



Open OnDemandの概要

中尾昌広（理化学研究所 計算科学研究センター）

背景

- HPCクラスタの問題点として、HPCクラスタを用いるための前提知識が多く、利用するまでの学習コストが大きい
 - シェルなどのコマンドラインインタフェース
 - SSH鍵ペアの生成と公開鍵の登録
 - ジョブスケジューラ
 - ・ スパコンによって異なる場合もある
- 計算ノード上でリモートデスクトップやJupyterLabなどのGUIアプリケーションを対話的に実行したい要望があるが、その手順は複雑
 - リモートデスクトップの設定や、SSHトンネリングの設定



```
top
Processes: 596 total, 2 running, 594 sleeping, 2418 threads      10:05:40
Load Avg: 2.15, 1.81, 1.63  CPU usage: 4.39% user, 3.8% sys, 92.51% idle
SharedLibs: 643M resident, 120M data, 40M linkedit.
MemRegions: 96880 total, 4270M resident, 425M private, 2994M shared.
PhysMem: 14G used (1893M wired, 364M compressor), 992M unused.
VM: 223T vsize, 4283M framework vsize, 160759(0) swapins, 248728(0) swapouts.
Networks: packets: 10719029/12G in, 2457839/525M out.
Disks: 8336373/93G read, 2912845/60G written.

PID    COMMAND      %CPU  TIME    #TH   #WQ   #PORT  MEM    PURG  CMPRS  PGRP
177    WindowServer 31.0  03:44:15 26    10    4438+  2247M  223M+  79M   177
0      kernel_task  6.0   01:53:35 484/8  0     0      2002M+ 0B     0B    0
78311  top          5.3   00:00.43 1/1    0     28     5761K  0B     0B    78311
78313  screencaptur 4.1   00:00.33 4      3     71     7618K  752K  0B    1803
1794   iTerm2       3.7   22:59.68 6      3     422-   226M+  13M-  83M   1794
```

Open OnDemand

<https://openondemand.org>

- HPCクラスタ用のWebポータルOpen OnDemandに着目
- オハイオ州立大学が中心となって開発
 - オープンソースソフトウェア
 - 250以上の世界中の機関で利用されている
- WebブラウザからHPCクラスタの操作が可能
 - 他のソフトウェアのインストールは不要
 - ファイルのアップロード・ダウンロード、ジョブ投入と監視、CLIによるシェルアクセス
 - 計算ノード上でGUIアプリケーションを簡易に利用可能
 - フレームワークによる新しいアプリ開発も可能

Welcome to the supercomputer Fugaku

RIKEN Center for Computational Science

Message of the day

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(vol0005)

Pending jobs

fugaku-small 17867	fugaku-large 915	prepost-gpu1 12	prepost-gpu2 0
prepost-mem1 22	prepost-mem2 3	prepost-ondemand 0	

Fugaku Schedule

Today: July 2023

Errors shown in time zone: Japan Standard Time

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
reccs-aet	/vol0400	5,120	608	4,512	11%	1,500,000	265,259	1,234,741	17%	527,360	81,665	445,695	15%
f-ccp	/vol0403	614,400	160,260	454,140	26%	180,000,000	144,261,340	35,738,660	80%	-	-	-	-
ra036002	/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%	-	-	-	-

Recently Used Apps

- Desktop
- OpenFOAM
- SCALE
- Jupyter

Passenger Apps

- Active Jobs
- Home Directory
- Gakunin RDM
- HPCI Storage
- Job Composer
- Fugaku Shell Access

アプリケーションの種類

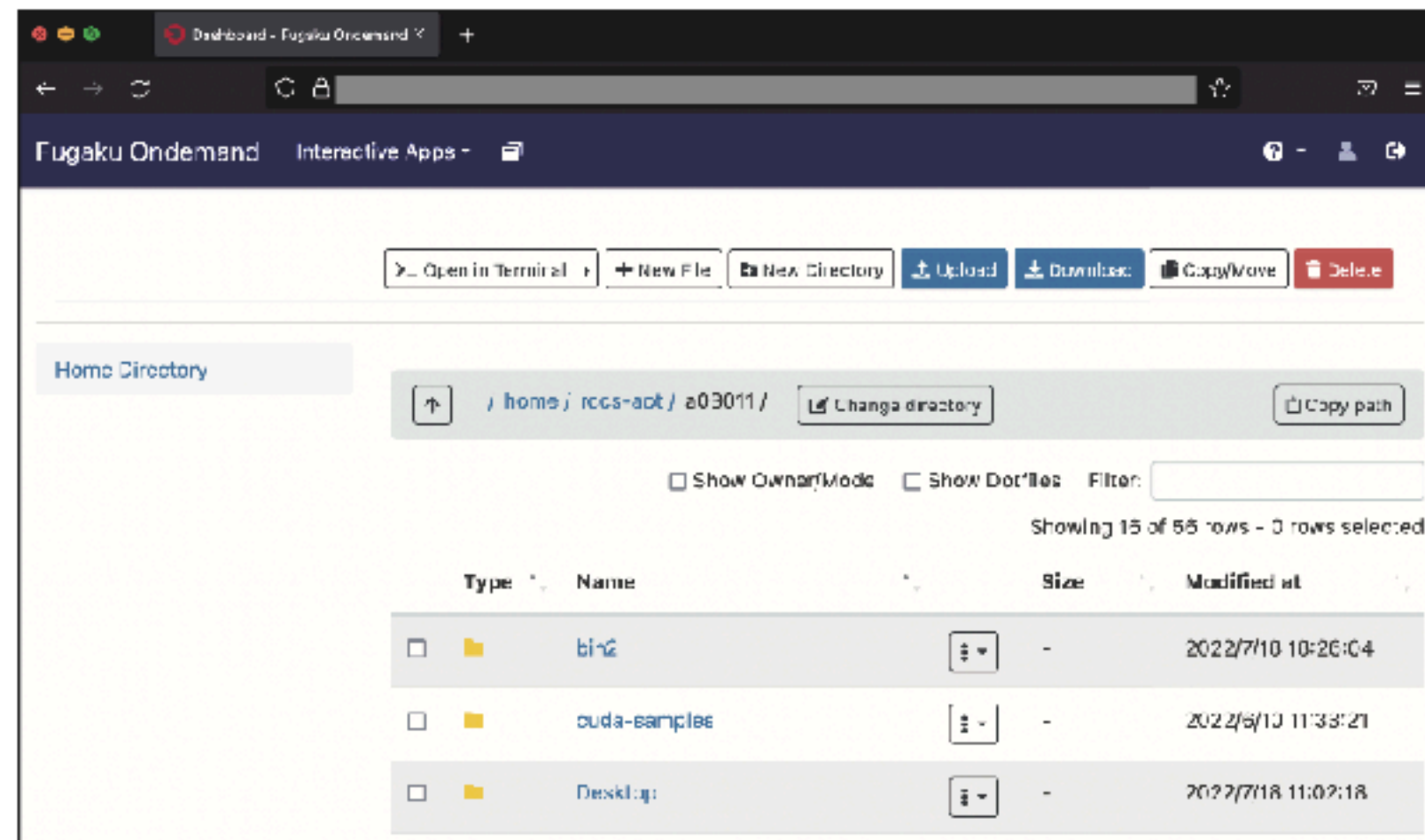
- Passenger Application : Open OnDemandがインストールされたWebサーバで動作するアプリケーション (ファイルのアップロードなど)
- Interactive Application : 計算ノード上で動作するアプリケーション
 - リモートデスクトップやJupyterLabなどの対話型GUIアプリケーション
 - 対話的操作が不要な通常のバッチジョブも可能

アイコンをクリックして
アプリケーションを起動する

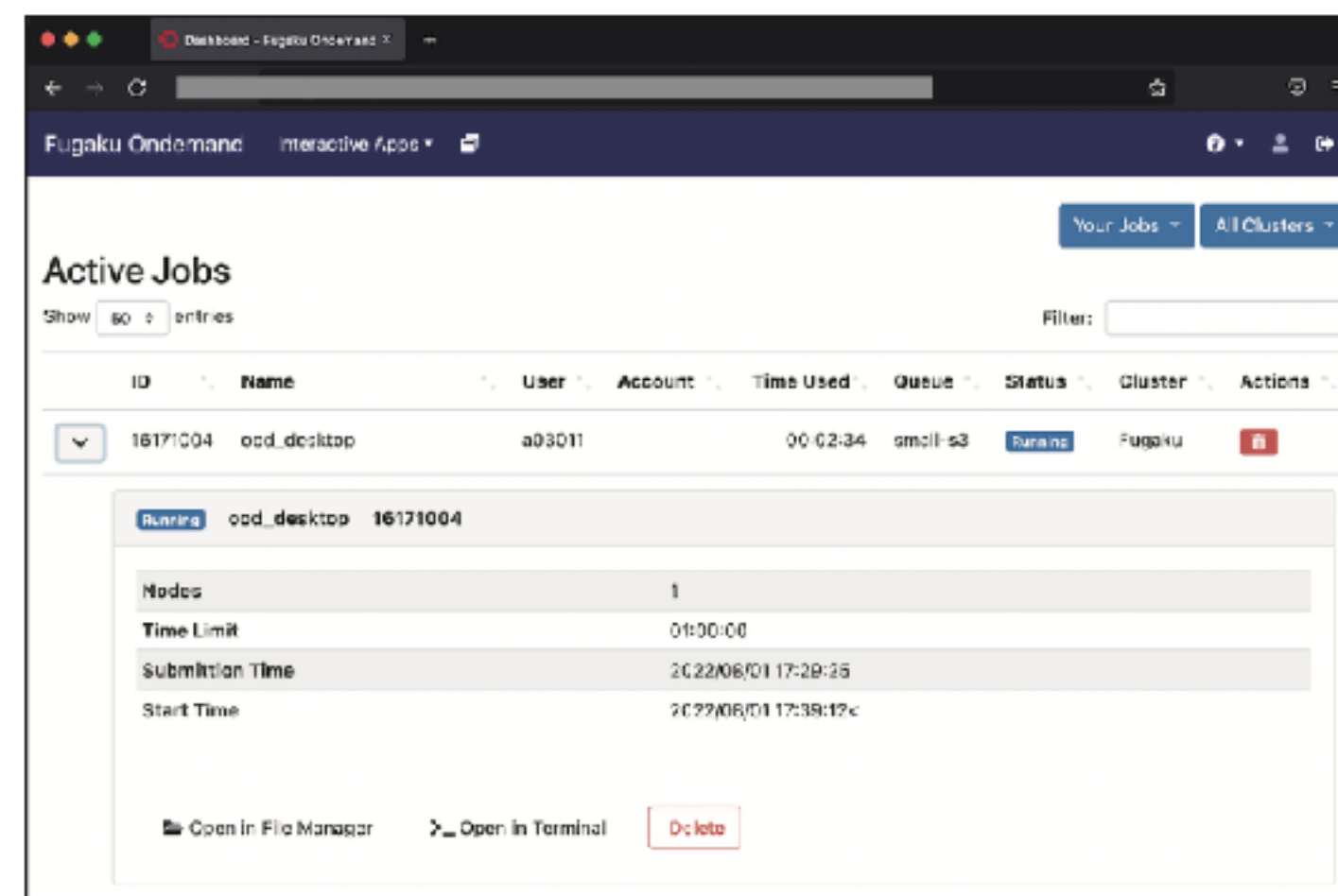
The screenshot displays the Fugaku OnDemand web interface. At the top, there is a navigation bar with links for 'Fugaku OnDemand', 'Batch Jobs', 'Interactive Apps', 'Passenger Apps', 'My Sessions', 'Develop', 'Help', and 'Log Out'. The main content area is divided into several sections:

- Welcome to the supercomputer Fugaku**: Includes the RIKEN R-CCS logo and a 'Link' menu with items like 'OnDemand Manual', 'Fugaku Portal', 'Fugaku Schedule', 'Fugaku Status', and 'Fugaku Support'.
- Message of the day**: A section for daily updates.
- Information**: A table of recent operations, including dates and descriptions of job execution periods and system maintenance.
- Pending jobs**: A grid of colored boxes representing different resource groups and their counts (e.g., 'fugaku-small' with 17867 jobs).
- Accounting**: A table showing disk usage and resource availability for various groups. The table has columns for Group, Volume, Disk (GiB) (Limit, Usage, Avail., Rate), Disk (inode) (Limit, Usage, Avail., Rate), and Resource (NH) (Limit, Usage, Avail., Rate).
- Recently Used Apps**: A row of four application icons: Desktop, OpenFOAM, SCALE, and Jupyter.
- Passenger Apps**: A grid of application icons including Active Jobs, Home Directory, Gakunin RDM, HPCI Storage, Job Composer, and Fugaku Shell Access.

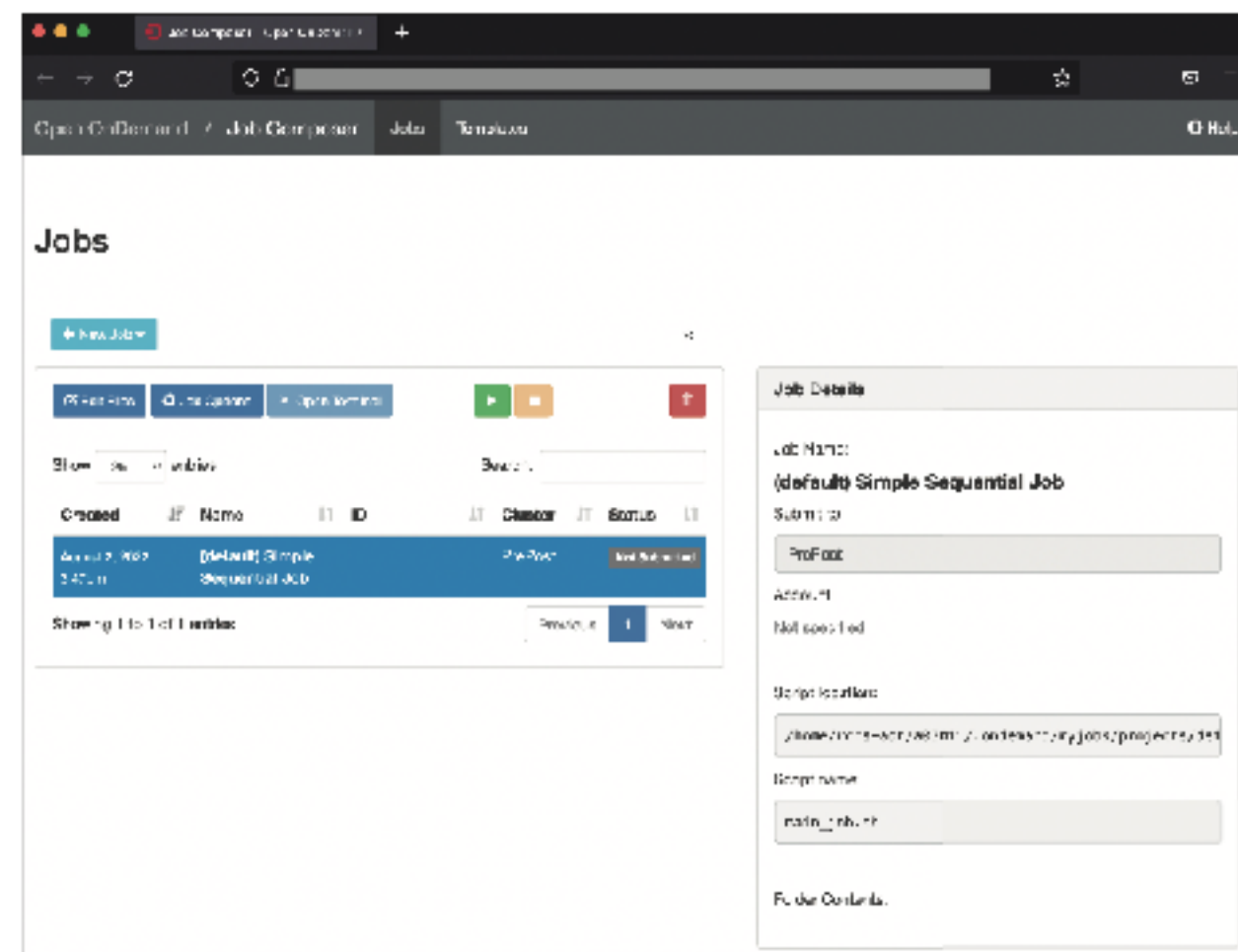
デフォルトのPassenger Applications



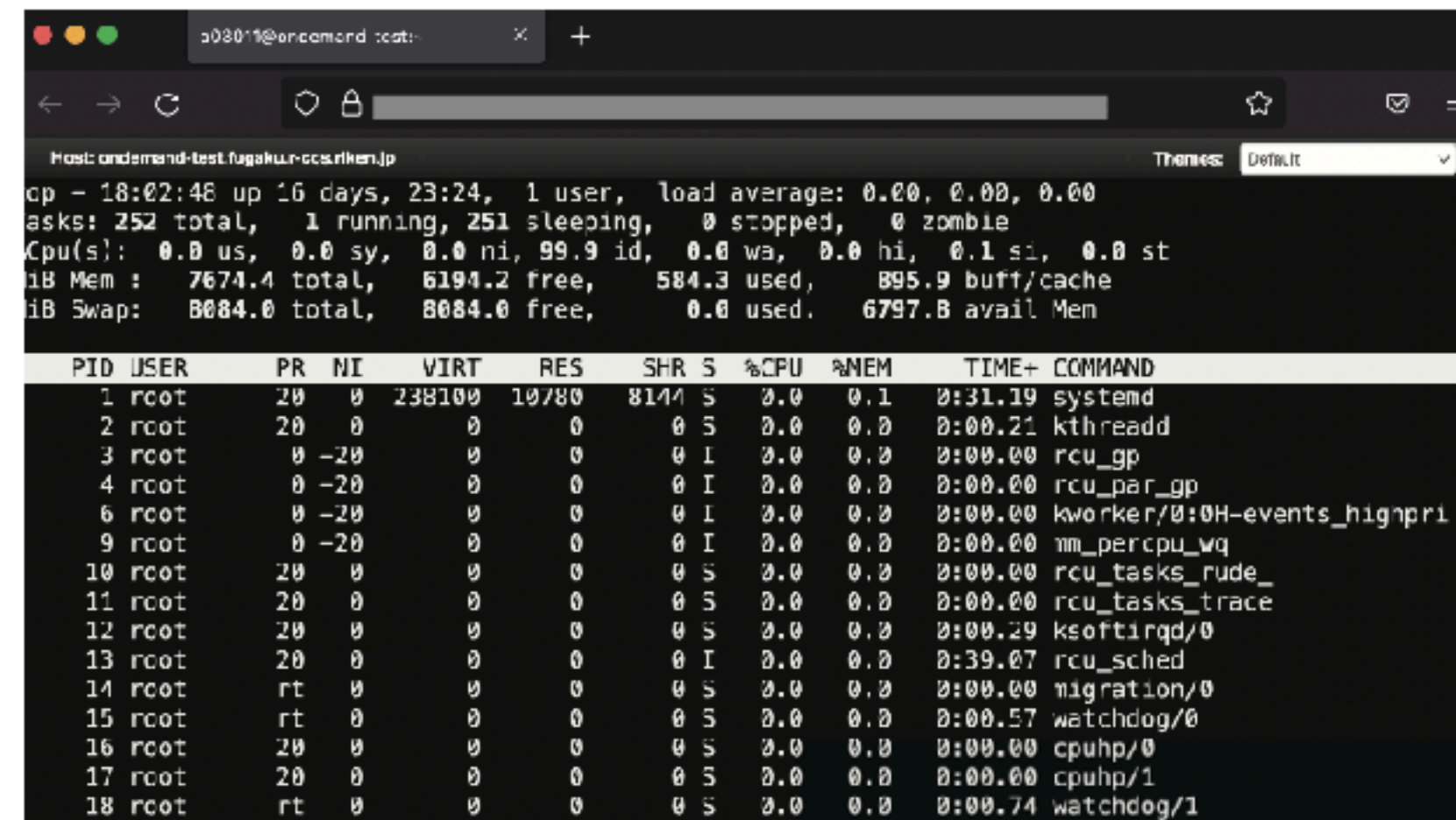
Home Directory



Active Jobs



Job Composer



Shell

- Home Directory : ファイルのアップロード・ダウンロード・編集が可能。rcloneと組み合わせることでAmazon S3などの外部ストレージとの連携も可能。

- Active Jobs : ジョブの監視

- Job Composer : ジョブの投入

- Shell : Webターミナル

フレームワークを用いることで新しいアプリケーションを開発もできる

Interactive Applicationの例（リモートデスクトップ）

Welcome to the supercomputer Fugaku

Message of the day

Information

Pending jobs

Group	Volume	Disk (GiB)	Disk (inode)	Resource (NH)									
		Limit	Usage	Avail.	Rate	Limit	Usage	Avail.	Rate	Limit	Usage	Avail.	Rate
rccs-aot	/vol0400	5,120	608	4,512	1%	1,500,000	255,259	1,234,741	7%	527,360	81,565	445,633	5%
fcp	/vol0403	814,400	160,260	654,140	20%	180,000,000	144,261,340	35,738,660	80%	-	-	-	-
ra030002	/vol0403	5,120	1	-	-	-	-	-	-	-	-	-	-
rhome	/vol0400	20	9	-	-	-	-	-	-	-	-	-	-

Recently Used Apps

Desktop

Passenger Apps

Active Jobs

Home Directory

Gakurin RDM

HPCI Storage

Job Composer

Fugaku Shell Access

Desktop

This app will launch an Xfce desktop.

Queue: fugaku-small

Group: rccs-aot

Elapsed time (1 - 72 hours): 1

Number of nodes (1 - 384): 1

Total number of processes (1 - 18,432): 1

Total number of processes <= Number of nodes x 48.

Execution mode: Normal

Please refer to the manual for details (English)

Email (You will receive an email when it starts)

Launch

必要事項を入力し、ジョブの投入を行う

Desktop (515) 1 node | 1 core | Running

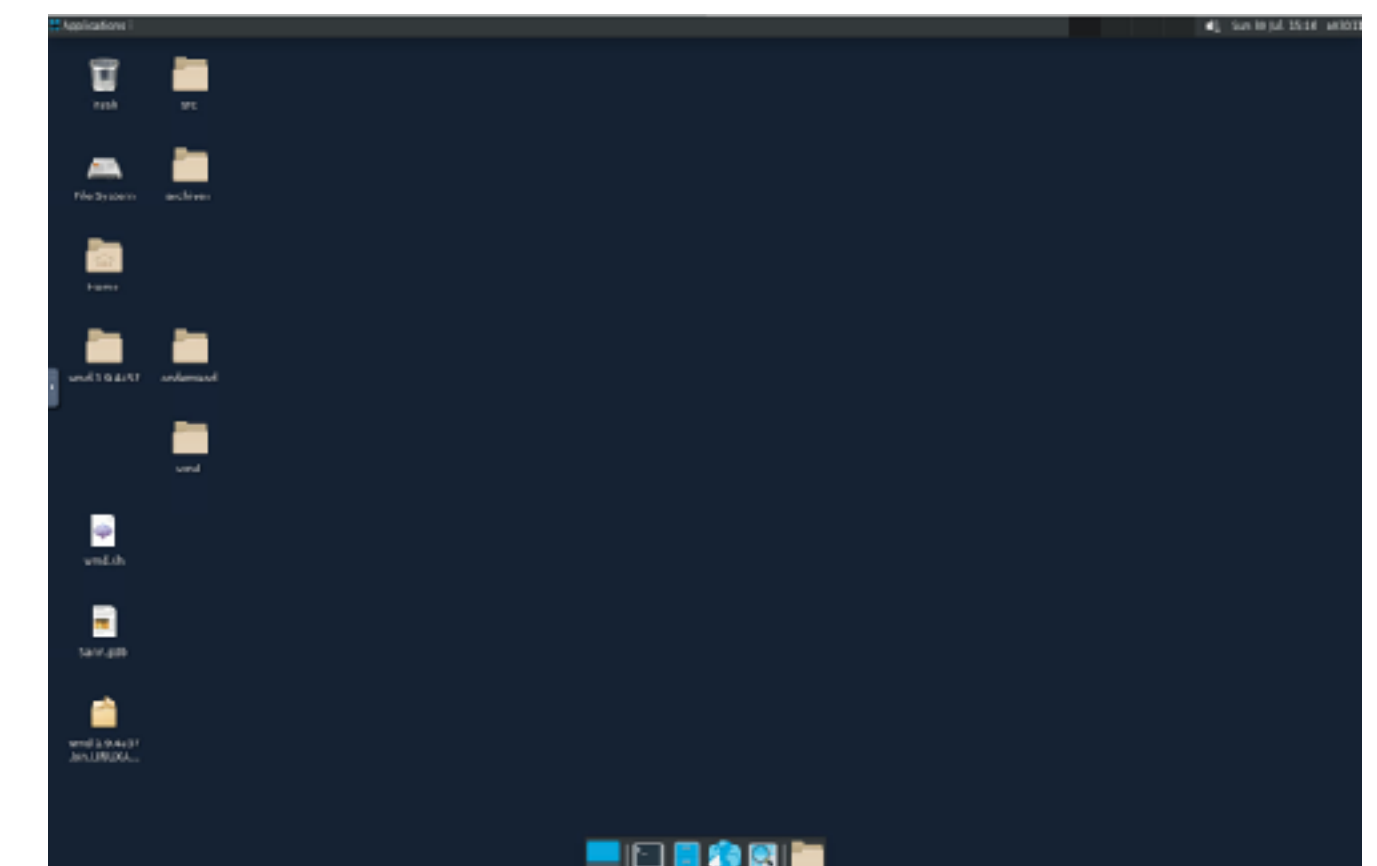
Host: >_globus.cloud.r-ccs.riken.jp

Created at: 2023-10-11 10:15:25 JST

Time Remaining: 59 minutes

Launch Desktop

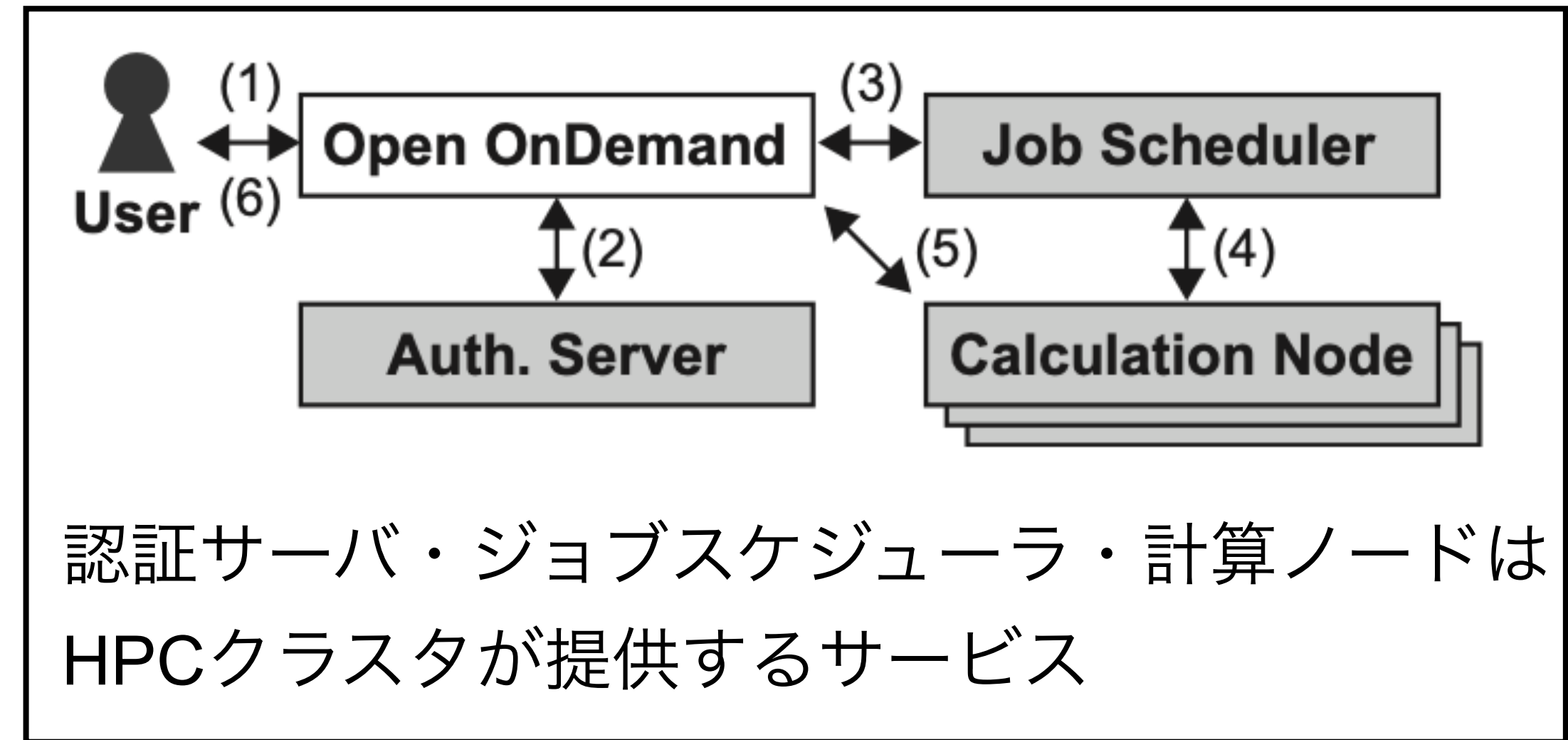
ジョブが起動するまで待つ



リモートデスクトップがブラウザから操作できる

Interactive Applicationの動作フロー

1. Webブラウザを用いてOpen OnDemandサーバにログイン
2. ログインのためのユーザ認証
3. アプリケーションの実行命令が発行されると、
計算ノードにジョブを投入
4. ジョブが実行されるまで待機
5. ジョブの実行時に、計算ノードのIPアドレスなどの情報をOpen OnDemandサーバに送信し、
リバースプロキシを設定（バッチジョブの場合は、ジョブの実行のみが行われ、ここで終了）
6. リバースプロキシ用のURLを用いて、WebブラウザからHPCシステム内部の計算ノードに接続



「富岳」の場合の従来のログイン手順との違い

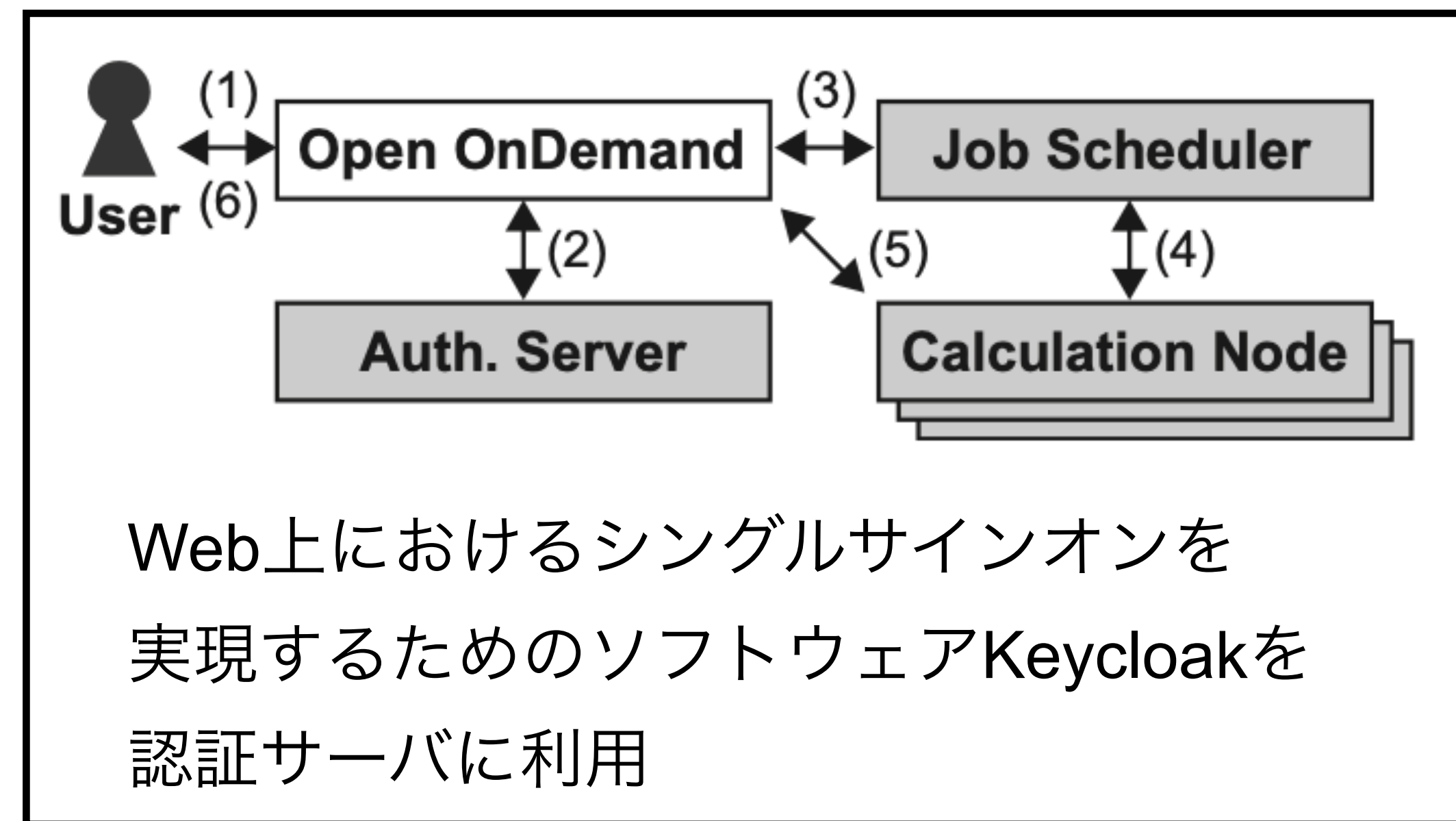
- 「富岳」の従来のログイン手順

1. クライアント証明書をWebブラウザにインストール
2. Webブラウザで「富岳」の利用者ポータルにログイン
3. (SSH鍵ペアがなければ) ローカルPCでSSH鍵ペアを作成
4. SSHの公開鍵を「富岳」の利用者ポータルに登録
5. ログインノードにSSHでログイン



- 「富岳」のOpen OnDemandを用いたログイン手順

1. クライアント証明書をユーザのWebブラウザにインストール (上と同じ)
2. WebブラウザでOpen OnDemandサーバにログイン



負荷について

- 2023年8月21日～25日に、高校生を対象としたスーパーコンピューティングコンテストを富岳を使って開催
- 昨年まではSSH + ターミナルを使っていたが、今年はOpen OnDemandを採用
 - PuTTYやWinSCPの利用、鍵の生成や登録の説明がなくなった。説明がOS非依存になった。
- 最大70名の同時利用だったが、問題なかった
- Intel Xeon Gold 6338 x 2、256GBのサーバ

