Chapter 23

Flagship 2020 Project

23.1 Project Overview

The Japanese government launched the FLAGPSHIP 2020 project [1] in JFYI 2014 whose missions are defined as follows:

- Building the Japanese national flagship supercomputer, the successor to the K computer, which is tentatively named the post K computer, and
- developing wide range of HPC applications that will run on the post K computer in order to solve the pressing societal and scientific issues facing our country.

RIKEN accepted a request by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to assume the responsibility of being the institution responsible for the research and development of the post K computer. In the summer of 2014, the Japanese Government selected nine Priority Issues to be targeted by the post K computer, and formulated project organizations that would be responsible for research toward solving them. The government also selected four research areas which are named Exploratory Challenging Issues to be developed. Table 23.1 lists the nine Priority Issues to be tackled by the FLAGSHIP 2020 project and institutions selected to lead the research for solving them. Table 23.2 lists four Exploratory Challenging Issues and selected institutions.

RIKEN AICS is in charge of co-design of the post K computer and development of application codes in collaboration with the Priority Issue institutes, as well as research aimed at facilitating the efficient utilization of the post K computer by a broad community of users. Under the co-design concept, AICS and the selected institutions are expected to collaborate closely. For the post K computer system development, in September 2015, Fujitsu Ltd. was selected to produce the basic design. A number of university partners are partially involved with the design.

23.2 AICS Teams

As shown in Table 23.3, four development teams are working on post K computer system development with the FLAGSHIP 2020 Planning and Coordination Office that supports development activities. The primary members are listed in Section 23.6.

The Architecture Development team designs the architecture of the post K computer in cooperation with Fujitsu and designs and develops a productive programming language, called XcalableMP (XMP), and its tuning tools. The team also specifies requirements of standard languages such as Fortran and C/C++ and mathematical libraries provided by Fujitsu.

The System Software Development team designs and specifies a system software stack such as Linux, MPI and File I/O middleware for the post K computer in cooperation with Fujitsu and designs and develops multi-kernel for manycore architectures, Linux with light-weight kernel (McKernel), that provides a noise-less runtime environment, extendability and adaptability for future application demands. The team also designs and develops a low-level communication layer to provide scalable, efficient and portability for runtime libraries and applications.

[1] FLAGSHIP is an acronym for Future LAtency core-based General-purpose Supercomputer with HIgh Productivity.

Table 23.1: Priority Issues

Health and longevity	1. Innovative computing infrastructure for drug discovery	RIKEN Quantitative Biology Center, and other institutions
	2. Personalized and preventive medicine using big data	Institute of Medical Science / the University of Tokyo, and five other institutions
Disaster prevention / Environment	3. Integrated simulation systems induced by earthquake and tsunami Earthquake	Research Institute, the University of Tokyo and 4 other institutions
	4. Meteorological and global environmental prediction using big data	JAMSTEC / Center for Earth Information Science and Technology of Japan, and three other institutions
Energy issues	5. New technologies for energy creation, conversion / storage, and use	Institute for Molecular Science / National Institute of Neural Science, and eight other institutions
	6. Accelerated development of innovative clean energy systems	School of Engineering / the University of Tokyo, and eight other institutions
Industrial competitiveness enhancement	7. Creation of new functional devices and high-performance materials	The Institute of Solid State Physics / the University of Tokyo, and eight other institutions
	8. Development of innovative design and production processes	Institute of Industrial Science / the University of Tokyo, and six other institutions
Basic science	9. Elucidation of the fundamental laws and evolution of the universe	Center for Computational Science / Tsukuba University, and seven other institutions

Table 23.2: Exploratory Challenges

10. Frontiers of basic science - challenge to the limits -	Institute for Material Research / Tohoku University, and nine other institutions
	Tokyo Woman's Christian University, and three other institutions
	School of Engineering / the University of Tokyo
11. Construction of models for interaction among multiple socioeconomic	RIKEN, Advanced Institute for Computational Science and twelve other institutions
	Tokyo University of Science and three other institutions
12. Elucidation of the birth of exoplanets (Second Earths) and the environmental variations of planets in the solar system	Kobe University Graduate School of Science and eight other institutions
13. Elucidation of how neural networks realize thinking and its application to artificial intelligence	Okinawa Institute of Science and Technology Graduate University and 4 other institutions
	The Research Center for Advanced Science and Technology / the University of Tokyo

Table 23.3: Development Teams

Team Name	Team Leader	Number of Members
Architecture Development	Mitsuhisa Sato	22
System Software Development	Yutaka Ishikawa	15
Co-Design	Junichiro Makino	19
Application Development	Hirofumi Tomita	39

Table 23.4: Target Applications

Program	Brief Description
GENESIS	Molecular dynamics simulation for protein in solvent (MD)
Genomon	Human Genome processing (Genome alignment)
GAMERA	Earthquake simulation system (FEM multigrid method)
NICAM+LETKF	Weather prediction system using Big data (structured grid & ensemble Kalman filter)
NTChem	Ab-initio molecular electronic structure (RI-MP2 method)
ADVENTURE	Computational mechanics system for large scale analysis and design (unstructured grid)
RSDFT	Ab-initio simulation based on density functional theory (real space DFT method)
FFB	FEM flow analysis around the complex geometry (Large Eddy Simulation)
LQCD	Lattice QCD simulation (structured grid Monte Carlo)

The Co-Design team leads to optimize architectural features and application codes together in cooperation with AICS teams and Fujitsu. It also designs and develops an application framework, FDPS (Framework for Developing Particle Simulator), to help HPC users implement advanced algorithms.

The Application Development team is a representative of nine institutions aimed at solving Priority Issues. The team figures out weakness of target application codes in terms of performance and utilization of hardware resources and discusses them with AICS teams and Fujitsu to find out best solutions of architectural features and improvement of application codes.

23.3 Co-design with applications from Priority Issues

During the development stage, emphasis will be placed on co-design concepts that allow the system and application development processes to work harmoniously together so that the societal and scientific issues can be solved effectively and early achievements can be attained. Since increasing power consumption is a critical issue in the design of the next-generation large-scale supercomputer, it will be important for the co-design to make trade-offs between energy/power, cost, and performance by taking application characteristics into consideration. It should be noted that we design the entire system using a proprietary processor and interconnect, in collaboration with Fujitsu, rather than choosing existing components to construct the system.

Table 23.4 shows a set of target applications provided from each area of the nine Priority Issues. System design is being performed using these applications as a target in the co-design process.

23.4 Target of System Development and Basic Design

The post K's design targets are as follows:

- A one hundred times speed improvement over the K computer is achieved in maximum case of some target applications. This will be accomplished through co-design of system development and target applications for the nine Priority Issues.
- The maximum electric power consumption should be between 30 and 40 MW.

In 2014, the basic design of the system was completed. The AICS decided to adopt a general-purpose many-core architecture rather than using accelerators, such as graphics processing units (GPUs), in order to support a wider range of applications. At the time of writing, Fujitsu has just announced that the node processor is based on an ARM architecture with Fujitsu HPC extensions that was invented for the K computer. To ensure the performance of large-scale parallel applications remained compatible with the K computer, we choose a similar topology for the interconnect, Tofu interconnect (6D mesh/torus).

The major components of system software shown in Figure 23.1 have been decided to be developed. These are summarized as follows:

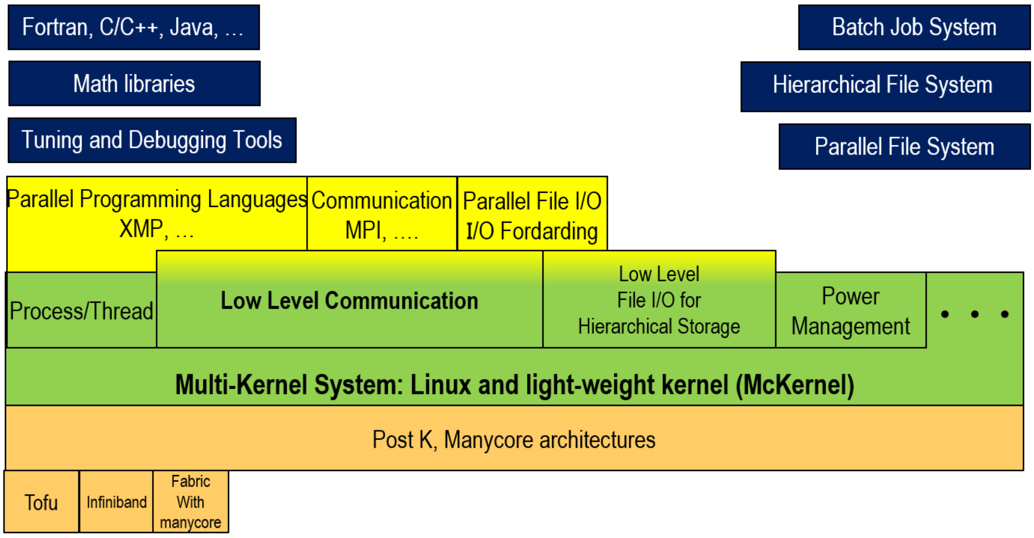

Figure 23.1: Software Architecture

- Highly productive programming language, XcalableMP
  XcalableMP (XMP) is a directive-based PGAS language for large scale distributed memory systems that combines HPF-like concept and OpenMP-like description with directives. Two memory models are supported: global view and local view. The global view is supported by the PGAS feature, i.e., large array is distributed to partial ones in nodes. The local view is provided by MPI-like + Coarray notation.

- Domain specific library/language, FDPS
  FDPS is a framework for the development of massively parallel particle simulations. Users only need to program particle interactions and do not need to parallelize the code using the MPI library. The FDPS adopts highly optimized communication algorithms and its scalability has been confirmed using the K computer.

- MPI + OpenMP programming environment
  The current de facto standard programming environment, i.e., MPI + OpenMP environment, is supported. Two MPI implementations are being developed. Fujitsu continues to support own MPI implementation based on the OpenMPI. RIKEN is collaborating with ANL (Argonne National Laboratory) to develop MPICH, mainly developed at ANL, for post K computer.

- New file I/O middleware
  The post K computer does not employ the file staging technology for the layered storage system. The users do not need to specify which files must be staging-in and staging-out in their job scripts in the post K computer environment. Asynchronous I/O and caching technologies are fully utilized in the post K computer in order to provide transparent file access with better performance.

- Application-oriented file I/O middleware
  In scientific Big-Data applications, such as real-time weather prediction using observed meteorological data, a rapid data transfer mechanism between two jobs, ensemble simulations and data assimilation, is required to meet their deadlines. Keeping a file I/O interface, such as netCDF, direct data transfer between two jobs without involving a storage system is being designed and implemented.

- Multi-Kernel for manycore architectures
  Multi-Kernel, Linux with light-weight Kernel (McKernel) is being designed and implemented. It provides: i) a noiseless execution environment for bulk-synchronous applications, ii) ability to easily adapt to new/future system architectures, e.g., manycore CPUs, a new process/thread management, a memory management, heterogeneous core architectures, deep memory hierar-

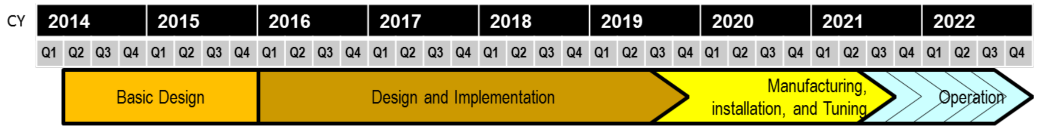

Figure 23.2: Schedule

chy, etc., and iii) ability to adapt to new/future application demands, such as Big-Data and in-situ applications that require optimization of data movement.

It should be noted that these components are not only for post K computer, but also for other manycore-based supercomputer, such as Intel Xeon Phi.

23.5 Schedule

At the time of writing, the new development plan has been announced by the government. As shown in Figure 23.2, the design and prototype implementations will be done before the end of 2019, and the system will be deployed after this phase. The service is expected to start public operation at the range from 2021 to 2022.

23.6 Members

Primary members are only listed.

23.6.1 System Software Development Team

Yutaka Ishikawa (Team Leader)

Masamichi Takagi (Senior Scientist)

Atsushi Hori (Research Scientist)

Balazs Gerofi (Research Scientist)

Masayuki Hatanaka (Research & Development Scientist)

Takahiro Ogura (Research & Development Scientist)

Tatiana Martsinkevich (Postdoctoral Researcher)

Fumiyoshi Shoji (Research & Development Scientist)

Atsuya Uno (Research & Development Scientist)

Toshiyuki Tsukamoto (Research & Development Scientist)

23.6.2 Architecture Development

Mitsuhisa Sato (Team Leader)

Yuetsu Kodama (Senior Scientist)

Miwako Tsuji (Research Scientist)

Hidetoshi Iwashita (Research & Development Scientist)

Jinpil Lee (Postdoctoral Researcher)

Tetsuya Odajima (Postdoctoral Researcher)

Hitoshi Murai (Research Scientist)

Toshiyuki Imamura (Research Scientist)

23.6.3 Application Development

Hirofumi Tomita (Team Leader)

Yoshifumi Nakamura (Research Scientist)

Hisashi Yashiro (Research Scientist)

Seiya Nishizawa (Research Scientist)

Hiroshi Ueda (Research Scientist)

Yukio Kawashima (Research Scientist)

Naoki Yoshioka (Research Scientist)

Yiyu Tan (Research Scientist)

Soichiro Suzuki (Research & Development Scientist)

Kazunori Mikami (Research & Development Scientist)

23.6.4 Co-Design

Junichiro Makino (Team Leader)

Keigo Nitadori (Research Scientist)

Yutaka Maruyama (Research Scientist)

Takayuki Muranushi (Postdoctoral Researcher)

23.7 Publications

Journal Articles

[1] A. Chambers et al. "Disconnected contributions to the spin of the nucleon". In: *Physical Review D* 91:113006 (2015).

[2] F-.K. Guo et al. "Electric Dipole Moment of the Neutron from 2+1 Flavor Lattice QCD". In: *Physical Review Letter* 115:062001 (2015).

[3] Y. Komura. "GPU-based cluster-labeling algorithm without the use of conventional: Appplication to the Swedsen-Wang muti-cluster spin flip algorithm". In: *Computer Physics Communication* 194 (2015), pp. 54–58.

[4] Y. Komura. "High-precision Monte Carlo Simulation of the Ising Models on the Penrose Lattice and Dual Penrose Lattice". In: *Journal of Physical Society of Japan* 185(4):044004 (2015).

[5] Y. Komura. "Improved CUDA programs for GPU computing of Swendsen-Wang multi-cluster spin flip algorithm: 2D and 3D Ising, Potts, and XY models". In: *Computer Physics Communication* 200 (2015), pp. 400–401.

[6] Y. Komura. "Multi-GPU-based Swendsen-Wang multi-cluster algorithm with reduced data traffic". In: *Computer Physics Communication* 195 (2015), pp. 84–94.

[7] Y. Komura. "OpenACC programs of the Swndsen-Wang muti-cluster spin flip algorithm". In: *Computer Physics Communication* 197 (2015), pp. 298–303.

[8] Takayuki Muranushi and Junichiro Makino. "Optimal Temporal Blocking for Stencil Computation". In: *Procedia Computer Science* (2015).

[9] J. Najiar et al. "Lattice determination of Sigma-Lambda mixing". In: *Physical Review D* 91:074512 (2015).

[10] P.E.Shanahan et al. "Charge symmetry violation in the eletromagnetic from factors of the nucleon". In: *Physical Review D* 91:113006 (2015).

Conference Papers

[11] Balazs Gerofi and Masamichi Takagi and Yutaka Ishikawa and Rolf Riesen and Evan Powers and Robert W. Wisniewski. "Exploring the Design Space of Combining Linux with Lightweight Kernels for Extreme Scale Computing". In: *International Workshop on Runtime and Operating Systems for Supercomputers (ROSS), held in conjunction with ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)*. Best Paper Award. 2015.

[12] Balazs Gerofi and Takagi Masamichi and Yutaka Ishikawa. "Toward Operating System Support for Scalable Multithreaded Message Passing". In: *21th European MPI Users' Group Meeting (EuroMPI)*. 2015.

[13] Masamichi Takagi and Norio Yamaguchi and Balazs Gerofi and Atsushi Hori and Yutaka Ishikawa. "Adaptive Transport Service Selection for MPI with InfiniBand Network". In: *International Workshop on Exascale MPI (ExaMPI)*. 2015.

[14] Min Si and Antonio J Pena and Jeff Hammond and Pavan Balaji and Masamichi Takagi and Yutaka Ishikawa. "Casper: An Asynchronous Progress Model for MPI RMA on Many-Core Architectures". In: *2015 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. 2015, pp. 665–676.

[15] Min Si and Antonio Pena and Jeff Hammond and Pavan Balaji and Yutaka Ishikawa. "Scaling NWChem with Efficient and Portable Asynchronous Communication in MPI RMA". In: *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*. Award finalist. 2015, pp. 811–816.

[16] Takayuki Muranushi, Keigo Nitadori, and Junichiro Makino. "Nanoblock Unroll: Towards the Automatic Generation of Stencil Codes with the Optimal Performance". In: *WOSC '14 Proceedings of the Second Workshop on Optimizing Stencil Computations*. 2015.

[17] Rolf Riesen and David N. Lombard and Kurt Ferreira and Robert W. Wisniewski and Arthur (Barney) Maccabe and John (Jack) Lange and Mike Lang and Ron Brightwell and Balazs Gerofi and Kevin Pedretti and Pardo Keppel and Todd Inglett and Yoonho Park and Yutaka Ishikawa. "What is a Lightweight Kernel?" In: *International Workshop on Runtime and Operating Systems for Supercomputers (ROSS), held in conjunction with ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)*. 2015.

Invited Talks

[18] Yutaka Ishikawa. *An Overview of the Next Flagship Supercomputer in Japan*. CCP 2015. 2015.

[19] Yutaka Ishikawa. *Codesign of System Software in Post K*. ISC 2015. 2015.

[20] Yutaka Ishikawa. *System software in post K supercomputer*. SC 2015. 2015.

[21] Mitsuhisa Sato. *Challenges for Parallel Programming Models and Languages of post-petascale and exascale computing*. IWOMP 2015. 2015.

[22] Mitsuhisa Sato. *FLAGSHIP 2020 project –Development of Japanese National Flagship supercomputer "post K" –*. US-Japan Joint Institute for Fusion Theory Workshop on Innovations and co-designs of fusion simulations towards extreme scale computing. 2015.

[23] Mitsuhisa Sato. *FLAGSHIP 2020 project –Development of Japanese National Flagship supercomputer "post K" –*. The satellite symposium of ICQC 2015. 2015.

Patents and Deliverables

[24] *McKernel: Light-weigh kernel with Linux.*

[25] *OFI/LLC: Communication library for high performance computing.*