

AI^{CS} • • • •

the K computer System overview

Atsuya Uno

System Development Team, Development Group
Next-Generation Supercomputer R&D Center
RIKEN (The Institute of Physical and Chemical Research)



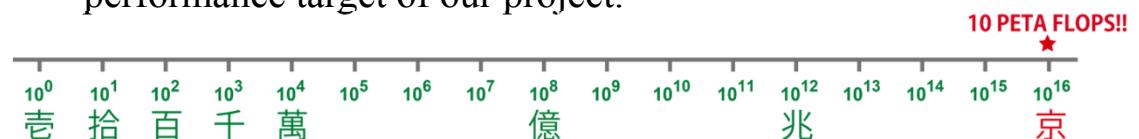
RIKEN Advanced Institute for Computational Science

SC'11

K computer, “京”, is ...



- “京 [kei]” is a nickname of the next-generation supercomputer system
 - The name was chosen from public applications in July last year.
 - “京” stands for a unit corresponding to 10 P, which is also the performance target of our project.



- No.1 system on the latest TOP500 list.
 - 10.51 PFlops in LINPACK benchmark (peak : 11.28PFlops)
 - Efficiency : 93.2%
 - Execution time : 29hours 28minutes.
 - Power consumption : 12.66 MW

AICS

SC'11

(1)

RIKEN Advanced Institute for Computational Science



Design targets (in 2006) and Schedule

- 10 peta-FLOPS in LINPACK benchmark
- Peta-FLOPS sustained performance in real applications
- Low power consumption system
- Highly reliable system

		FY2006	FY2007	FY2008	FY2009	FY2010	FY2011	FY2012
System		Conceptual design	Detailed design	Prototype, evaluation	Production, installation, and adjustment	Tuning and improvement		
Buildings	Computer building		Design	Construction				
	Research building		Design	Construction				

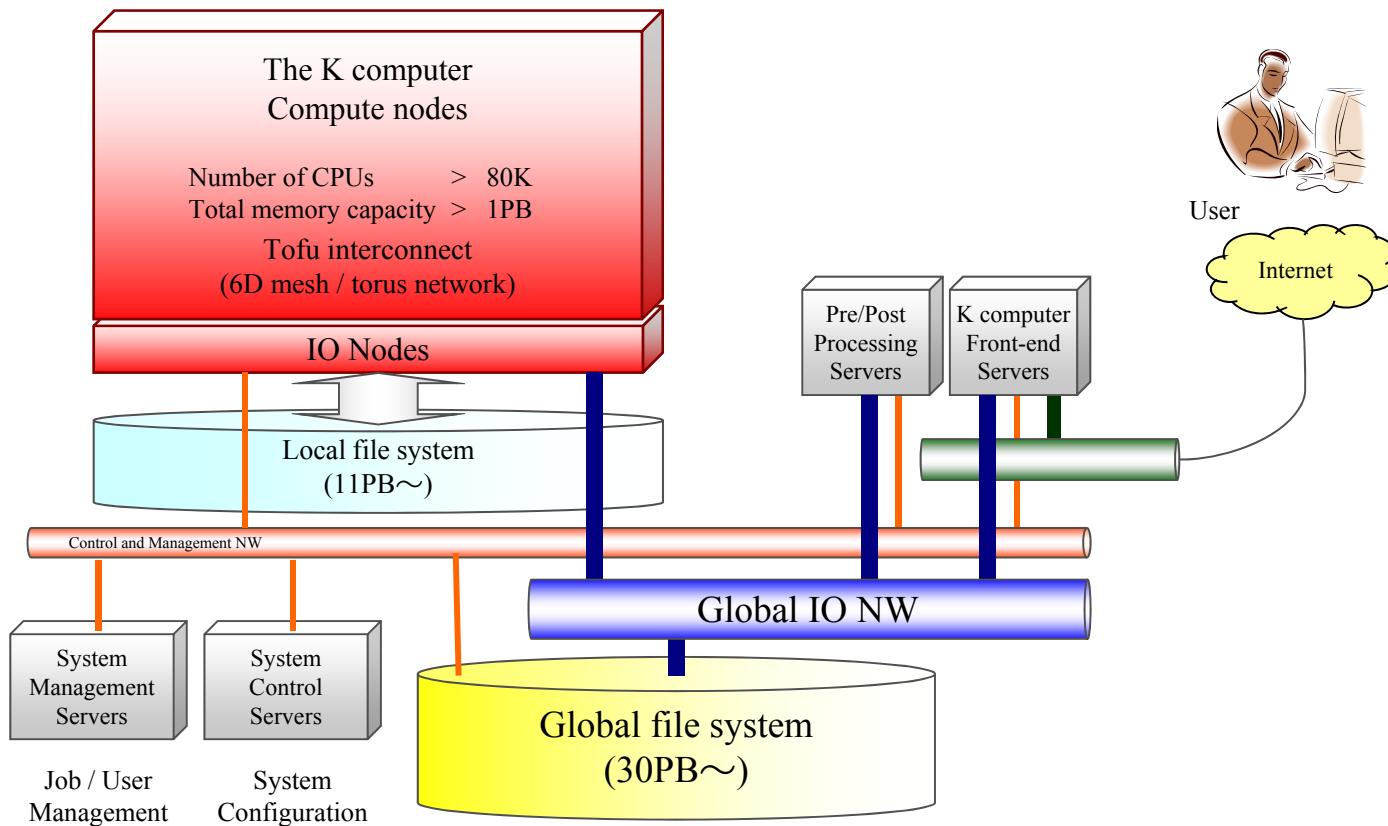
the present



System Overview of the K computer

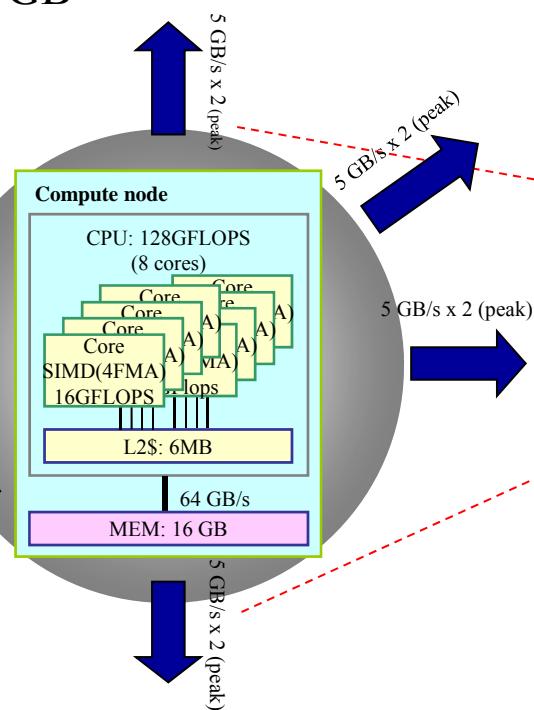
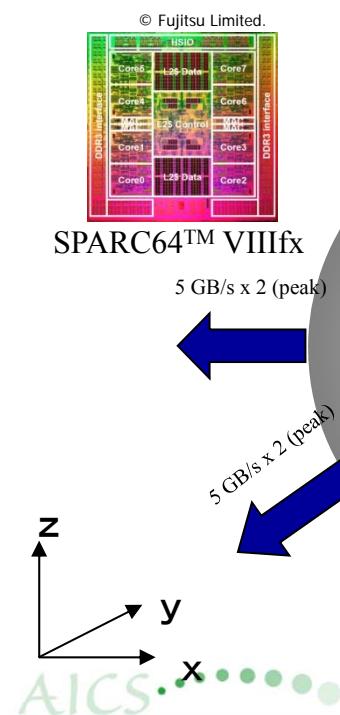


System Configuration

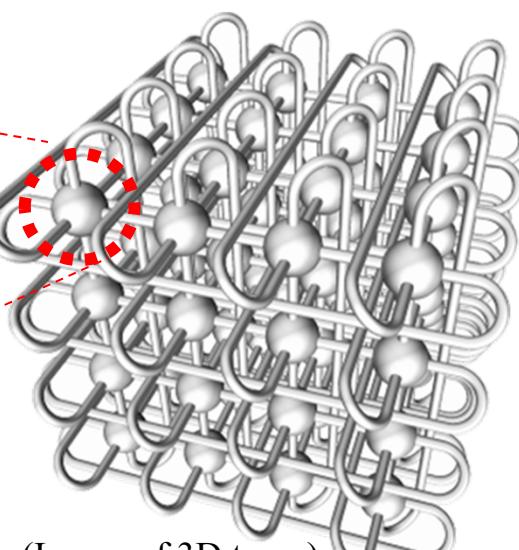


Compute Nodes

- CPU : SPARC64™ VIIIIfx (8 core)
 - 128 Gflops
 - 2.0 GHz
- Memory : 16 GB
- Network : Tofu interconnect (6D mesh / torus network)
 - provide logical 3D torus network for each job
 - 5 GB/s x 2 bandwidth of each link



SC'11



(Image of 3D torus)

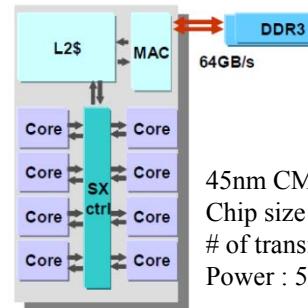
© Fujitsu Limited.



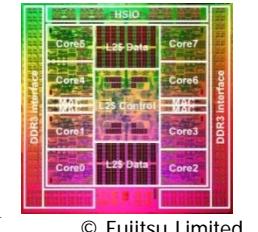
CPU – SPARC64™ VIIIfx

- Superscalar Multi-core processor (8 cores)
 - 128 GFLOPS (16 GFLOPS/core)
 - 2 GHz
 - 256 FP registers (double precision) / core
 - 2 SIMD (FMA x 4) / core
- Shared 6MB L2\$ (12way)
 - Software controllable cache (Sector cache)
- Memory throughput
 - L2\$ - Main memory : 0.5B/FLOP (64GB/s)
- Hardware barrier
 - High-speed synchronization of on-chip cores
- Power consumption
 - 58W at 30°C by water-cooling

You can get more detailed information from ‘SPARC64™ VIIIfx Extensions’
<http://img.jp.fujitsu.com/downloads/jp/jhpc/sparc64viiifx-extensions.pdf>



45nm CMOS process
 Chip size : 22.7mm x 22.6mm
 # of transistors : 760M
 Power : 58W (TYP, 30°C), Water Cooling

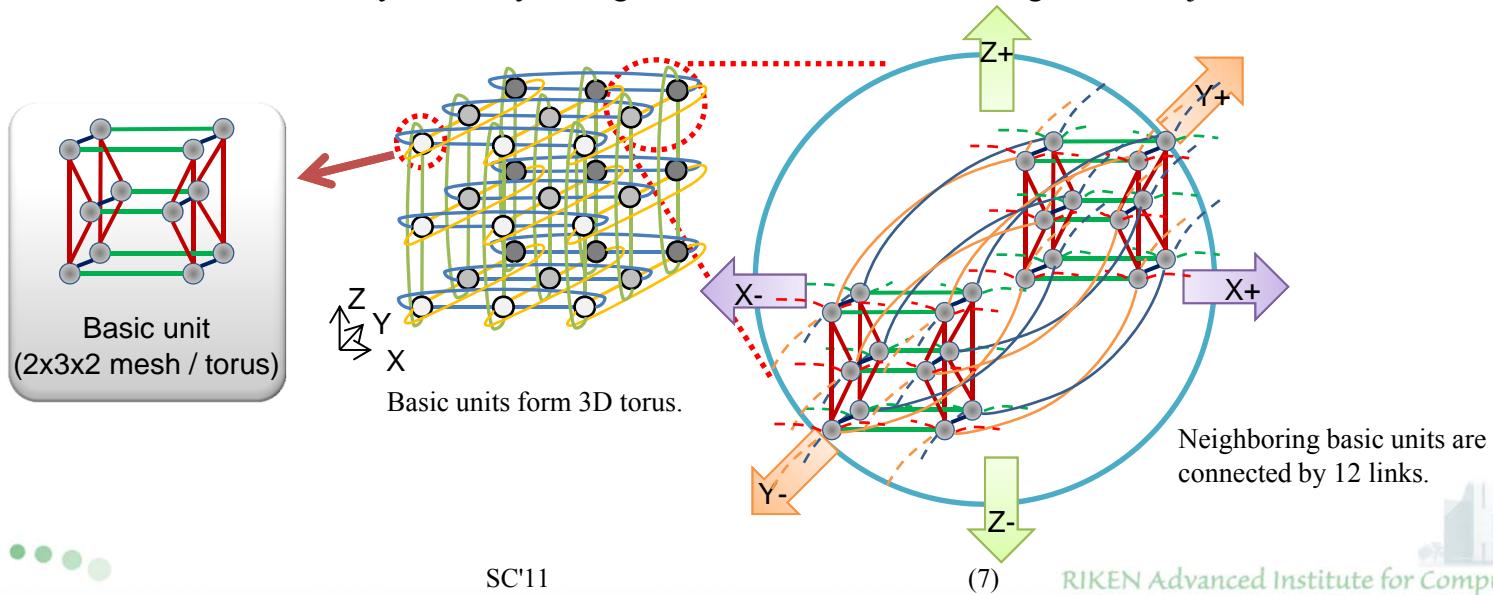


© Fujitsu Limited.

	Specification
Performance (peak)	128 GFLOPS (16 GFLOPS x 8 cores)
Core	8
Clock	2.0 GHz
Floating-point Execution units (Core spec)	FMA x 4 (2 SIMD) DIVIDE x 2 COMPARE x 2 Floating-point register (64bit) : 256 General purpose register (64bit) : 188
Cache	L1I\$: 32 KB (2way) L1D\$: 32 KB (2way) L2\$: Shared 6 MB (12way)
Memory throughput	64 GB/s (0.5B/F)

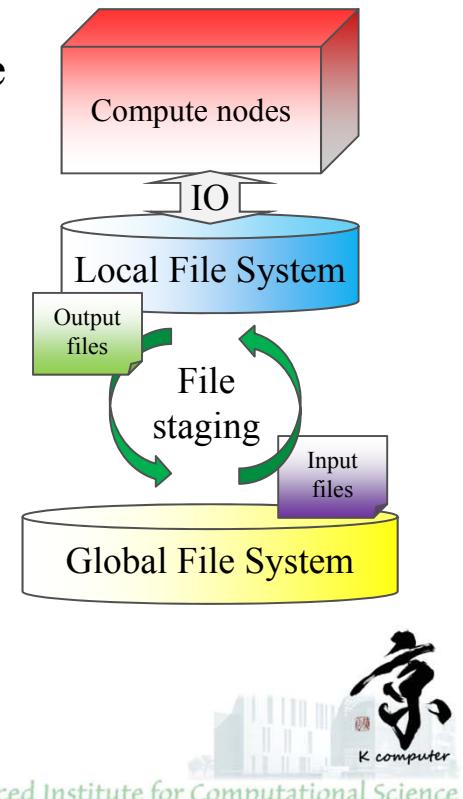
Tofu (Torus fusion) Interconnect

- High communication performance and fault-tolerant network
- Network topology : 6D mesh / torus network
 - Each node has 10 links (5 GB/s x 2 bandwidth)
 - 4 links for making a basic unit (2x3x2 mesh / torus) and 6 links for XYZ link (XZ:torus / Y:mesh)
 - Multi-path routing by a combination of XYZ link and 2x3x2 mesh / torus network enables to make a detour of faulty nodes
- From user's point of view, network topology of the job is 1,2 or 3D torus network.
 - This torus network is dynamically configured when the nodes are assigned to the job.



System environments

- OS: Linux based OS on compute nodes
- File system : FEFS (Fujitsu Exabyte File System) based on Lustre file system
 - Two-level large-scale distributed file system
 - Local file system for temporary files used by jobs
 - Global file system for user's permanent files
 - Staging functions
 - Stage-in
Input files on the global file system which are used in a job are copied on the local file system.
 - Stage-out
Output files generated during a job execution are moved back to the global file system after the job finished.
- Batch job-oriented system
 - Interactive environments are available for debugging.



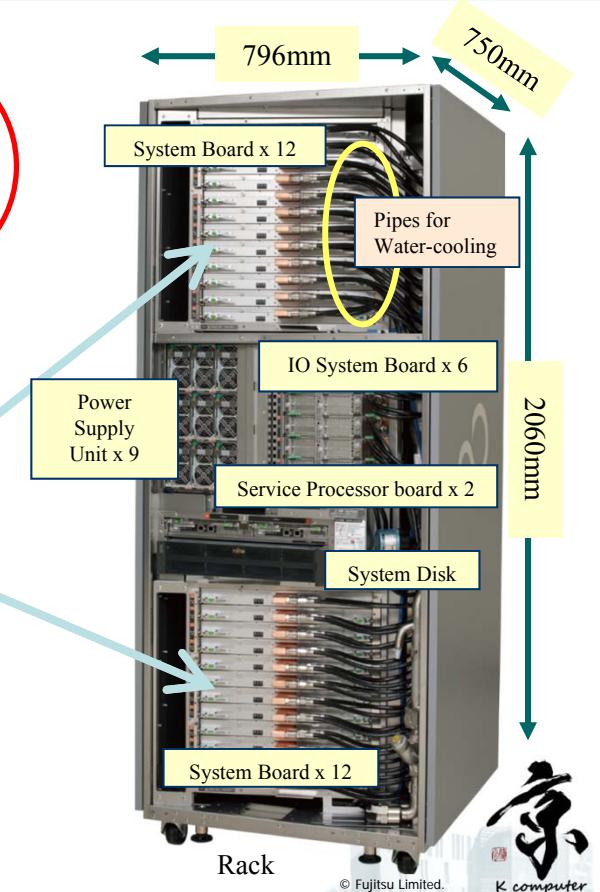
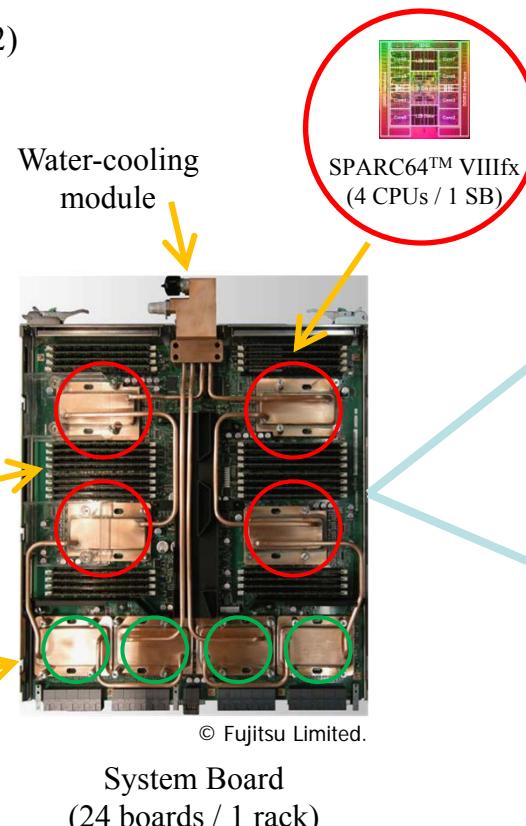
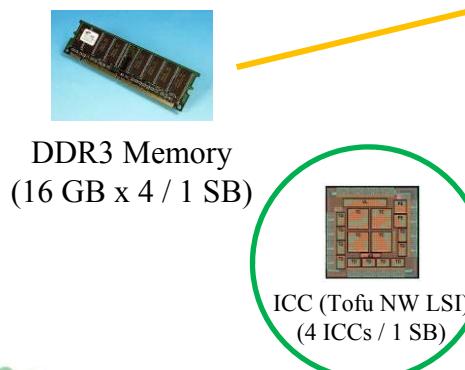
Programming environments

- Programming languages and compilers
 - Fortran 95 and subset of Fortran 2003, XPFortran
 - C99 and C++2003 including several extensions to GNU C and C++
 - Optimized compilers to obtain full capabilities of SPARC64™ VIIIIfx
 - SIMD, 256 FP registers, programmable L2 cache (sector cache), and so on.
 - OpenMP 3.0 is supported
- MPI library based on MPI-2.1 specification
 - Low-latency and high-throughput
 - Collective communications are optimized for the Tofu interconnect
 - Broadcast, Allgather, Alltoall, and Allreduce
- Numerical libraries
 - BLAS, LAPACK, FFTW, Fujitsu scientific numerical library SSL II, and so on will be provided.



Hardware Implementation

- Compute rack houses :
 - System board x 24 (top:12, bottom:12)
 - IO system board x 6
 - Power supply unit x 9
 - Service processor board x 2
 - System disk (RAID 5)
- Peak performance : 12.3TF / Rack
- Total memory : 1.5TB / Rack
- Weight: 1,500Kg (max.)



Hardware view

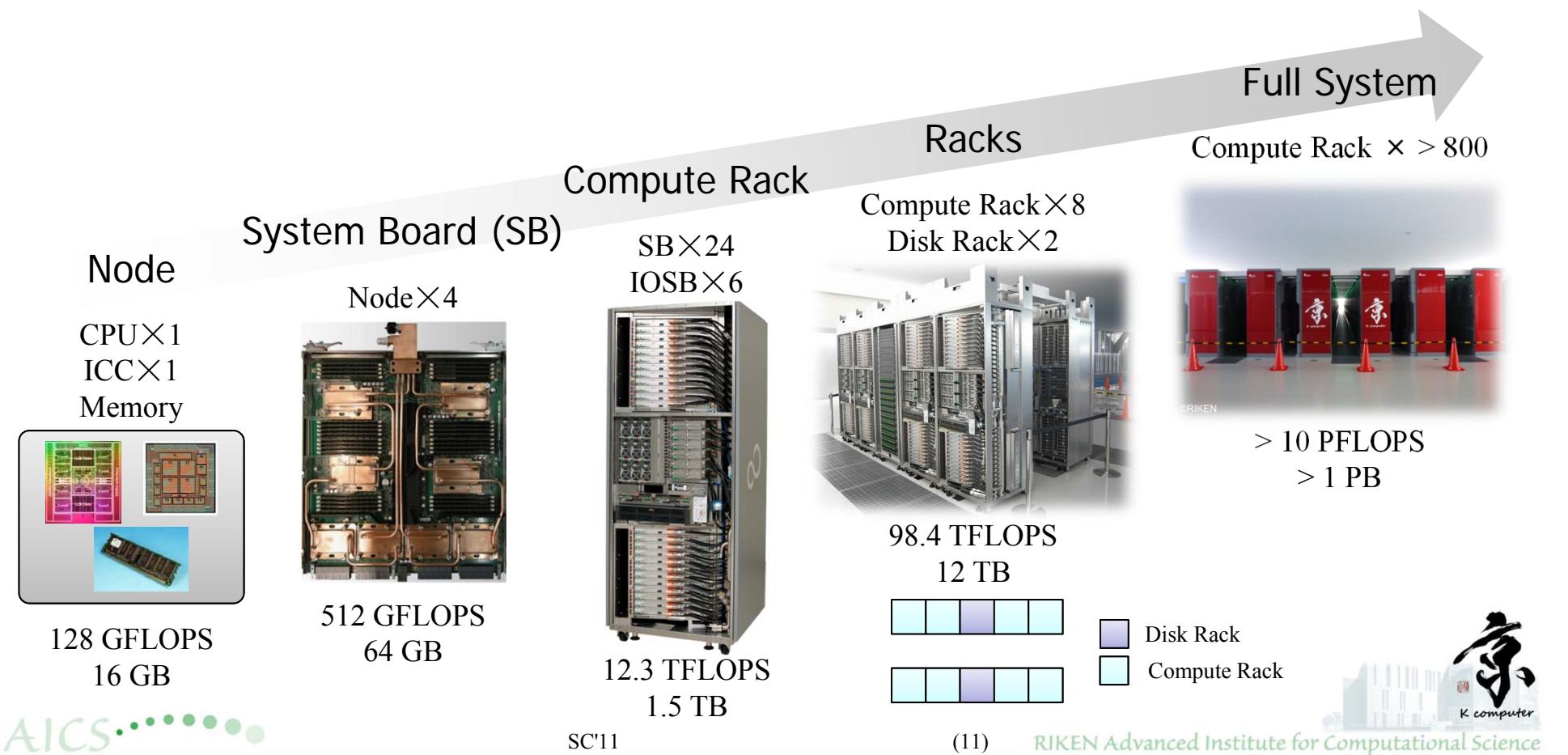
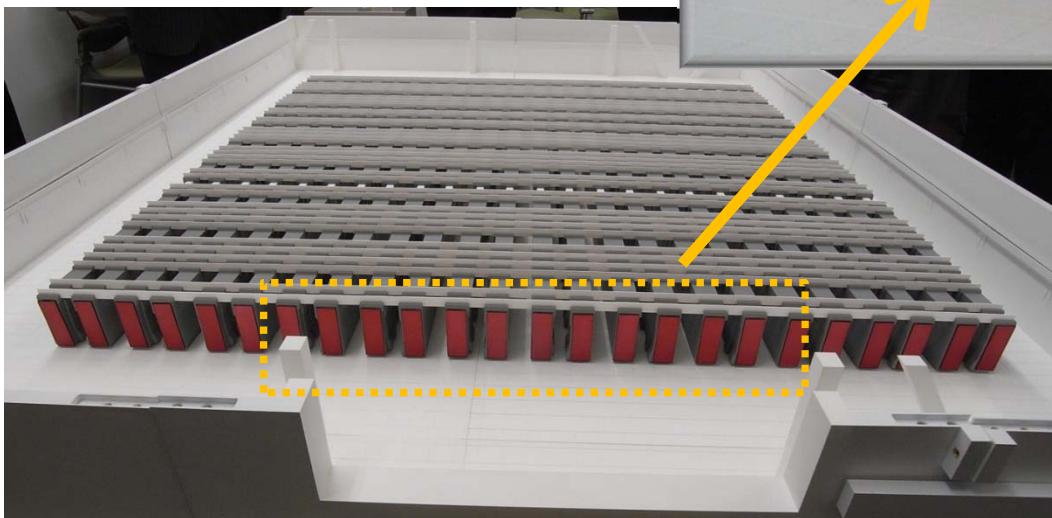


Image of the K computer

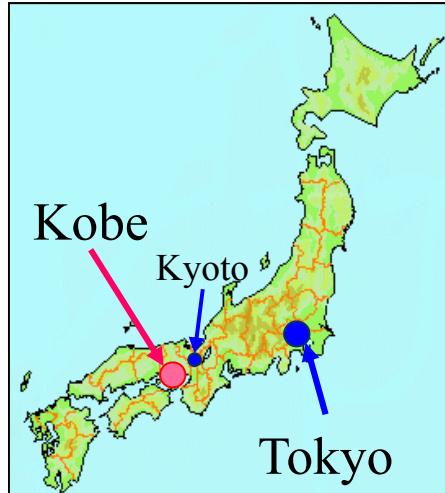
864 racks are housed in the computer room.



Facilities for the K computer



Location of the K computer in Kobe

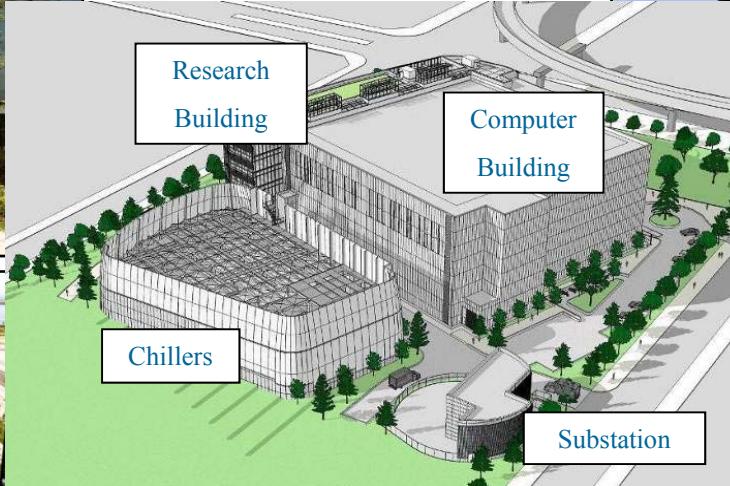


450km (280miles)
west from Tokyo

AICS (Advanced Institute for Computational Science) was established at RIKEN in July, 2010.



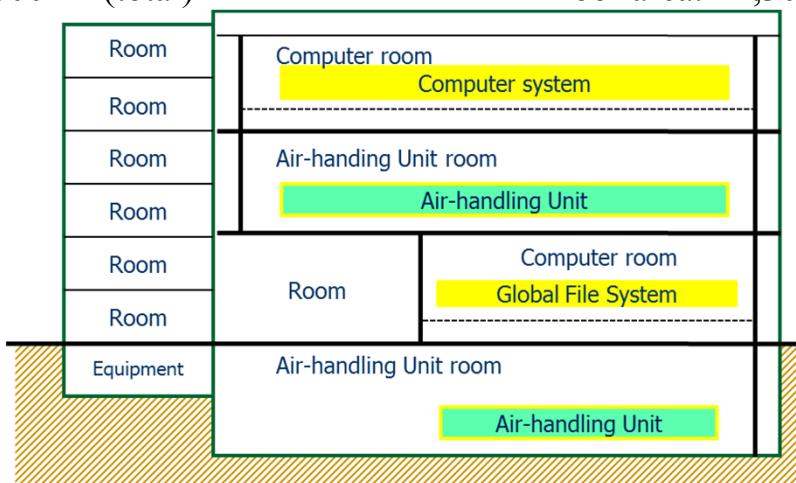
Layout of buildings



Buildings and Cooling System

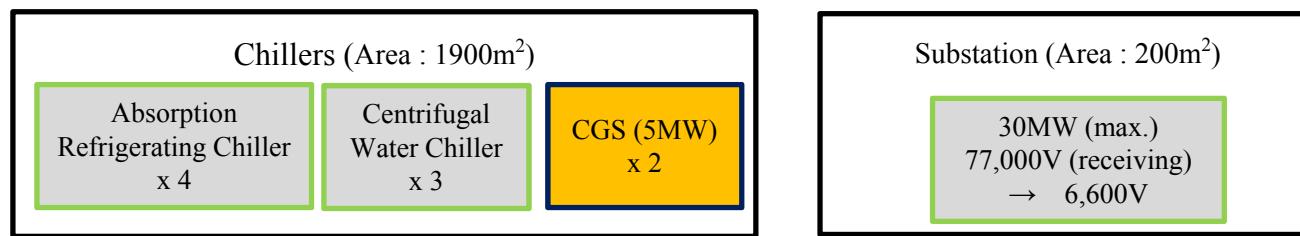
Research Building

- Six-story above ground and one below
- Floor area : ~1,800 m², ~9,000 m² (total)



Computer Building

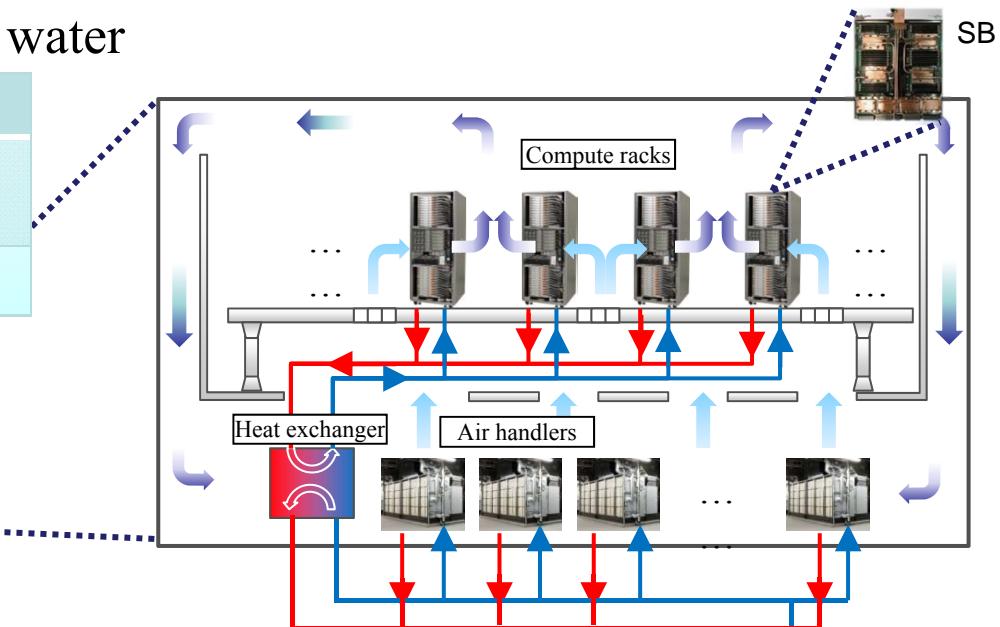
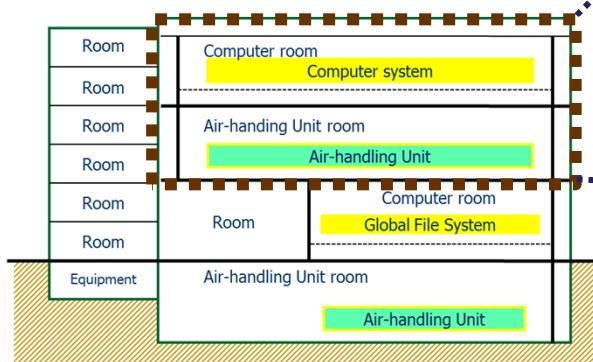
- Three-story above ground and one below
- Floor area: ~4,300 m², ~10,500 m² (total)



Cooling System

- K computer is cooled by air and water

	Parts	Temperature
Water	CPU and ICC	Input : $\sim 15^{\circ}\text{C}$ Output : $\sim 17^{\circ}\text{C}$
Air	the others	Room temp. : $\sim 20^{\circ}\text{C}$

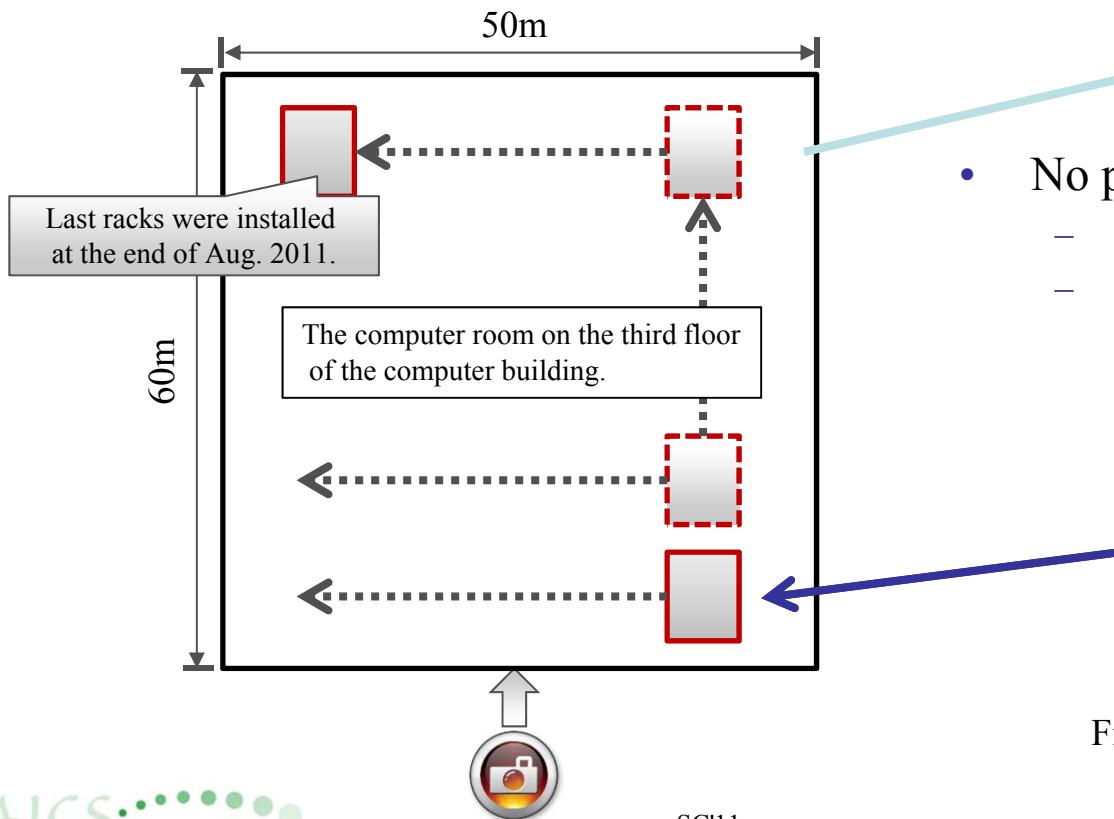


Absorption Refrigerating Chillers
&
Centrifugal Water Chillers



Installation of the K computer

Period of Installation : 2010.9 ~ 2011.8





Copyright (c) RIKEN AICS 2011

Thank you for your attention !

AICS Booth #112

A photo in the early evening

AICS

SC'11

(20)

RIKEN Advanced Institute for Computational Science

