

SGI®

Many core computing Solutions

Gábor Lehoczki
Silicon Computers Kft.
leho@silicon.hu



Nvidia Tesla series

Features	M2075	M2090	K10	K20	K20X
Number and Type of GPU	1 Fermi GPU	1 Fermi GPU	2 Kepler GK104s	1 Kepler GK110	
Peak double precision floating point performance	515 Gigaflops	665 Gigaflops	190 Gigaflops (95 Gflops per GPU)	1.17 Tflops	1.31 Tflops
Peak single precision floating point performance	1030 Gigaflops	1331 Gigaflops	4577 Gigaflops (2288 Gflops per GPU)	3.52 Tflops	3.95 Tflops
Memory bandwidth (ECC off)	150 GBytes/sec	177 GBytes/sec	320 GB/sec (160 GB/sec per GPU)	208 GB/sec	250 GB/sec
Memory size (GDDR5)	6 GB	6 GB	8 GB (4 GB per GPU)	5 GB	6 GB
CUDA cores	448	512	3072 (1536 per GPU)	2496	2688

Intel® Xeon Phi™ Coprocessor

	<u>5110P</u>	<u>3120A</u>	<u>3120P</u>	<u>7120P</u>	<u>7120X</u>	<u>5120D</u>
Code Name	<u>Knights Corner</u>					
# of Cores	60	57	57	61	61	60
Clock Speed	1.053 GHz	1.1 GHz	1.1 GHz	1.238 GHz	1.238 GHz	1.053 GHz
Peak double precision floating point performance						1.011 TFlops
Peak single precision floating point performance						2.022 TFlops
Cache	30 MB	28.5 MB	28.5 MB	30.5 MB	30.5 MB	30 MB
Max TDP	225 W	300 W	300 W	300 W	300 W	245 W
Max Memory Size	8 GB (16 channels)	6 GB	6 GB	16 GB	16 GB	8 GB
Max Memory Bandwidth	320 GB/s	5 GB/s	5 GB/s	5.5 GB/s	5.5 GB/s	5.5 GB/s

SGI Development Suite

- For Linux software development
- Supports 2-4 developers



- Technical Computing Performance for SGI Systems
- Consists of SGI Accelerate, SGI MPI, SGI REACT, SGI UPC
- Tech support via SGI Support contract



- Software development tool suite tuned for application performance and code robustness
- Consists of C++ & Fortran compilers, Math Kernel Library, Integrated Performance Primitives, Threading Building Blocks
- Includes 1 yr tech support



- Dynamic source code, thread, and memory debugging for C, C++ and Fortran HPC applications
- Consists of TotalView, MemoryScope, ReplayEngine, and CUDA debugging
- Includes 1 yr tech support

Intel® Xeon Phi™ compared to a GPU

- Intel® Xeon Phi™ is a multi-core architecture with the following characteristics :
 - based on Pentium 4 in-order cores with a vector extension
 - OpenMP programming
 - “UV on a chip”
 - PCIe x16 Gen 2 (~ 6.7 GB/s)
 - Ease-of-Use
- NVIDIA Fermi GPU use streaming multi-processors with currently 32 SIMD engines
 - A very high number of threads (10s of thousands) hides latency to memory
 - Minimal context switching time between threads
 - HW thread dispatcher
 - PCIe x16 Gen (~ 11.3 GB/s) [Kepler 10]

Software Development Ecosystem for Intel Xeon Phi

Open Source

Commercial

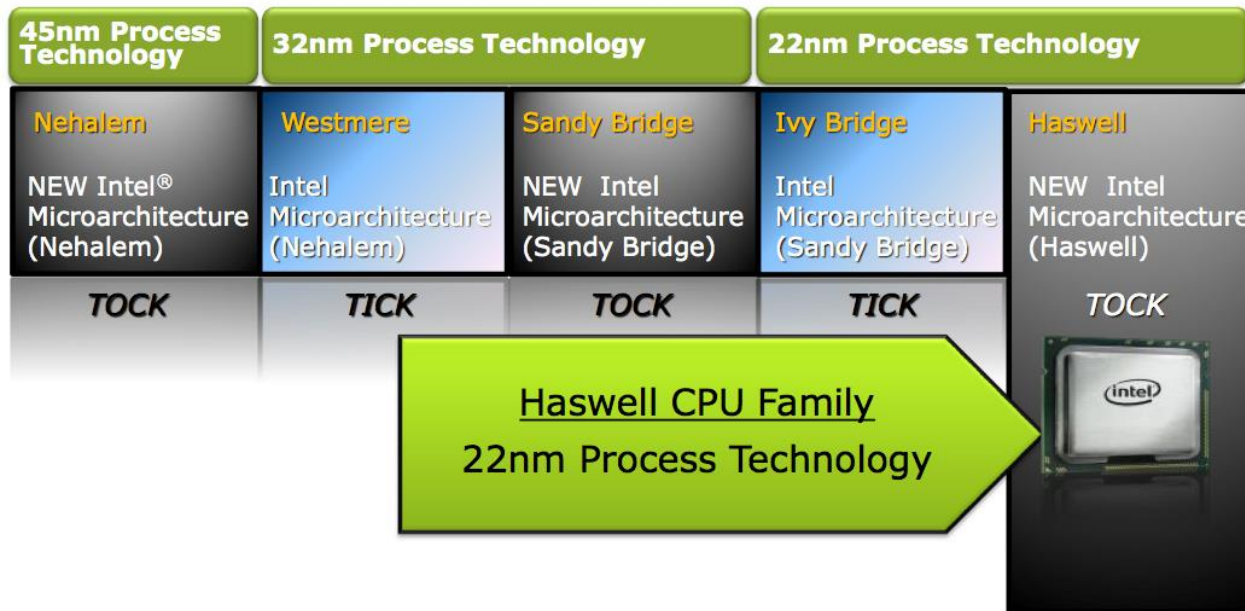
	Open Source	Commercial
Compiler	gcc (kernel build only, not for applications), Python	Intel Parallel Studio XE (C++ & Fortran) Intel Cluster Studio XE (C++ & Fortran) CAPS HMPP compiler (beta)
Debugger	gdb	Intel Debugger Rogue Wave TotalView (beta) Allinea DDT
Libraries	TBB (in Intel Studio XE) MPICH2 FFTW NetCDF	Intel MKL, Intel MPI, OpenMP, Intel IPP, Cilk™ Plus (in Intel Studio XE products), NAG Rogue Wave IMSL
Profiling & Analysis Tools		Intel VTune Amplifier XE Intel Trace Analyzer & Collector Intel Inspector XE
Workload Scheduler		Altair PBS Professional, Adaptive Computing Moab
System Management		SGI Management Center Bright Cluster Manager (beta)

For more information:

<http://software.intel.com/en-us/articles/intel-and-third-party-tools-and-libraries-available-with-support-for-intelr-xeon-phi>

Next generation XEON: Haswell

Tick/Tock Development Model



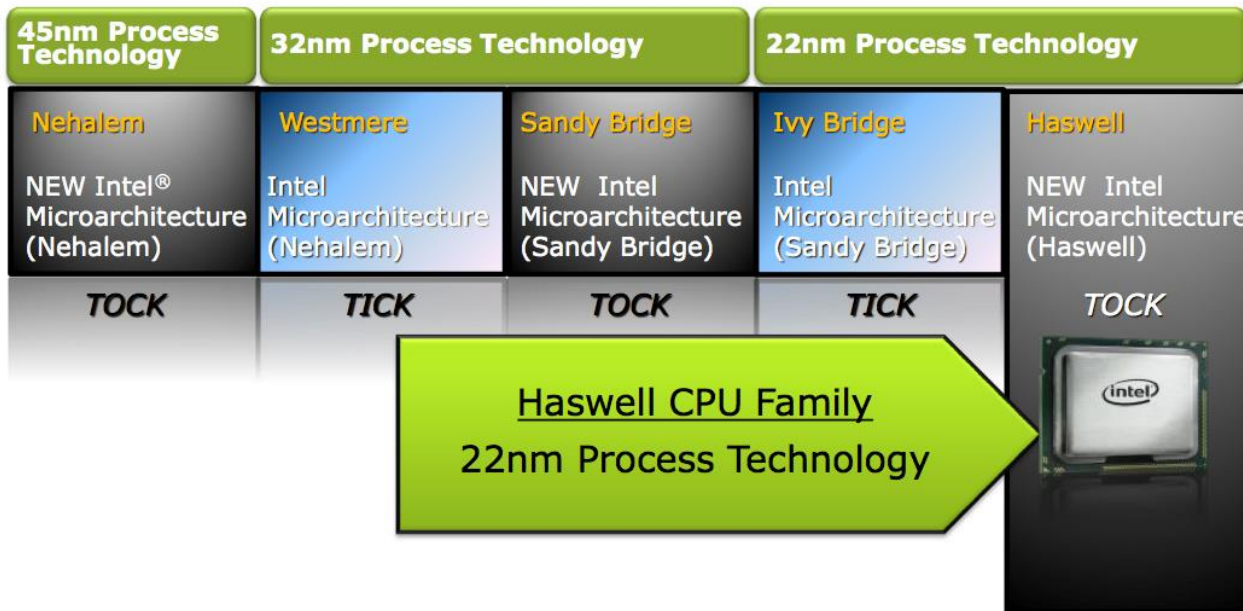
*Builds upon innovations in the
2nd and 3rd Generation Intel® Core™ i3/i5/i7 Processors*

IDF2012
INTEL DEVELOPER FORUM

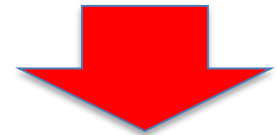
4

Next-next generation HPC XEON

Tick/Tock Development Model



PHI



Builds upon innovations in the 2nd and 3rd Generation Intel® Core™ i3/i5/i7 Processors

IDF2012
INTEL DEVELOPER FORUM



“Pyramid 1U”



NVIDIA GPU accelerator support:

- 3x K10, K20, K20X

NVIDIA graphics GPU support:

- Quadro Plex 7000

Intel accelerator support:

- 3x Phi

Model	C1104G-RP5
Chassis Profile	1U standard-depth
Servers/System	One dual-socket
Chipset	Intel® C600
Max. Processors	Two Intel® Xeon® processor E5-2600 series
Max. CPU TDP	115W
Memory Slots	8 DIMM slots
Memory Type	1600/1333/1066/800 MHz DDR3 ECC Reg
Max. Hard Disk Drives	4 x 2.5" drives
Expansion Slot	Three PCIe 3.0 x16 double-width slots One external PCIe 3.0 x8 low-profile slot
Networking (Onboard)	Dual-Port GigE controller (Intel® I350)
IPMI Remote Management	Integrated IPMI 2.0
Power Supply	1800W Redundant* Platinum Level

“Pyramid 2U”



NVIDIA GPU accelerator support:

- 4x K10, K20, K20X

NVIDIA graphics GPU support:

- Quadro Plex 7000

Intel accelerator support:

- 4x Phi

Model	C2110G-RP5
Chassis Profile	2U standard-depth
Servers/System	One dual-socket
Chipset	Intel® C600
Max. Processors	Two Intel® Xeon® processor E5-2600 series
Max. CPU TDP	130W
Memory Slots	8 DIMM slots
Memory Type	1600/1333/1066/800 MHz DDR3 ECC Reg
Max. Hard Disk Drives	10 x 2.5" drives
Expansion Slot	Four internal PCIe 3.0 x16 double-width (optional riser card for 2 additional external PCIe 3.0 x16 double width slots), One external PCIe 3.0 x8 low profile and one PCIe 2.0 x4 full height
Networking (Onboard)	Dual-Port GigE controller (Intel® I350)
IPMI Remote Management	Integrated IPMI 2.0
Power Supply	1800W Redundant** Platinum Level

SGI UV 2000

Within a Single Standard 19" Rack



Up to:

- 64 socket/512 cores/1024 threads
 - Or **34 CPU + 30 Intel[®] Xeon[®] Phi[™]**
 - Or **36 CPU + 28 NVIDIA[®] Tesla[®] K20, K20X, or K40 (2 partitions)**
- 16TB memory
- 63 x16 PCIe Gen 3 Links

SGI UV and Scale-out Compared

Feature	Standard Scale-out Servers	Sgi UV
Architecture Reference Terms	Scale Out Cluster Distributed Memory	Scale Up Single System Image Shared Memory
System Limit	16 cores, 0.5 TB memory	2048 cores, 64 TB memory
CPU	x86 Intel® Xeon® or AMD Opteron™	X86 Intel® Xeon®
Memory, Storage, networking	Industry standard	Industry Standard
Interconnect or System Fabric	Ethernet or InfiniBand	Sgi NUMalink
Hardware Package or Building Block	Blades or Rackmount, 19" rack	Blades (UV2000) or Rackmount(UV20), 19" rack
Software	Off the Shelf	Off the Shelf
Applications	Small scale single apps cluster or "MPI" apps	Small or large single apps Cluster or "MPI" apps also
Cost	Lowest cost x86 architecture	20-75% > scale-out, ~1/3 the cost of 'Big Iron'.

SGI® Value Propositions

- Long history of working with accelerators
 - “Home-brewed” – Geometry Engine, TPU (Tensor Processor Unit)
 - FPGA’s (RASC™ technology), GPUs, SOCs
 - 50 application experts focused on it
- Accelerators in both a scale-up and scale-out environment
 - 32 Intel® Xeon® Phi™ coprocessors in SGI® UV™ 2000 (COSMOS and TGAC)
 - SGI: the only vendor able to deliver hybrid scale-out and scale-up solutions
- Everything you need in a powerful solution
 - Factory-integrated, tested, rack level delivery – plug in and go
 - Starter kits
 - Worldwide customer support
- Fully-managed, with SGI Management Center and Performance Suite software

SGI® Integrated Clusters for Intel Xeon Phi

❑ A complete, managed GPU solution of software and hardware, all you need for a powerful deployment

- Hardware stack

- Rackable™ C2108 head node
- Either/or Rackable C1104G or C2110G compute nodes
- Infiniband NICs and Switch
- Ethernet switch for management
- Add SGI InfiniteStorage solutions, either direct-attached to the head NAS

- ❑ Software stack

- SGI Management Center software
- Altair PBSpro Load Management software

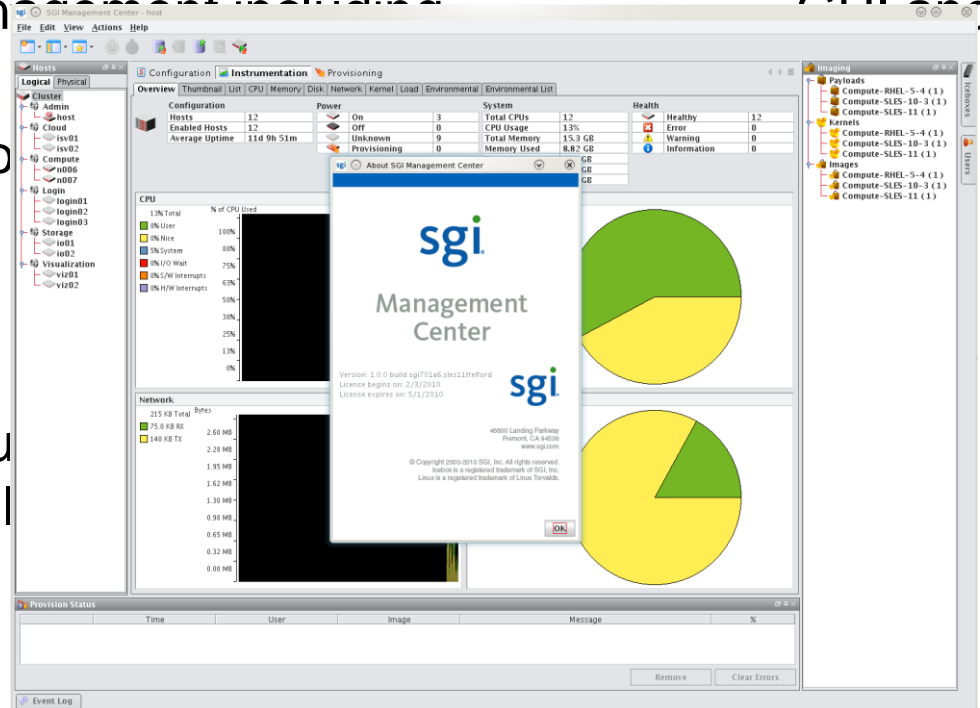
- ❑ Complete Factory Test & Integration



SGI® Management Center

Premium Edition offers Accelerator Monitoring and Management!

- Single system management console with remote server monitoring and control
- Ease of use, full system management including CLI
- Policy driven, fine grained power computing
- Advanced fault, event, improved reliability
- Advanced capabilities include management, and high avail



SMC Premium Edition– Accelerator Monitoring

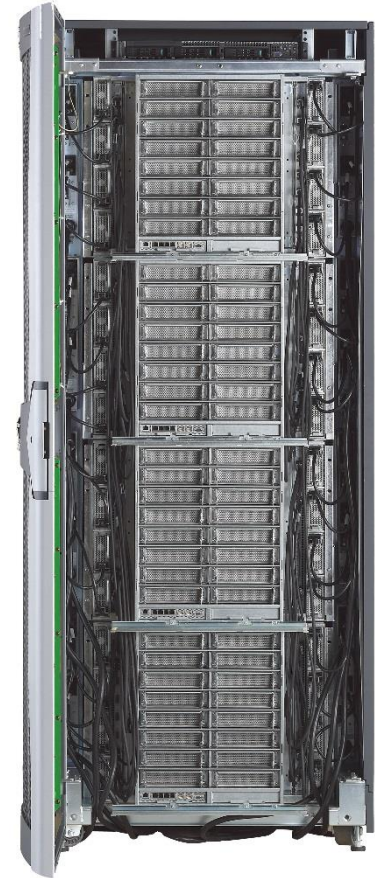
The screenshot shows the SGI Management Center interface with the 'Instrumentation' tab selected. The 'GPU' sub-tab is active, displaying a grid of monitoring data for eight nodes (n001 to n008). Each node's data is organized into two sections for GPU 1 and GPU 2. The metrics for each GPU include Name, Temperature, Fan speed, Usage percentage, Memory Usage, GPU Clock, SM Clock, and Mem Clock. For example, node n001 shows GPU 1 with 96.76% Mem Usage and GPU 2 with 287.0% Usage. The interface also includes a sidebar for 'Hosts' and 'Nodes', and a bottom bar with 'Event Log' and 'Provision Status' buttons.

Node	GPU	Name	Temp	Fan	Usage	Mem Usage	GPU Clock	SM Clock	Mem Clock
n001	GPU 1	Tesla C2050	64.0 °C	30 %	33.0 %	96.76 %	950 MHz	201 MHz	235 MHz
	GPU 2	Tesla C2050	68.0 °C	30 %	287.0 %	14.89 %	500 MHz	501 MHz	235 MHz
n004	GPU 1	Tesla C2050	64.0 °C	30 %	33.0 %	38.11 %	2500 MHz	191 MHz	238 MHz
	GPU 2	Tesla C2050	68.0 °C	30 %	287.0 %	0.19 %	50 MHz	101 MHz	135 MHz
n005	GPU 1	Tesla C2050	91.0 °C	100 %	99.0 %	0.19 %	50 MHz	101 MHz	135 MHz
	GPU 2	Tesla C2050	52.0 °C	42 %	50.0 %	0.19 %	50 MHz	101 MHz	135 MHz
n006	GPU 1	Tesla C2050	61.0 °C	30 %	0.0 %	25.31 %	350 MHz	201 MHz	350 MHz
	GPU 2	Tesla C2050	46.0 °C	15 %	0.0 %	14.89 %	500 MHz	501 MHz	235 MHz
n007	GPU 1	Tesla C2050	91.0 °C	100 %	99.0 %	38.11 %	2500 MHz	191 MHz	238 MHz
	GPU 2	Tesla C2050	68.0 °C	30 %	287.0 %	73.76 %	503 MHz	326 MHz	325 MHz
n008	GPU 1	Tesla C2050	43.0 °C	10 %	1.0 %	73.58 %	50 MHz	101 MHz	135 MHz
	GPU 2	Tesla C2050	46.0 °C	15 %	0.0 %	73.76 %	503 MHz	326 MHz	325 MHz

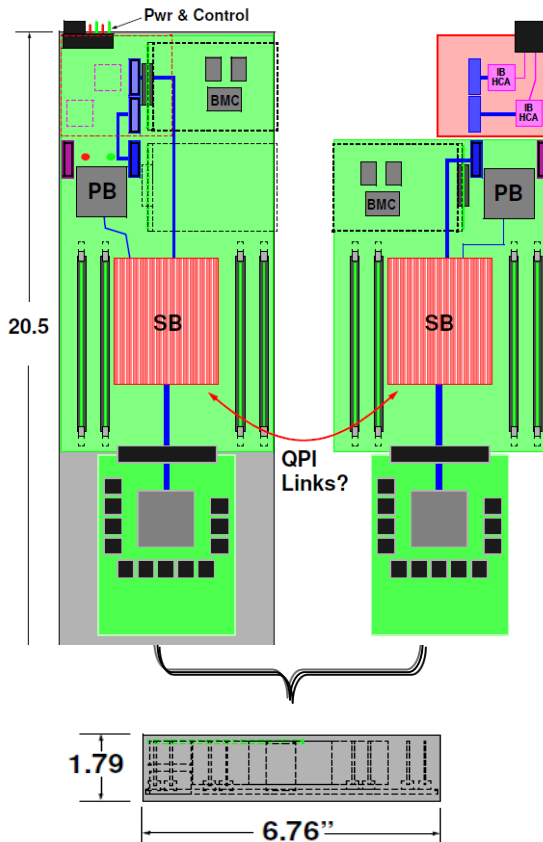
Detailed GPU reporting enables high performance systems to remain at their peak performance

Optimized for HPC & Big Data Solutions

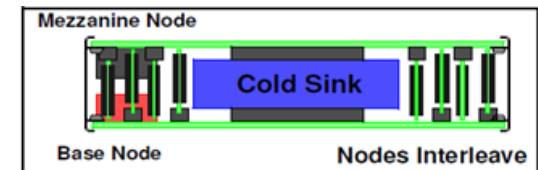
- SGI® ICE™ X
- SGI UV™ 2000



SGI ICE X Compute Node with GPU



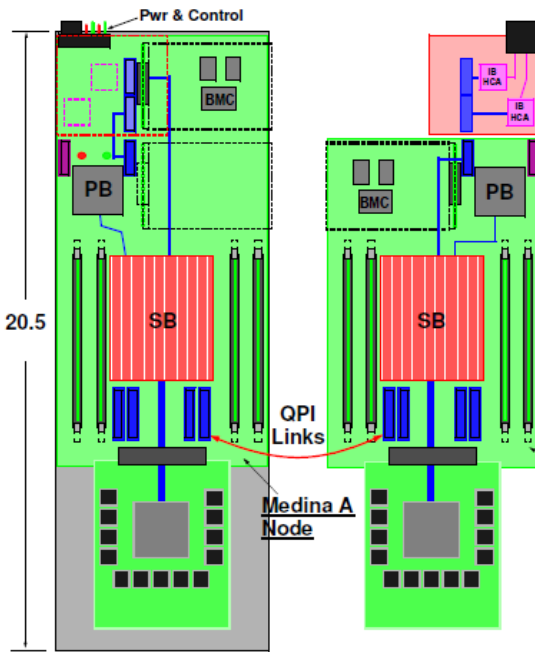
- Two Intel® Xeon® processor E5-2600 series
- **Two NVIDIA® TESLA® K20**
- FDR InfiniBand - single or dual plane
- Four DDR3 DIMMs per socket @ 1600 MT/s
- Up to one 2.5" SATA HDD/ SSD drives
- Liquid cold sinks



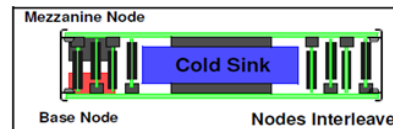
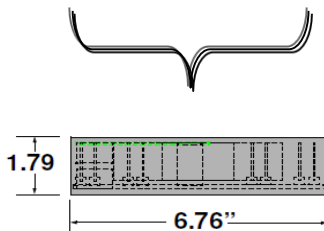
One dual socket node w/
two coprocessors in
one blade slot!

FDR + 1:1 processor to coprocessor ratio
= **Balanced Throughput**

SGI ICE X Compute Node with Phi



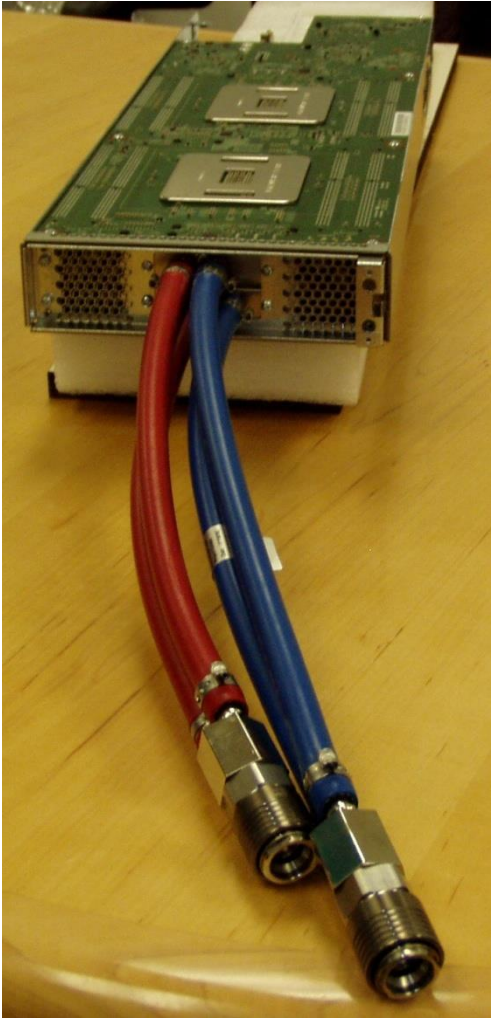
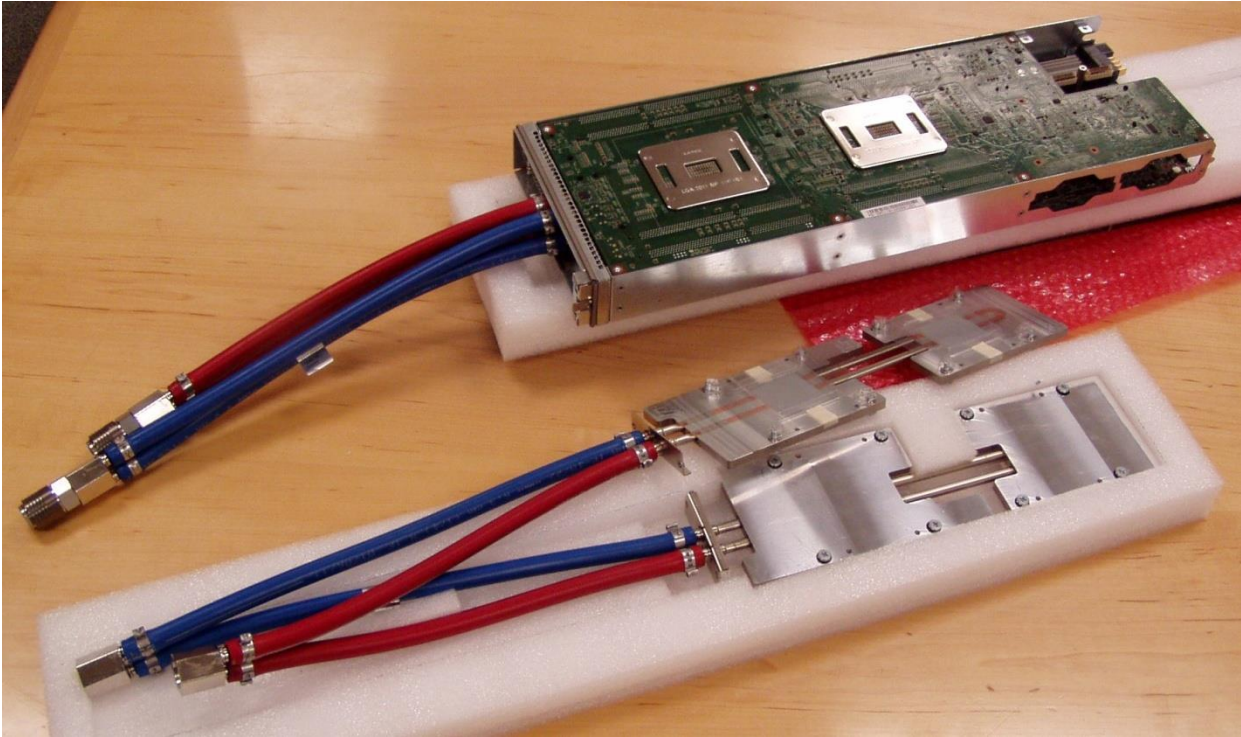
- Two Intel® Xeon® processor E5-2600 series
- **Two Intel® Xeon Phi™ 5120D Coprocessor**
- FDR InfiniBand - single or dual plane
- Four DDR3 DIMMs/ socket @ up to 1600 MT/s
 - Expect 1866 MT/s support w/1 DPC populated (w/ Ivy Bridge-EP)
- Up to one 2.5" SATA HDD/ SSD drive
- Liquid cold sinks



One dual socket node w/
two coprocessors in
one blade slot!

FDR + 1:1 processor to coprocessor ratio
= Balanced Throughput

SGI Cold Sink Technology – Twin Blades



SGI Meets All Accelerator Needs!

- SGI has the longest history with accelerators
- Accelerators have gone mainstream, we are supporting them across our product line
- The only vendor to be able to deploy scale-up, scale-out and hybrid landscapes with accelerators everywhere.

SGI ICE: Sample ICE Customers

BAW - Bundesanstalt für
Wasserbau

Bridgestone

Central Research Inst of
Electric Pwr Ind

CSIR - Centre for
Mathematical Modeling and
Compute

Exeter University

GENCI

HLRN Berlin

HLRN Hannover

ICHEC

ICR

Idaho National Laboratory

IFREMER

Imperial College

IMSc

INMET

JAMSTEC

Korean Air Force

KU Leuven

Mazda Motor Corporation

McLaren Motor Racing Ltd

Mercedes-Benz GP

MPO

NASA Ames

NASA Langley

NIIFI

NIMS

NMCAC

NOAA (CSC)

NTNU

Onera

ORNL

Pontificia Universidade
Catolica – PUC-Rio

Rosshydromet

Scientific Analysis Group

Semiconductor Energy Labs

Sikorsky

Skoda Auto

TI-09

Tokyo University – ISSP

Total

Toyota

U.S. Air Force - Arnold AFB

U.S. Navy/NRL

Universidade Catolica

Universite Paul Sabatier –
CICT/Cal mip

University of Arizona

University of Hyderabad

University of Oxford

University of São Paulo –
USP/IAG

University of Rio de Janeiro –
NACAD

University of Rio Grande do
Sul – CESUP

US Army TACOM

Hungary 2011-2014

NIIFI supercomputers for non-profit customers:

- Cluster: SGI ICE
Debrecen (18TFlops, 1536 core; 6TB ram)
- Cluster + GPU: HP
Szeged (14TFlops, 2304 core; 5,6TB ram, 6xM2070)
- SMP/ccNUMA: SGI UV
Pécs (10TFlops, 1152 core, 6TB ram)

http://www.niif.hu/szolgalatasok/szuperszamitastechnika/altalanos_ismerteto

Hungary 2015

- **Cluster: 200+ Tflops**
(200pc GPU or PHI)
- **Cluster: 30+ GPU or PHI**
- Cluster: SGI ICE
Debrecen (18TFlops, 1536 core; 6TB ram)
- Cluster + GPU: HP
Szeged (14TFlops, 2304 core; 5,6TB ram, 6xM2070)
- SMP/ccNUMA: SGI UV
Pécs (10TFlops, 1152 core, 6TB ram)

sgi