

## Progress of "Plasma and Fusion Cloud" Research Data Ecosystem Orienting "Open Science"

A new cloud platform to realize plasma/fusion experimental data ecosystem, named "Plasma and Fusion Cloud," has been technically verified on some fundamental issues. Enormous amount of diagnostic data require a high-performance computing (HPC) platform not only for the LHD physics data analyses and also for the next-generation experiments, such as ITER and JT-60SA. Performance evaluation have been made at NIFS by using the HPC supercomputer "Raijin" and the LHD primary data storage system, both of which are directly connected by 100 Gbps Ethernet optical link. The test results show that almost full bandwidth can be used by means of multiple parallel streams.

Commercial or academic clouds are also very promising as a high-performance data computing platform not only for the physics data analyses but also for the real-time plasma and plant controls. AWS (Amazon Web Service) S3 cloud storage has accepted the proposal to store all the 2.0 peta-byte of compressed LHD physics data for open access, under the AWS Open Data Sponsorship Program (ODP) [1]. AWS is also one of the commercial providers of computing clouds in the framework of NII's Research Data Cloud (RDC) [2] in Japan, allowing LHD data users to increase or decrease computing power they need on demand, simply by paying for the CPU hours used.

In order to make plasma and fusion diagnostic data "FAIR" [3], all the LHD's diagnostic data objects are now under way to be registered with the digital object identifiers (DOI) for each acquisition node and plasma pulse, e.g. <https://doi.org/10.57451/lhd.bolometer.123456.1>. In 2023, more than 1.2 million DOIs have been issued for the LHD diagnostic data. There still remain more than 20 million unregistered diagnostic and analyzed data objects, the registration work would continue for at least the next few years.

To provide API access methods to the data users on the Internet, the APIs must be able to properly control read and write access privileges for privileged and non-privileged users and groups, even though all the LHD data are open publicly. The necessary modifications on the data server have been implemented applying the "gRPC" secure framework [4]. The new gRPC-based data servers have successfully demonstrated their reliability and operability during the most recent 25th LHD campaign in 2024.

Those demonstrations and technical verifications done in this study clearly suggest what the next-generation fusion data research center should be based on the cloud technology.

This research was supported by MEXT as "Developing a Research Data Ecosystem for the Promotion of Data-Driven Science".

[1] Open Data on AWS, "NIFS LHD Experiment", <https://registry.opendata.aws/nifs-lhd/> .

[2] NII Research Data Cloud, <https://rcos.nii.ac.jp/en/service/> .

[3] FORCE11, "The FAIR Data Principles", <https://force11.org/info/the-fair-data-principles/> .

[4] gRPC, <https://grpc.io/> .

### Speaker's Affiliation

National Institute for Fusion Science, Toki

### Member State or IGO

Japan

**Primary author:** Dr NAKANISHI, Hideya (National Institute for Fusion Science)

**Co-authors:** Dr TAKAYAMA, Arimichi (National Institute for Fusion Science); YAMANAKA, Kenjiro (National Institute of Informatics); EMOTO, Masahiko (National Institute for Fusion Science); OHSUNA, Masaki (National

Institute for Fusion Science); Prof. OSAKABE, Masaki (NIFS); YOSHIDA, Masanobu (National Institute for Fusion Science); Ms NONOMURA, Miki (National Institute for Fusion Science); Prof. SAKAMOTO, Ryuichi (National Institute for Fusion Science); IMAZU, Setsuo (National Institute for Fusion Science); Prof. URUSHIDANINI, Shigeo (National Institute of Informatics); SATAKE, Shinsuke (National Institute for Fusion Science, Japan); Dr YAMAMOTO, Takashi (National Institute for Fusion Science); TODO, Yasushi (National Institute for Fusion Science)

**Presenter:** Dr NAKANISHI, Hideya (National Institute for Fusion Science)

**Track Classification:** Data Storage and Retrieval, Distribution and Visualization