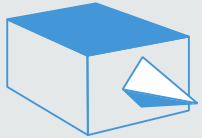# A Status Update on HP's Solutions for HPC

Illustrated with examples from Research and Operational Weather

Dr. Michael F. Lough

High-Performance Computing is everywhere

- Computer-Aided Engineering
- Electronic Design Automation

- Research & Development
- Life Sciences
- Pharmaceutical

- Geophysical Sciences
- Energy Research & Production
- Meteorological Sciences

- Entertainment
- Media Production
- Visualization & Rendering

- Government
- Academia
- Finance

You are here!

Institutes using HP Hardware for Geophysical and Meteorological Research

Accelerate your innovation

Tackle any challenge, with HPC solutions from HP

Overcome barriers to Innovation and Scale

- Realized system performance and throughput
- Power capacity and cost
- Infrastructure complexity and inflexibility

Faster > **Better** > **Smarter** >

Speed advancements with a converged infrastructure, purpose-built for scale.

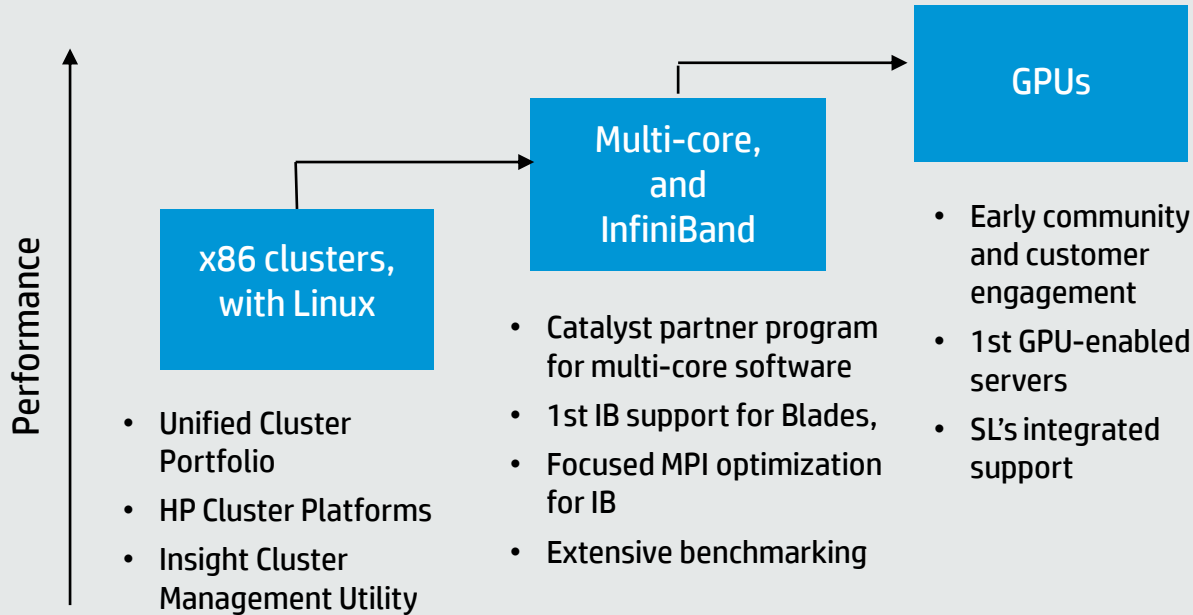Optimize your performance footprint with the world's most efficient systems.

Deploy easily, adapt quickly to change, and improve quality of service.

Affordable performance and technology transforming HPC

HP leading this transition

Performance →

GPUs

Multi-core, and InfiniBand

x86 clusters, with Linux

- Unified Cluster Portfolio
- HP Cluster Platforms
- Insight Cluster Management Utility

- Catalyst partner program for multi-core software
- 1st IB support for Blades,
- Focused MPI optimization for IB
- Extensive benchmarking

- Early community and customer engagement
- 1st GPU-enabled servers
- SL's integrated support

HP's HPC solutions

Mostly clusters (70%+ of market)

Leveraging our Unified Cluster Portfolio and Cluster Platforms

Flexible design allows us to configure to meet specific workload needs

Standalone systems

Large SMP systems, used for large memory applications (e.g., pre-processing)

Workstations

Services

Factory integration, on-site installation and start-up services, outsourcing

Data Center infrastructure including PODs

Emerging – HPC cloud capability

HP Unified Cluster Portfolio

HP Technical and Enterprise Services

HPC application, development
and cloud software portfolio

Advanced and specialty options

Scalable data management

Cluster **management** layer

Operating systems and extensions
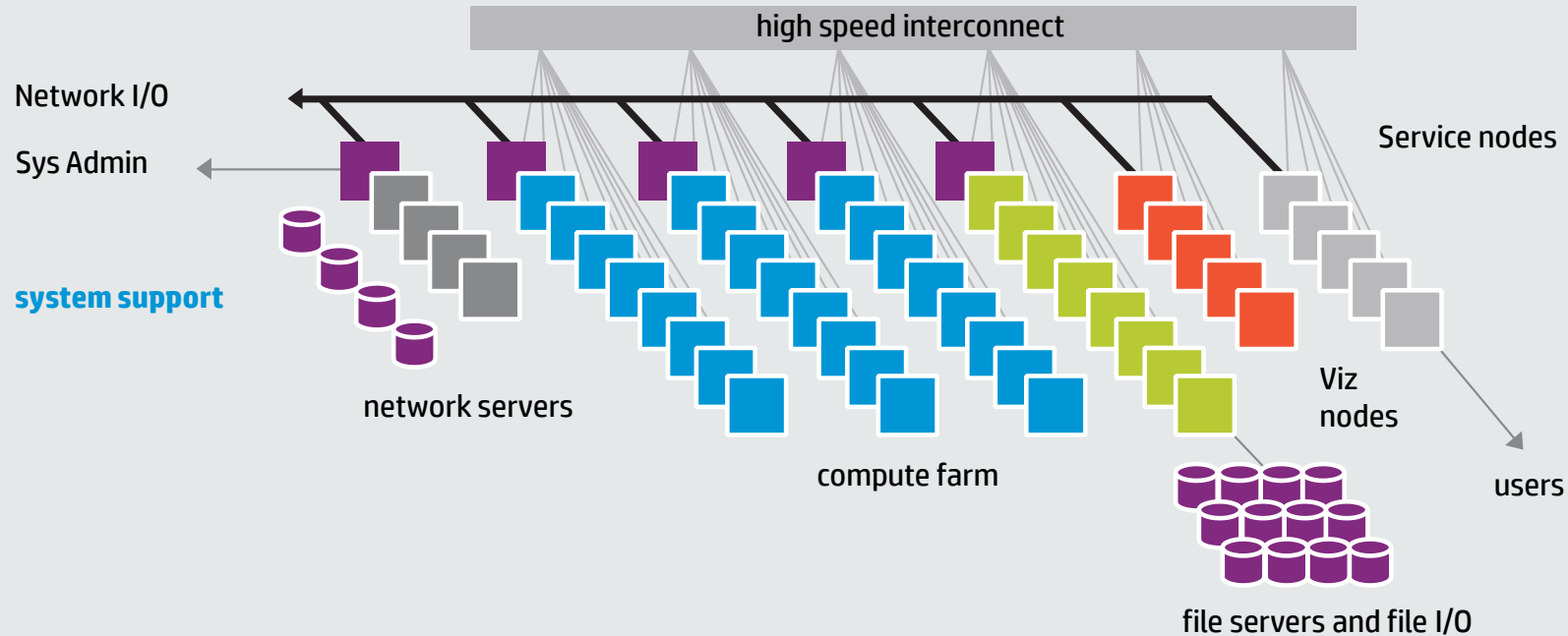
HP Cluster Platforms

HP Datacenter Products and Services

Simplified cluster design and deployment

- Base of HP Cluster Platforms with system nodes and networking, and choice of software
- Storage options for scalable I/O
- Springboard for new technologies
- Reference platform for ISV qualification

A typical HPC cluster architecture

high speed interconnect

Network I/O

Sys Admin

system support

network servers

compute farm

Service nodes

Viz nodes

users

file servers and file I/O

Industry's most complete portfolio for HPC

Workload optimized, engineered for any demand

ProLiant DL Family

Versatile, rack-optimized servers with a balance of efficiency, performance and management

ProLiant BL Family

Cloud-ready converged infrastructure engineered to maximize every hour, watt and dollar

ProLiant SL Family

Purpose built for the world's most extreme data centers

ProLiant BL Blade Servers

Generally speaking, the blade form factor remains popular in HPC

Blades offer some interesting advantages

Manageability and sharing of resources

Coordinated management of all servers in a blade chassis (onboard administrator)

Shared power and cooling

Ability to "mix" different types of server blades in the same enclosure

The HP C7000 blade enclosure provides 8 full-height or 16 half-height bays

Increased density

The HP C7000 with 16 BL2x220c blades comprises 32 servers (2P) in 10 U

Examples of HP Customers using HP ProLiant Blade Servers for Weather (Operational or Research)

Serbian Hydrometeorological Service (activities described in later talk)

Clusters based on C7000 and BL2x220c servers:

2007: C7000 with 8 BL2x220c G5 using Xeon x5450

2010: Added 8 BL2x220c G6 using Xeon x5620

2012: Added C7000 with 16 BL2x220c G7 using Xeon x5645

CSIR Centre for Mathematical Modelling and Computer Simulation (C-MMACS)

360 TF cluster based on C7000 with 1084 BL460c Gen8 servers (Xeon E5-2670 w/ FDR)

"Climate and Environmental Modelling Program" & "Multiscale Modelling and Simulation" are two of C-MMACs main research areas

DL1xx series – 1U servers – the "Pizza Box" Server

Were very popular for HPC, including geophysical and meteorological science sites

In 2008, DL160 G5 clusters were installed at both the Irish Marine Institute (IMI) and the Swedish Meteorological and Hydrological Institute (SMHI)

IMI runs its suite of applications (ROMS) on 70 DL160 G5 servers clustered with IB.

SMHI runs its suite of applications (HIRLAM/HARMONIE) on systems hosted by the National Supercomputer Centre (NSC) at Linköping – the DL160 G5 comprised 140 nodes clustered with IB.

Density considerations usually makes half-width nodes a more attractive choice

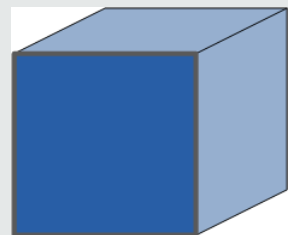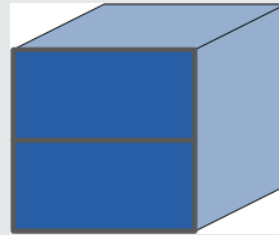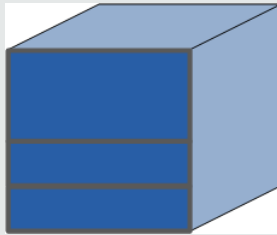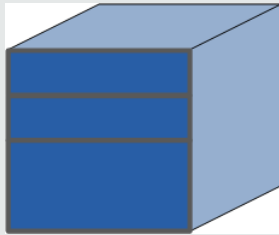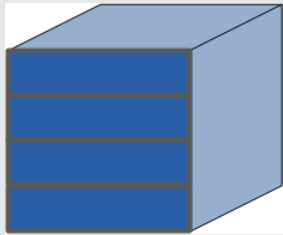SMHI added (2010) 128 additional nodes of DL170h G6 servers (4 server nodes / 2U)

The HP ProLiant SL6500 Scalable System:
A Versatile 4U Rack-mounted Chassis

Shared power and cooling for up to 8 server nodes

Can be configured with 4 full-width 1U servers or <u>each half</u> can be independently configured with half-width servers in up to 5 different ways:

SL6500 with the Previous Generation: SL390 G7

Predecessors of current SL2x0 series of systems were known as SL390

Half-width servers available as 1U, 2U (w/ 3 GPU) or 4U (w/ 8 GPU) variants

TSUBAME 2.0 (November 2010)

2.4 PFLOP/s peak: 1408 SL390 with 2 Intel x5670 and 3 Nvidia M2050 per node

ASUCA (JMA) GPU enabled version – 145 TFLOP/s on 3990 GPU of TSUBAME 2.0

See Masami Narita's presentation from 14th ECMWF workshop on HPC

More details on TSUBAME 2.0 and GPU version of ASUCA from Tokyo Tech site:

http://www.sim.gsic.titech.ac.jp/DL/ESJ/TSUBAME_ESJ_02en.pdf

HP eliminates the barriers to scale

HP ProLiant Gen8 systems driven new levels of performance, efficiency and agility

Next gen performance with the SL6500 Gen8 portfolio

The ProLiant SL6500 Gen8 portfolio, purpose-built for HPC, enables scientific and engineering innovation
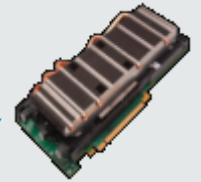
Integrated accelerators boost performance

Family of integrated accelerator offerings enables explosive growth in performance and efficiency

New levels of scalability with FDR InfiniBand

Mellanox 56 Gb/s FDR InfiniBand establishes the basis for new levels of performance and scalability

ProLiant SL6500

NVIDIA Tesla GPUs

ProLiant SL230s Gen8 and SL250s Gen8

Mellanox 56 Gb/s FDR IB

Modular configurations to meet any requirement

Mix and match within the shared SL6500 chassis

Balanced HPC GPU Performance

- Balanced GPU/CPU performance for a broad set of apps

Scaling HPC Performance

- Scalable CPU performance/$/watt/ft2 without porting

New HP ProLiant SL230s & SL250s Gen8

Built on ProActive Insight Architecture

Integrated Lifecycle Automation

- Enhanced performance, quality of service with HP Active Health, Agentless Management

Dynamic Workload Acceleration

- More memory capacity and greater performance
- Increased storage performance with SSD-optimized Smart Array Controllers

Automated Energy Optimization

- Energy optimized technology, with 3D Sea of Sensors, automated power discovery, rack level power management and new 94% Platinum Plus power supplies

Proactive Service and Support

- Quality of service innovations including Smart Socket guides and Smart Drives

Purpose-built design for HPC

- Expanded application accelerator support with optional NVIDIA Tesla GPUs and PCIe IO Accelerators
- Over 30% increased InfiniBand performance with PCI-e gen3 and Mellanox CX3 Flexible LOM

SL230s

SL250s

Examples of HP Customers using HP ProLiant SL230 Gen8 Servers

First Intel SandyBridge with IB FDR system to appear in Top500 was the "Carter" system installed at Purdue University

215 TFLOP/s cluster comprising 648 SL230 Gen8 with E5-2670 and IB FDR

System is available for general research – no specific weather work

SMHI will complete an upgrade to a cluster with SL230 Gen8 nodes in 2012

67 TFLOP/s cluster comprising 240 SL230 Gen 8 with E5-2660 and IB FDR (2:1)

First phase of Krypton system (96 nodes) under test since July

Some configuration considerations

Memory

- GB/core can be important… Many HPC applications use a lot of memory
- It's all about speed… Choose the fastest memory possible (usually)

Interconnects

- Gigabit Ethernet dominates in total volume; cost effective for workloads not latency and bandwidth bound
- 10 GigE emerging but InfiniBand is the top choice for top HPC systems
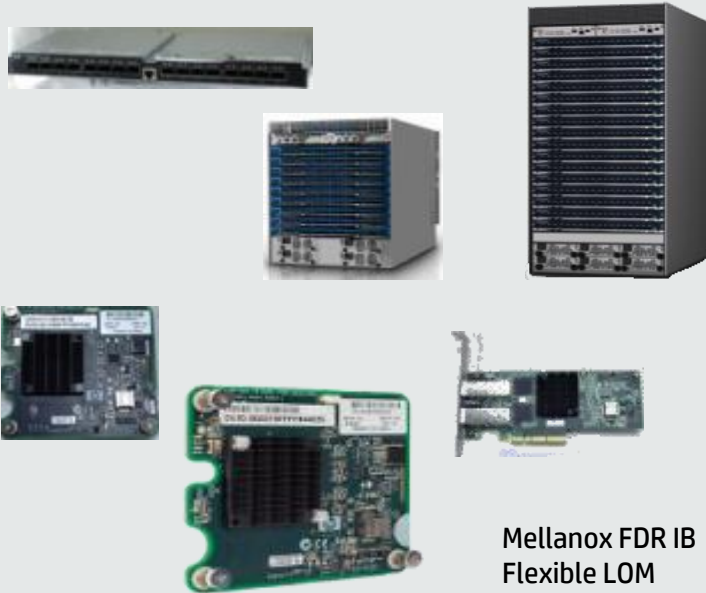
Other networking

- Typically, one admin network and one ilo out-of-band network—10/100 ethernet or GigE
- Head node/file node connection to the LAN/WAN and Storage

New levels of fabric performance

Mellanox 56Gb/s FDR* InfiniBand supported at full speed

Mellanox FDR IB
Flexible LOM

Next generation ConnectX-3 Flexible LOM

- The industry's first FDR 56Gb/s InfiniBand and 10/40 gigabit Ethernet multi-protocol adapter, with PCIe-gen3 for full bandwidth
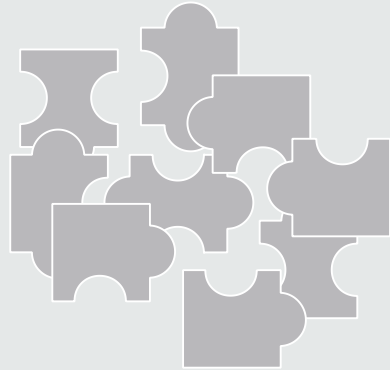
End-to-End FDR solutions

- A complete solution for FDR 56Gb/s InfiniBand consisting of adapter cards, switch systems, software and cables

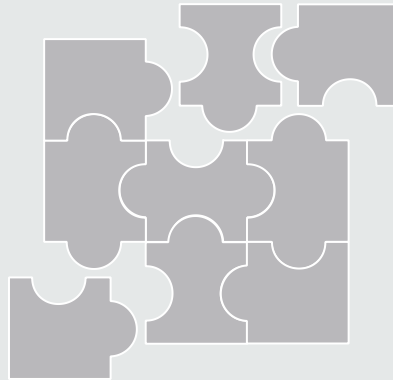## Largest FDR system on the Nov'12 TOP500 list

- "Carter" cluster at Purdue University
- based on 648 SL230s Gen8 servers

* Fourteen Data Rate
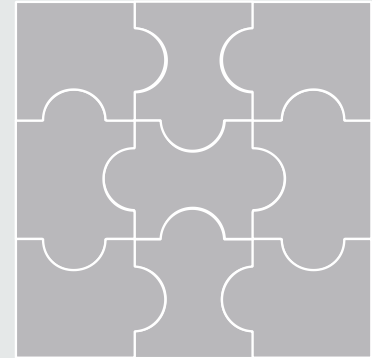
HP Cluster Platforms: Simple and robust

Typical cluster delivery

HP Cluster Platforms delivered

HP Cluster Platforms installed

Lots of choice but:
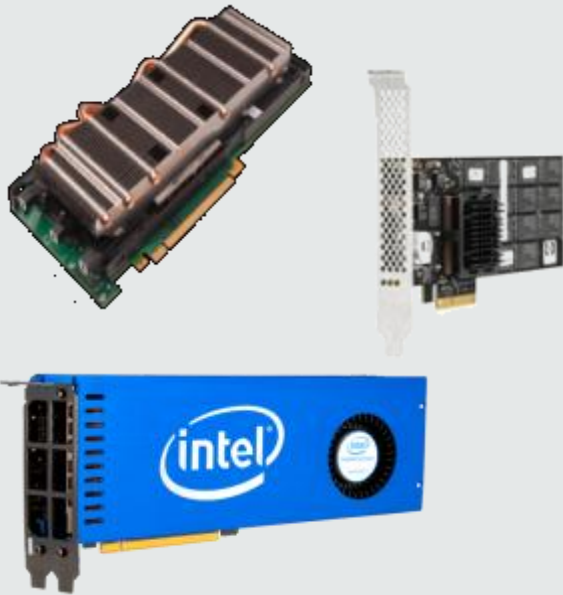
- Work • Time • Risk

Not a reference design

HP = choice without the pain and risk

Integrated accelerator solutions for the SL200s family

Driving new levels of performance/$/watt/ft

Next generation NVIDIA Tesla performance

Up to 30% higher performance with M2090, combined computation and visualization with M2070Q

Optional HP PCIe IO Accelerator

Integrated solid state storage device to accelerate I/O bound applications

Future: Intel® Xeon Phi (MIC)

Accelerate highly parallel applications, using the standard IA instruction set

Future: Nvidia Kepler

Choice of best of breed software to support your needs

Qualified range of options, available installed on HP Clusters

Operating systems: Red Hat, SUSE, or customer-supported community distributions – as well as Windows HPC

Cluster management: HP Insight Cluster Management Utility (CMU), or third party, via HP Software and Licensing Management Solutions (SLMS) or customer installed

MPI: Proprietary third party/open source

Workload manager: Altair PBS Professional (HP SKU), Adaptive Computing Moab (HP user unit SKU), Platform LSF (HP SKU), SLURM

ScaleMP – Virtual SMP software – for large memory, large SMP

HP HPC Linux Value Pack: HP UPC, SHMEM, Platform MPI & UPC

HP Insight Cluster Management Utility (CMU)

Hyperscale cluster lifecycle management software

Provision

- Simplified discovery, firmware audits
- Fast and scalable cloning

Monitor

- 'At a glance' view of entire system; zoom to component
- Customizable
- Lightweight and efficient

Control

- GUI and CLI options
- Easy, friction-less control of remote servers

- 10 years+ in deployment, included Top500 sites with1000's of nodes
- Built for Linux, with support for multiple Linux distributions
- HP supported, available as factory-integrated cluster option

What's new: Insight Cluster Management Utility V7

Improve your system performance and control
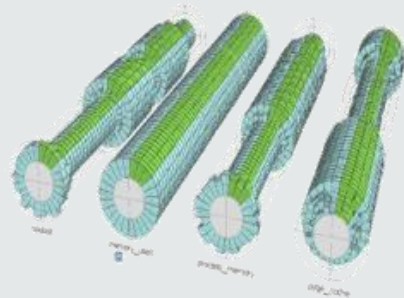
Faster

- Simplified cluster configuration and updates through integration with the new iLO Management Engine
- Increased cluster performance by offloading sensor traffic to the out-of-band network via Agentless Management

Smarter

- Improved RAS through integration with SIM event management
- Unique 3-D history displays for performance analysis

Simpler

- 'At a glance' view of entire system and 'zoom' to component level
- Easy, friction-less remote management of servers

Scalable storage solutions for HPC

X9000 Network Storage System and Fusion File System

- Shared datacenter multipurpose storage
  - Desktops, clusters, farms, clouds
  - Windows, Linux; CIFS/NFS
- Shared datacenter multipurpose storage
- High performance and scale with distributed metadata
- Data tiering
- Will address most HPC needs, including; bioInformatics, FSI, GeoScience apps with many small files

Cluster File System with DataDirect Networks

- Tightly coupled HPC storage
  - For large HPC Linux clusters with large single files/single stream requirements (traditional HPC, such as CAE)
- High parallel bandwidth
- High capacity
- High scalability and reliability
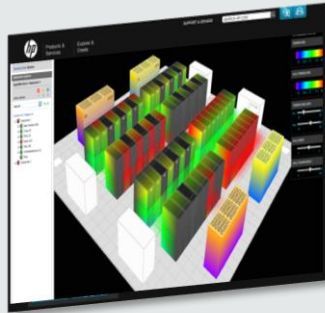- Lustre-open source technology

HP innovations continue to green the data center

Power distribution

Power usage

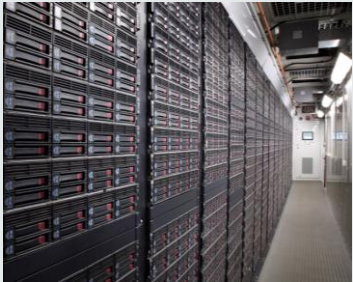**Operating environment**

Extreme low-energy servers

Scale the data center – fast and efficiently

HP PODs providing rapid expansion of capacity, great efficiency, less expensive than brick and mortar

AIRBUS

- Doubled performance
- 40% less power
- Deployed in 4 months
- HP manages this as turnkey HPC datacenter

MIT

- Needed more performance, fast
- Deployed a 20-foot, water-cooled, HP POD 20c, at MIT, to be redeployed near a Hydroelectric dam on the Connecticut River.

Picture courtesy of MIT

Building the "World's Greenest Production Supercomputer"[2]

Tokyo Institute of Technology – Tsubame 2.0

Goals

- Over 1PetaFLOPS of sustained performance
- Fit in 200 meters^2 and 1.8 MW power
- Support the broad research agenda of Tokyo Tech

Solutions

- Gigabit Ethernet dominates in total volume; cost effective for workloads not latency and bandwidth bound
- 10 GigE emerging but InfiniBand the top choice for top HPC systems

Results

- Typically, one admin network and one ilo out-of-band network—10/100 ethernet or GigE
- Head node/file node connection to the LAN/WAN and Storage

Best HPC Collaboration Between Government and Industry

HPC wire

Readers' Choice Awards 2011

#5 on TOP500[1]

#2 on Green500[2]

[1] www.top500.org, Nov'11
[2] www.green500.org, Nov'10

Expert design, deployment, and support

Designed-in expertise; dedicated design and support services

ISV Engineering

- Dedicated ISV Engineering Team
- Qualification and characterization lab

Integration and deployment

- Integration Center & Factory Express services
- Customized configuration & testing
- Onsite installation & cabling

HPC Consulting

- Cluster Startup services to implement HPC software
- HPC training/knowledge transfer
- Regional competency centers

HP Financial Services

- Leasing, Asset Recovery, Refresh
- Selective "Shared Risk" Instruments Available for Service Providers

Managed Services

- Facility & technology assessment and design services with Critical Facility Services
- Managed HPC and outsourcing services

Datacenter Care

- Support calls handled by HPC experts
- Flexible reactive support and proactive services

Accelerate innovation with HP

HP delivers high-performance innovation at any scale.

HP
Converged
Infrastructure

Faster

Speed advancements with a converged infrastructure, purpose-built for scale.

Better

Optimize your performance footprint with the world's most efficient systems.

Smarter

Deploy easily, adapt quickly to change, and improve quality of service.

Thank you

For more information, visit www.hp.com/go/hpc