



Interconnect Your Future  
Enabling the Best Datacenter Return on Investment

TOP500 Supercomputers, November 2016

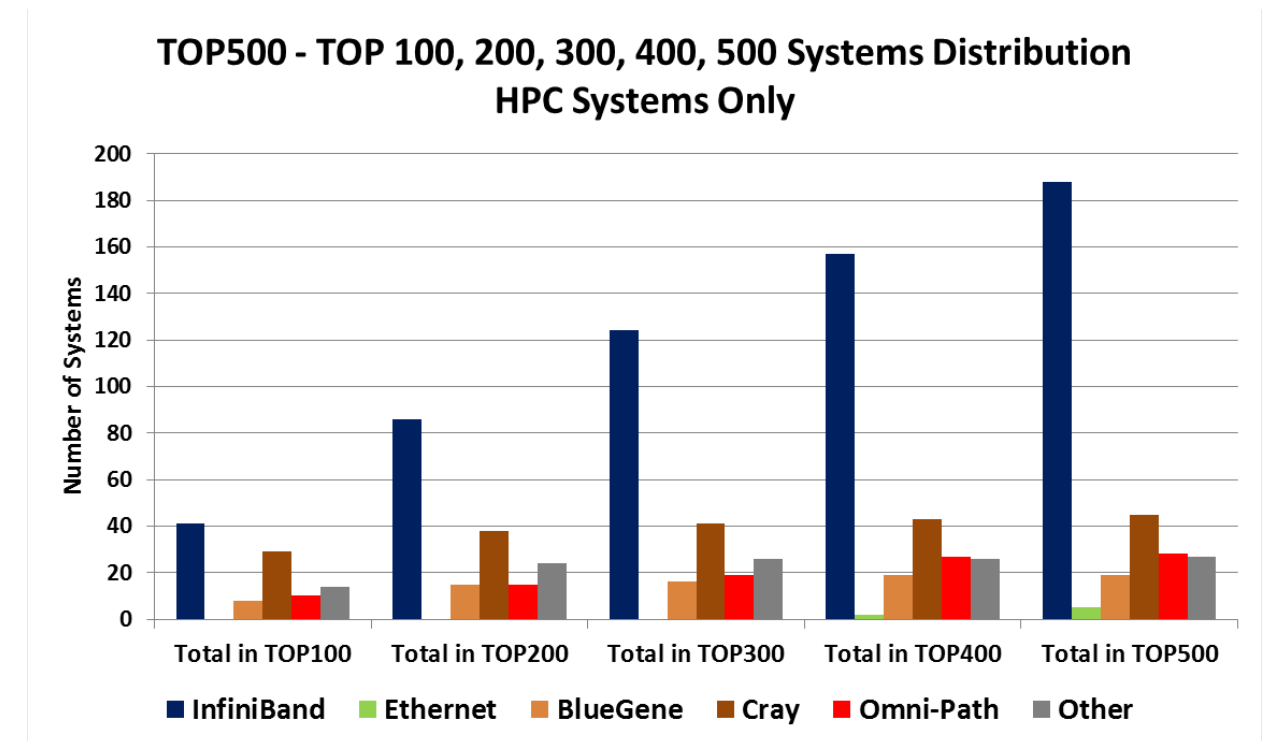
 **Mellanox**  
TECHNOLOGIES  
Connect. Accelerate. Outperform.™



# Mellanox Accelerates The World's Fastest Supercomputers



- Accelerates the #1 Supercomputer
- 39% of Overall TOP500 Systems (194 Systems)
- InfiniBand Connects 65% of the TOP500 HPC Platforms
- InfiniBand Connects 46% of the Total Petascale Systems
- Connects All of 40G Ethernet Systems
- Connects The First 100G Ethernet System on The List (Mellanox End-to-End)
- Chosen for 65 End-User TOP500 HPC Projects in 2016, 3.6X Higher versus Omni-Path, 5X Higher versus Cray Aries



**InfiniBand is the Interconnect of Choice for HPC Infrastructures**  
Enabling Machine Learning, High-Performance, Web 2.0, Cloud, Storage, Big Data Applications

## National Supercomputing Center in Wuxi, China #1 on the TOP500 List



- 93 Petaflop performance, 3X higher versus #2 on the TOP500
- 41K nodes, 10 million cores, 256 cores per CPU
- Mellanox adapter and switch solutions

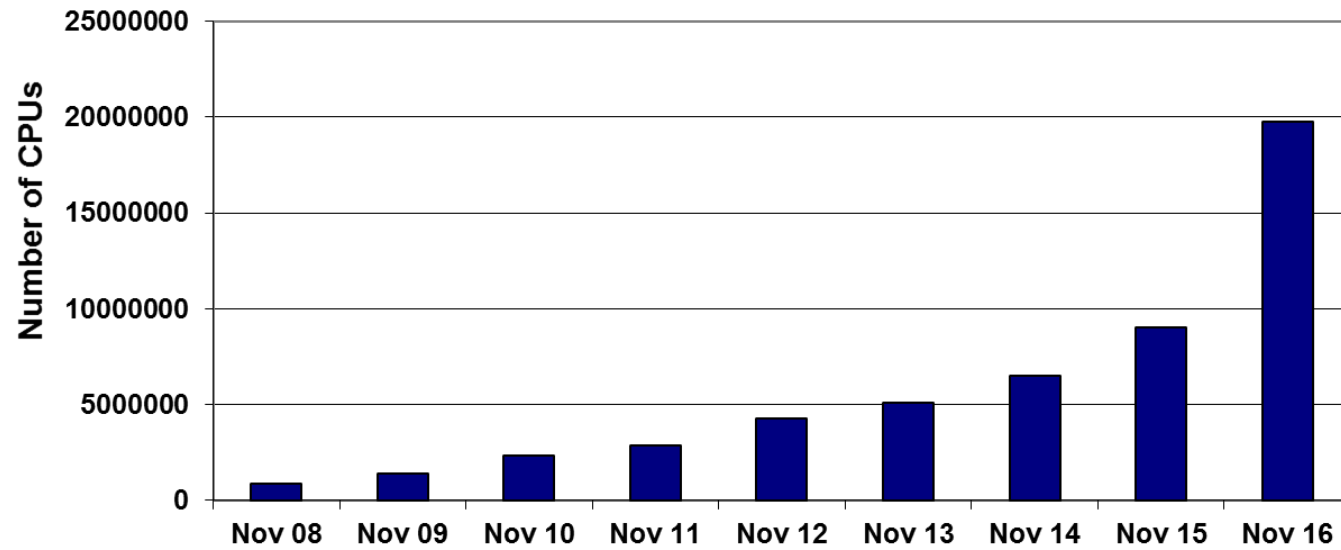


\* Source: "Report on the Sunway TaihuLight System", Jack Dongarra (University of Tennessee), June 20, 2016 (Tech Report UT-EECS-16-742)

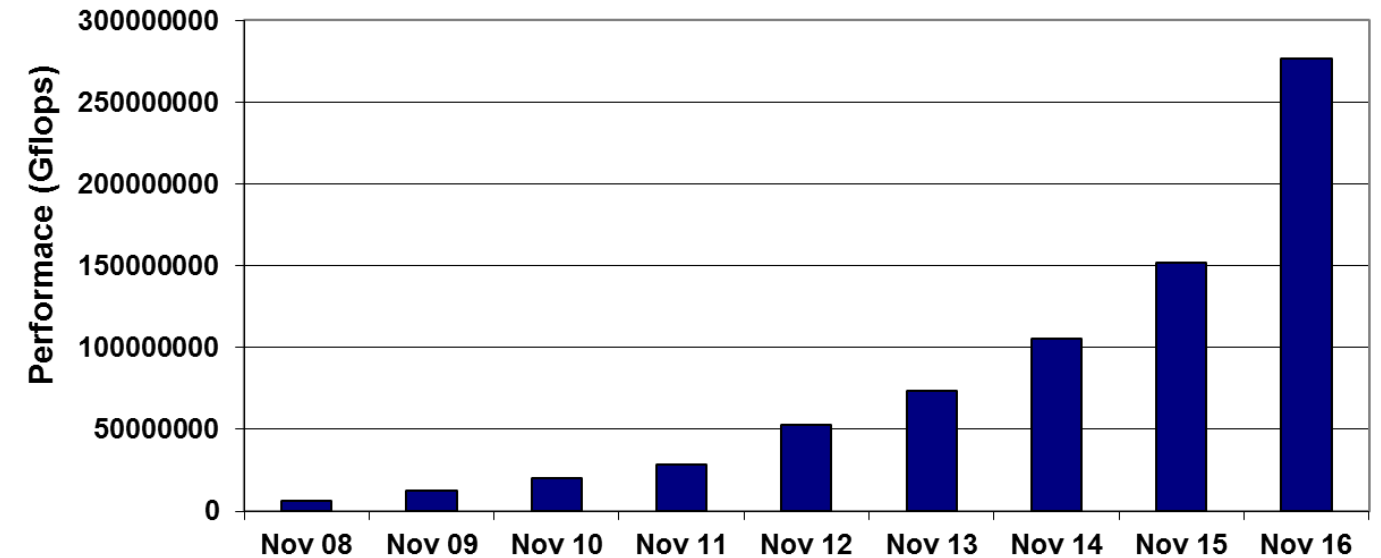
- Connects the world fastest supercomputer, 93 Petaflops, 41 thousand nodes, and more than 10 million CPU cores
- Fastest interconnect solution, 100Gb/s throughput, 200 million messages per second, 0.6usec end-to-end latency
- Broadest adoption in HPC platforms , connects 65% of the HPC platforms, and 39% of the overall TOP500 systems
- Preferred solution for Petascale systems, Connects 46% of the Petascale systems on the TOP500 list
- Connects all the 40G Ethernet systems and the first 100G Ethernet system on the list (Mellanox end-to-end)
- Chosen for 65 end-user TOP500 HPC projects in 2016, 3.6X higher versus Omni-Path, 5X higher versus Cray Aries
- Most Used Interconnect on the TOP500 for TOP100, 200, 300 and TOP400 systems
- Enabling Machine Learning, High-Performance, Web 2.0, Cloud, Storage, Big Data and more applications

**Paving The Road to Exascale Performance**

### InfiniBand Systems - CPU Count



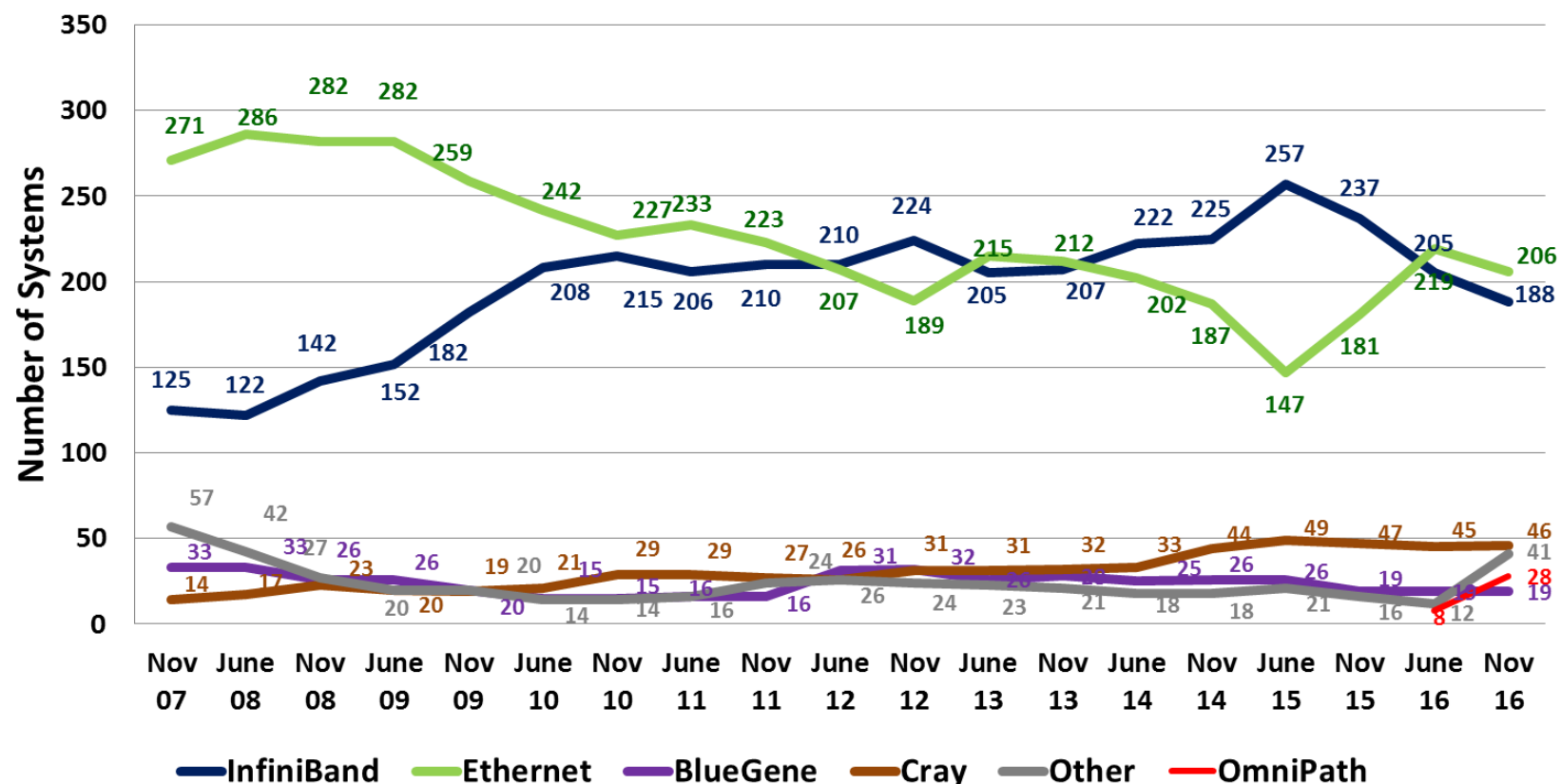
### InfiniBand Systems - Performance



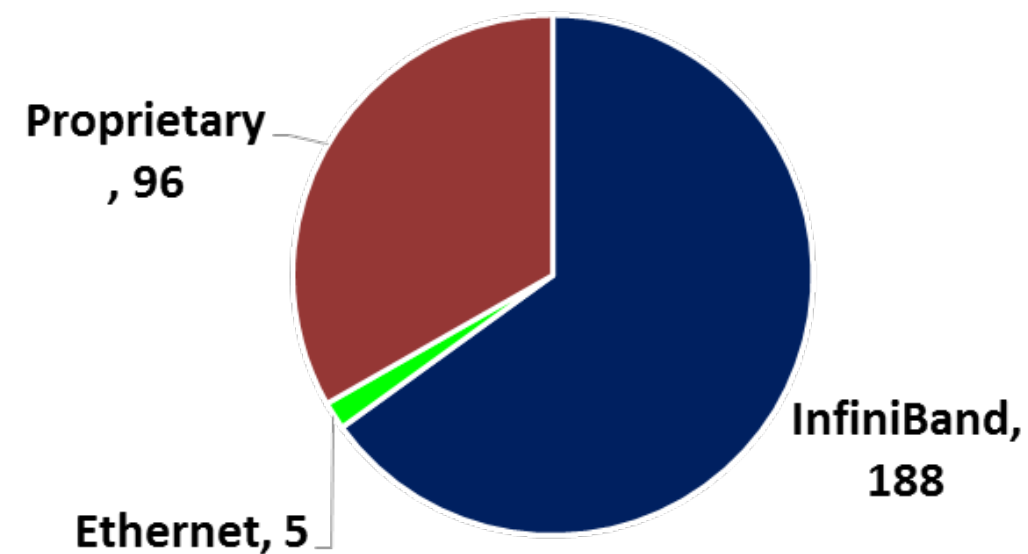
- Explosive market growth

**Mellanox InfiniBand Solutions Provide the Highest Systems Return on Investment**

## TOP500 Interconnect Trends



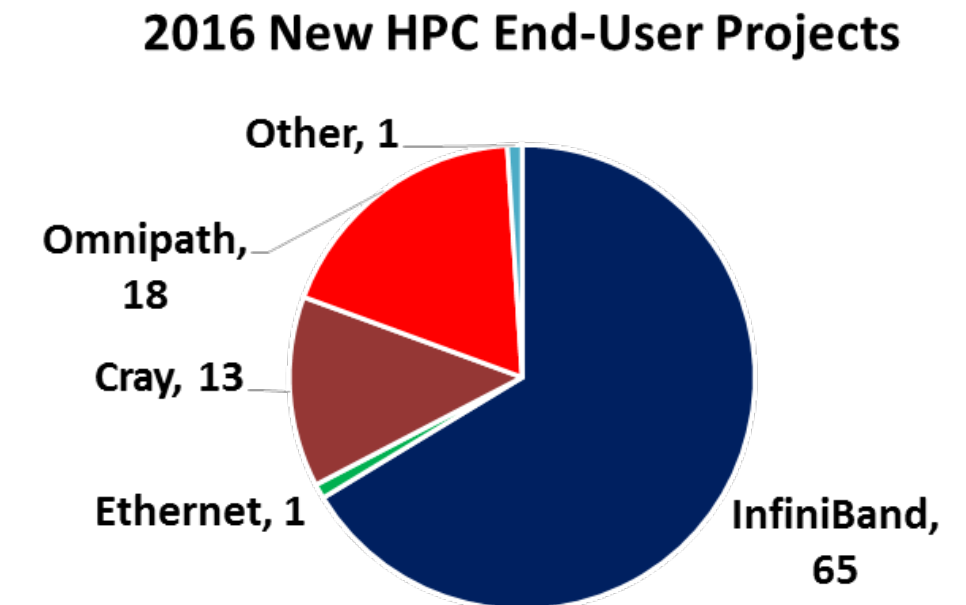
## TOP500 - HPC Systems



The TOP500 List has Evolved to Include Both HPC and Cloud / Web2.0 Hyperscale Platforms. For the HPC Platforms, InfiniBand Continues it's Leadership as the Most Used Interconnect Solution for High-Performance Compute and Storage Infrastructures



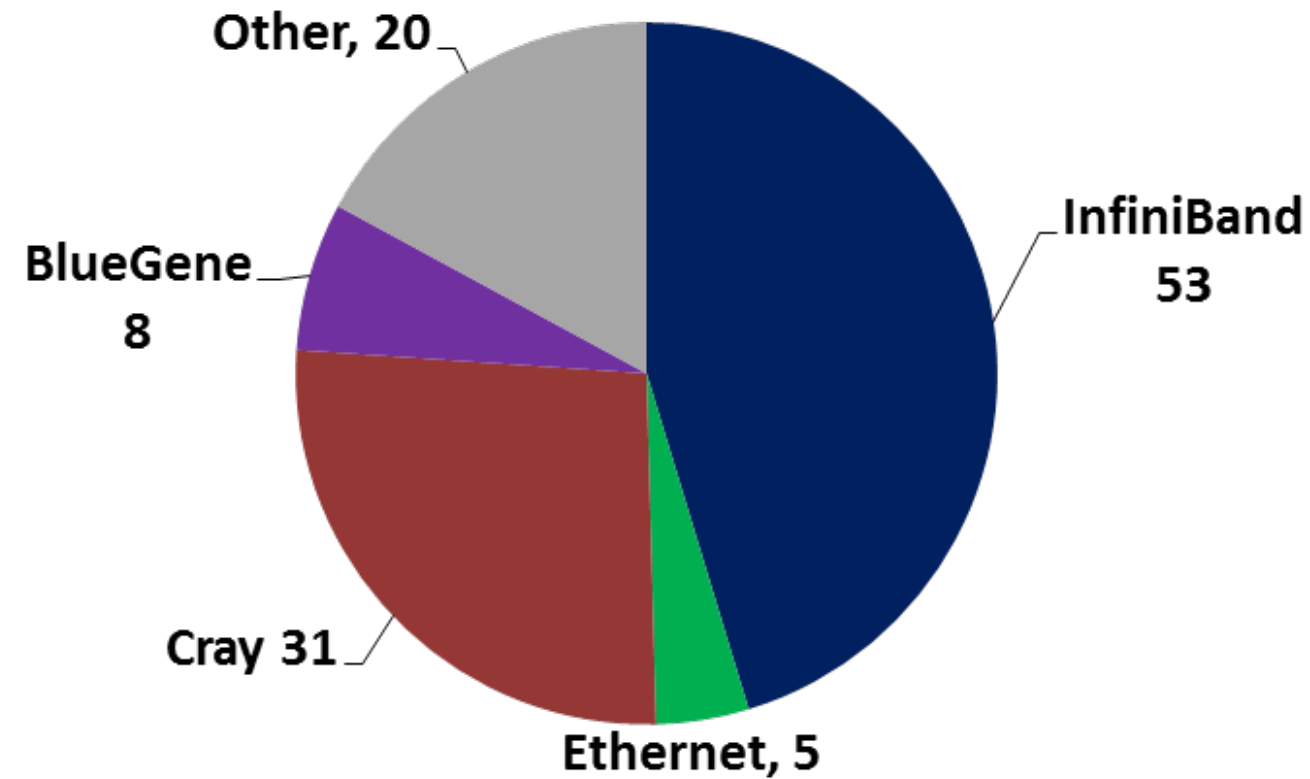
- InfiniBand is The Preferred HPC Interconnect Technology
- Accelerates 3.6X More End-User Projects Versus Omni-Path
- Accelerates 5X More End-User Projects Versus Cray



**InfiniBand is the Interconnect of Choice for HPC Infrastructures**

\* End-User projects exclude internal vendor benchmark systems and combine same project submissions into one

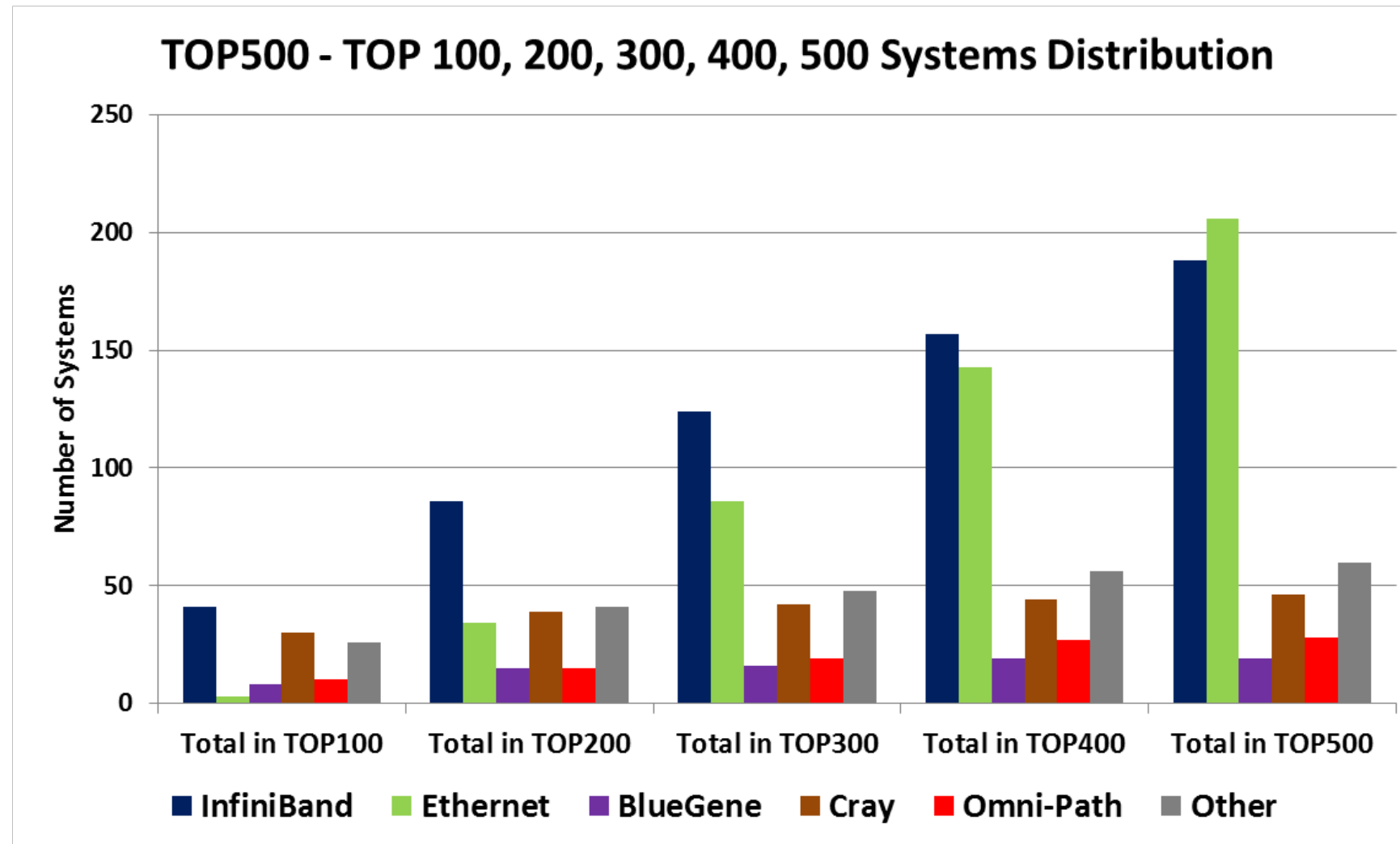
## PetaFlops Systems on the TOP500 list



**InfiniBand is the Interconnect of Choice for Petascale Computing**

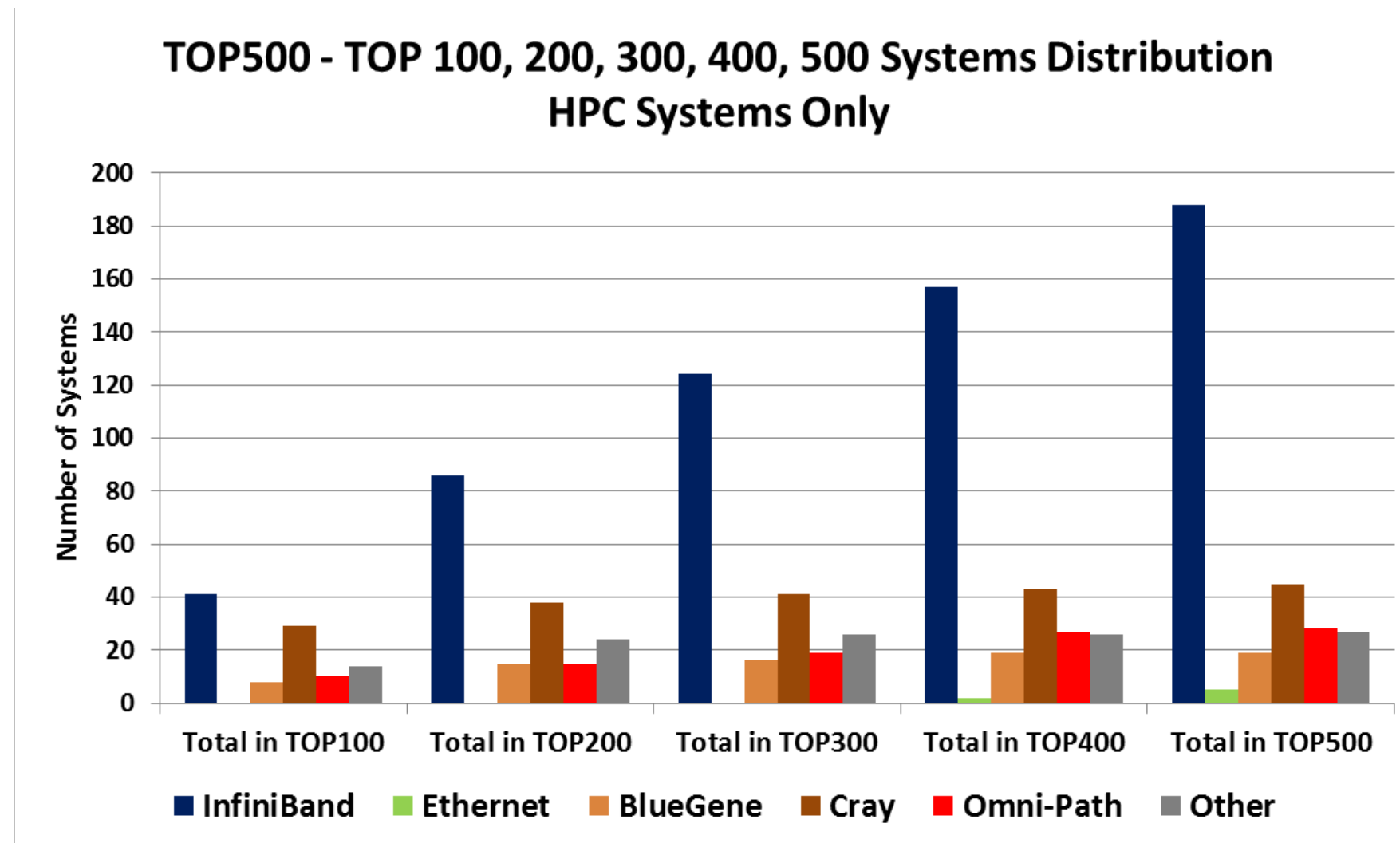
**Accelerates 46% of the Petaflop Systems**





**InfiniBand is The Most Used Interconnect of the TOP100, 200, 300, 400 Supercomputers**  
**Superior Performance, Scalability, Efficiency and Return-On-Investment**

# InfiniBand Solutions – TOP100, 200, 300, 400, 500 HPC Systems Only (Excluding Cloud, Web2.0 etc. Systems)



**InfiniBand is The Most Used Interconnect For HPC Systems**  
**Superior Performance, Scalability, Efficiency and Return-On-Investment**

# Maximum Efficiency and Return on Investment



- Mellanox smart interconnect solutions enable In-Network Computing and CPU-Offloading
- Critical with CPU accelerators and higher scale deployments
- Ensures highest system efficiency and overall return on investment

## InfiniBand

**System: JULICH, System Efficiency: 84%**

**System: Intel, System Efficiency: 81%**

**System: ISM, System Efficiency: 84%**

**53% Higher System Efficiency!**

## Omni-Path

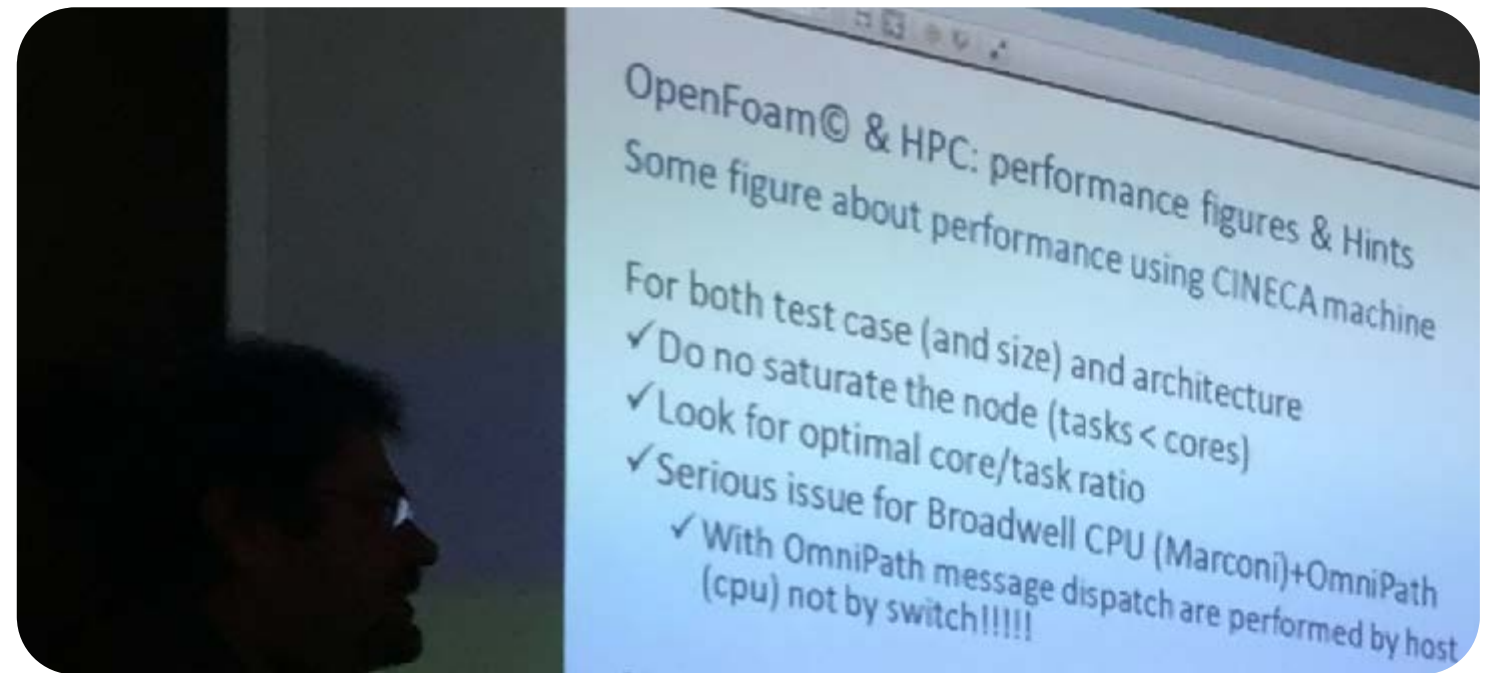
**System: JCAHPC, System Efficiency: 54%**

**System: CINECA, System Efficiency: 57%**

**System: TACC, System Efficiency: 54%**

**45% System Resources not Utilized!**

## User testimonial on Omni-Path at CINECA



- Performance issues with Omni-Path due to Omni-Path non-offload Architecture
- Inability to use all CPU cores, resulted in degraded data center performance
  - CPU cores are needed to handle communications instead of application tasks
- As a result some future end-user projects mandate InfiniBand solutions

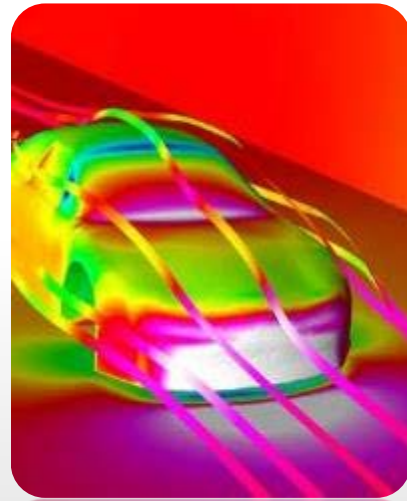
**Mellanox Delivers Highest and Proven Performance and Scalability**



# InfiniBand Delivers Higher Performance Over Competition



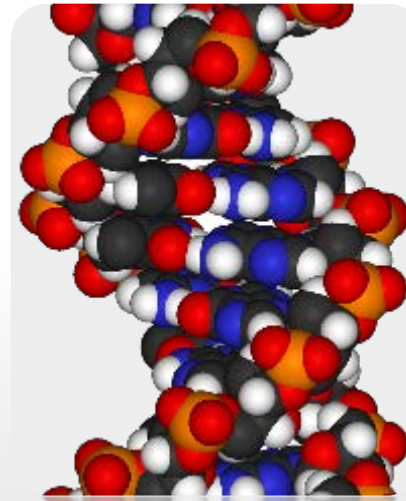
**Weather  
Simulations  
28% Better**



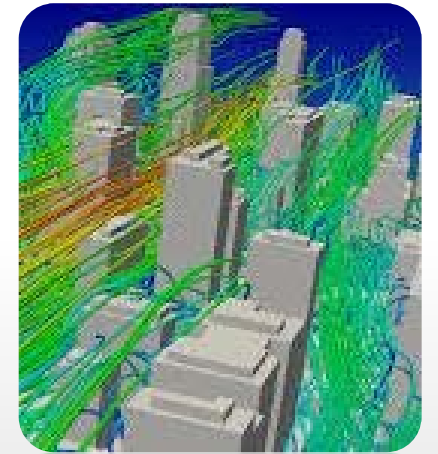
**Automotive  
Simulations  
48% Better**



**Material  
Modeling  
42% Better**



**DNA  
Modeling  
24% Better**



**Computational  
Fluid Dynamics  
17% Better**

- Performance comparisons based on ConnectX-4 / SwitchIB-2
- ConnectX-5 / SwitchIB-2 will increase performance gap by 5-20%

**Mellanox Deliver Highest Data Center Return on Investment**

# High can Accelerate



- **Higher** performance
- **Higher** efficiency
- **Higher** scalability
- **Higher** speed

**World's First  
200G Switch**



**World's First  
200G Adapter**



**40-Ports 200G**  
**80-Ports 100G**

**4.6X Higher Scalability**

**128K Nodes**  
**3-Level Fat Tree**

**15.6 Billion**  
**Messages / Sec**

**2X Higher Throughput**

**200G Data Speed**

**Collective Operations**  
**5-10X Faster**

**In-Network Computing**

**MPI Tag Matching**  
**2X Faster**









**World's First**  
**200G Switch**



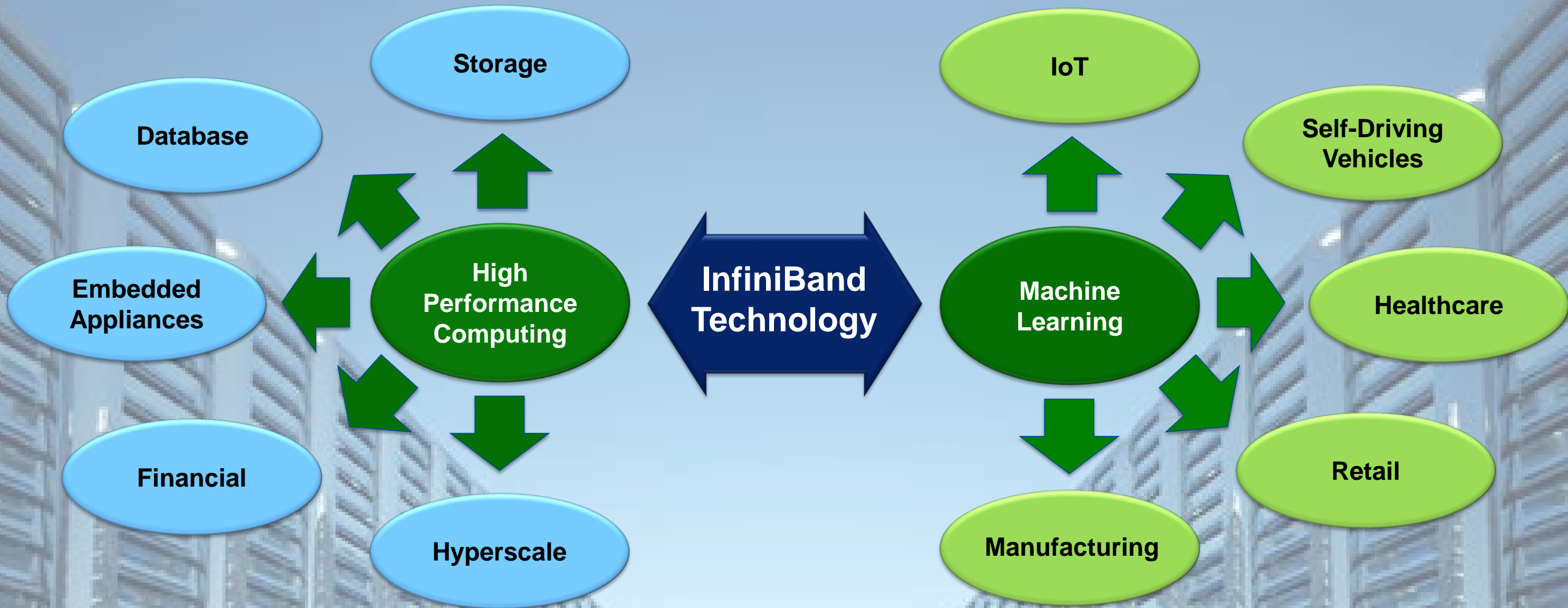
**World's First**  
**200G Adapter**

# Highest-Performance 100Gb/s and 200Gb/s Interconnect Solutions



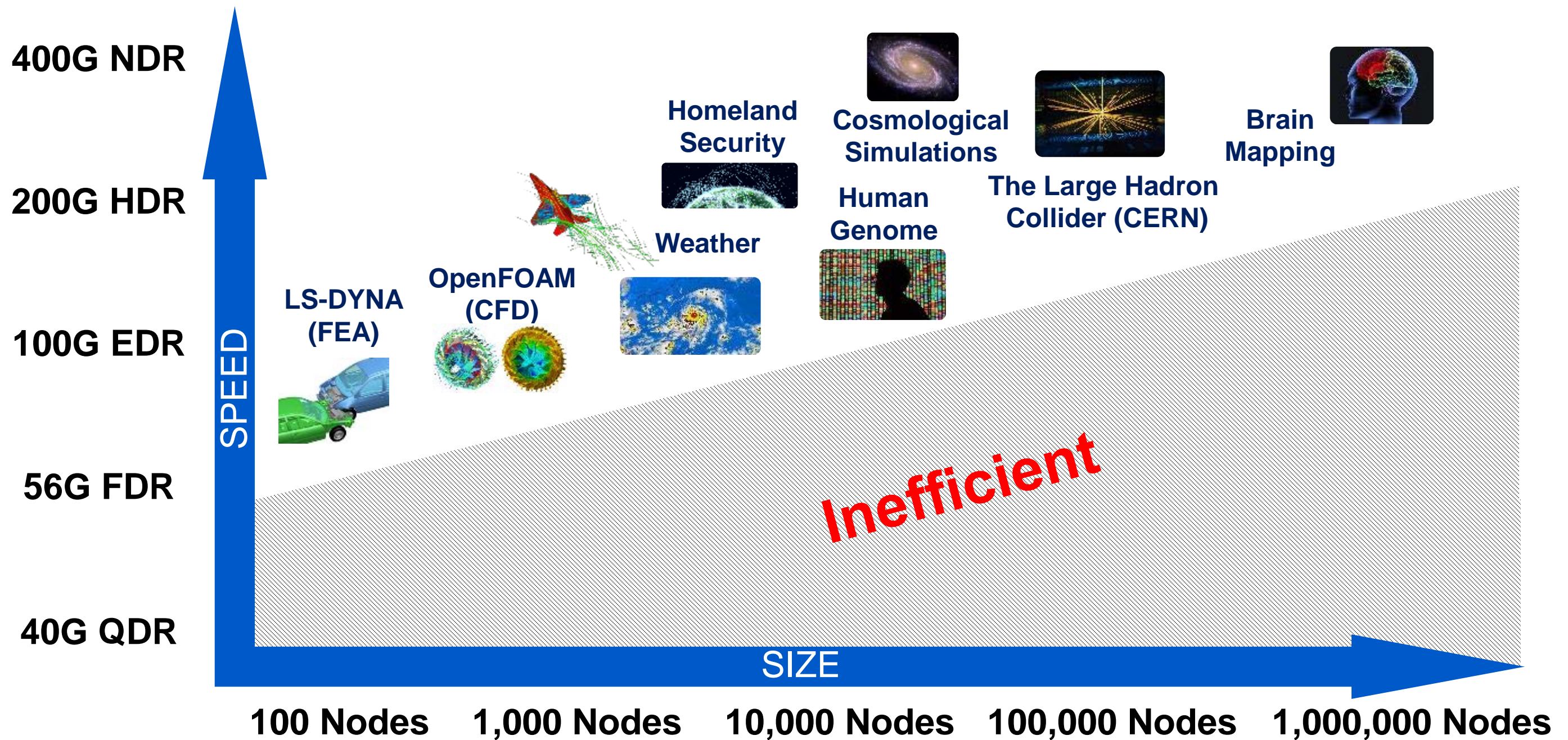
<b>Adapters</b>		<p><b>200Gb/s Adapter, 0.6us latency</b> <b>200 million messages per second</b> <b>(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)</b></p>	
<b>Switch</b>	 <p>The smartest switch, became smarter</p>	<p><b>40 HDR (200Gb/s) InfiniBand Ports</b> <b>80 HDR100 InfiniBand Ports</b> <b>Throughput of 16Tb/s, &lt;90ns Latency</b></p>	
<b>Switch</b>		<p><b>32 100GbE Ports, 64 25/50GbE Ports</b> <b>(10 / 25 / 40 / 50 / 100GbE)</b> <b>Throughput of 6.4Tb/s</b></p>	
<b>Interconnect</b>		<p><b>Transceivers</b> <b>Active Optical and Copper Cables</b> <b>(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)</b></p>	 <p><b>VCSELs, Silicon Photonics and Copper</b></p>





**HPC and Machine Learning Share Same Interconnect Needs**

# Interconnect Technology: The Need for Speed and Intelligence



# Mellanox to Connect Future #1 HPC Systems (Coral)



**OAK RIDGE**  
National Laboratory

**“Summit” System**



**Lawrence Livermore**  
National Laboratory

**“Sierra” System**

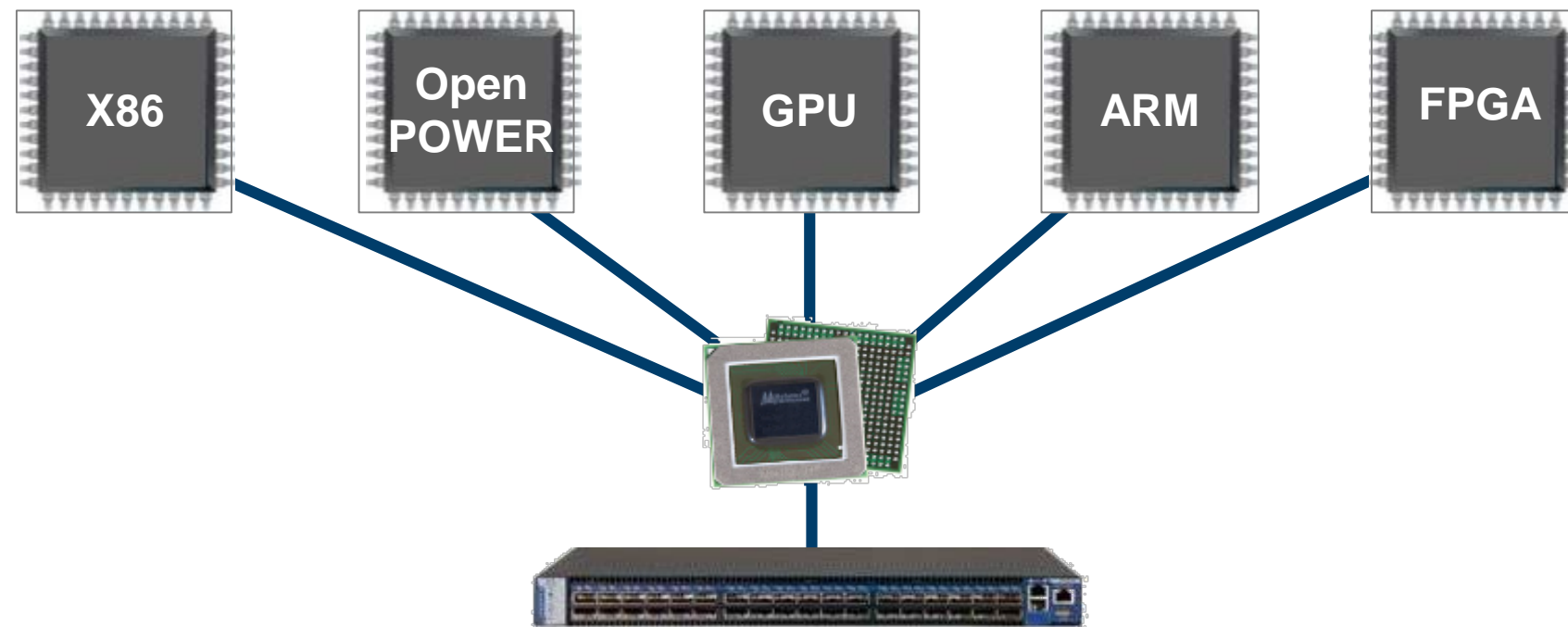


**Paving the Path to Exascale**



## Highest Performance and Scalability for X86, Power, GPU, ARM and FPGA-based Compute and Storage Platforms

10, 20, 25, 40, 50, 56, 100 and 200Gb/s Speeds



**Smart Interconnect to Unleash The Power of All Compute Architectures**



- Scalable, intelligent, flexible, high performance, end-to-end connectivity
- Standards-based (InfiniBand, Ethernet), supported by large eco-system
- Supports all compute architectures: x86, Power, ARM, GPU, FPGA etc.
- Offloading architecture: RDMA, application acceleration engines, etc.
- Flexible topologies: Fat Tree, Mesh, 3D Torus, Dragonfly+, etc.
- Converged I/O: compute, storage, management on single fabric
- Backward and future compatible

**The Future Depends On Smart Interconnect**

# Wuxi Supercomputing Center – World's Fastest Supercomputer



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

**Petaflop**  
Mellanox Connected



## “Pangea” system

- SGI Altix X system, 220K cores
- Mellanox InfiniBand
- 5.3 Petaflops (sustained performance)
- 80% efficiency



**TOTAL**  
EXPLORATION & PRODUCTION

**Petaflop**  
Mellanox Connected



## “Stampede” system

- Mellanox InfiniBand
- 5.2 Petaflops
- 6,000+ Dell nodes
- 462462 cores, Intel Phi co-processors

# TACC

**Petaflop**  
Mellanox Connected





## Pleiades system

- 20K Mellanox InfiniBand nodes
- 4 Petaflops (sustained performance)
- SGI Altix ICE
- Supports variety of scientific and engineering projects
  - Coupled atmosphere-ocean models
  - Future space vehicle design
  - Large-scale dark matter halos and galaxy evolution



**Petaflop**  
Mellanox Connected



## “HPC2” system

- IBM iDataPlex DX360M4
- NVIDIA K20x GPUs
- 3.2 Petaflops
- Mellanox InfiniBand

**Petaflop**  
Mellanox Connected



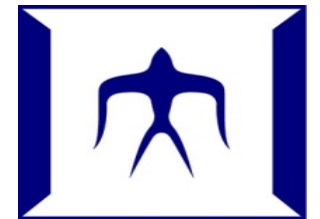
## IBM iDataPlex and Intel Sandy Bridge

- 147456 cores
- Mellanox InfiniBand solutions
- 2.9 sustained Petaflop performance
- The fastest supercomputer in Europe
- 91% efficiency

**Petaflop**  
Mellanox Connected



- TSUBAME 2.0, first Petaflop system in Japan
- 2.8 PF performance
- HP ProLiant SL390s G7 1400 servers
- Mellanox InfiniBand



**Petaflop**  
Mellanox Connected





## “Cascade” system

- Mellanox InfiniBand
- 2.5 sustained Petaflop performance
- Atipa Visione IF442 Blade Server
- Intel Xeon Phi 5110P accelerator



**Pacific Northwest**  
NATIONAL LABORATORY

**Petaflop**  
Mellanox Connected

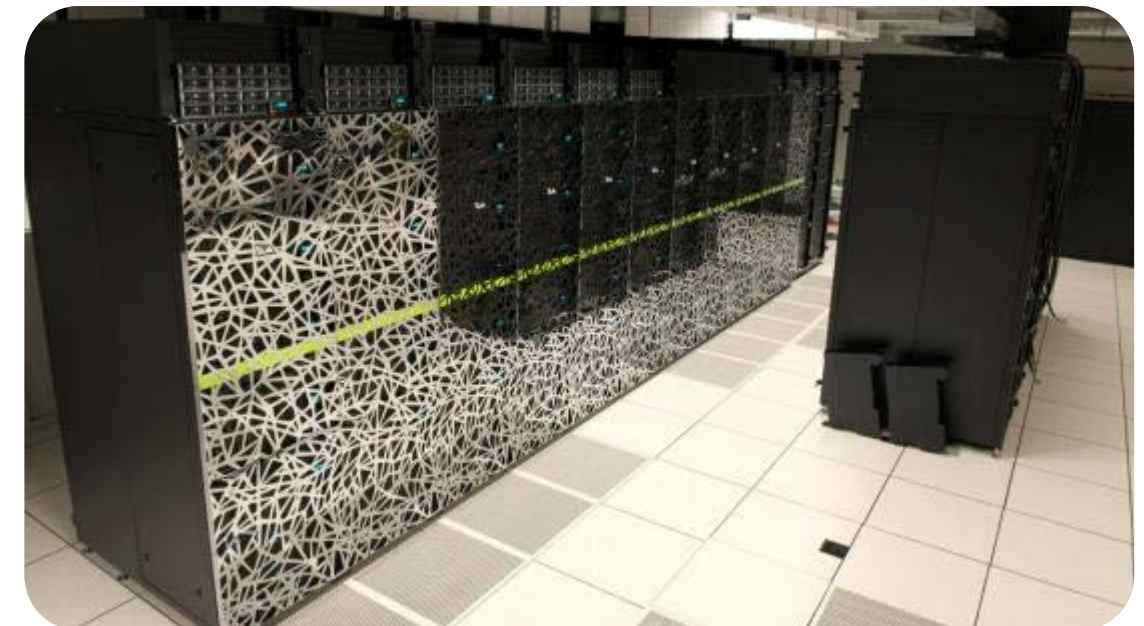


## Occigen system

- Mellanox InfiniBand
- 1.6 sustained Petaflop performance
- Bull bullx DLC



**Petaflop**  
Mellanox Connected



## “Spirit” system

- Mellanox InfiniBand
- 1.4 sustained Petaflop performance
- SGI Altix X system, 74584 cores
- 92.5% efficiency



**Petaflop**  
Mellanox Connected



- Mellanox InfiniBand solutions
- 1.36 sustained Petaflop performance
- Bull Bullx B510
- 77184 cores Intel Sandy Bridge



**Petaflop**  
Mellanox Connected





- Mellanox InfiniBand solutions
- 1.3 sustained Petaflop performance
- IBM iDataPlex DX360M4



**Petaflop**  
Mellanox Connected



## The first Petaflop systems in China

- Mellanox InfiniBand solutions
  - ConnectX-2 and IS5000 switches
- 1.27 sustained Petaflop performance
- Dawning TC3600 Blade Supercomputer
- 5200 nodes, 120640 cores, NVIDIA GPUs



Shenzhen Institutes of Advanced Technology

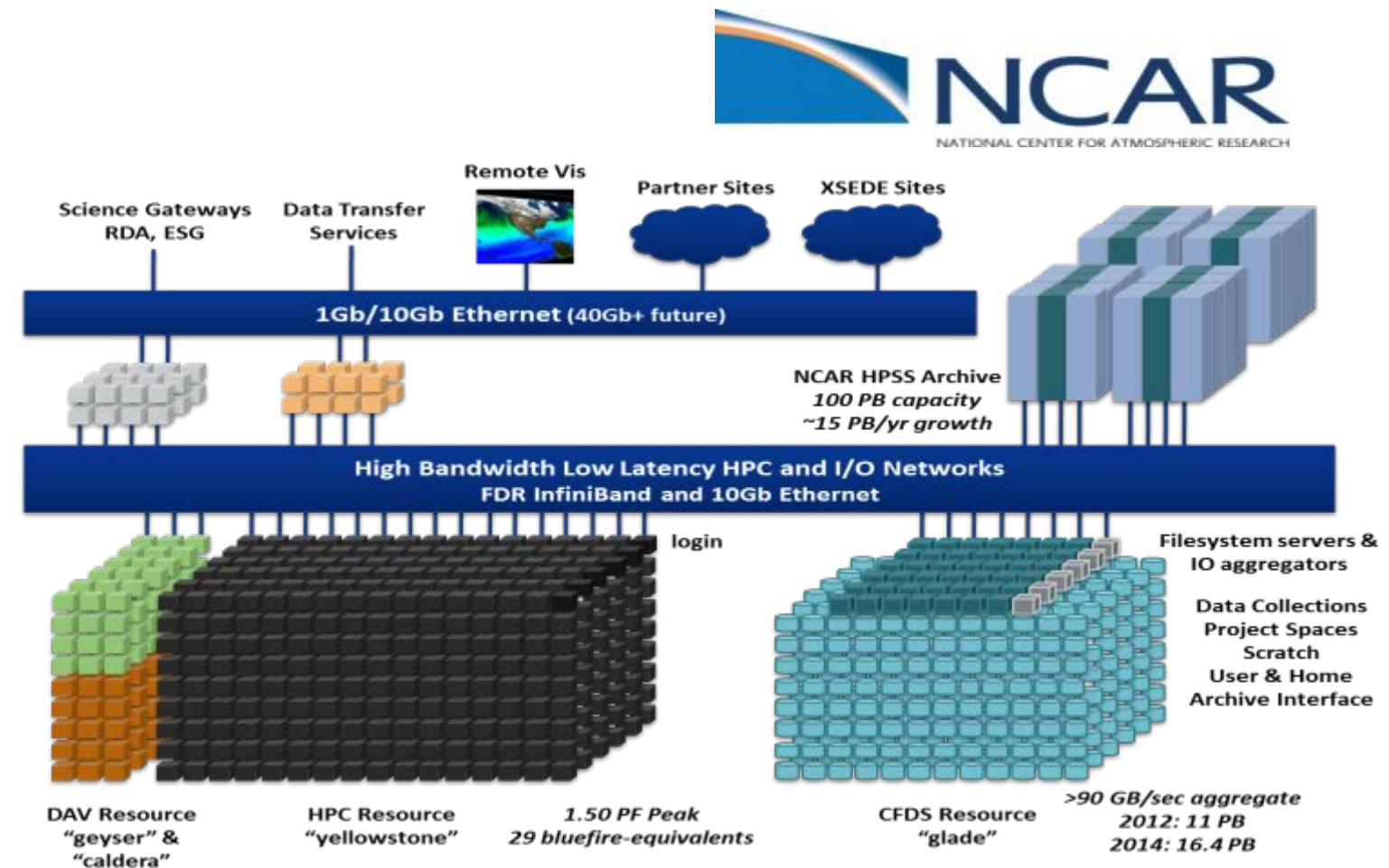


## “Yellowstone” system

- Mellanox InfiniBand, full fat tree, single plane
- 1.3 sustained Petaflop performance
- 72,288 processor cores, 4,518 nodes (IBM)



**Petaflop**  
Mellanox Connected





# International Fusion Energy Research Centre (IFERC), EU(F4E) - Japan Broader Approach collaboration



- Mellanox InfiniBand solutions
- 1.24 sustained Petaflop performance
- Bull Bullx B510
- 70560 cores Intel Sandy Bridge



**Petaflop**  
Mellanox Connected



The “Cartesius” system - the Dutch supercomputer

- Mellanox InfiniBand solutions
- 1.05 sustained Petaflop performance
- Bull Bullx DLC B710/B720 system



Tera 100, first Petaflop system in Europe

- Mellanox InfiniBand
- 1.05 PF performance
- 4,300 Bull S Series servers
- 140,000 Intel® Xeon® 7500 processing cores
- 300TB of central memory, 20PB of storage



**Petaflop**  
Mellanox Connected



## Fujitsu PRIMERGY CX250 S1

- Mellanox InfiniBand
- 980 Tflops performance



THE AUSTRALIAN NATIONAL UNIVERSITY





## “MareNostrum 3” system

- Mellanox InfiniBand
- 1.1 Petaflops peak performance
- ~50K cores, 91% efficiency



**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*





Thank You