

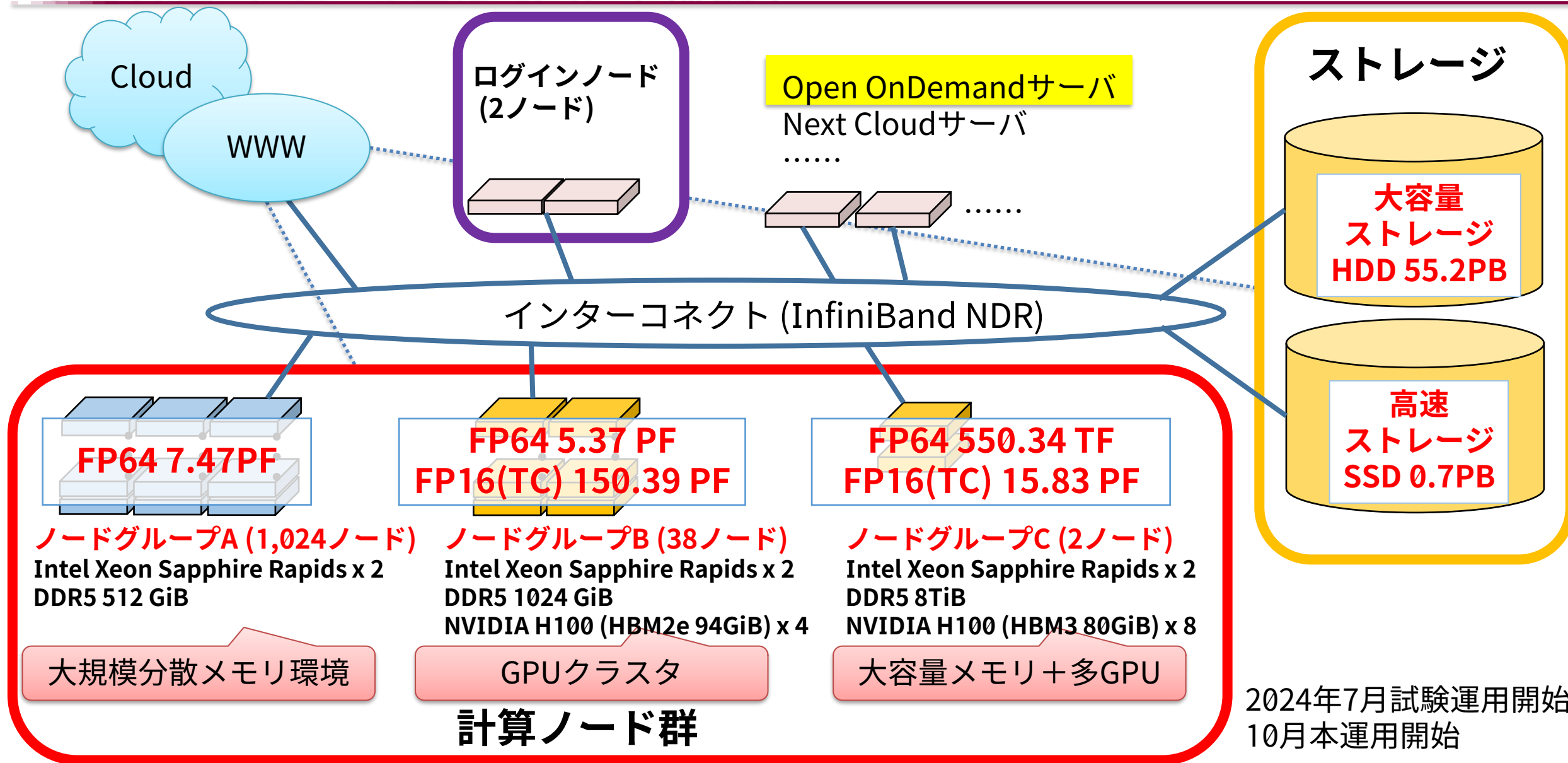
2025.01.27-28 PCCC HPC OSS ワークショップ in 九州大学
01.28 PCCC 2nd Open OnDemand workshop
13:00-13:20 活用事例(2) by 九大

大島聡史（九州大学 情報基盤研究開発センター 准教授）

玄界におけるOpen OnDemandの利用状況

スーパーコンピュータ 玄界 全体概要

消費電力：システム1.5MVA+空調0.9MVA、総計2.4MVA
納入業者：富士通株式会社



九州大学の歴代スーパーコンピュータ等

導入年	停止年	機種名	備考
1969	1974	FACOM 230-60	
1974	1977	FACOM 230-75	
1977	1979	FACOM M-190	
1979	1983	FACOM M-200	
1983	1985	FACOM M-382	
1986	1987	FACOM VP100	250MFLOPS
1987	1992	FACOM VP200	500MFLOPS
1992	1995	FUJITSU M1800	
1992	1996	FUJITSU VP2600	5GFLOPS
1993	?	Thinking Machines CM5-16	
1997	2001	FUJITSU VPP700/56	123GFLOPS
2000	2005	FUJITSU GP7000F/900	38GFLOPS
2001	2006	FUJITSU VPP5000/64	614GFLOPS
2001	2006	COMPAQ GS320	

導入年	停止年	機種名	備考
2005	2008	IBM eServer p5	3.3TFLOPS
2007	2008	HITACH SR11000 model J1, K2	3TFLOPS
2007	2011	FUJITSU PRIMEQUEST 580	13TFLOPS
2007	2011	FUJITSU PRIMERGY RX200S3	18TFLOPS
2009	2013	HITACH SR16000 model L2	25TFLOPS
2011	2012	FUJITSU PRIMERGY RX200S6	55TFLOPS
2012	2017	FUJITSU PRIMEHPC FX10	182TFLOPS
2012	2017	FUJITSU PRIMERGY CX400	510TFLOPS
2013	2017	HITACHI HA8000 tc/HT210	713TFLOPS
2013	2017	HITACHI SR16000 VM1	8TFLOPS
2017	2023	FUJITSU PRIMERGY CX2550	6.9PFLOPS
2017	2023	FUJITSU PRIMERGY CX2570	3PFLOPS
2017	2023	HPE DL380 Gen10	420TFLOPS
2017	2023	SGI UV300	50TFLOPS
2024	(2030)	FUJITSU PRIMERGY CX2550 M7	7.47PFLOPS
2024	(2030)	FUJITSU PRIMERGY GX2560 M7	5.37PFLOPS
2024	(2030)	Supermicro SYS-821GE-TNHR	0.68PFLOPS

ITO

玄界

※ ITOと玄界のGPUシステムの性能はCPUとGPUのFP64合算値
(機械学習性能は考慮されていません)

過去のジョブ実行インターフェース（SSH以外、Webベース）

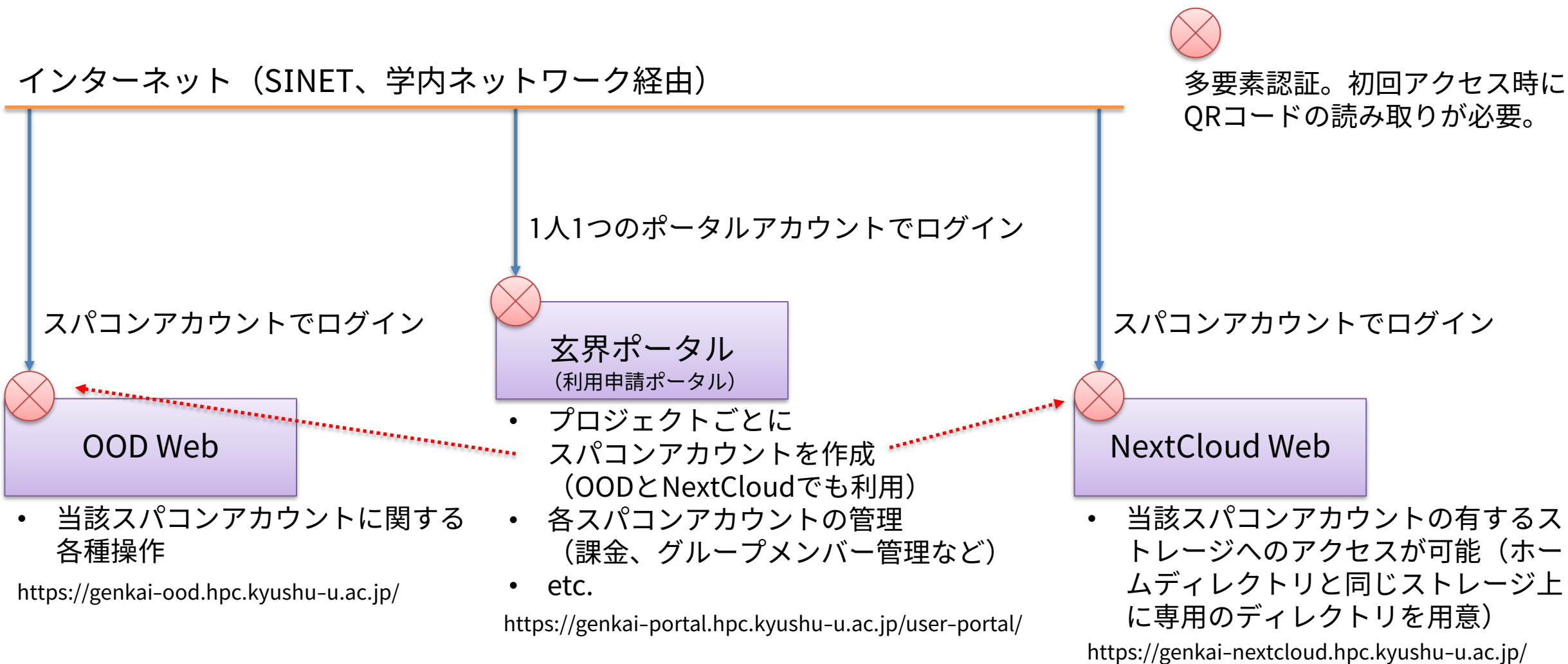
- 過去の資料（すぐにアクセスできるデータ）があまり残っておらず……
- FX10/CX400世代
 - Webブラウザ上でファイル操作やコンパイルが可能（おそらくベンダー独自）
 - ファイル操作はある程度便利だったが、コンパイルは細かい設定ができず不便だった記憶がある
- ITO
 - フロントエンドノードの予約システムあり（富士通UNCAI）
 - 予約のみ、予約した資源に別途SSH
 - バックエンドノードとは独立した運用
 - バックエンドは定額使い放題



玄界の状況

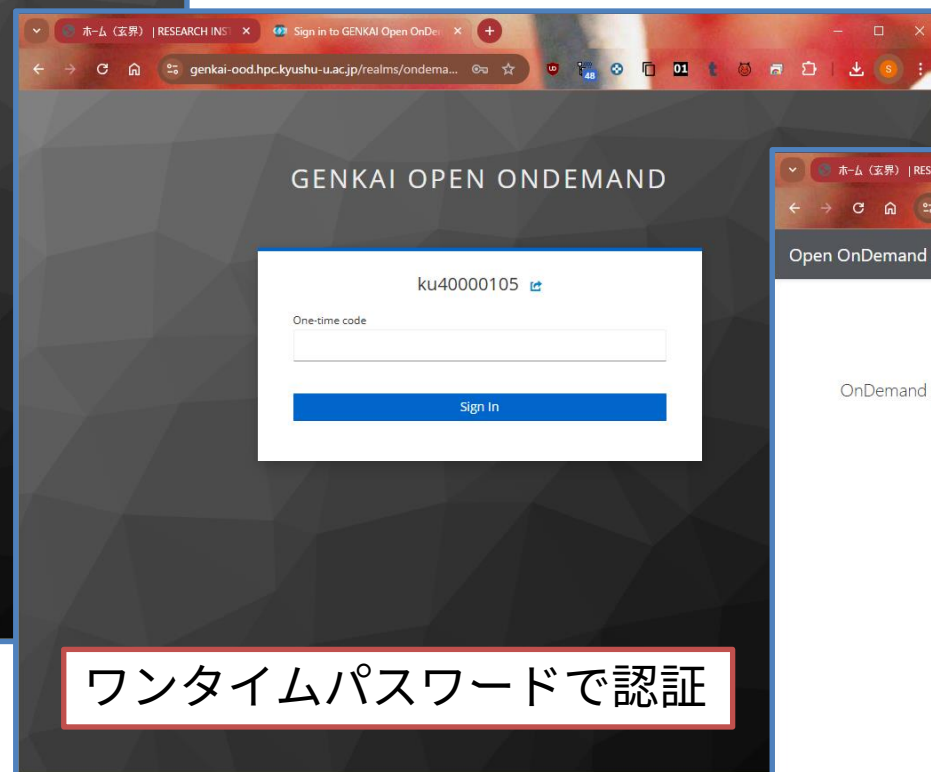
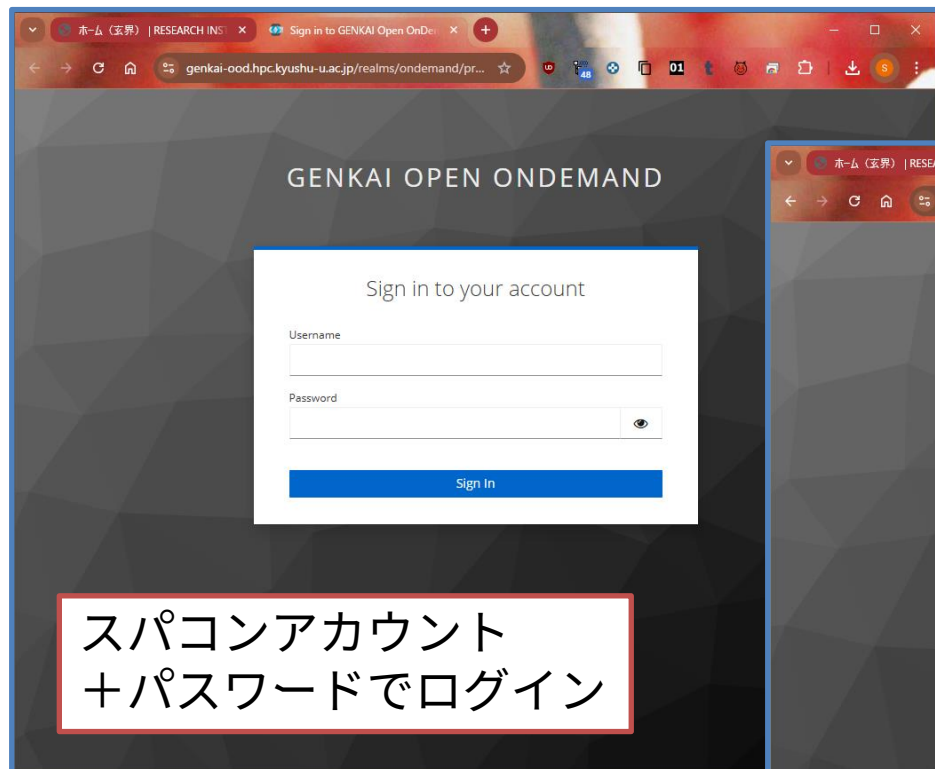
- Open OnDemandの導入を仕様書で要求、試験運用時から利用可能に
 - 他の運用ソフトウェア等と同様にベンダーによるインストール
 - 九大側でソフトウェアの整備が全くできないのは不便なため、テスト環境を準備中
- 玄界のWebポータルとは別にOODを用意
 - アカウントの連携はするが、認証自体は別
 - 多要素認証を複数管理する必要があり若干めんどくさい（仕方がない気はしている）
 - 玄界ポータル、OOD、NextCloudがそれぞれワンタイムパスワードを要求

アカウントと認証の関係

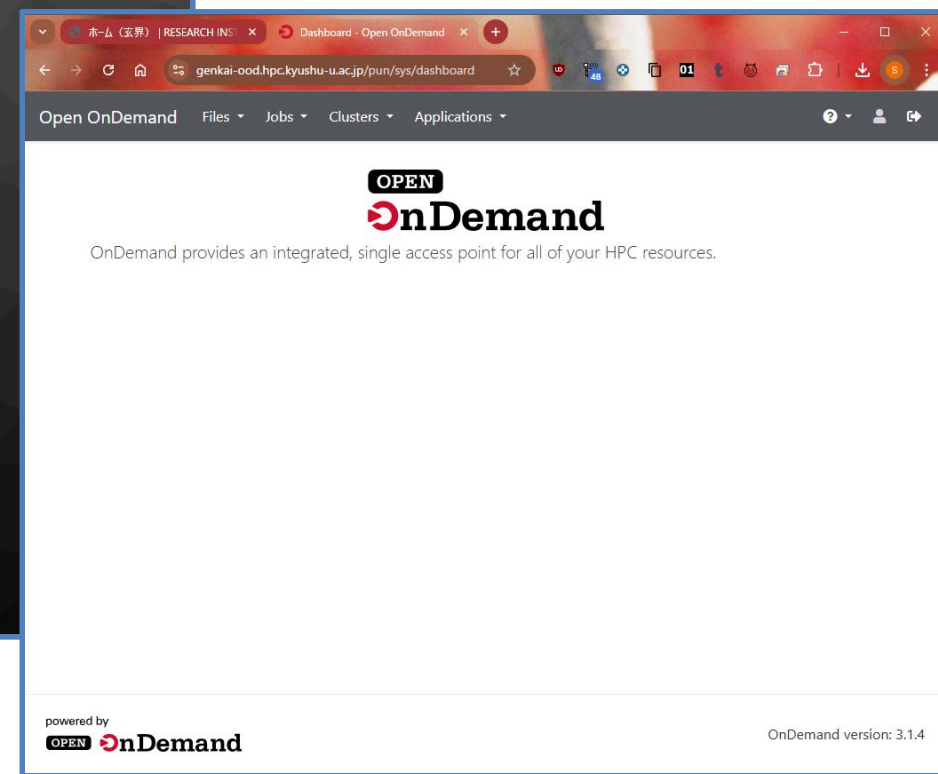


Open OnDemandへのログイン

- 実際の画面



↓例の「味気ないログイン画面」

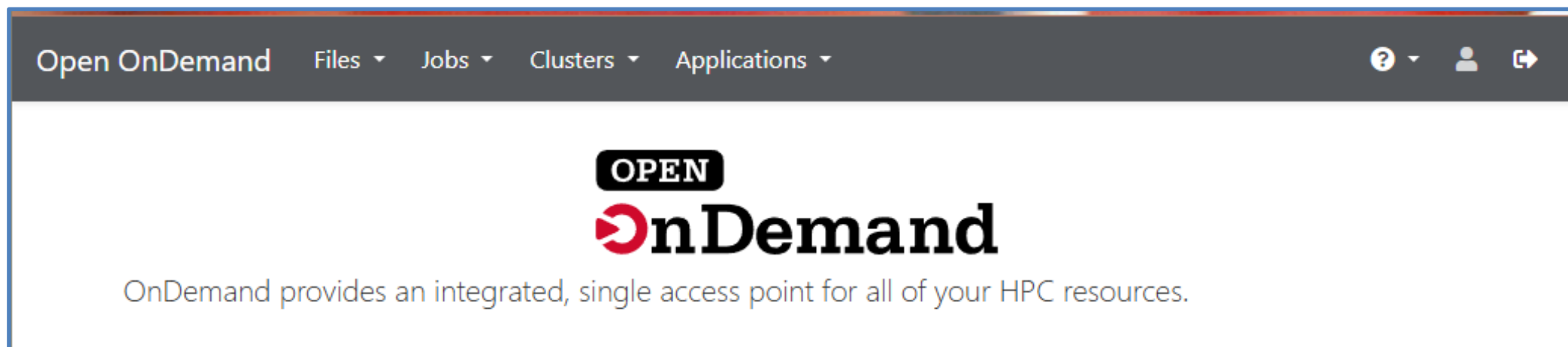


現在提供している機能

- ① ファイル操作
- ② ジョブの実行と管理
- ③ Webターミナル
- ④ アプリケーション
 - ① Jupyter Notebook
 - ② RStudio
 - ③ WHEEL



- 基本的にOODが提供する基本的な機能（デフォルトのPassenger Apps）
- 他で使っている便利な機能を積極的にパクって参考にして充実させたい
 - どんなアプリに需要がある？
- 玄界における利用で便利に感じている点や不便な点をいくつか紹介する



各種機能について

① ファイル操作

- Linuxのファイル操作に慣れていないユーザには便利

② ジョブの実行と管理 →

- 玄界の計算ノードグループにあわせて設定、どのノードグループにもジョブを投入可能
 - MIGで7分割してある資源も利用可能
- (ノード固定などの特殊な例はあるが) OODから自由に計算資源を利用可能
- 本格的に利用するには機能不足か
 - あらかじめ用意しておいたジョブスクリプトを利用するような機能がない
 - 過去に実行したジョブの参照ができない
- 実はOOD上で実行した以外のジョブを監視できるのが便利？

Open OnDemand Files ▾ Jobs ▾ Clusters ▾ Applications ▾

Home / My Interactive Sessions / Job Submitter

Interactive Apps

Job Submitter

Jupyter

RStudio

WHEEL

Job Submitter

Simple form for a batch job.

Queue

b-batch-mig-gpu ▾

Number of GPUs (1 - 28)

4

Elapsed time *hours* (0 - 168 hours)

1

Elapsed time *minutes* (0 - 59 minutes)

0

Elapsed time *seconds* (0 - 59 seconds)

0

Output statistical information

(none) ▾

Working directory

リソースグループ（キュー）構成

- ノード単位利用とコア数・GPU数指定利用でノードを分けている

A

グループ名	CPUコア数	ノード数	メモリ量	最長実行時間
a-batch	1 - 120	1	3.7GiB * コア数	168時間
		1 - 128	453GiB * ノード数	6 - 168時間
a-inter	1 - 120	1	3.7GiB * コア数	6時間
		1 - 128	453GiB * ノード数	6時間

B

グループ名	GPU数	ノード数	CPUコア数	メモリ量	最長実行時間
b-batch	1 - 4	1	30 * GPU数	226.7GiB * GPU数	168時間
		1 - 4	120 * ノード数	907GiB * ノード数	48 - 168時間
b-inter	1 - 4	1	30 * GPU数	226.7GiB * GPU数	6時間
		1	120	907GiB	6時間
b-batch-mig	1 - 28 サブGPU		4 * サブGPU数	32.3GiB * サブGPU数	168時間
b-inter-mig	1 - 28 サブGPU		4 * サブGPU数	32.3GiB * サブGPU数	6時間

C

グループ名	GPU数	ノード数	CPUコア数	メモリ量	最長実行時間
c-batch	1 - 8	1	14 * GPU数	930.6GiB * GPU数	24時間
		1 - 8	120	7.1TiB	24時間
c-inter	1 - 8	1	14 * GPU数	930.6GiB * GPU数	6時間
		1	120	7.1TiB	6時間

各種機能について

③ Webターミナル

- できることはSSHアクセスと変わらない（操作的にはむしろ劣化）だが、初心者が一番躓くポイントであるSSHアクセスを考えなくて済む
 - SSHクライアントとSSH鍵の準備が不要
 - 講習会がスムーズに始められる
 - 多要素認証で躓く人もいるが……

④ アプリケーション

- もっと充実させたいとは思っているが、これから

Jupyter Notebookの起動例

Open OnDemand Files Jobs Clusters Applications

Home / My Interactive Sessions / Jupyter

Interactive Apps

- Job Submitter
- Jupyter**
- RStudio
- WHEEL

Jupyter

Jupyter is a web-based interactive computing platform.

Queue

b-batch-mig-gpu

Number of GPUs (1 - 28)

1

Elapsed time *hours* (0 - 168 hours)

Elapsed time *minutes* (0 - 59 minutes)

0

Elapsed time *seconds* (0 - 59 seconds)

0

Output statistical information

-S : Output the statistical information containing the n

Working directory

任意の計算ノードグループを利用可能

Open OnDemand Files Jobs Clusters Applications

Session was successfully created.

Home / My Interactive Sessions

Interactive Apps

- Job Submitter
- Jupyter
- RStudio
- WHEEL

Jupyter (1529985) Running

Host: >_ b0022

Created at: 2025-01-25 11:06:22 JST

Time Remaining: 59 minutes

Session ID: 19dc6ee2-8d14-466b-90b1-91cb89de1872

Connect to Jupyter

簡単に接続できる

jupyter

File View Settings Help

Files Running

Open Download Rename Duplicate Delete

/

Name	Last Modified
miniconda3-py39_23.11.0	7 months ago
miniconda3.bak	7 months ago
ondemand	7 months ago
OpenFOAM_HMM.hmm_v2312	7 months ago
OpenFOAM-BenchmarkTest	7 months ago
opt	3 months ago
Rpackages	5 months ago
singularity-ce-4.2.0	17 days ago
spack	3 months ago
tmp	2 months ago
ubuntu	7 months ago
venv	7 months ago

New Upload

- Python 3 (ipykernel)
- Terminal
- Console
- New File
- New Folder

講義での利用事例

- 学部生向けの並列アルゴリズムの講義で利用（美添先生）
 - sshは教えず、OODで操作を完結
 - ログイン、ファイルアップロード、コンパイル、ジョブ実行
 - 「sshログインを教えるのと比べて格段に便利」

講習会での利用事例

- 10月21日に「CUDA-Qハンズオン講習会」を実施
https://dx.kyushu-u.ac.jp/news/cudaq_handson2024/
 - 玄界をハンズオン環境として利用
 - OODからCUDA Quantumのコンテナを起動できるように準備
 - 適切に準備をすれば
 - SSH接続環境の整備
 - ソフトウェア実行環境の準備
 (module loadやpythonモジュールの準備など)
- を省略して「本質的なこと」だけに集中することが可能であることが確認できた
- ※スパコンの利用講習会という意味では
 良し悪しな気もしなくもない

量子古典ハイブリッド計算プラットフォーム

CUDA-Q ハンズオン講習会

日時	10/21 (月) 13:00-17:00	場所	九州大学 情報基盤研究開発センター
開催方式	対面・オンライン ハイブリッド、ノートPC持参		
申し込み	 https://forms.office.com/r/RhWsYS18F5 申し込み締切: 10月14日(月) 定員: 40名程度		
受講資格	国内在住の国公立大学・高専の教員・学生・研究生、研究機関研究員、企業所属の研究員・技術者。実施内容は、学術研究、教育および社会貢献に関連するものに限る		

主旨

「CUDA-Q」はNVIDIAが開発している量子古典ハイブリッド計算のためのオープンソースプラットフォームです。量子コンピュータ向けのアルゴリズム研究やアプリケーション開発には、量子回路シミュレーションが大きな役割を果たしています。CUDA-Qは複数GPU、複数ノードを用いた量子回路シミュレーションをサポートしているため、玄界が搭載するNVIDIA H100の豊富なGPU・メモリ資源を活用することができます。CUDA-Qは実機を含めたバックエンドを選択し量子計算を実行でき、古典(従来型)高性能計算と組み合わせたハイブリッド計算を実行できます。本講習会では、量子計算プラットフォームに興味のある大学・研究所・企業の研究者及び学生を対象に、ハンズオンを通して玄界におけるCUDA-Qの利用方法を理解することを目的とします。

講習内容予定

- NVIDIA Quantum の紹介
- CUDA-Q入門
- 発展的な利用例とハンズオン
- 持ち込み課題に関する相談も可

開催場所詳細

九州大学 伊都キャンパス
 情報基盤研究開発センター
 2階 多目的教室 (203号室)
 キャンパスマップの43番
 入り口は建物の北側です

九州大学 伊都キャンパス
 アクセス マップ (PDF)



<https://www.kyushu-u.ac.jp/ja/campus/ito/>

運営体制

問い合わせ先: qda-cudaq_handson@googlegroups.com

主催: 九州大学 情報基盤研究開発センター

共催: 九州大学 量子コンピューティングシステム研究センター、
 九州大学 データ駆動イノベーション推進本部 (DX推進本部)、エヌビディア合同会社

その他、今後の展望

- 現在のOOD利用状況
 - ログインユーザ（アカウント）数
 - 10月 77 アカウント
 - 11月 47 アカウント
 - 機能の充実が足りないこともあり控えめな数
- 足りないもの
 - マニュアルの充実
 - 利用者への周知