

National Institute of Advanced Industrial Science and Technology Accelerating the societal implementation of AI with AI Bridging Cloud Infrastructure (ABCI) Realizing world-class performance and energy saving with 4,352 GPUs

Under the slogan "Bring Technology to society," the National Institute of Advanced Industrial Science and Technology (AIST) aims to contribute to the development of Japanese society and industry. In August 2018, AIST launched its AI Bridging Cloud Infrastructure (ABCI), a large-scale, AI cloud computing system that general companies can utilize to accelerate the societal implementation of AI. ABCI consists of 4,352 graphics processing units (GPUs), making both world-class performance and energy saving a reality. The adoption of off-the-shelf hardware has also enabled AIST to keep service costs low while allowing users to deploy their software as is. By applying the know-how of Fujitsu in the construction and operation of large-scale systems and the AI expertise of Fujitsu Laboratories Ltd., it offers the ABCI construction, operational and maintenance support required to respond to corporate challenges.

Main product FUJITSU Server PRIMERGY CX2570 M4

Client feedback

"At AIST, unlike university supercomputer centers, we do not have an organization that specializes in the operation of supercomputers. In April 2017, we began operating AIST Artificial Intelligence Cloud (AAIC), which focused mainly on research with medium-scale HPC systems, but was limited to internal use and joint research use. Providing a large-scale HPC system such as ABCI to thousands of users across industry, academia and government was an unknown area for us. Fujitsu has plenty of experience in deploying HPC systems, so we were able to launch the ABCI service on schedule, despite the short construction schedule."

Hiroataka Ogawa, Team Leader, Artificial Intelligence Cloud Research Team, Artificial Intelligence Research Center (AIRC), AIST

Challenges and benefits

- 1 Realize an environment that offers world-class performance and is easy for general companies to use.
 - » Off-the-shelf hardware installed on large scale, allowing users to use in-house software as is. Keeps costs down, offers world-class performance.
- 2 Construct a high-quality, large-scale AI cloud computing system in a short period of time.
 - » High quality and short delivery time ensured by utilizing Fujitsu's vast experience in deploying HPC systems and AI expertise of Fujitsu Laboratories Ltd.
- 3 Promote advancement of technological challenges by providing an environment where cutting-edge AI research can be carried out.
 - » Fujitsu Laboratories Ltd. achieved the world's fastest speed for deep learning software using image data with ABCI*, and ranks among the world's best in various other benchmarks.

*Current as of March 26, 2019 based on in-house research



Background

"AI Bridging Cloud Infrastructure" aims to return new value to society via AI

Since its establishment in 2001, AIST has been carrying out research activities in a wide range of industrial technologies under the slogan "Bring technology to society." As one of Japan's largest public research institutes, it is focusing on the creation and practical realization of technologies useful to Japanese industry and society, and on "bridging" the gap between innovative technological seeds and commercialization. AIST plays a central role in constructing the national system required for Japan to continue creating innovation.

AIST hopes to build a sustainable society by focusing on its three research and development pillars—"Green Technology for a rich and eco-friendly society," "Life Technology for healthy, safe, and secure living," and "Information Technology for a super smart society." One of the key themes under Information Technology is artificial intelligence (AI) technology that creates value from big data. Hiroataka Ogawa, leader of the Artificial Intelligence Cloud Research Team at AIST's Artificial Intelligence Research Center (AIRC), explains some of the challenges related to the promotion of AI in Japan.

"According to results in the Survey Report on Promotion of AI Implementation in Society published by the Information-technology Promotion Agency, Japan (IPA) in 2018, about 10% of companies have already deployed AI, including proof of concept, into their business operations. Meanwhile, 80% of companies say they still want to do so. The problem is that there has been no response to this 80%. Many data centers and other operators in Japan are currently refraining from investing heavily because the number of users they can expect is unknown." Dr. Ogawa continues, "AIST constructed ABCI* in line with the Project for Development of Global Research Centers for Artificial Intelligence promoted by the Ministry of Economy, Trade and Industry."

*ABCI: <https://abci.ai/>



Hiroataka Ogawa, Team Leader, Artificial Intelligence Cloud Research Team, Artificial Intelligence Research Center (AIRC), AIST Director of AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory (RWBC-OIL)

Key points

Procurement with a practical perspective

ABCI is intended not only for research institutes, but also general companies, and services place great importance on providing a user-friendly system to humans anytime, anywhere. Low costs and a user-friendly environment promote ease of use. "Thanks to its general-purpose, commercial off-the-shelf (COTS) hardware and software, ABCI allows users to apply in-house software while taking advantage of the system, and for a relatively low fee," explains Dr. Ogawa. The challenge is to maintain low costs and high performance while conducting world-class AI research. Dr. Ogawa says that, in the ABCI procurement, the specifications were prepared so that better proposals within the published estimated

costs could get rated higher.

"In machine learning, in addition to performing iterative calculations until convergence occurs, repeating trial and error is required because we need to improve the accuracy of machine learning predictions. Computing power is the key. In the specifications, we prescribed the minimum requirements, but we newly developed the benchmarks through investing the big data processing and machine learning requirements expected of ABCI based on actual usage scenarios. We always planned to select a vendor with high technical abilities, but also based on the actual usage conditions of ABCI. In the case where the proposed specifications are the same, we planned to select the vendor that achieved the higher score in the benchmarks. We were able to make our decision from a practical perspective."

Based on the comprehensive evaluation of ABCI specifications by AIST and competitive bidding, Fujitsu was selected as the provider to fill the order for AIST.

System overview

4,352 GPUs, world-class performance and energy saving achieved

After Fujitsu was awarded the contract in September 2017, a data center was completed in January 2018. The system was completed in June of that year. Trial operations began in July, full operations in August.

ABCI combines high integration and a power-saving design with high-performance architecture thanks to Fujitsu's PRIMERGY CX2570 M4 high-density multi-node server. Two Intel® Xeon® Gold processors installed per server (totaling 2,176 CPUs) and four NVIDIA® V100 GPU computing cards installed per server (totaling 4,352 GPUs) produce a total FP16 (half precision floating point) theoretical peak performance of 550 PFLOPS, which is effective in AI or big data computing and a total FP64 (double precision floating point) theoretical peak performance of 37 PFLOPS, which is crucial in conventional computer simulations. Furthermore, high-temperature processing units are directly cooled by water close to outdoor temperatures supplied by the AI data center to achieve world-class energy saving. Seventeen chassis and 34 nodes in one rack save space and increase the cooling effect.

Easy to use software environment is also one of the major characteristics of ABCI. "While considering their existing software environment, regardless of field or version, users can freely select and use what they prefer. Container support is also available so user software environments can be easily deployed."

It was the first deployment of the Tesla® V100 in Japan when ABCI was being constructed, and the number of GPUs is so large as 4,352. Fujitsu's Development, Support, Sales, SE, CE, and Fujitsu Laboratories teams worked together on a special response system. "At AIST, unlike university supercomputer centers, we do not have an organization that specializes in the operation of supercomputers. In April 2017, we began operating AIST Artificial Intelligence Cloud (AIC), which focused mainly on research with medium-scale HPC systems, but was limited to internal use and joint research use. Providing a large-scale HPC system such as ABCI to thousands of users across industry, academia and government was an unknown area for us. However, we were able to launch the ABCI service on schedule in collaboration with Fujitsu, who has plenty of experience in introducing HPC systems," Dr. Ogawa comments on Fujitsu's support.

Benefits and future outlook

Meeting the world's best high-speed and energy-saving benchmarks

ABCI's design, development, and operations are handled jointly with AIRC and the AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory (RWBC-OIL). In addition to construction, Fujitsu assists in operation and maintenance support to ensure the stable running of ABCI.

Regarding benchmark results, the ABCI system demonstrated the world's best high-speed and energy-saving abilities by placing fifth in the world in the TOP500 international performance ranking of supercomputers (June 2018), third in the world in the Green500, which ranks supercomputers in terms of power consumption performance (June 2019), and fifth in the world in the HPCG (High Performance Conjugate Gradients) list (November 2018). In addition, using ABCI, Fujitsu Laboratories Ltd. achieved the world's highest training speed with ResNet-50¹, a deep neural network for image recognition using image data from ILSVRC2012², a contest of image recognition accuracy (as of March 26, 2019 based on company research). ABCI continues to challenge the most important AI issues, and its open-ended program offers all its computing nodes for up to 24 hours of use to any single research group expecting groundbreaking results in their projects.

According to Dr. Ogawa, because it is available to general companies via the internet, the realization of security at the cloud service level has become an important point in terms of operations. "AIST maintains a security white paper that summarizes ABCI's security management system, security implementation, demarcation points between users and AIST, etc. which are all made public. In addition to their own security concerns, companies view ABCI's ability to meet security requirements as a criterion for using our service. Of course, we also request that Fujitsu support operations in accordance with this white paper."

Currently, ABCI is used in roughly 100 projects and by some 1,000 users in a wide range of fields including manufacturing, electrical manufacturers, IT, AI startups, medical care and universities. Dr. Ogawa explains the future outlook for the system.

"AIST developed ABCI not only as a tool for technology development, but also as the first step to introduce technologies such as AI and big data to every corner of industry and society, and to create innovation and realize Society 5.0 to resolve pressing issues. After an "ideal" AI for society has been designed, the next step is to vertically integrate the edge (device) with the cloud (ABCI). I hope that Fujitsu to contribute to the growing, sophisticated needs of users through operational and maintenance support."

Fujitsu continues to contribute to the sustainable development of society and industry through technical support for AI Bridging Cloud Infrastructure.

Glossary

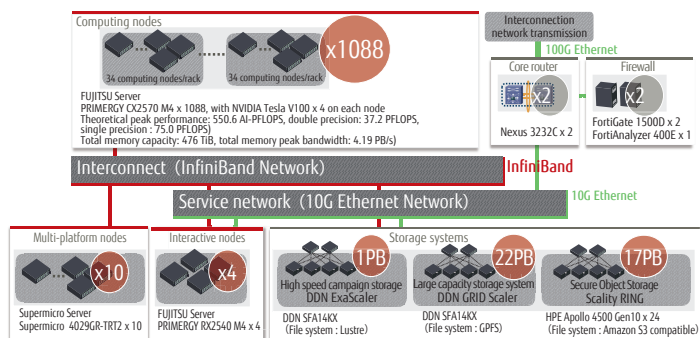
1: ResNet-50

High-performance, image recognition deep neural network developed by Microsoft.

2: ILSVRC2012

ImageNet Large Scale Visual Recognition Challenge 2012

System Diagram



- Names of companies and products mentioned in this catalog are trademarks or registered trademarks of their respective companies.
- Information in this document is subject to change without notice for improvement or other reasons.

Customer profile

National Institute of Advanced Industrial Science and Technology



President Ryoji Chubachi
Address AIST Tokyo Headquarters: 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8921 Japan
 AIST Tsukuba Headquarters: AIST Tsukuba Central 1, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8560 Japan
Established April 1, 2001
Employees 3,030 (2,331 researchers, 699 administrative employees) *As of July 1, 2018
Main business activities Research related to industrial technology
Website <https://www.aist.go.jp/>