

# Fujitsu Petascale Supercomputer PRIMEHPC FX10



4x2 racks (768 compute nodes) configuration

Copyright 2012 FUJITSU LIMITED

# **PRIMEHPC FX10** Highlights



- Scales up to 23.2 PFLOPS
- Improves Fujitsu's supercomputer technology employed in the FX1 and "K computer".
  PRIMEHPC FX10

#### Tofu interconnect

Hybrid parallel (VISIMPACT)

#### Collective SW FX1



K computer\*

**ISA extension (HPC-ACE)** 

#### SPARC64<sup>™</sup> VIIIfx 8C/128GF 11PFLOPS, CY2010∼

SPARC64<sup>™</sup> VII 4C/40GF 121TFLOPS, CY2008~

# SPARC64<sup>™</sup> IXfx 16C/236.5GF ~ 23PFLOPS, CY2012 ~

\* "K computer" is the name of a next-generation supercomputer developed by RIKEN in July 2010

NA DECEM

### **K** computer





#### Linpack 10.51 PF (88,128 nodes)

#### Specifications

Calculation time 29 hours and 28 minutes

Efficiency **93.2%** 

 SPARC64<sup>TM</sup> VIIIfx (8 cores, 128 GF), 16GB memory, 64GB/s
 Interconnect: Tofu 6D mesh/torus with 4 RDMAs and 5GB/s x bidirectional x 10 links

Copyright 2012 FUJITSU LIMITED



#### High-speed and ultra-large-scale computing environment

Up to 23.2 PFLOPS (98,304 nodes, 1,024 racks, 6 PB of memory)

#### ■ SPARC64<sup>TM</sup> IXfx- high performance w/ low power consumption

■ 236.5 GFLOPS, and over 2GFLOPS/W

#### High execution performance with massively parallel apps

Tofu interconnect, Technical Computing Suite, and VISIMPACT

#### High reliability and high operability for a large-scale system

## **PRIMEHPC FX10** Configuration





System	Maximum configuration	Node	
Number of nodes	98,304 (= 32x32x8x2x3x2) nodes	Number of CPUs	1 CPU
Peak performance	23.2 PFlops	Peak performance	236.5 GFlops
Total memory bandwidth	8.3 PB/s	Memory bandwidth	85 GB/s
Total interconnect bandwidth	3.1 PB/s	Interconnect link bandwidth	5 GB/s x bidirectional
Total IO bandwidth	49.1 TB/s	IO bandwidth	8 GB/s per IO node

# SPARC64<sup>™</sup> IXfx

Binary compatible with K computer

- Double the number of cores of K's SPARC64 VIIIfx
- SPARC V9 + HPC-ACE
  - # of registers: 256 double precision registers
  - SIMD instructions: 2-way, 2-wide SIMD w/ FMA
  - Software-controllable cache (Sector cache)
- VISIMPACT
  - Automatic parallelizing compiler
  - Inter-core hardware barrier
  - Shared secondary cache
- Integrated memory controller
  - 1.333GHz DDR3



Frequency	1.848 GHz	
Secondary cache	12 MB (Shared cache)	
Theoretical peak	236.5 GFLOPS	
Memory throughput	85 GB/s	
Power consumption	110 W	
Process technology	40 nm	



# **Tofu interconnect**



Highly scalable and usable direct network (6D mesh/torus)

- 10 redundant high BW links, 4 RDMA engines (4x2 simultaneous transfer)
- Good collective communication performance with Tofu original algorithms
- Tofu barrier for barrier & reduction in H/W
- Direct attached interconnect controller



Tofu realizes scalable systems beyond 100,000 nodes With low power consumption, low latency, and high BW

# K computer and **PRIMEHPC FX10**



Fujitsu supercomputer w/ enhanced technology introduced for K computer

	K computer	PRIMEHPC FX10	Note
CPU	SPARC64 VIIIfx	SPARC64 IXfx	SPARC V9 + HPC-ACE
Peak perf.	128 GFLOPS	236.5 GFLOPS	
# of cores	8	16	
Memory	16GB	32GB/64GB	2GB/core~
BW	64GB/s	85GB/s	
Interconnect	6D mesh/torus	$\leftarrow$	Tofu interconnect
System size	X x Y x 17	X x Y x 9	Z=0 is I/O node
link BW	5GB/s x bidirectional	$\leftarrow$	

PRIMEHPC FX10 supports water cooling of ordinary chilling facility and resolves layout restrictions



Improves Fujitsu's supercomputer technology employed in the "K computer".

- Newly developed SPARC64 IXfx processor (236.5GFLOPS)
- Tofu interconnect (scales up to 98,304 nodes)
- Fujitsu provides all hardware and software
  - Processor and interconnect
  - Technical Computing Suite HPC middleware, as well as the FEFS highperformance distributed file system

# FUJITSU

#### shaping tomorrow with you