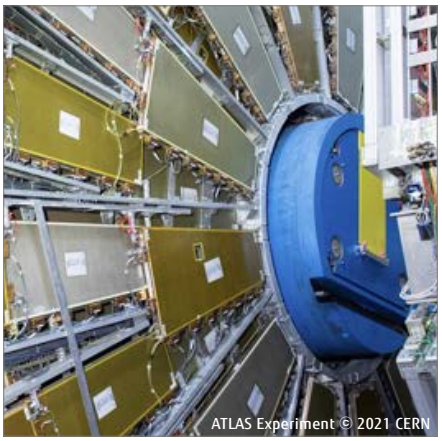
International Center for Elementary Particle Physics (ICEPP), The University of Tokyo

Exploring the Possibilities of Utilizing Supercomputers for Data Analysis with the PRIMEHPC FX700

"In the world of particle physics, researchers across the globe are striving to discover new phenomena and overwrite the Standard Model of particle physics. Meanwhile, the amount of data and computation required is growing rapidly. Introducing PRIMEHPC FX700 will not only address this challenge but will also be a milestone for the future of computing."

Professor Tanaka Junichi,
ICEPP, The University of Tokyo


ATLAS Experiment © 2021 CERN

Summary of Case Study

- **Hardware** FUJITSU Supercomputer PRIMEHPC FX700
- **Software** OS / Red Hat Enterprise Linux 8
  HPC Middleware / FUJITSU Software Compiler Package

The International Center for Elementary Particle Physics (ICEPP) at the University of Tokyo has been promoting research through international collaboration with the aim of unraveling the mysteries of elementary particles and the universe. The current research primarily focuses on ATLAS experiment, MEG experiment, and International Linear Collider (ILC) projects. Analyzing a large amount of data involved in these projects requires extensive computing power: in particular, the ATLAS experiment already has managed massive amounts of data. The size of data is expected to grow according to upgrades of accelerators and detectors. The ICEPP has introduced the PRIMEHPC FX700 as part of an effort to expand its computational resources to prepare for such a growth, as well as for the possible utilization of supercomputers such as "Fugaku" in the future. Currently, operational verification of software used in data analysis is in progress.

Challenge	Selection Key Point
A computational resource capable of performing a significant number of calculations for data analysis in the ATLAS experiment was required. Preparation for future data growth was desired.	PRIMEHPC FX700, which is a system similar to Fugaku, can handle the necessary amount of computations. PRIMEHPC FX700 also offers the opportunity to open up the possibility of utilizing Fugaku and other supercomputers, for example, at the Information Technology Center of the University of Tokyo.
Conventionally, there had been no general-purpose CPUs to choose from other than those from x86 series. Moreover, GPU supercomputers can be quite difficult to handle and, in some cases, impractical for software applications used in the ATLAS experiment.	The A64FX processor, which Fugaku and PRIMEHPC FX700 are based on, may be an option because it is a general-purpose processor that combines ease of use and high performance.

Background

The ATLAS Experiment Requires Massive Numbers of Computations

The European Organization for Nuclear Research (CERN) is an international particle physics laboratory located in Switzerland. There, experiments in which protons are collided using the Large Hadron Collider (LHC) placed 100m below ground are conducted. The particles produced from the collisions are measured with the ATLAS detector.

The ATLAS experiment is a project that conducts studies based on this. Approximately 3,000 researchers from around the world, including the ATLAS Japan group that consists of Japanese universities and researchers, 35 of whom are members of the ICEPP, are involved in the project.

Professor Tanaka Junichi, one of the researchers, explains, "We are aiming to discover new phenomena by analyzing data obtained from experiments and simulations."

The challenge is the massive number of computations required, with an accumulated data volume amounting to 500 petabytes. Research institutes around the world share this large volume of data, 10 petabytes of which is managed by the ICEPP. "Upgrades of the LHC and ATLAS detectors have being conducted between 2019 and 2021, and we expect that, starting 2022, the volume of data obtainable will be two to three times that of the present. We also expect the High Luminosity LHC will start operation in 2027, which will lead to a further data increase, resulting in the volume of data necessary to be handled expanding by an order of exabytes. Therefore, we hope to prepare computational resources about 100 times greater than what we have today." (Tanaka)

The ICEPP operates an ATLAS regional analysis center, one of the computing centers for ATLAS experiments, and provides the Worldwide LHC Computing Grid (WLCG) with computation resources (Note 1). However, even these computational resources will eventually be insufficient considering how rapidly the volume of data is increasing.

Note 1: The computing system at the ATLAS regional analysis center has been in operation since 2007, and is available to the entire ATLAS Japan Group, including non-ICEPP, members. WLCG is a computing grid that consists of computing centers around the world for the purpose of analyzing huge amounts of data produced in ATLAS experiments and the like.

Key Points

Hopes to Expand Research Possibilities by Utilizing Supercomputers

Professor Tanaka Junichi, IICEPP, The University of Tokyo

"Supercomputers have not been used so much in the ATLAS experiment" explains Project Assistant Professor Kishimoto Tomoe. Unlike other research fields where large amounts of supercomputer resources and time are used to solve one single problem, the data analysis of the ATLAS experiment involves a large number of smaller independent computations. Each computation can be conducted with only several gigabytes of memory and takes just 10 minutes or so, but such computations must be conducted several hundreds of millions or billions of times.

In other words, each computation in this field require fewer cores and less memory, but do require a massive number of total jobs to be conducted. It is this characteristic that makes it incompatible with conventional supercomputers. The ICEPP has gathered a large number of standard commercial servers at the ATLAS regional analysis center and has been operating them efficiently.

However, Tanaka and Kishimoto began to think they should introduce supercomputers to this field too.

There are two main reasons for this; one is that current supercomputers have begun to focus on the integration of AI and big data. Adopting this idea, the Information Technology Center at the University of Tokyo (Note 2), has already installed the Wisteria/BDEC-01 built with the Fujitsu PRIMEHPC FX1000. The other reason is their desire to have more options in computation resources as they do not know which technology will become the mainstream in computing around the world in the future.

In this context, Tanaka and Kishimoto came to focus on the PRIMEHPC FX700. "I first saw it in an advertisement from an SIer (or system integrator) of HPC Systems. I thought that if we were to purchase something available on the market, this was it." (Kishimito) The A64FX processor is suitable for data analysis, and has sufficient computing power, as it has been designed for the supercomputer Fugaku.

Another attraction was that it was built with Arm architecture. "I have been interested in Arm architecture for a long time. I think it was fortunate that the Fugaku, which uses it, appeared and became the world's fastest supercomputer. Many overseas researchers are interested in the Fugaku, and using PRIMEHPC FX700, which is of the same system, will have a positive impact on collaboration with researchers, too." (Tanaka)

Note 2: A facility for basic research to promote information processing related to education, research, and social contribution inside and outside of the University of Tokyo. It consists of four divisions: the Campus-wide Computing Research Division, the Academic Information Science Research Division, the Network Research Division, and the Supercomputing Research Division.

Software Validation Is an Immediate Aim

The immediate purpose of purchasing the PRIMEHPC FX700—with the long-term purpose being to expand computational resources—was to verify the operation of the analysis software programs used in the ATLAS experiment. The configuration of PRIMEHPC FX700, therefore, was kept to a minimum, such as the non-use of InfiniBand.

"We are currently running the PRIMEHPC FX700 ourselves, and optimizing and compiling the analysis software programs while checking their compatibility with A64FX." (Kishimoto)

The hard part of this task is that there are dozens of sections of the analysis software that depend on external libraries, all of which must be compiled for use with A64FX.

Having said that, the PRIMEHPC FX700 is very easy to use as it consists only of the A64FX general-purpose CPU, making it fast and easy to support software. "With supercomputers with a huge number of GPUs, their structure is different so we need to change the software. This makes it almost impossible to compile existing analysis programs. But with A64FX, it's possible. In addition, A64FX-based supercomputers have a high potential to be adopted in many universities in the future." (Kishimoto)

Explains Project Assistant Professor Tomoe Kishimoto, ICEPP, The University of Tokyo

Tanaka and Kishimoto say they find benefits in using PRIMEHPC FX700 other than computing performance. "It's compact enough to fit in a standard rack, easy to use, and reasonably priced." (Kishimoto)

They particularly have high expectations of its power consumption. As the world strives to create sustainable societies, low power consumption is essential. Kishimoto is testing its power consumption as well as the operational verification of software.

Effect and Future Outlook

A64FX-Based Computers Expected to Spread Globally and Establish a Solid Presence

There is a piece of software that has already completed an operational verification test. That is Geant4—a simulation program for radiation simulation jointly developed by researchers from international research institutes including CERN. "Now, it has become possible to visualize its performance in numbers. We have shared this information with other researchers involved in the ATLAS experiment, including those overseas." (Kishimoto)

Tanaka and Kishimoto are also considering using Fugaku in the future. Since "Fugaku" has attracted much interest of researchers as the world's fastest number one supercomputer, they have high hopes in connecting it to the international computing grid and put to use in collaborative researches.

"We hope Fujitsu will spread their A64FX-based products to the world." (Tanaka) There are no national borders in the world of physics. Fujitsu's cultivation of technologies, that are used widely around the globe, will lead to it being a presence in the world of research too.

Profile

International Center for Elementary Particle Physics (ICEPP)

Established in 1974 (as the high energy physics laboratory of the university's faculty of science; reorganized as a campus-wide center in 2004)

Overview : Functions as a research center in Japan for international collaborative projects including the ATLAS experiment, the MEG experiment, and the ILC projects.
Website : https://www.icepp.s.u-tokyo.ac.jp/en/

HPC SYSTEMS Inc

Address : 8F LOOP-X, 3-9-15 Kaigan, Minato-ku, Tokyo
Website : https://www.hpc.co.jp

FUJITSU

Website: www.fujitsu.com