



Introducing Open OnDemand to Supercomputer Fugaku

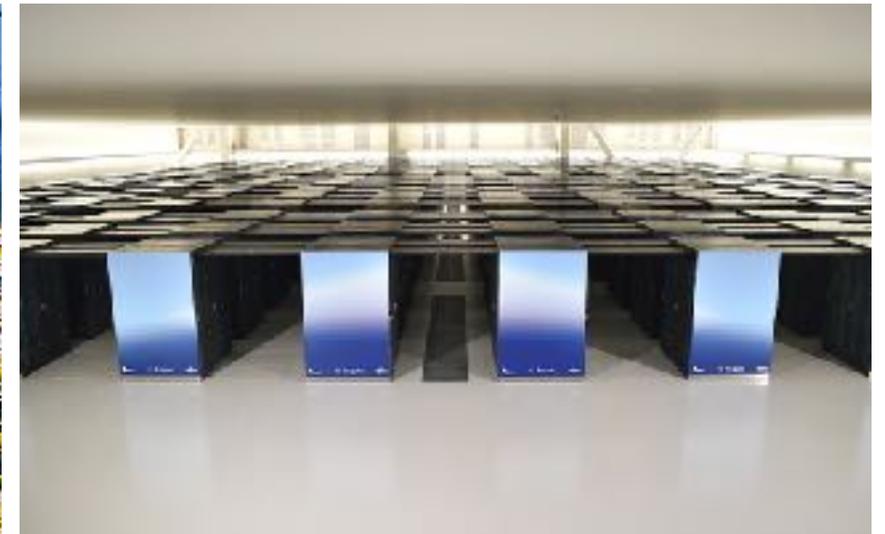
Masahiro Nakao¹, Hidetomo Kaneyama¹, Masaru Nagaku², Ikki Fujiwara²
Atsuko Takefusa², Shin'ichi Miura¹, Keiji Yamamoto¹

1. RIKEN Center for Computational Science, Japan
2. National Institute of Informatics, Japan



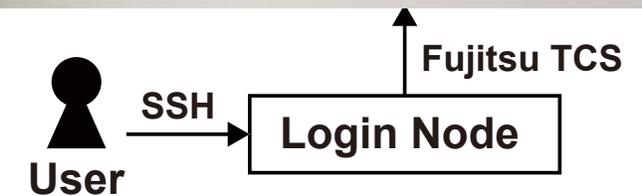
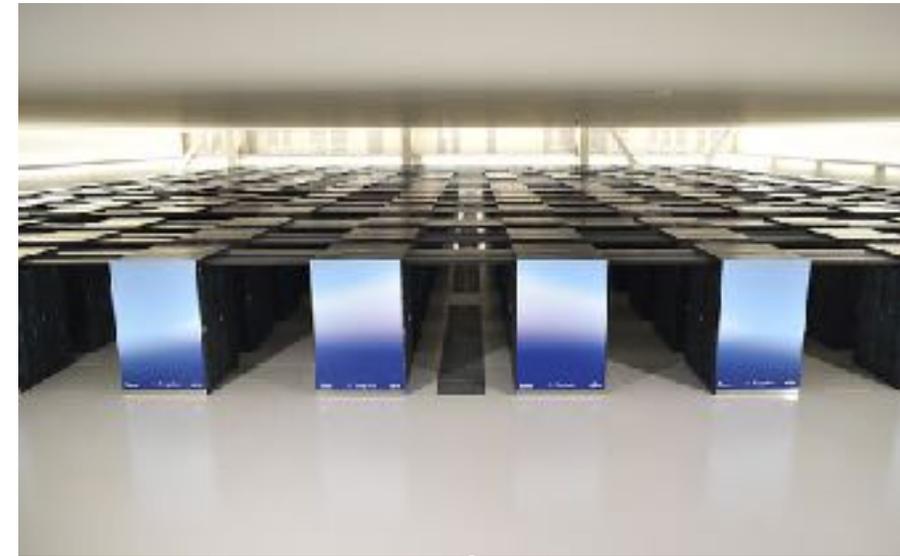
Supercomputer Fugaku

- We have been operating the supercomputer Fugaku, a Japanese flagship super-computer, since 2021
- Located in RIKEN Center for Computational Science in Kobe, Hyogo, Japan
- About 160,000 compute nodes
- Graph500 and HPCG: 1st, Top500: 2nd, HPL-AI: 3rd (Jun. 2023)



Conventional usage of HPC clusters

- The conventional method of using Fugaku and other HPC clusters is to enter the login node via **SSH** and submit a job through **its job scheduler**
- There is a lot of prerequisite knowledge for using HPC clusters, so the learning cost is high for beginners
 - Generate SSH key pair and register public key
 - Command line interface (bash, zsh, tcsh)
 - Job scheduler (Fujitsu TCS in Fugaku)
- Users want to run GUI applications (Remote Desktop, JupyterLab) on compute nodes, but this procedure is complicated
 - Requires VNC client and SSH tunneling settings



Open OnDemand

<https://openondemand.org>

- Web portal for HPC clusters
- Developed mainly by Ohio State University
 - Open source software
 - Used by over 250 institutions worldwide
- Operate HPC clusters from user's web browser
 - No other software installation required
 - File upload/download, job submission and monitoring, shell access with CLI
 - GUI applications on compute nodes can be launched easily

Accounting

Group	Volume	Disk [GB]				Disk [blocks]				Revenue [USD]			
		Used	Usage	Avail	Rate	Used	Storage	Avail	Rate	Used	Usage	Avail	Rate
research	jv0600	5,100	40%	4,910	1%	1,000,000	249,750	1,749,750	1%	800,000	21,900	44,600	1%
eng	jv0600	814,400	90,74%	476,140	58%	180,000,000	64,740,000	115,260,000	36%	-	-	-	-
RESEARCH	jv06100	6,700	1%	5,770	9%	1,400,000	2	1,399,998	0%	-	-	-	-
IS08	jv06100	20	0%	0	0%	400,000	11,000	388,999	3%	-	-	-	-

Objectives

- We present how to introduce Open OnDemand to Fugaku
 - Display of useful information for users on the dashboard
 - Support for Fujitsu TCS, Fugaku job scheduler
 - Install about 50 applications
 - Develop applications to connect with external storages
- The configuration files can be downloaded from https://github.com/RIKEN-RCCS/ondemand_fugaku

Message of the day

Information

- Jul 24, 2023 Operation: July 2023 Large-scale job execution period
- Jul 21, 2023 Operation: Resource groups starting the large-scale job execution
- Jul 18, 2023 Operation: Occurrence of inaccessibility and poor response at both nodes and jobs due to a file system crash (severity:R0004)

Pending jobs

Fugaku-01-01	Fugaku-01-02	Fugaku-01-03	Fugaku-01-04
107/107	8/8	10/10	0/0
0/0	0/0	0/0	0/0

Accounting

Group	Volume	Disk (GB)				Disk (node)				Resource (VM)			
		Used	Usage	Avail	Rate	Used	Usage	Avail	Rate	Used	Usage	Avail	Rate
vmware	vmw001	5,100	40%	4,900	1%	1,000,000	249,708	1,750,291	1%	800,000	21,995	44,000	1%
hpc	vmw001	814,400	90,7%	496,182	6%	182,000,000	84,740,120	97,259,880	4%	-	-	-	-
RESCORE	vmw001	6,000	1	5,78	0%	1,400,000	2	1,399,998	0%	-	-	-	-
IS001	vmw001	20	8	18	4%	800,000	11,000	788,999	1%	-	-	-	-

Recently Used Apps

- Desktop
- Open OnDemand
- SCC-CE
- Avast

Password Apps

- Active Ads
- Home Directory
- Sakana B2M
- HPC Storage
- Job Composer
- Fugaku shell access

Dashboard of Fugaku OnDemand

We modified the default template

- A. External links (Fugaku manual, etc.)
- B. Failure information, operation information, etc.
- C. Number of waiting jobs in each queue using Grafana
- D. Operational calendar using Google Calendar
- E. User disk and budget utilization
- F. Recently used apps on compute nodes
- G. Utilities on Open OnDemand web server
 - Other apps are selected from a navigation bar
 - File upload, job monitoring, etc.

The screenshot shows the Fugaku OnDemand dashboard with the following components and annotations:

- A:** A navigation bar at the top right containing links for 'Default Manual', 'Fugaku Portal', 'Fugaku Schedule', 'Fugaku Status', and 'Fugaku Support'.
- B:** A 'Message of the day' section with a blue header and text about system operations.
- C:** A 'Pending jobs' section with a Grafana-style bar chart showing job counts for various queues: 'fugaku-queue1' (17407), 'fugaku-queue2' (895), 'compute-queue1' (10), 'compute-queue2' (0), 'compute-queue3' (24), 'compute-queue4' (0), and 'compute-queue5' (0).
- D:** A 'Fugaku Schedule' section with a Google Calendar interface showing a weekly view.
- E:** An 'Accounting' section with a table showing disk and budget utilization for various users.
- F:** A 'Recently Used Apps' section with four icons: Desktop, Open OnDemand, SGE, and Jupyter.
- G:** A 'Personal Apps' section with four icons: Active Jobs, Home Directory, Submit BDN, and HPC Storage. Below these are two more icons: Job Composer and Fugaku Shell access.

Prj	Volume	Disk [GB]				Disk [blocks]				Resource [MB]			
		Used	Usage	Avail	Rate	Used	Usage	Avail	Rate	Used	Usage	Avail	Rate
user	job000	1100	40%	4,910	1%	1,000,000	249,750	1,749,250	1%	800,000	21,900	778,100	1%
log	job000	814,000	90,74%	496,142	1%	182,000,000	84,740,100	97,259,900	0%	-	-	-	-
RESCORE	job000	63,000	1	5,78	0%	1,400,000	2	1,399,998	0%	-	-	-	-
BDN	job000	20	0	10	0%	800,000	11,000	788,999	0%	-	-	-	-

A

B

C

D

E

F

G

Support for Fujitsu TCS

- Open OnDemand supports various job schedulers
 - Slurm, Torque, PBS, and so on
 - Fujitsu TCS did not be supported
- Open OnDemand provides an adapter interface to support new job schedulers

```
submit Submit a job  
delete Delete a job  
status Get status of a job  
hold Hold a job  
release Release a held job  
info Get information for a job  
info_all Get information for all jobs  
cluster_info Get system information for an HPC cluster  
supports_job_arrays Bulk job support availability
```

We developed an adapter for Fujitsu TCS by implementing these methods in Ruby language (about 400 lines).

This addition has been merged into the Open OnDemand GitHub repository, so Open OnDemand is available for all HPC clusters using Fujitsu TCS.

Application Type

- Interactive Application
 - Run on compute nodes
 - Remote Desktop, JupyterLab
 - Batch Jobs (no-interactive): Real applications, user-developed applications
- Passenger Application (Utility)
 - Run on Open OnDemand web server
 - Upload a file, monitor a job

Message of the day

Information

- Jul 24, 2023 Operation: July 2023 Large-scale job execution period
- Jul 21, 2023 Operation: Resource groups starting the large-scale job execution
- Jul 18, 2023 Operation: Occurrence of inaccessibility and poor response at both nodes and jobs due to a file system crash (severity:R0004)

Pending jobs

compute-101	compute-102	compute-103	compute-104
15/10	8/5	10	0
compute-mem1	compute-mem2	compute-on-demand	
2/4	1	0	

Accounting

Group	Version	Disk (GB)				Disk (blocks)				Resource (VR)			
		Used	Usage	Avail	Rate	Used	Usage	Avail	Rate	Used	Usage	Avail	Rate
compute	compute	1,100	40%	1,600	1%	1,000,000	249,708	1,750,291	1%	800,000	21,900	14,000,000	1%
log	compute	214,400	90,742	496,142	1%	182,000,000	64,740,140	90,756,860	0%	-	-	-	-
MEMORY	compute	6,000	1	5,750	0%	1,400,000	2	1,400,000	0%	-	-	-	-
HOME	compute	20	8	78	0%	800,000	11,000	109,000	0%	-	-	-	-

Recently Used Apps

- Desktop
- Open OnDemand
- SCC-CE
- Jupyter

Passenger Apps

- Active Jobs
- Home Directory
- Submit Batch
- HPD Storage
- Job Composer
- Fugaku Shell access

How to use Remote Desktop

Welcome to the supercomputer Fugaku

Message of the day

Information

Pending jobs

Accounting

Recently Used Apps

Desktop

Desktop

Home Directory

Submit Job

Job Composer

Fugaku shell access

Desktop

This app will launch an Xfce desktop.

Queue

fugaku-small

Group

rcos-std

Elapsed time (1 - 72 hours)

1

Number of nodes (1 - 384)

1

Total number of processes (1 - 18,432)

1

Total number of processes \leq Number of nodes x 48.

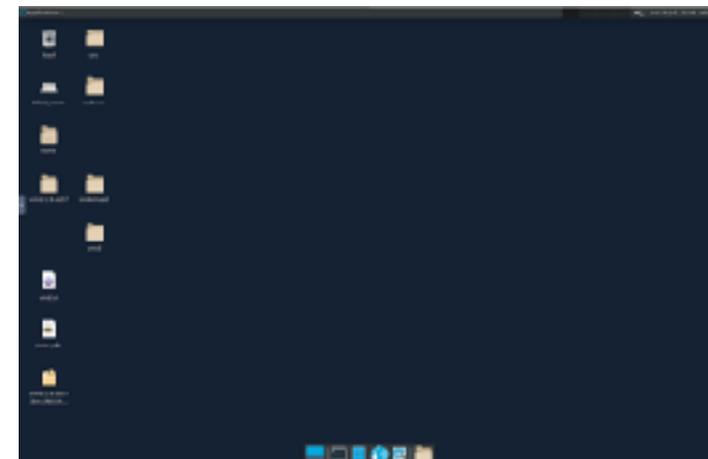
Execution mode

Normal

Please refer to the manual for details (English or Japanese).

Email (You will receive an email when it starts)

Launch

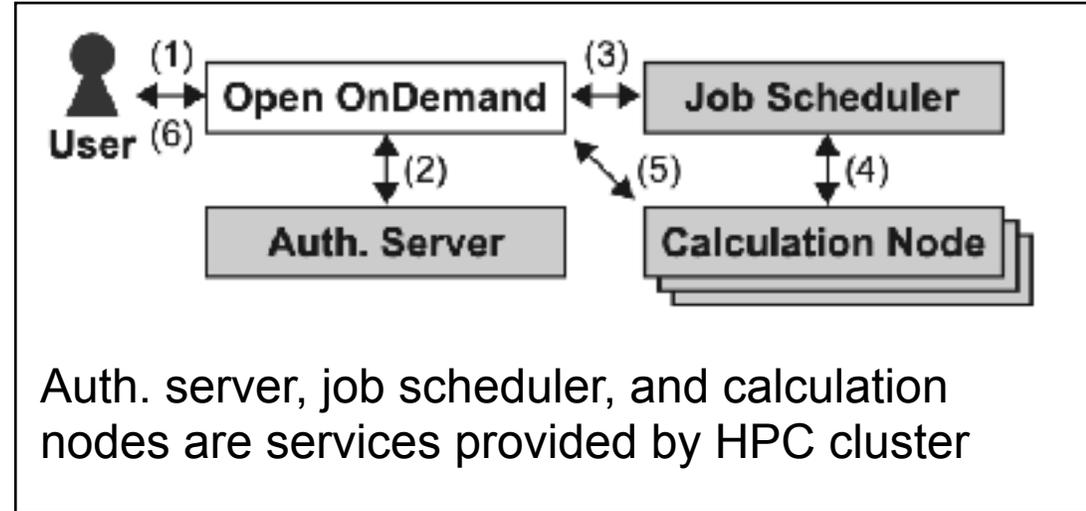


Remote Desktop runs on calculation node, and you can access it in your web browser

Input parameters (Elapsed time, Number of processes, and so on)

Operation Flow of Interactive Application

1. Log in to the Open OnDemand server using user's web browser
2. User authentication for login
3. When issuing an execution command, the job is submitted to the calculation node
4. Wait until job is executed
5. When a job runs, information of the calculation node is sent to the Open OnDemand and a reverse proxy is set
6. Connect to the calculation node inside the HPC cluster from user's web browser using the reverse proxy URL



Interactive Applications

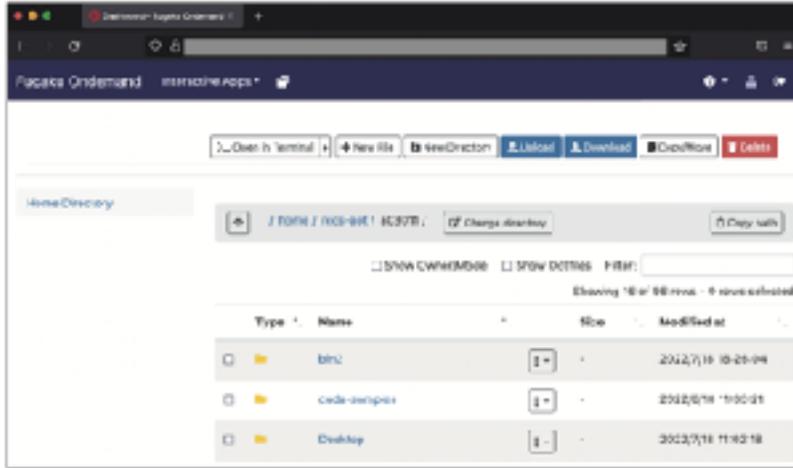
- Interactive Application

Category	Application
Development	Remote Desktop, JupyterLab, MATLAB, RStudio, VSCode
Profiler	NVIDIA Visual Profiler, NVIDIA Nsight Compute*, NVIDIA Nsight Systems, Vampir
Viewer	AVS/Express, C-Tools, GaussView, ImageJ, OVITO, Paraview, PyMOL, SALMON view, Smokeview, VESTA, VMD, VisIt, XCrySDen
Workflow	WHEEL

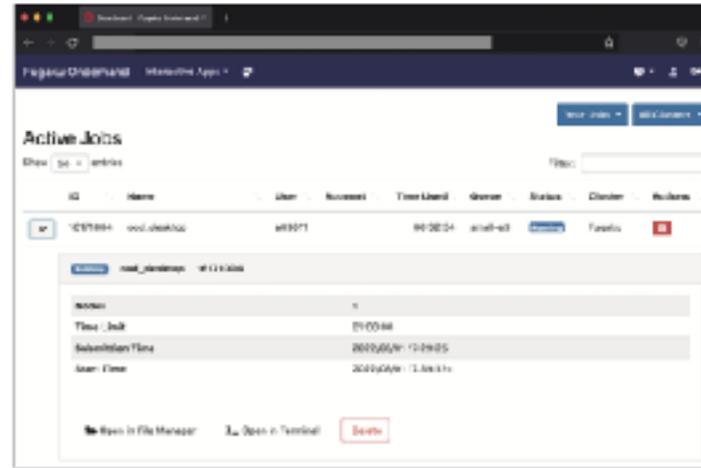
- Batch Job (Not operated interactively)

Category	Application
Climate	SCALE
Computer Aided Engineering	FDS, FrontFlow (blue/X), FrontISTR, OpenFOAM (Foundation/OpenCFD)
Condensed Matter Physics	ALAMODE, AkaiKKR, H Φ , mVMC, OpenMX, PHASE/0, Quantum Espresso, SALMON
Molecular Dynamics	GENESIS, GROMACS, LAMMPS, MODYLAS
Quantum Chemistry	ABINIT-MP, Gaussian, NTCChem, SMASH
Quantum Simulation	braket

Passenger Applications pre-installed in Open OnDemand



Home Directory



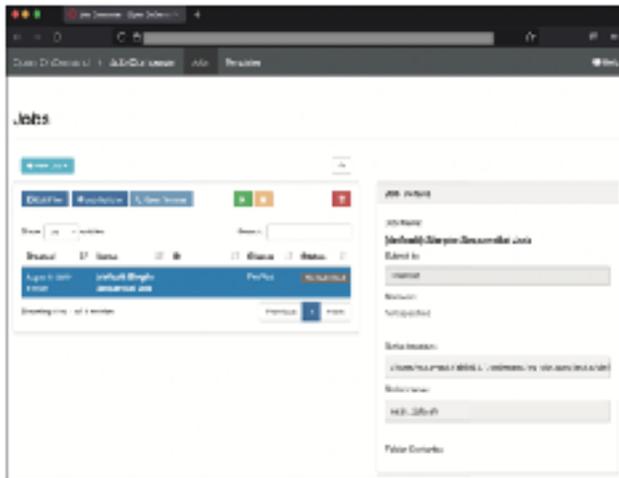
Active Jobs

Home Directory : Files can be uploaded, downloaded, and edited. By combining with rclone, it is also possible to share with external storage such as Amazon S3.

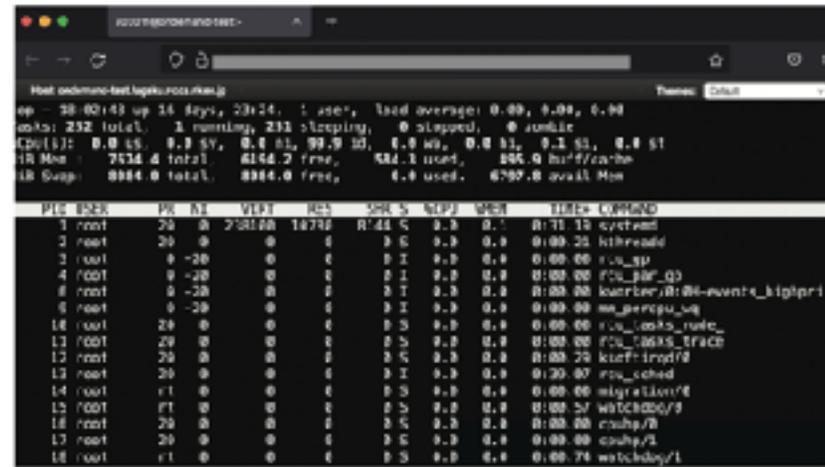
Active Jobs : Monitor jobs

Job Composer : Submit jobs

Shell : Terminal based on Web



Job Composer



Shell

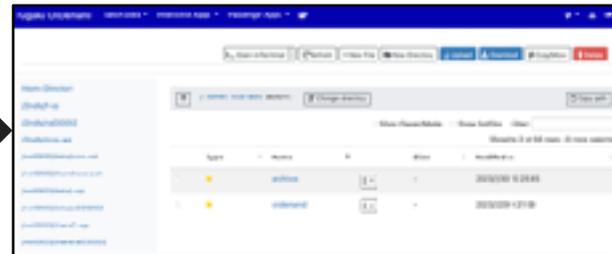
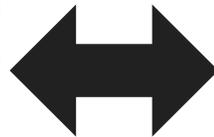
You can also develop new applications by using frameworks of Open OnDemand

Transferring data to external storage

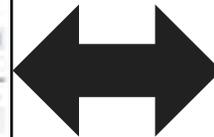
- Applications on Open OnDemand for **HPCI Shared Storage** and **GakuNin RDM**
 - **HPCI Shared Storage** is a large-scale data sharing infrastructure for high-speed of research data among Japanese research organizations
 - **GakuNin RDM** is a research data management service for sharing research data with collaborators
- Users can share data between Open OnDemand and these storages in your web browser by a high-speed network called **SINET6**



HPCI Shared Storage



Open OnDemand on Fugaku

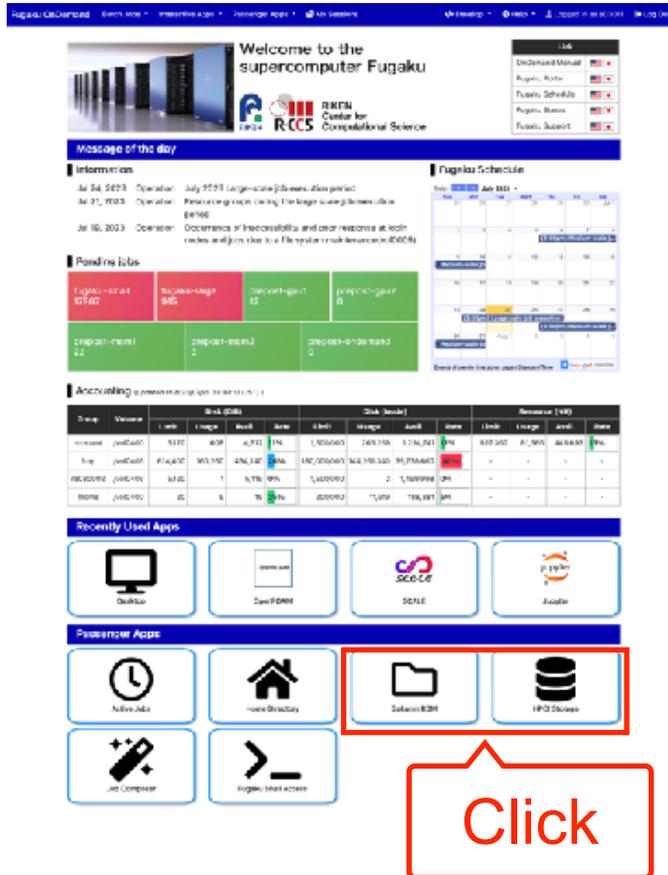


GakuNin RDM

The rclone does not support these storages.

Transferring data to external storage

- Developed using the framework provided by Open OnDemand
- <https://osc.github.io/ood-documentation/latest/tutorials/tutorials-passenger-apps.html>



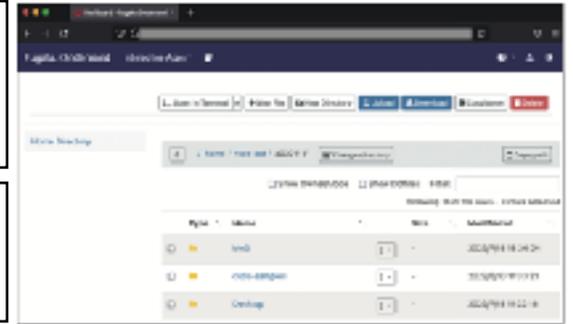
HPCI Storage

Please issue a proxy certificate with the HPCI Certificate Issuing System before mounting the HPCI storage.

Action	HPCI ID	Elapsed time (up to 168 hours)	Passphrase	Mount Path
mount	<input type="text"/>	12	<input type="text"/>	<input type="text"/>

GakuNin RDM

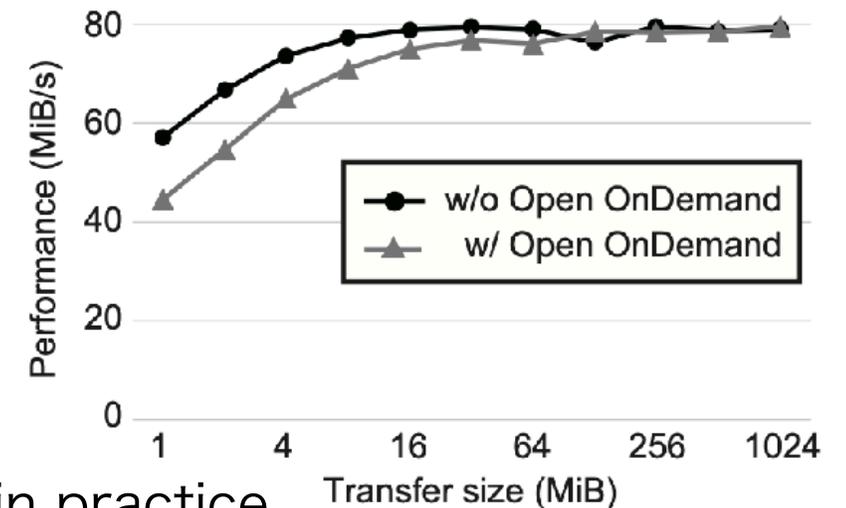
Available Actions	MOUNT_PATH	RDM_NODE_ID	RDM_TOKEN
mount	<input type="text"/>	<input type="text"/>	<input type="text"/>



- After inputting the required information, mount on each storage
- Launch the home directory application to upload files to each storage

Transferring data to external storage

- Evaluate transfer speed
 - To examine the overhead of Open OnDemand, compare data transfer speeds with and without Open OnDemand
 - Transfer data from Open OnDemand web server to HPCI shared storage
 - Open OnDemand web server and HPCI shared storage are in the same building
 - Open OnDemand web server CPU: Xeon Gold 6338 x2, Memory: 256GB, network: 100Gbps
- Result
 - Up to 28% faster without Open OnDemand
 - As the data size increases, the performance difference decreases. There is no difference in performance above 64MiB, so there is no problem in practice



Summary

- To improve user convenience, we install Open OnDemand on Fugaku
- We have made various modifications of Open OnDemand
 - Support for Fujitsu TCS, Fugaku job scheduler
 - Display of useful information for users on the dashboard
 - Install about 50 applications
 - Development of data transfer applications for external storages