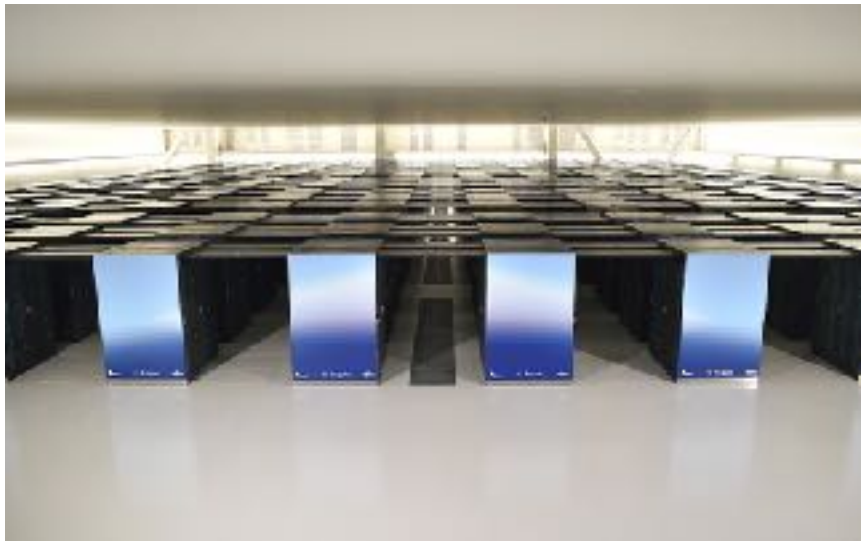
Building the Open OnDemand Community at RIKEN R-CCS

Masahiro Nakao (RIKEN R-CCS)

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 city, Japan
- About 160,000 compute nodes
- Graph500, HPCG: 1st, HPL-AI: 3rd, Top500: 4th



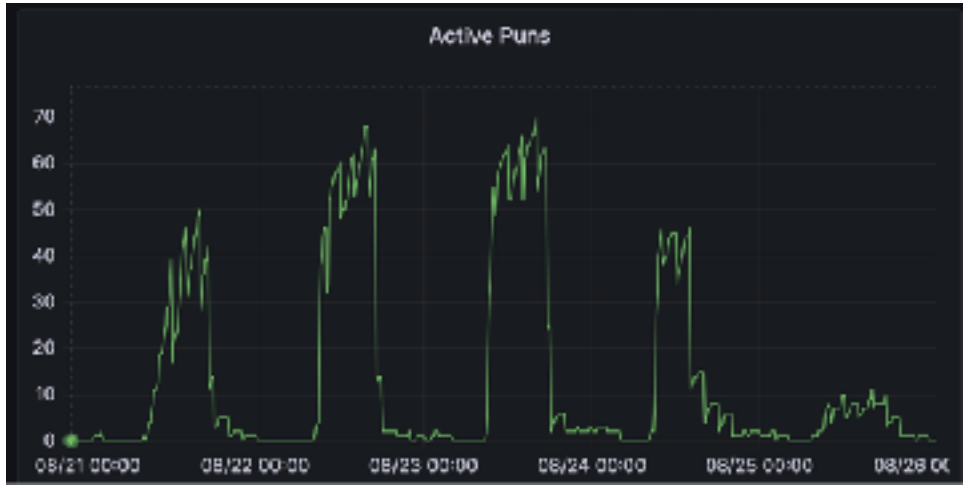
Open OnDemand in Japan

- We had run test operations of Open OnDemand on Fugaku since Aug. 2022, and have run production operations since May 2023
- Open OnDemand is also attracting attention at other national research institutes and supercomputer centers in Japan
- To spread our experience, we held some events in Japan
  - Programming contest for high school students [1]
  - Open OnDemand Workshop in PC Cluster Consortium [2]
  - Meeting for application code tuning on A64FX computer systems [3]
  - Other application courses, etc.

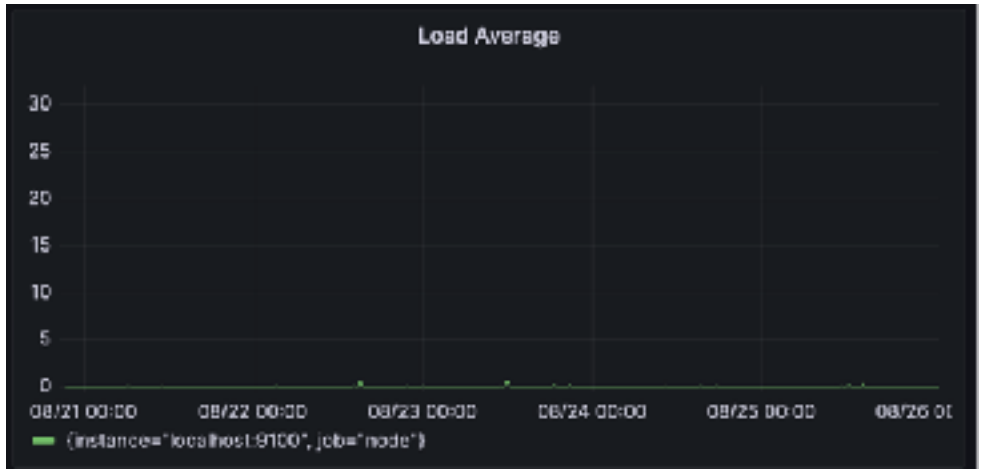
[1] https://www.gsic.titech.ac.jp/supercon/main/attwiki/index.php
[2] https://www.pccluster.org/ja/event/2023/09/231011-ws-openondemand.html
[3] https://www.hpci-office.jp/en/events/symposia/meeting_A64FX_231023

Supercomputing contest on Fugaku

- A supercon for high school students was held from August 21st to 25th, 2023
- Fugaku was be used from 2021
- This year we use Open OnDemand instead of SSH
- There is no longer an explanation about using PuTTY and WinSCP, and how to generate and register SSH keys
- The description is OS independent

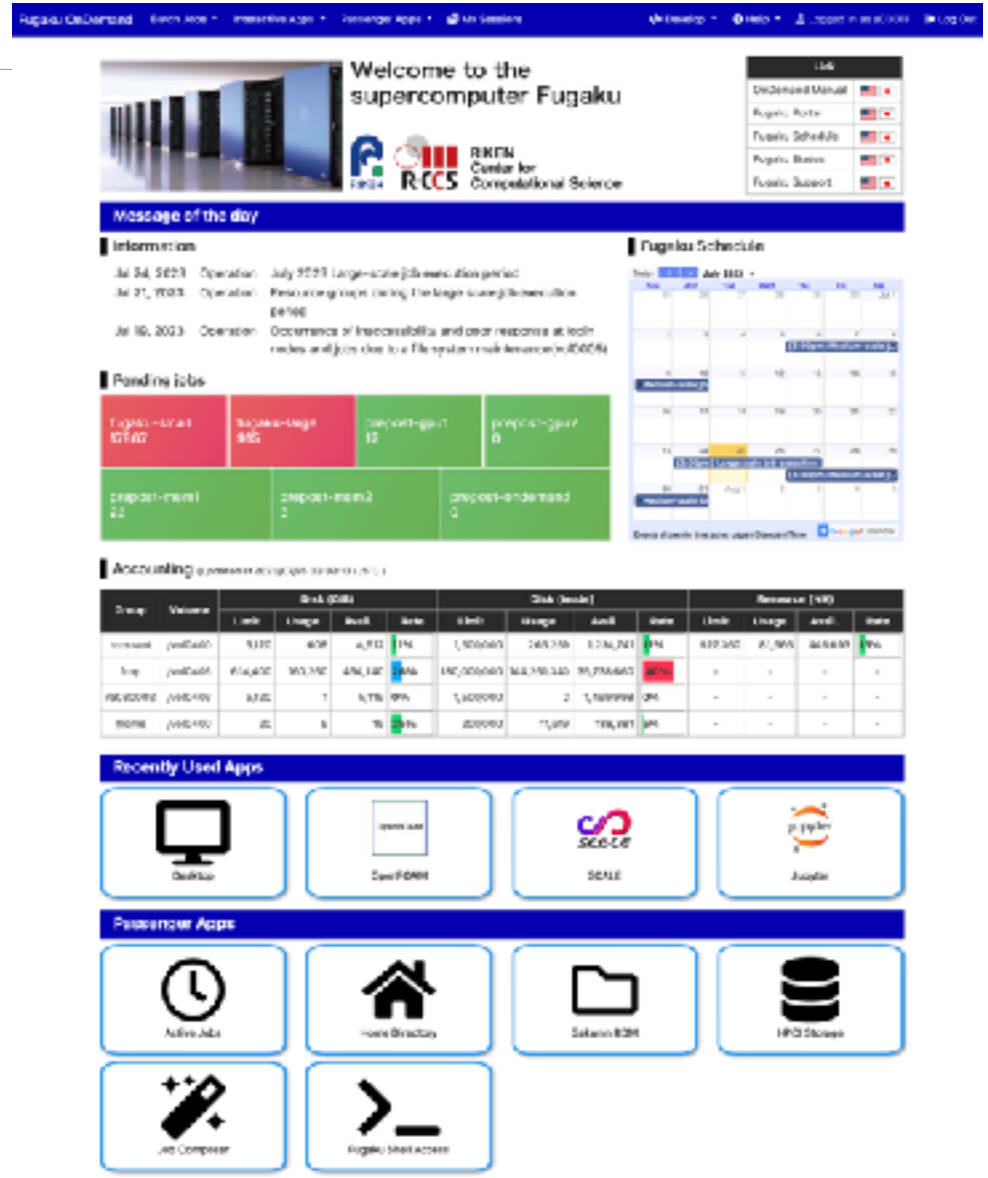


Open OnDemand has a very low load



Open OnDemand on Fugaku

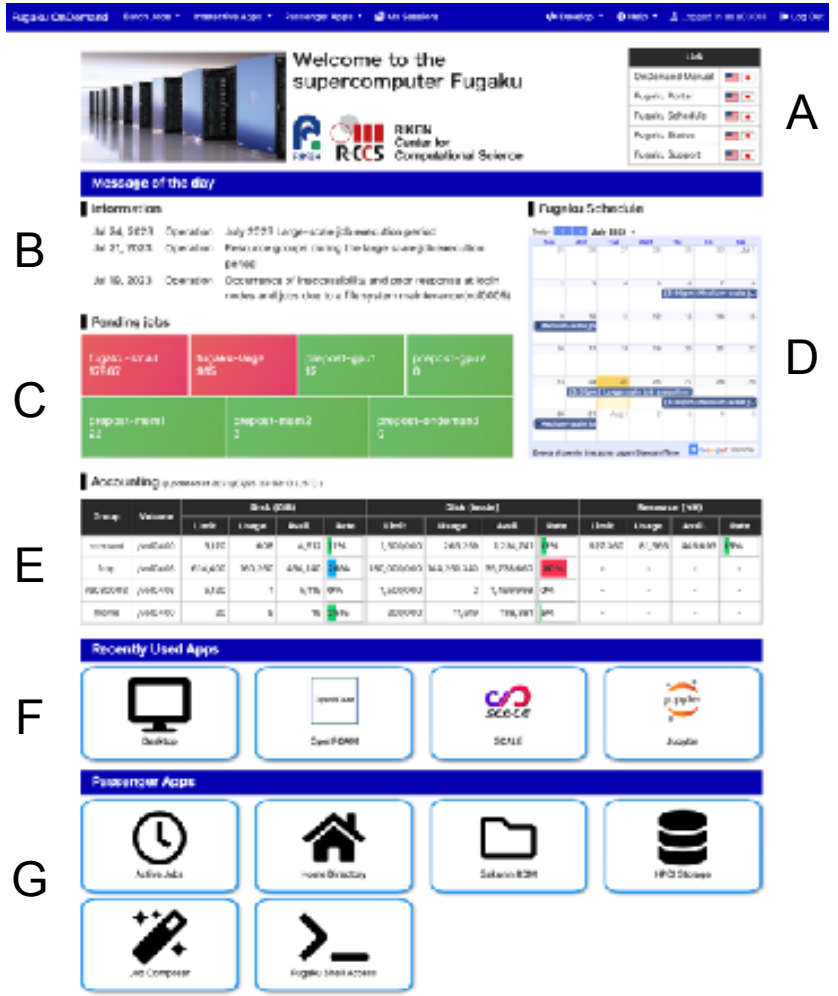
- This presentation describes how we introduced Open OnDemand to Fugaku
  - Display of useful information for users on the dashboard
  - Support for Fujitsu TCS, Fugaku job scheduler
  - About 50 applications are now available from Open OnDemand
  - Develop applications to connect with external storages
- The configuration files can be downloaded from https://github.com/RIKEN-RCCS/ondemand_fugaku



Dashboard of Open OnDemand on Fugaku

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. Apps that run on recently used compute nodes
   - Interactive application (GUI jobs and Batch Jobs)
   - Other apps can be selected from a navigation bar
G. Utilities that work with Open OnDemand server
   - Passenger application
   - File upload, job monitoring, etc.

https://ondemand.fugaku.r-ccs.riken.jp



Dashboard of Open OnDemand on Fugaku

```
<%-
  require 'utils.rb'
  info = dashboard_info("/system/ood/motd/info.txt")
-%>
                :
<h4>Information</h4>
<table>
<%- info.each do |i| -%>
  <tr>
    <%- c = i.split(":") -%>
    <td><%= c[0] %></td><td><%= c[1] %></td><td><%= c[2] %></td>
  </tr>
<%- end -%>
</table>
```

B

Information

Jul 24, 2023	Operation	July 2023 Large-scale job execution period
Jul 21, 2023	Operation	Resource groups during the large scale job execution period
Jul 19, 2023	Operation	Occurrence of inaccessibility and poor response at login nodes and jobs due to a file system maintenance(vol00005)

The info.txt is updated once a day by cron.
The info.txt is generated from the information on the Fugaku portal site.
Disk usage and budget information are the same.

E

Accounting (Updated at 2023/07/25 02:32:10 (JST))

Group	Volume	Disk (GiB)				Disk (inode)				Resource (NH)			
		Limit	Usage	Avail.	Rate	Limit	Usage	Avail.	Rate	Limit	Usage	Avail.	Rate
rccs-adl	/vol0400	5,120	608	4,512	11%	1,500,000	265,259	1,234,741	17%	527,360	81,666	445,693	16%
f-op	/vol0403	614,400	160,260	454,140	26%	180,000,000	144,261,340	35,738,660	80%	-	-	-	-
ra000002	/vol0403	5,120	1	5,119	0%	1,500,000	2	1,499,998	0%	-	-	-	-
/home	/vol0400	20	5	15	25%	200,000	11,619	188,381	5%	-	-	-	-

Dashboard of Open OnDemand on Fugaku

```
<%-
  require 'time'

  # URL of Dashboard
  dashboard_url = "https://status.fugaku.r-ccs.riken.jp/d-solo/cf06d886-
e672-41d8-a587-85ccb32fce7e/
5a6f6KGM5b6F44Gh44K444On44OW5pWw?
orgId=1&theme=light&panelId=2"

  # Get the current time as epoch milliseconds.
  now = (Time.now.to_f * 1000).to_i

  # Get time one month ago as epoch milliseconds
  one_month_ago = now - (30 * 24 * 60 * 60 * 1000)
-%>
                :
<h4>Pending Jobs</h4>
<iframe src="<%= dashboard_url %>&to=<%= now %>&from=<%=
one_month_ago %>" frameborder="0" width="100%" height="225px">
</iframe>
```

C

D

Grafana and Google Calendar have the abilities to port graphs to external HTML.

Development of adapter 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 various 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 defined in the parent class in Ruby language (about 400 lines).

This addition has been merged into the Open OnDemand GitHub repository, so Open OnDemand is now available for Fujitsu TCS.

Add applications

- In Fugaku, applications are managed by spack
  - Set the environment variables using **spack load** command from Open OnDemand
- Some interactive applications (Remote desktop etc.) are not installed on Fugaku
  - Create a container image using Singularity
  - Run the applications using **singularity run** command from Open OnDemand

Applications

- Interactive Application

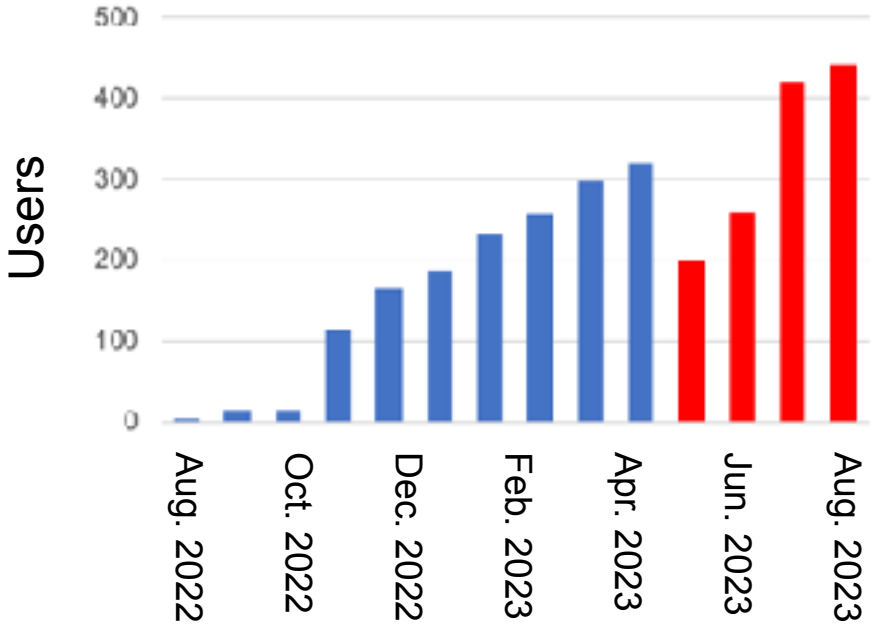
Category	Application
Development	**Remote Desktop**, **JupyterLab**, MATLAB, **RStudio**, **VSCode**
Profiler	NVIDIA Visual Profiler, NVIDIA Nsight Compute*, NVIDIA Nsight Systems, Vampir
Viewer	AVS/Express, **C-Tools**, GaussView, **ImageJ**, **OVITO**, **Paraview**, **PyMOL**, SALMON view, **Smokeview**,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

- Batch Job (Not operated interactively)

Managed by singularity

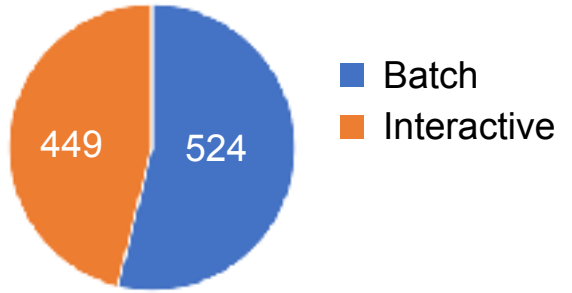
Category	Application
Climate	SCALE
Computer Aided Engineering	FDS, FrontFlow (blue/X), FrontISTR, OpenFOAM (Foundation/OpenCFD)
Condensed Matter Physics	ALAMODE, AkaiKKR, HΦ, mVMC, OpenMX, PHASE/0, Quantum Espresso, SALMON
Molecular Dynamics	GENESIS, GROMACS, LAMMPS, MODYLAS
Quantum Chemistry	ABINIT-MP,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

Status on Fugaku

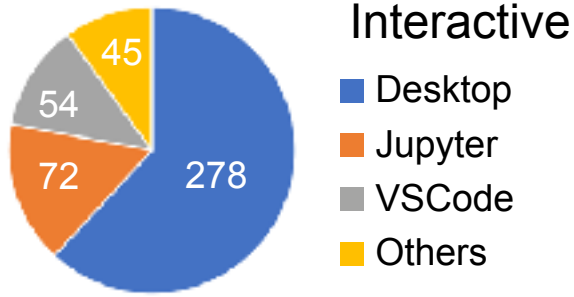


Since we switched to a new server in May 2023,
we have been re-counting the number of new users.
The number of users as of today (15th Nov. 2023) is **548**.
The number of Fugaku accounts is about 2,500.
The graph above is a cumulative number,
so it also includes deleted users.

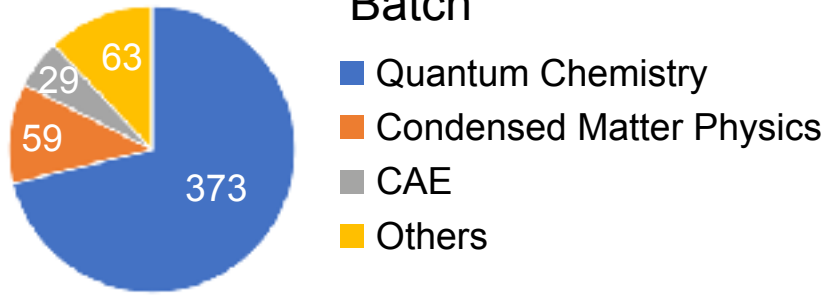
Job type from May to Aug. 2023



Interactive



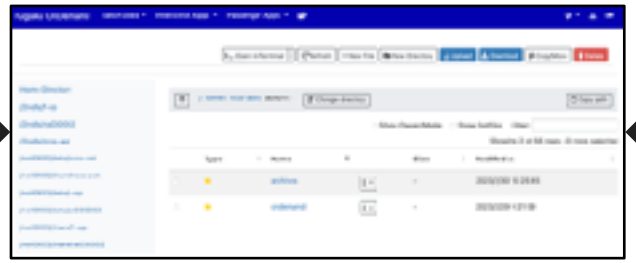
Batch



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



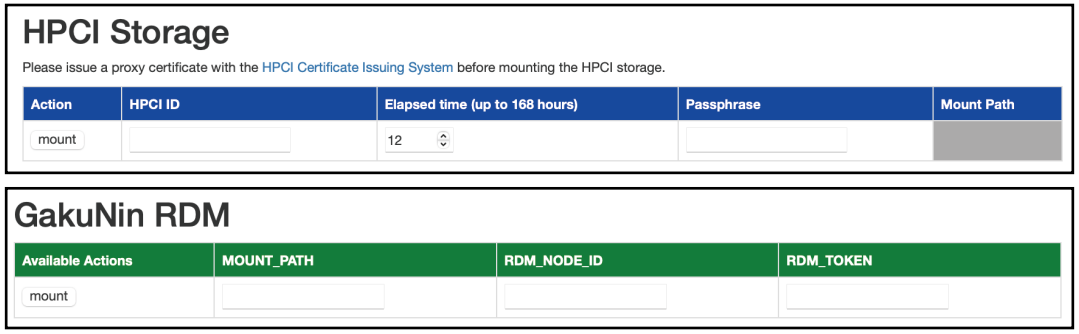
Open OnDemand on Fugaku

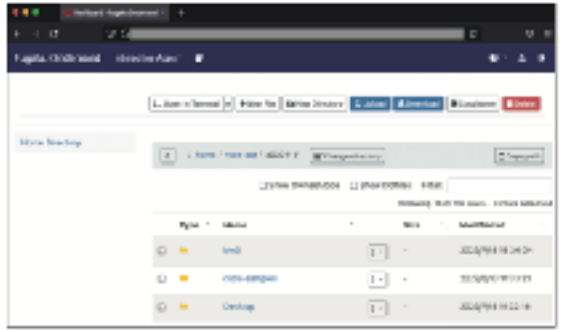
GakuNin RDM

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

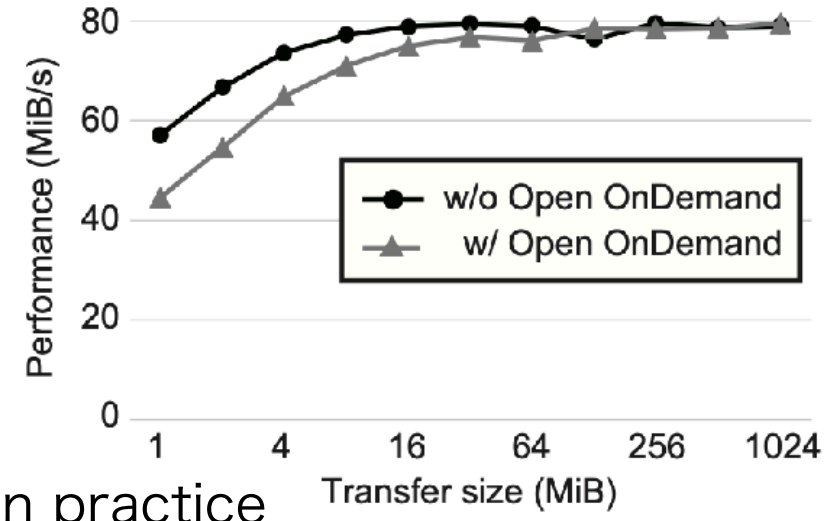




- After inputing the required information, mount on each storage
- Launch the home directory application to upload files to each storage

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

- Introduce the spread of Open OnDemand in Japan
- Hold some Open OnDemand events
- Install Open OnDemand on Fugaku
  - Development an adapter for Fujitsu TCS
  - Display of useful information for users on the dashboard
  - Install about 50 applications
  - Development of data transfer application for external storages