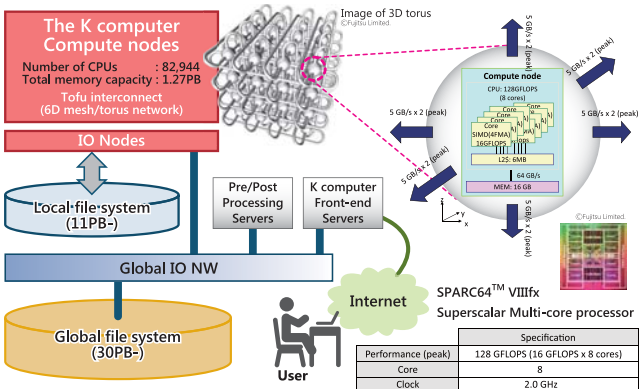
The K computer Hardware and Operations

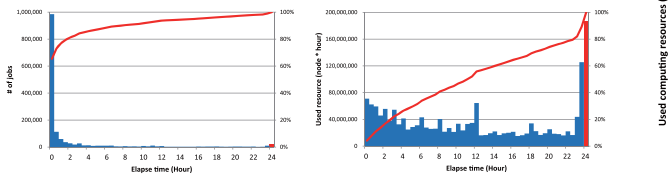
System Configuration and Operation Statistics (Oct. 2012 - Mar. 2016)



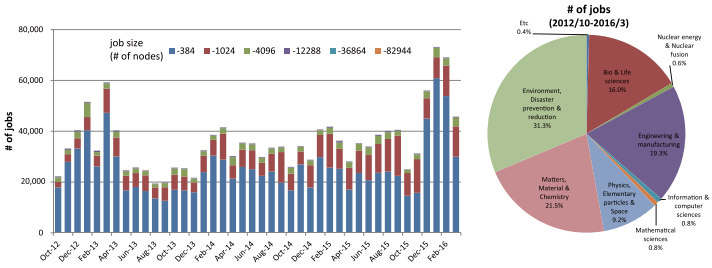
SPARC64™ VIIIfx
Superscalar Multi-core processor

	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)

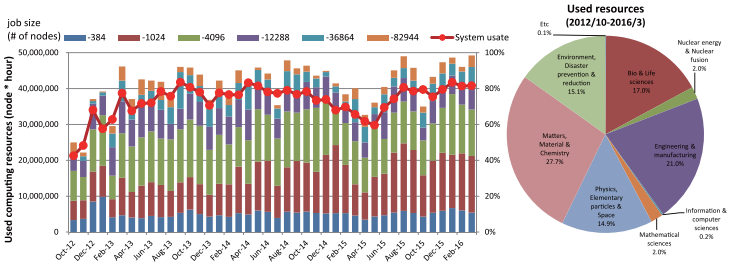
- Peak Performance: 10.62 PFLOPS
- Total Memory Capacity: 1.26 PiB (16 GiB/ node)
- Interconnect: Tofu
  - 6D mesh/torus, Logical-1,2 or 3D torus
  - Peak bandwidth: 5 GiB/s x 2
  - Bisection bandwidth: 30 TiB/s



Number of executed jobs (upper) and amount of resources used (lower) by elapsed job time (rightmost red rectangle denotes jobs over 24 hours elapsed time; solid lines are cumulative values)

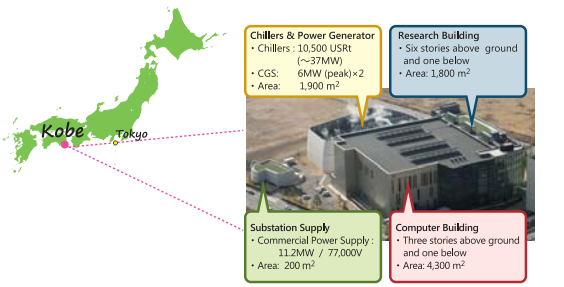


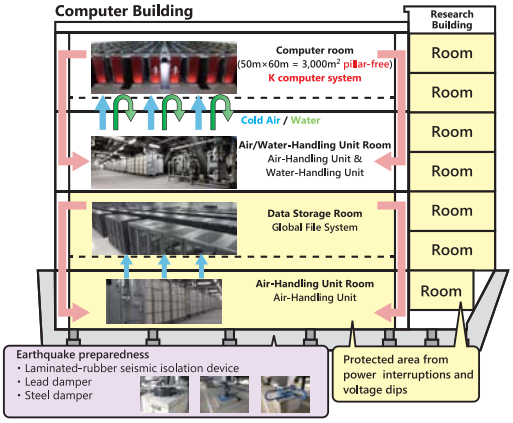
Number of executed jobs of each month (left) and a breakdown of jobs classified by research areas (right) between October 2012 and March 2016.



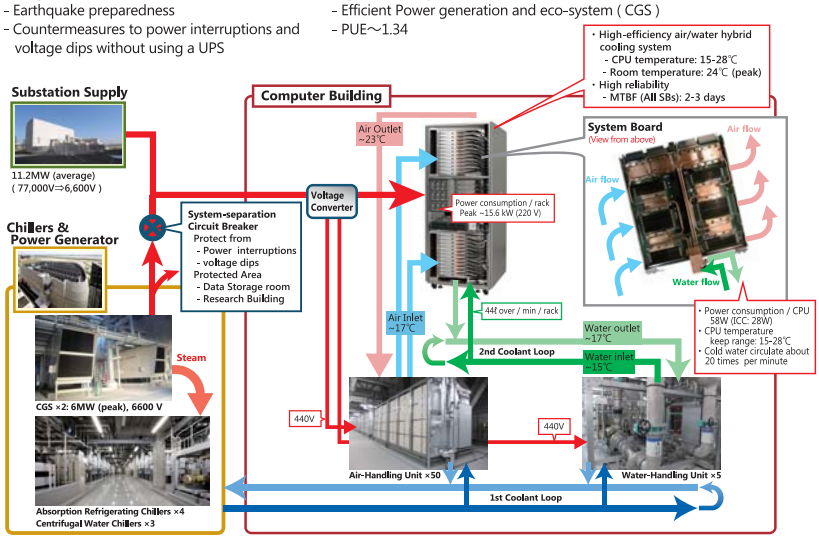
Amount of used computing resources (product of number of used nodes and elapsed time) and system usage of each month (left) and a breakdown of used resources classified by research areas (right) between October 2012 and March 2016.

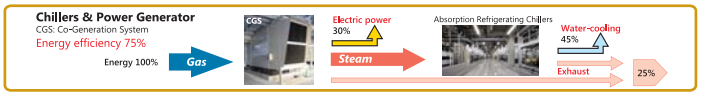
Hybrid Cooling System





- **Safety**
  - Earthquake preparedness
  - Countermeasures to power interruptions and voltage dips without using a UPS
- **High-efficiency**
  - Efficient Power generation and eco-system (CGS)
  - PUE~1.34





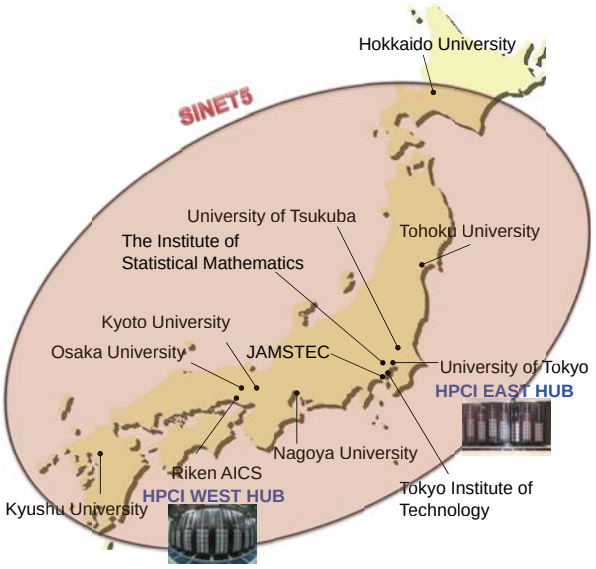
K computer

HPCI System Development Team
Software Development Team

HPCI System Development Team

HPCI : High Performance Computing Infrastructure

- **One-Stop service for nationwide super computer resources**
  - Support from project proposal to user report.
  - The K computer and other major super computers in japan, large scale MPP, GPU Cluster, PC Cluster, and Vector machines are available.
  - The HPCI helpdesk is available for HPCI users, and provides consultation before applying project proposal, technical support for each computer resources, and guidance of user report after finishing the project.
- **HPCI Shared Storage**
  - Provide 22PBytes single view file space for all the HPCIuser
  - Available from all the HPCI super computer resources.
  - HPCI single sign on authentication, if a user wants to transfer the computational results to the HPCI shared storage after the computational session, the user does not have to log in again to the HPCI shared storage.
  - Automatic file replication, HPCI shared storage adopts 2 file replicas in default. Secure network communication, HPCI user can specify data encryption method such as “gsi” in ~/.gfarm2fs.

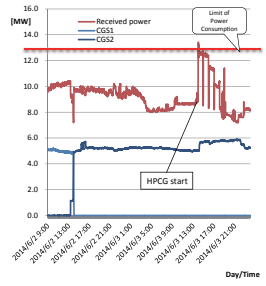


- **Single sign On Authentication**
  - Allow efficient use of nation wide super computer resources.
  - Users of the HPCI can utilize all the granted HPC resources seamlessly after signing on one of the resources.

Software Development Team

Collaborate between the system and the application to improve system and usability

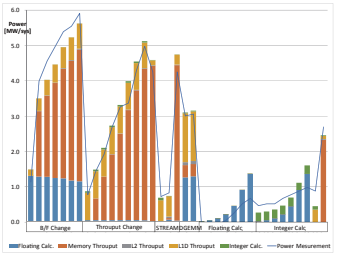
Improvement of Operation Efficiency and Effort for Power Consumption Reduction on the K computer



Recently jobs with large power consumption have increased and total power consumption exceeded the power consumption limit regulated by contract with provider.
If the total power consumption exceed the power limit, we have to pay additional charge as a penalty.

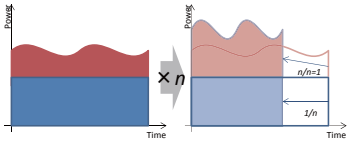
Analysis of correlation with power consumption and application performance

- Standby power and memory throughput has a greater impact on the power consumption of the K computer.



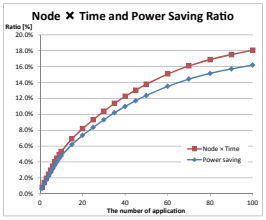
Simple kernel loops

$$\delta\text{Power}[\text{MW/sys}] = 8.8128 + 1.3659 \times \text{Floating Calc.}[\%] + 4.3906 \times \text{Memory Throughput}[\%] + 0.0857 \times \text{L2 Throughput}[\%] + 2.3299 \times \text{L1D Throughput}[\%] + 0.2429 \times \text{Integer Calc.}[\%]$$



A trial of improving of power consumption and system operation

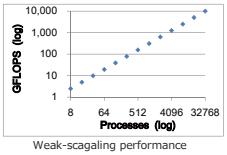
- A evaluation model of power consumption of each performance tuned applications.



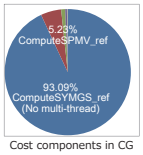
- The performance of each big user’s applications has estimated how much differs compared to the field average.
- If both the CPU performance and the memory performance improved to field average, NODE×TIME can be reduced as below.
  - ✓ When we improve performance about the 10(50) major user application, we can use more 25M(66M) NODExTIME.
- The power consumption is
  - ✓ When we improve performance about 10(50) major user application, we can achieve energy savings of 2.7GWh(7.0GWh).

HPCG Performance Tuning on the K computer

Evaluate Original HPCG on the K



Weak-scagaling performance



Cost components in CG

The original HPCG code evaluation on the K computer gave these informations.

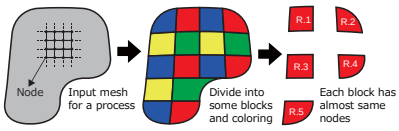
- Linear scalability was obtained, so parallelization tuning is not ncessary.
- Single CPU performance was low since SYMGS is not multi-thread.

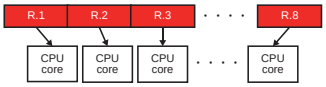
Therefore we have aimed the. single CPU.

Tune: Coloring for SYMGS

In the original code, SYMGS is not able to be multi-thread since there are data dependencies between rows.

To avoid data dependencies, we employed the colored blocking that divide the mesh into some blocks and do the coloring to blocks.





Nodes in same colored blocks are reordered to be continuous in memory space.

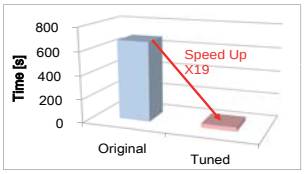
There are no data dependencies between same colored blocks, so the blocks are able to be processed by multi-threading.

Significant Improving Obtained

We tried these additional ways for improving.

- Memory serialization for matrix
- Data access ordering improvement for SYMGS
- Loop optimization for SPMV, SYMGS
- Parameter adjustment
- Refinement miscellaneous routines

We obtained 19 times improvement.



Comparing CG run time for 5 CG calls

Result for SC15

Rank	Computer	HPL PFLOPS	HPCG PFLOPS	Ratio to HPL %
1	Tianhe-2	33.86	0.580	1.7%
2	K computer	10.51	0.461	4.4%
3	Titan	17.59	0.322	1.8%
4	Trinity	8.10	0.183	2.3%
5	Mira	8.59	0.167	1.9%

http://www.hpcg-benchmark.org/custom/index.html?lid=155&slid=282

Using V2.4, We obtained the HPCG score 0.461 PFLOPS using 82944 processors.This score was ranked 2nd on the list at SC15.

Under the measurement rule of the V3.0, the performance of computational kernel is same to V2.4, but the HPCG score will be decreased.

K computer