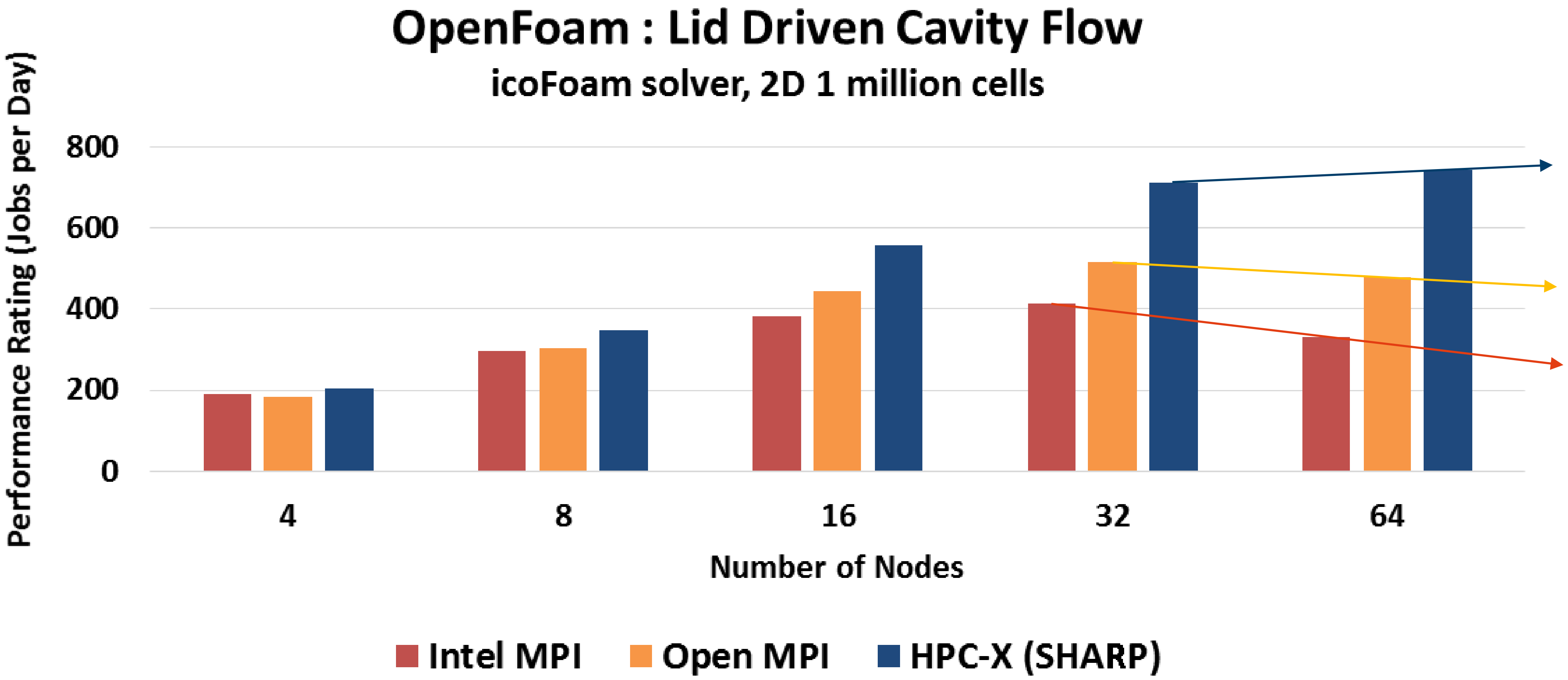


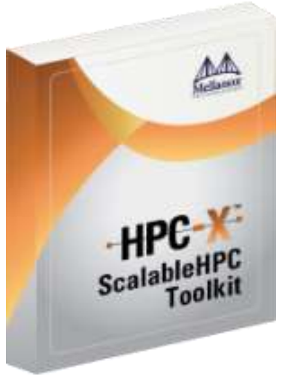

CPU-based versus Switch Collectives Offloads
MiniFE Application - Latency Ratio (8 Bytes)
Ratio
30
25
20
15
10
5
0
32
64
128
256
512
1024
2048
4096
8192
Number of Nodes
AllReduce MPI Collective

10X to 25X Performance Improvement

HPC-X with SHARP Technology

OpenFOAM is a popular computational fluid dynamics application



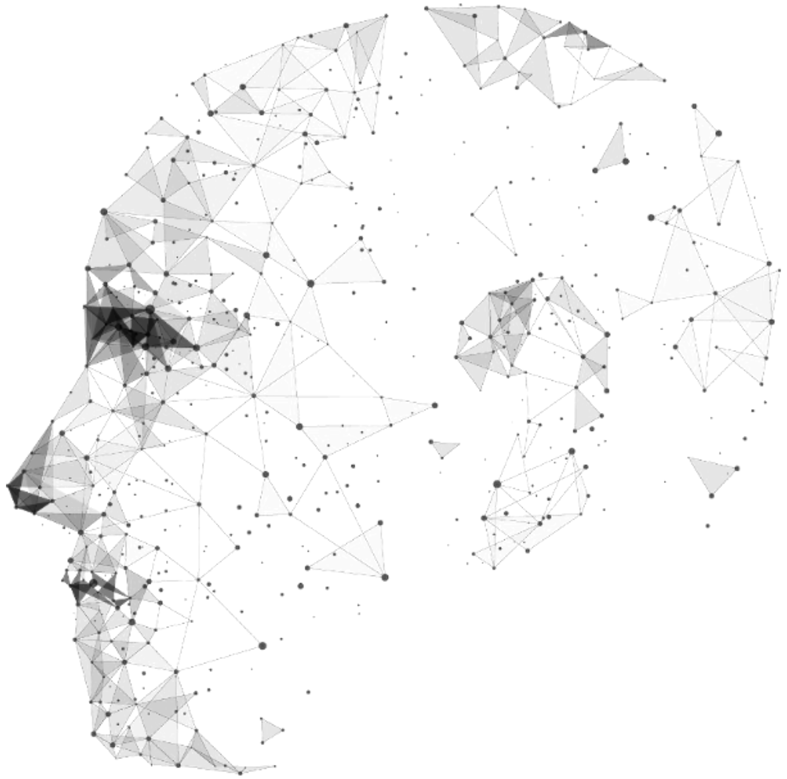


Scalable Hierachical Aggregation and Reduction Protocol

HPC-X with SHARP Delivers 2.2X Higher Performance over Intel MPI

Proven Advantages

- Scalable, flexible, high performance, high bandwidth, end-to-end connectivity
- Standards-based and supported by the largest eco-system
- Supports all compute architectures: x86, Power, ARM, GPU, FPGA etc.
- Native Offloading architecture
- RDMA, GPUDirect, rCUDA, SHARP and other accelerations
- Backward and future compatible



Scalable HPC Depends on Mellanox

Media Resources

OrionX Reports Position InfiniBand as the Leading HPI Technology and Mellanox the Leading Vendor

July 7, 2016 by staff   Leave a Comment

In this special guest feature, Peter ffoulkes from OrionX outlines a series of new reports that show how InfiniBand continues to dominate the market for High Performance Interconnects.

The OrionX Constellation reports published June 29th address the evolution, environment, evaluation and excellence ratings for the High Performance Interconnect (HPI) market. Defined as the very high end of the networking equipment market where high bandwidth and low latency are non-negotiable, HPI technologies support the most demanding workloads that are typical of extreme-scale systems in high performance computing (HPC), artificial intelligence, cloud computing, and web-scale deployments.

Peter ffoulkes, OrionX

Link to the article

InfiniBand Enables Intelligent Networks

January 13, 2016 by staff

In this special guest feature, Gilad Shainer from Mellanox writes that the network is the key to future scalable systems.

HPC Frequently Reinvents Itself to Keep Pace

In the world of high-performance computing, there is a constant and ever-growing demand for even higher performance. Technology providers have worked ceaselessly to keep up with that demand, with each new generation of faster, more reliable, and more efficient systems.

Ultimately, though, every technology reaches its limits, and progress can therefore stall unless there is a

Gilad Shainer, VP of Marketing, Mellanox

Link to the article

Slidecast: Advantages of Offloading Architectures for HPC

April 19, 2016 by Rich Brueckner

Interconnect Your Future with InfiniBand
SM
SB7780 Router 1U
Supports up to 6 Different Subnets

Link to the article

Interview: Why Co-design is the Path Forward for Exascale Computing

March 4, 2016 by Rich Brueckner

Link to the article

April 12, 2016

The Ultimate Debate – Interconnect Offloading Versus Onloading

Gilad Shainer, Mellanox

The high performance computing market is going through a technology transition – the Co-Design transition. As has already been discussed in many articles, this transition has emerged in order to solve the performance bottlenecks of today's infrastructures and applications, performance bottlenecks that were created by multi-core CPUs and the existing CPU-centric system architecture.

How are multi-core CPUs the source for today's performance bottlenecks? In order to understand that, we need to go back in time to the era of single-core CPUs. Back then, performance gains came from increases in CPU frequency and from the reduction of networking functions (network adapter and switches). Each new generation of product brought faster CPUs and lower-latency network adapters and

Link to the article

June 18, 2016

Offloading vs. Onloading: The Case of CPU Utilization

Gilad Shainer, Mellanox

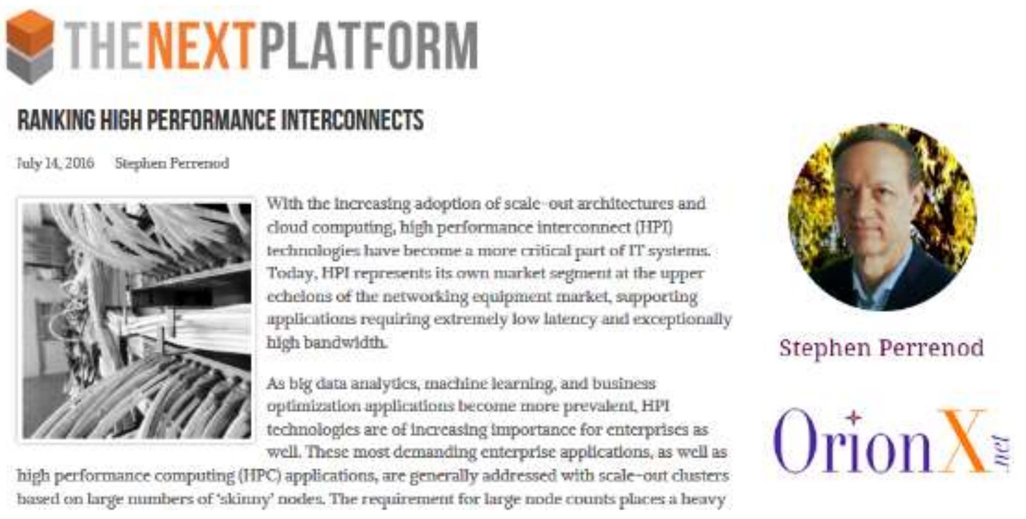
One of the primary conversations these days in the field of networking is whether it is better to onload network functions onto the CPU or better to offload these functions to the interconnect hardware.

Onloading interconnect technology is easier to build, but the issue becomes the CPU utilization; because the CPU must manage and execute network operations, it has less availability for applications, which is its primary purpose.

Offloading, on the other hand, seeks to overcome performance bottlenecks in the CPU by performing the network functions, as well as complex communications operations,

Link to the article

Media Resources


THENEXTPLATFORM

RANKING HIGH PERFORMANCE INTERCONNECTS

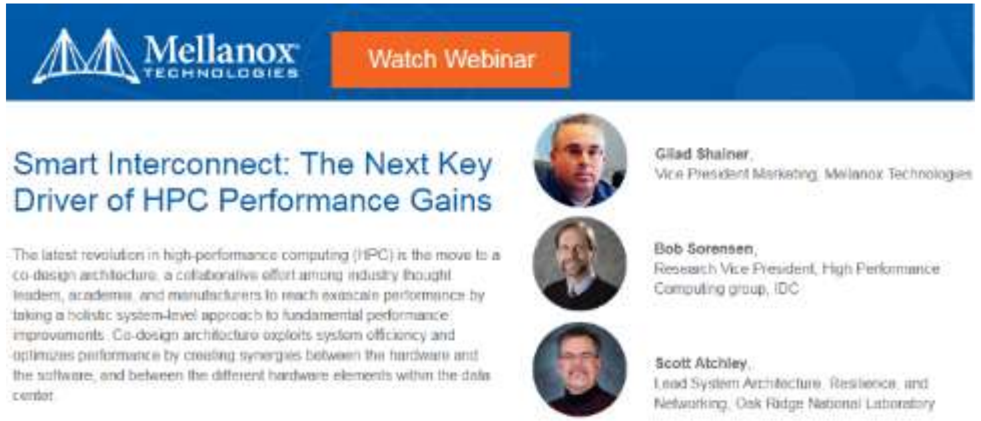
July 14, 2016 Stephen Perrenod

With the increasing adoption of scale-out architectures and cloud computing, high performance interconnect (HPI) technologies have become a more critical part of IT systems. Today, HPI represents its own market segment at the upper echelons of the networking equipment market, supporting applications requiring extremely low latency and exceptionally high bandwidth.

As big data analytics, machine learning, and business optimization applications become more prevalent, HPI technologies are of increasing importance for enterprises as well. These most demanding enterprise applications, as well as high performance computing (HPC) applications, are generally addressed with scale-out clusters based on large numbers of 'skinny' nodes. The requirement for large node counts places a heavy

Stephen Perrenod

Link to the article



Smart Interconnect: The Next Key Driver of HPC Performance Gains

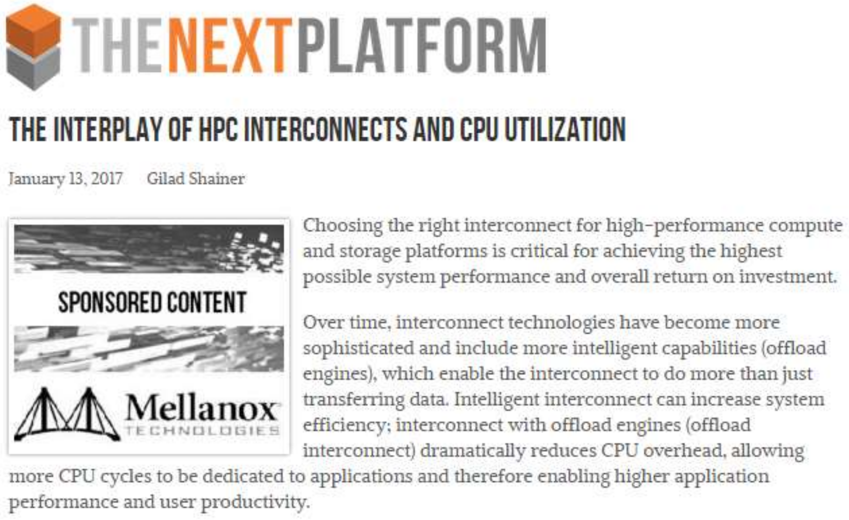
The latest revolution in high-performance computing (HPC) is the move to a co-design architecture, a collaborative effort among industry thought leaders, academia, and manufacturers to reach exascale performance by taking a holistic system-level approach to fundamental performance improvements. Co-design architecture exploits system efficiency and optimizes performance by creating synergies between the hardware and the software, and between the different hardware elements within the data center.

Gilad Shainer, Vice President Marketing, Mellanox Technologies

Bob Sorensen, Research Vice President, High Performance Computing group, IDC

Scott Atchley, Lead System Architecture, Resilience, and Networking, Oak Ridge National Laboratory

Link to the webinar


THENEXTPLATFORM

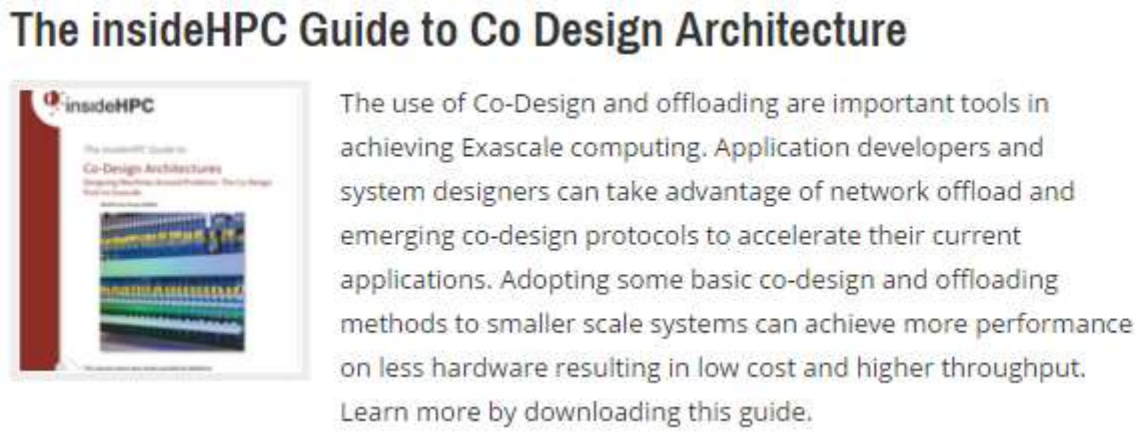
THE INTERPLAY OF HPC INTERCONNECTS AND CPU UTILIZATION

January 13, 2017 Gilad Shainer

Choosing the right interconnect for high-performance compute and storage platforms is critical for achieving the highest possible system performance and overall return on investment.

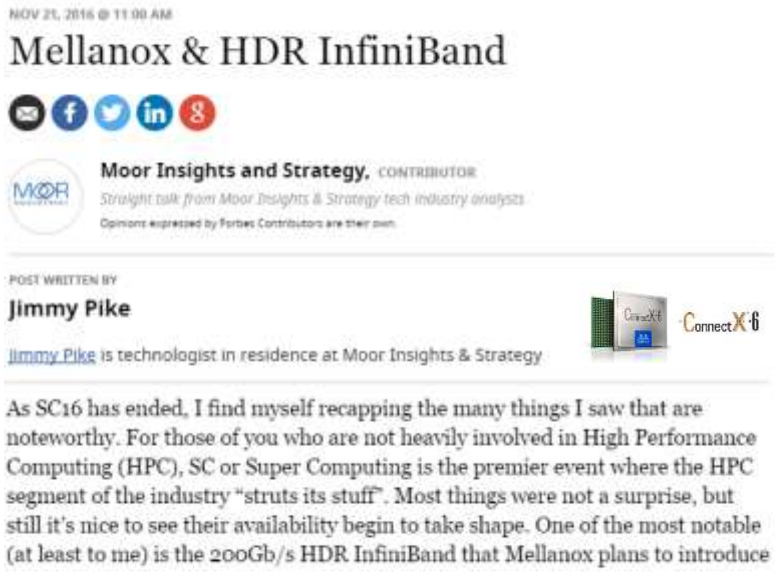
Over time, interconnect technologies have become more sophisticated and include more intelligent capabilities (offload engines), which enable the interconnect to do more than just transferring data. Intelligent interconnect can increase system efficiency; interconnect with offload engines (offload interconnect) dramatically reduces CPU overhead, allowing more CPU cycles to be dedicated to applications and therefore enabling higher application performance and user productivity.

Link to the article


The insideHPC Guide to Co Design Architecture

The use of Co-Design and offloading are important tools in achieving Exascale computing. Application developers and system designers can take advantage of network offload and emerging co-design protocols to accelerate their current applications. Adopting some basic co-design and offloading methods to smaller scale systems can achieve more performance on less hardware resulting in low cost and higher throughput. Learn more by downloading this guide.

Link to the article


NOV 21, 2016 @ 11:00 AM

Mellanox & HDR InfiniBand

Moor Insights and Strategy, CONTRIBUTOR

Straight talk from Moor Insights & Strategy tech industry analysts

Opinions expressed by Forbes Contributors are their own.

POST WRITTEN BY

Jimmy Pike

Jimmy Pike is technologist in residence at Moor Insights & Strategy

As SC16 has ended, I find myself recapping the many things I saw that are noteworthy. For those of you who are not heavily involved in High Performance Computing (HPC), SC or Super Computing is the premier event where the HPC segment of the industry "struts its stuff". Most things were not a surprise, but still it's nice to see their availability begin to take shape. One of the most notable (at least to me) is the 200Gb/s HDR InfiniBand that Mellanox plans to introduce

Link to the article

Supercomputing 2016 Conference Session Next Generation of Co-Processors Emerges: In-Network Computing

SC16
Salt Lake City, Utah | hpc matters
Mellanox TECHNOLOGIES
OAK RIDGE National Laboratory

Link to the Session Video

Thank You

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™