Chapter 23

HPC Usability Development Unit

23.1 Members

Fumiyoshi Shoji (Unit Leader)

Masaaki Terai (Research & Development Scientist)

Hitoshi Murai (Research Scientist)

Jorji Nonaka (Technical Scientist)

Tomohiro Kawanabe (Technical Staff I)

Hiroshi Harada (Technical Scientist)

Hidetomo Kaneyama (Technical Staff I)

Megumi Takeuchi (Technical Staff I)

Kenji Ono (Senior Visiting Scientist)

Naohisa Sakamoto (Visiting Scientist)

Kazuki Koiso (Student Trainee)

Yoshiaki Yamaoka (Student Trainee)

23.2 Overview of Research Activities

The HPC Usability Development Unit has covered a wide range of topics including software, service, and infrastructure, aiming to improve the overall usability of the K computer environment, and of its successor (Fugaku). In addition, we also have some activities dedicated to one of the HPCI shared storage sites, located at the R-CCS, which is a commissioned project funded by the MEXT (Ministry of Education, Culture, Sports, Science and Technology). The official operation of the K computer was finished at August 16, 2019. As a consequence, we have joined in several projects related to the Fugaku operation and its application, and have aimed to prepare a useful environment for the Fugaku users. Targeting the start of the official service of Fugaku, we continued working on the already ongoing projects, and have focused on other related and relevant activities.

The summaries of our activities are listed below.

- **Software Center:** We have been running the *RIKEN R-CCS Software Center* to support and promote a set of software, also called "R-CCS software", developed at the R-CCS.
- **Large Data Visualization and Analysis:** We have worked on tools and applications for assisting the visualization and analysis of large data sets generated from the K computer environment, which include traditional simulation results and the log data sets from the K computer facility and system. We have also worked on a scalable display system, named ChOWDER, for displaying high-resolution visualization results or simultaneous display of multiple data sets.

- **Workflow Management Software (WHEEL):** We have been working on WHEEL, a workflow management software. WHEEL was originally developed by the former Advanced Visualization Research Team, and continues being developed to enable more efficient capacity computing. In FY 2019, we have improved WHEEL in its usability, and it was used in trials by multiple research institutes, and verified the usefulness for capacity computing workflows in combination with practical application software.

- **K Pre-Post Cloud (Data Analysis Server using Virtualization Technology):** Data analysis has become an important topic in the field of HPC/Data Centers. However, the existing data analysis servers in the K computer environment are disproportionately small compared to the K compute nodes. Nowadays, the virtualization technology has sufficiently matured and can provide an ideal software environment for the users. Therefore, to accelerate the data analysis without compromising the usability, we have considered the application of OpenStack-based virtualization technology for data analysis and worked on the design and deployment of an experimental cloud system.

- **HPCI Shared Storage:** In the commissioned project of the HPCI shared storage, we have collaboratively worked with the University of Tokyo, which maintains the other shared storage. We have operated in the data-duplication mode to provide high availability, and to avoid the interruption of the service even if a serious failure occurs at one of the sites.

- **Oracle Cloud FastConnect Service:** We have developed a new secure and high-speed network infrastructure by using the "Oracle Cloud FastConnect Service" under the collaboration with Oracle. This infrastructure will enable the Fugaku users to transfer their data between the Fugaku/HPCI shared storage and their users' environment on the Oracle Cloud Infrastructure (OCI), which is a public cloud service, via SINET (Science Information NETwork). This service will be available in FY 2020.

Furthermore, we have other activities that have been conducted in collaboration with universities and research institutes, in various fields of science and technology (e.g., Econophysics, Bioinformatics, and an analysis of the "hot-water" cooling method), and what they have in common is the investigation and exploration on how to enhance the usability, and to improve the productivity of the K computer environment. Although we have a wide variety of missions, we can say that our common objective is to improve the usability of the K computer environment as well as of the HPCI shared storage. Under this main goal, we have investigated and explored new usage for the K computer and the HPCI storage, trying to embrace not only the traditional HPC users but especially new users from various other fields. In the rest of this chapter, we will detail these aforementioned research and development projects carried out in FY 2019.

23.3 Research Results and Achievements

23.3.1 Software Center

We have been running the *RIKEN R-CCS Software Center*, since FY 2017, to support and promote *R-CCS software*, which is a set of applications, libraries, programming tools, and other software developed at the R-CCS.

The major achievements of the Center, in FY 2019, are as follows:

- Peformance tuning of NTChem

- Development of Spack recipes for some of the R-CCS software, which include FrontFlow/red–HPC, SCALE, GENESIS, and EigenEXA

In addition, an overview and some use case examples of FrontFlow/red–HPC were also provided on the portal site, to promote its recognition and increase its usability.

23.3.2 Large Data Visualization and Analysis

The K computer environment includes pre- and post-processing oriented systems (K Pre/Post Server and K Pre-Post Cloud), in addition to the K computer itself, as shown in the Fig. 23.1. The target data for visualization and analysis includes those from the large-scale simulation as well as those generated from the K computer system and its facility, which are stored as log data [2]. Thanks to the availability of Mesa3D library, with its OSMesa (Off-screen Mesa) functionality, which was provided via OSS (Open-Source Software) porting effort done in FY 2018, we started working on some applications and tools by using the KVS (Kyoto Visualization Software) library for both simulation data visualization and log data analysis. In order to facilitate the visual perception of high-resolution rendering results, we have extended our work on the ChOWDER, a VDA (Virtual Display Area) based scalable display system, for increasing its functionality and usefulness.

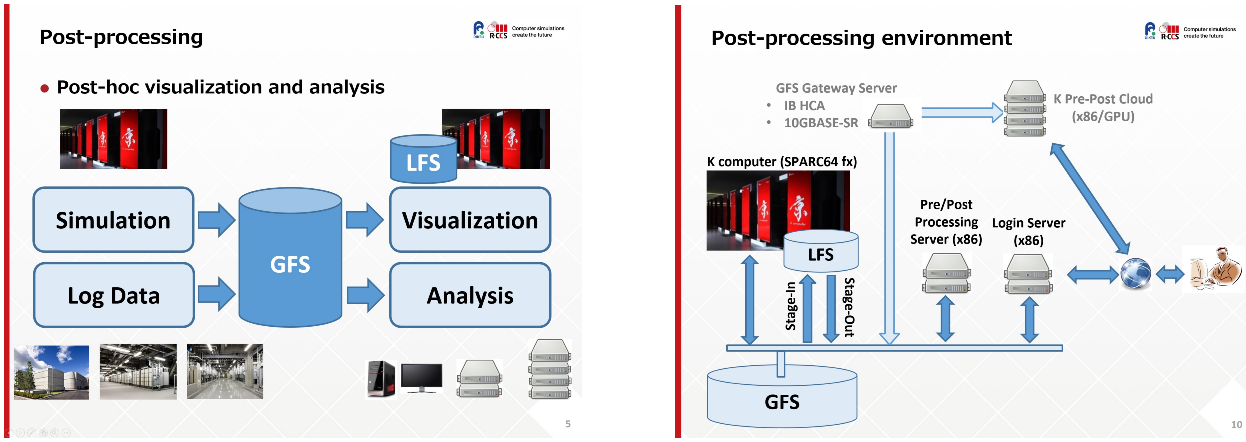

Figure 23.1: Post-processing environment for large data visualization and analysis.

23.3.2.1 Simulation Data

In addition to the traditional post-hoc visualization approach by reading and processing the data sets already stored in the disk, we have also developed a KVS-based in-situ visualization API for coupling to the simulation codes, in order to enable in-situ visualization processing simultaneously with the simulation. We have also proposed an adaptive time-step sampling method, based on the simulation data results, in order to improve the rendering throughput and reduce unnecessary data I/O, and we evaluated this in-situ API by integrating with OpenFOAM CFD simulation code [3, 4, 5, 11]. As an approach for analyzing ensemble computational climate simulation results, we have worked on a visualization approach which uses stochastic isosurface technique [6, 12].

23.3.2.2 Log Data

We have also worked on some visual data analysis techniques focusing on the operational improvements of the system and its facility. We have developed a KVS-based visual causal exploration framework to investigate the applicability of transfer entropy technique for analyzing the causality of hardware failure on the K computer system [13, 15]. We have also investigated the applicability of this visual exploration framework for the causal analysis of a severe rainfall event, simulated on the K computer, and a flash flood in a river situated at the Kobe-city [7, 14]. We have also analyzed some log data sets generated from experiments and from the K computer system and its facility, by using traditional Python-based data science tools, aiming to investigate the temperature changing behavior on air-cooled CPU [9] and compute rack [17, 18].

23.3.2.3 Scalable Display System (ChOWDER)

We are advancing the research and development of a scalable display system which was named ChOWDER (COoperative Workspace DrivER) [23]. It is also known as tiled display, which is a method that tiles multiple physical display devices to create one large pixel space. In recent years, it becomes possible to build tiled displays that connects multiple physical displays onto a single PC by using high-performance graphics cards. However, it can have limitations on the maximum achievable resolution, since it depends on the hardware and its middleware. In contrast, since ChOWDER is a software-based system, there is no logical upper limit regarding the maximum resolution it can provide. Since ChOWDER is a web-based software and requires only standard web browser, it can provide high availability. In addition, the server has a virtual pixel space called VDA (Virtual Display Area), and it enables the coexistence of tiled displays with physical displays possessing different resolutions and aspect ratios. It is also possible to use the ChOWDER as a remote collaboration tool for displaying the same content on the tiled displays placed at multiple sites thanks to the web-based high availability functionality.

In FY 2019, to improve the display efficiency of ultra-high-resolution images on ChOWDER, we have introduced a new feature that split the visualization data at the sender-side, and send them to each corresponding display device. By using this feature, and in our tiled display environment, we have observed that the displaying speed of 16K resolution images can become up to 2.7 times faster [10, 20].

23.3.3 Workflow Management Software (WHEEL)

To easily handle many computational tasks on the HPC resources, we have developed a scientific workflow management system named WHEEL (Workflow in Hierarchical distributEd parallEL) [24] jointly developed with Kyusyu University. WHEEL can support typical HPC simulation with a workflow consisted of pre-processing, simulation, and post-processing and can work on various HPC computing resources including not only the K computer but also other computing resources. In addition, this system is a web-based application written in JavaScript, and runs on common web browsers available for Windows, macOS, and Linux. WHEEL employs ordinary SSH and login shell to communicate with remote computers when it performs job submitting, monitoring, and file transfer. Therefore, the additional middleware is not required on the remote computers. In addition, WHEEL has the iteration/conditional branch functionality (e.g., For, Foreach, While, and If), and a parameter study functionality that allows multiple simulations simultaneously with the given parameter range specified by the user. By both features, the users can construct more complex workflows.

In FY 2019, WHEEL has been experimentally used in several research institutes and has been improved by their feedbacks. For example, at Kyushu University, WHEEL has been used for the VASP structural relaxation calculations for inorganic materials, and was performed with for capacity computing using the supercomputer ITO. At the University of Tokyo, WHEEL was also used to implement a shape optimization workflow for designing rotating machines. This workflow realizes two-objective optimization to maximize the performance and noise minimization of the rotating machine based on evolutionary calculations using genetic algorithms. In this workflow, WHEEL itself runs on a Linux PC, and the wing geometry creation and mesh generation programs are executed on a cluster computer owned by the University of Tokyo, and fluid dynamics and acoustic calculations are run on Kyushu University's ITO. Due to the limitation of the time period, the workflow running has not finished, however, this use case indicated the capability of WHEEL which is able to handle complex workflow among multiple computational resources.

23.3.4 K Pre-Post Cloud (Data Analysis Servers using Virtualization Technology)

The K Pre-Post Cloud is a private cloud system/service, and an experimental platform in the K computer and Fugaku environment, build to enhance the pre/post data processing features including data analysis and visualization. This system consists of 11 compute nodes and some auxiliary servers. Each compute node has 48 x86-based physical cores, and 384 GiB of RAM, and connects to other nodes with a 25 GiB Ethernet network switch. In addition, the system has an external Ceph-based 150 TiB storage. To improve the usability regarding OSS environment, we applied the server virtualization technology using the OpenStack framework. By using this framework, the users are able to install and uninstall any software on their own virtualized environment, including not only the OSS but even the operating system they want to use. Furthermore, we also added GPUs on some of the compute nodes to enhance and accelerate the data processing capabilities.

In the K computer environment, there exist auxiliary x86-based data-analysis servers, known as "Pre/Post Processing Servers," where the users are allowed to submit jobs using large memory, and with longer duration than compared to those allowed on the K computer. The primary target usage of these servers is data preparation, analysis, and visualization, which are some of the essential processes needed in the early and/or late phases of a regular end-to-end scientific workflow. In addition, these servers were built using the traditional x86 architecture, in order to provide a readily available environment for the users to exploit various OSS packages.

On the other hand, the Pre/Post Processing Servers can currently be considered as outdated since they were installed several years ago. In addition, although data analysis is one of the significant topics in HPC/Data Centers, their computational resources have gradually become insufficient for the data analysis. As a result, there was a strong demand from the academic/industrial users to enhance the pre/post processing capabilities in the K computer environment. In FY 2018, to address the problem of the pre/post data processing by using OSS in the environment, we have launched a new experimental IaaS platform which was named "K Pre-Post Cloud." In FY 2019, to improve the inbound and outbound connections between the system and external sites, we changed the routing configuration that directly connects the center's network infrastructure as well as the

K computer, and enhanced the network speed with 10GbE fiber cable in the inbound. Finally, we continued to provide the service for the users.

23.3.5 HPCI Shared Storage

In FY 2019, we operated the HPCI storage system with the clear operational goals listed below.

- 99.3 % or higher read-write service utilization
- No security incidents
- No data loss

In October 2019, due to the Fugaku installation, the HPCI shared storage was taken out of the operations for a month. At the time, by using the multiplexed data mechanism between the University of Tokyo and R-CCS, the entire service continued while the R-CCS equipment was out of service. In addition, the metadata server operated by the R-CCS maintained its continuous operation by failing over to the University of Tokyo. Thanks to the effectiveness of the data multiplexing operation, there was no service outage, no data loss, no security incidents in FY 2019 [21, 22]. As a result, we achieved all operational goals planned for FY 2019.

As a new attempt, the usability of the login servers was improved, and the monitoring function was strengthened. In order to occupy/submit batch jobs, a PBS batch scheduler was installed on the four login nodes installed at the R-CCS side. The PBS batch scheduler provides the queue optimization feature for MPI and other parallel execution jobs. Further, we improved the environment for using shared storage from Singularity containers. We also enhanced the service-level monitoring to provide a visual overview of the operational status. We will continue enhancing the monitoring capabilities so that users can quickly be informed of the service status.

Figure 23.2 shows the monthly operation rate of the HPCI shared storage system.

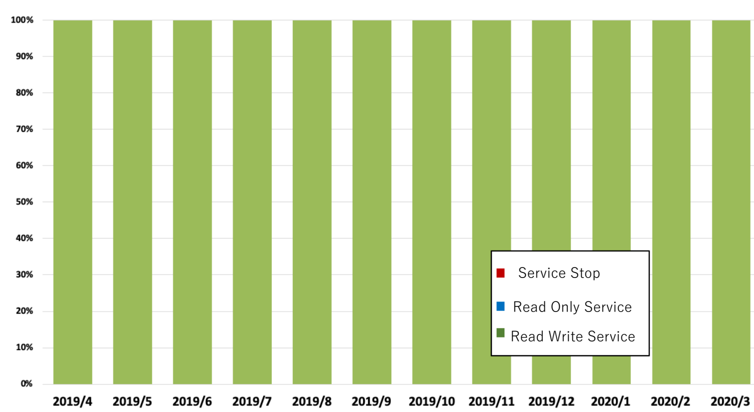

Figure 23.2: FY 2019 monthly utilization rate of the HPCI shared storage.

23.3.6 Oracle Cloud FastConnect Service

We have developed a new secure and high-speed network infrastructure by using the "Oracle Cloud FastConnect Service" under the collaboration with Oracle. This infrastructure enables the Fugaku users to transfer their data, without additional fee, between the Fugaku/HPCI shared storage and their users' environment on the Oracle Cloud Infrastructure (OCI), which is as a public cloud service, via SINET (Science Information NETwork). This service will be available in FY 2020.

As a background, various cloud services (e.g., data store, visualization) has recently matured sufficiently, and in the next decade, we expect that the Fugaku users will normally utilize the services in their research. Thus, one of the objectives for this infrastructure development is to promote the use of cloud services by the Fugaku users in order to improve usability in the supercomputing environment.

In FY 2019, we built and tested the service. The activities include many tasks: designing the service environment; designing the configuration of R-CCS main routers; setting up OCI tenant; create a user manual; and testing the network and service. Finally, we have connected the R-CCS and OCI data center with a private line, and have provided OCI tenant for service management. The overall of the new infrastructure is shown in Fig.23.3. This service is expected to be available to the users in FY 2020.

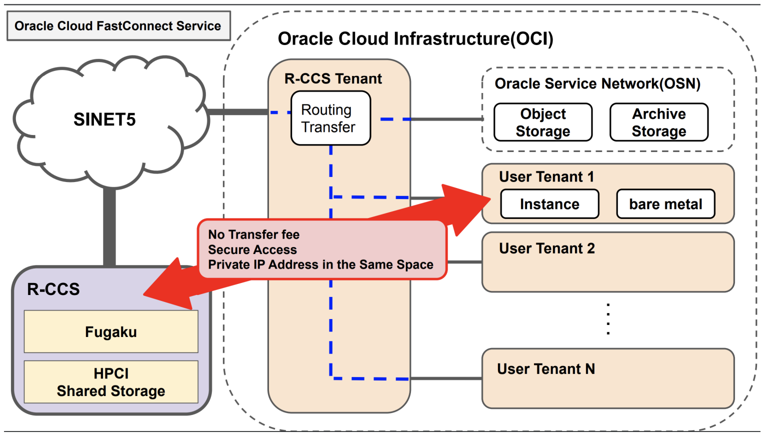

Figure 23.3: Oracle Cloud FastConnect Service Overrall

23.3.7 Other Activities

In addition to those activities detailed in the previous subsections, we have also contributed to the following projects:

- Graph-based Economic Simulation under the Exploratory Challenges on Post-K computer (Studies of Multi-level Spatiotemporal Simulation of Socioeconomic Phenomena, Macroeconomic Simulations) [1]
- To share knowledge and actual usage regarding cutting-edge cloud resources in Bioinformatics and HPC, we started to collaborate with the Laboratory for Bioinformatics Research at the RIKEN Center for Biosystems Dynamics Research (BDR)
- Systematic and quantitative analysis of the “hot-water” cooling method, trying to better understand the impact of the higher cooling liquid temperature on the system and its facility. The investigations include the use of a portable water circulation equipment with chiller and heater [16], and Modelica-based cooling facility simulations [8, 19]

23.4 Schedule and Future Plan

FY2019 was a milestone year for us because the official operation of the K computer was finished in August, and was started the installation of the Fugaku, which is the successor of the K computer. Many of our members have been involved in the Fugaku project to prepare an easy-to-use environment and services for the Fugaku. In FY2020, we will have to accelerate our activities because the early access program for the Fugaku is planned to start before the beginning of the official operation. We will continue doing our best to investigate and explore new usage modes for the Fugaku environment, focusing not only on the traditional HPC users, but especially on potential new users from different fields. We also hope that Fugaku will produce many remarkable outcomes from the researchers and engineers, in the same manner as the K computer's impressive achievements.

23.5 Publications

23.5.1 Articles

[1] Hazem Krichene, Yoshi Fujiwara, Abhijit Chakraborty, Yoshiyuki Arata, Hirasu Inoue, Masaaki Terai, “The emergence of properties of the Japanese production network: How do listed firms choose their partners?,” Social

Networks, Vol.59, pp.1-9, (2019).

23.5.2 Oral Talks

[2] Jorji Nonaka, Yuichi Tsujita, "An Overview of the Storage and Post-Processing Environment at RIKEN R-CCS", HPC I/O in the Data Center Workshop (HPC-IODC 2019) held in conjunction with ISC 2019, Frankfurt, Germany, 2019. (Oral Talk)
[3] Yoshiaki Yamaoka, Kengo Hayashi, Naohisa Sakamoto, Jorji Nonaka, "In Situ Adaptive Timestep Control and Visualization based on the Spatio-Temporal Variations of the Simulation Results", In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV 2019) held in conjunction with SC'19, Denver, USA, 2019.
[4] Yoshiaki Yamaoka, Kengo Hayashi, Naohisa Sakamoto, Jorji Nonaka, Tsukasa Yoshinaga, Kazunori Nozaki, "Adaptive Time-Step Sampling for In-Situ Visualization", JSFM 33th CFD Symposium, Sapporo, Japan, 2019. (in Japanese)
[5] Yoshiaki Yamaoka, Kengo Hayashi, Naohisa Sakamoto, Jorji Nonaka, Tsukasa Yoshinaga, Kazunori Nozaki, "Adaptive In-Situ Visualization Considering Spatio-temporal changes of Large-scale Numerical Simulation Results", JSME 97th FED Meeting, Toyohashi, Japan, 2019. (in Japanese)
[6] Go Tamura, Naohisa Sakamoto, Yasumitsu Maejima, Jorji Nonaka, "Stochastic Isosurface Visualization for Meteorological Ensemble Data Visual Analysis", NIFS Visualization Research Meeting (VR 2019), Toki, Japan, 2019. (in Japanese)
[7] Naohisa Sakamoto, Jorji Nonaka, Yasumitsu Maejima, Koji Koyamada, "Transfer Entropy based Visual Causality Analysis of a Severe Rainfall Event", VizAfrica Data Visualization Symposium 2019, Gaborone, Botswana, 2019.
[8] Motohiko Matsuda, Hiroya Matsuba, Jorji Nonaka, Keiji Yamamoto, Hiroshi Shibata, Toshiyuki Tsukamoto, "Modeling the Existing Cooling System to Learn its Behavior for Post-K Supercomputer at RIKEN R-CCS", Energy Efficient HPC State of the Practice Workshop (EE HPC SOP 2019) held in conjunction with ICPP 2019, Kyoto, Japan, 2019.
[9] Angelo N. C. Vieira, Paulo S. S. Souza, Wagner S. Marques, Marcelo S. Conterato, Tiago C. Ferreto, Marcelo C. Luizelly, Arthur F. Lorenzon, Antonio C. S. B. Filho, Fabio D. Rossi, Jorji Nonaka, "The Impact of Parallel Programming Interfaces on the Aging of a Multicore Embedded Processor", IEEE International Symposium on Circuits and Systems (ISCAS 2019), Sapporo, Japan, 2019.
[10] Tomohiro Kawanabe, Jorji Nonaka, Daisuke Sakurai, Kazuma Hatta, Shuhei Okayama, Kenji Ono, "Showing Ultra-High-Resolution Images in VDA-Based Scalable Displays", International Conference on Cooperative Design, Visualization and Engineering (CDVE 2019), LNCS Vol. 11792, pp. 116-122, Mallorca, Spain, 2019.

23.5.3 Posters

[11] Yoshiaki Yamaoka, Naohisa Sakamoto, Jorji Nonaka, "Adaptive Spatial and Temporal Sampling for In-Situ Visualization", 2nd R-CCS International Symposium, Kobe, Japan, 2020. (Poster)
[12] Go Tamura, Yasumitsu Maejima, Naohisa Sakamoto, Jorji Nonaka, "Visual Analysis of Meteorological Ensemble Data Sets by using Stochastic Isosurface Visualization Technique", 2nd R-CCS International Symposium, Kobe, Japan, 2020. (Poster)
[13] Kazuki Koiso, Naohisa Sakamoto, Jorji Nonaka, Fumiyoshi Shoji, Keiji Yamamoto, "BiClustering and Transfer Entropy for the Visual Analysis of Critical Hardware Failures on the K computer", IEEE Pacific Visualization Symposium (PacificVis 2019), Bangkok, Thailand, 2019. (Poster)
[14] Naohisa Sakamoto, Jorji Nonaka, Yasumitsu Maejima, Koji Koyamada, "Visual Causal Exploration with Transfer Entropy Applied to a Severe Rainfall Event", IEEE Pacific Visualization Symposium (PacificVis 2019), Bangkok, Thailand, 2019. (Poster)
[15] Kazuki Koiso, Naohisa Sakamoto, Jorji Nonaka, Keiji Yamamoto, Fumiyoshi Shoji, "A Visual Causality Exploration System for HPC Hardware Failure Analysis", 2nd R-CCS international Symposium, Kobe, Japan, 2020. (Poster)
[16] Fumiyoshi Shoji, Jorji Nonaka, Motohiko Matsuda, Hiroya Matsuba, Keiji Yamamoto, Yasumitsu Maejima, Toshiyuki Tsukamoto, "CPU Water Cooling Temperature Effects on the Performance and Energy Consumption", ISC 2019 HPC in Aisa Poster, Frankfurt, Germany, 2019. (Poster)
[17] Jorji Nonaka, Keiji Yamamoto, Akiyoshi Kuroda, Toshiyuki Tsukamoto, Kazuki Koiso, Naohisa Sakamoto, "View from the Facility Operations Side on the Water/Air Cooling System of the K Computer", SC'19 Research Poster, Denver, USA, 2019. (Poster)

[18] Jorji Nonaka, Toshiyuki Tsukamoto, Motohiko Matsuda, Keiji Yamamoto, Akiyoshi Kuroda, Atsuya Uno, Naohisa Sakamoto, “A Brief Analysis of the K Computer by using the HPC Facility’s Water Cooling Subsystem”, 2nd R-CCS International Symposium, Kobe, Japan, 2020. (Poster)
[19] Motohiko Matsuda, Hiroshi Shibata, Jorji Nonaka, Toshiyuki Tsukamoto, Keiji Yamamoto, Hajime Naemura, “R-CCS Facility Simulation Modeling for Assisting Operation Planning and Decision Making”, 2nd R-CCS International Symposium, Kobe, Japan, 2020. (Poster)
[20] Tomohiro Kawanabe, Jorji Nonaka, Kenji Ono, “ChOWDER: A VDA-Based Scalable Display System for Displaying High-Resolution Visualization Results”, HPC Asia 2020, Fukuoka, Japan, 2020. (Poster)
[21] Hiroshi Harada, Osamu Tatebe, Toshihiro Hanawa, Isamu Koseda, Hidetomo Kaneyama, Noriyuki Soda, Akira Kondo, Takahiro Yugawa, “Introduction of HPCI shared storage that has achieved year-round non-stop operation”, HPC Asia 2020, Fukuoka, Japan, 2020. (Poster)
[22] Hiroshi Harada, Hidetomo Kaneyama, Chihiro Shibano “Incident and operation analysis of HPCI Shared Storage System R-CCS hub”, 2nd R-CCS International Symposium, Kobe, Japan, 2020. (Poster)

23.5.4 Software

[23] WHEEL : Scientific workflow Management System. (https://gitlab.com/aicshud/WHEEL) (Currently managed as a private repository).
[24] ChOWDER : Scalable Display System. (https://github.com/SIPupstreamDesign/ChOWDER).