



Science and
Technology
Facilities Council

Scientific Computing

Data Analysis as a Service (DAaaS)

Group Leader: Jeremy Spencer



Background

- Scientific Computing
- DAaaS
 - Team of (currently) 10 developers
 - Running for over 5 years
 - Supporting STFC Facilities:
 - ISIS Neutron and Muon Spallation Source
 - Central Laser Facility (CLF)
 - Dedicated training platform released in 2023
- We also support Machine Learning in STFC through a different technology stack, also using IRIS resources



ISIS Neutron and
Muon Source



Central Laser Facility



Scientific Computing




What is DAaaS?

- Remote data analysis
- It brings together the data, software and compute resources for a variety of scientific workflows
 - >1PB data stored
 - >100 scientific software applications available
 - 15,000 CPUs, 50TB RAM, 100 GPU cards
- In summary:
 - A handful of clicks gets a user a remote desktop to a computer with everything they need readily available



Setting up a training course

- <https://training.analysis.stfc.ac.uk/>
- Course creation is currently open to STFC staff only
- Log in, create a course by filling in a form
- My team configures your Workspaces
- You send a sign-up link to your attendees
- They sign-up and receive a unique login link

 DAaaS Training

Demo Sign Up

Course Details

Name	Demo
Purpose of Training	To showcase our training functionality
Start Date	2023-06-28
End Date	2023-06-29

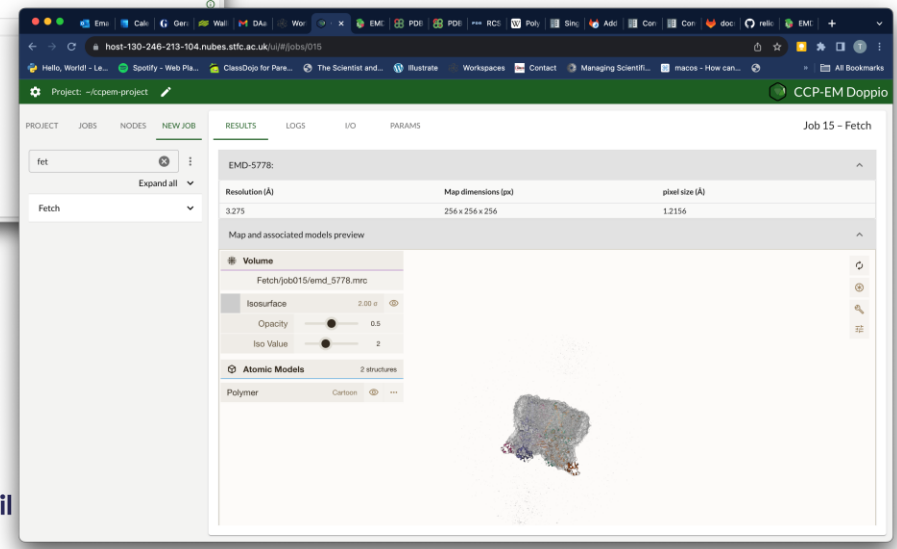
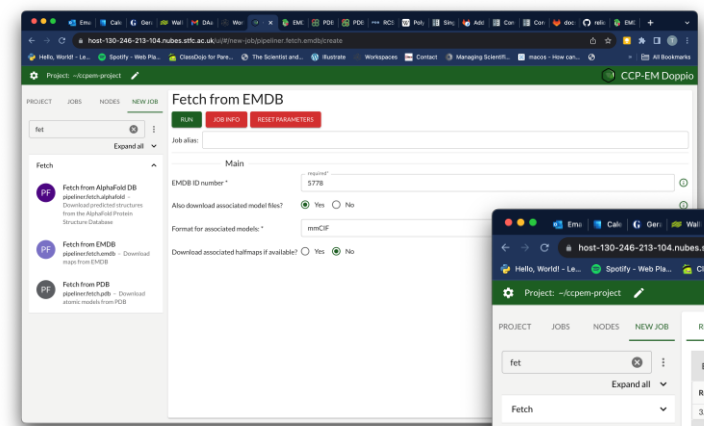
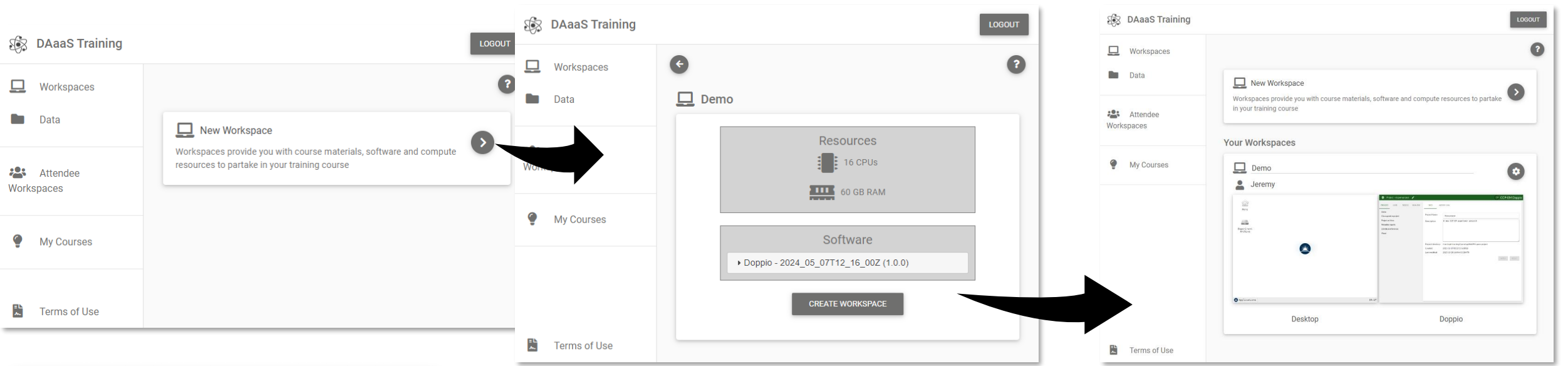
Name

Email

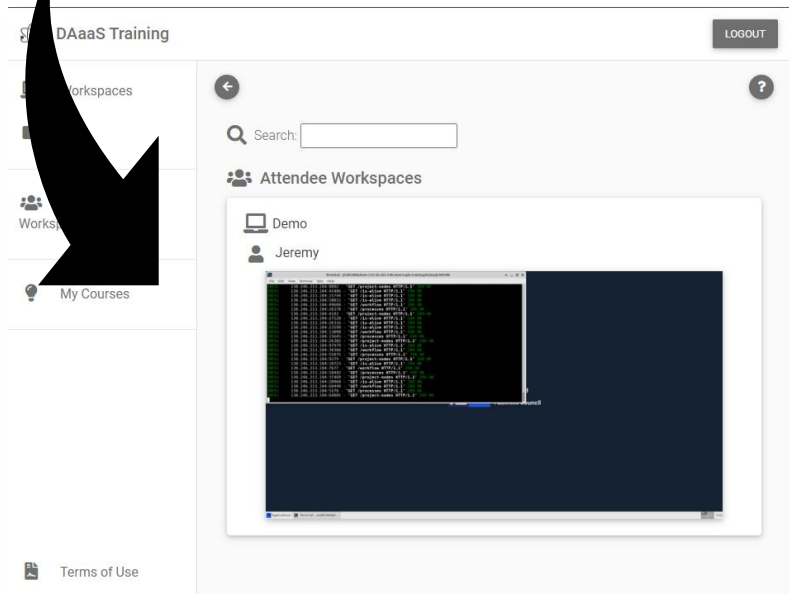
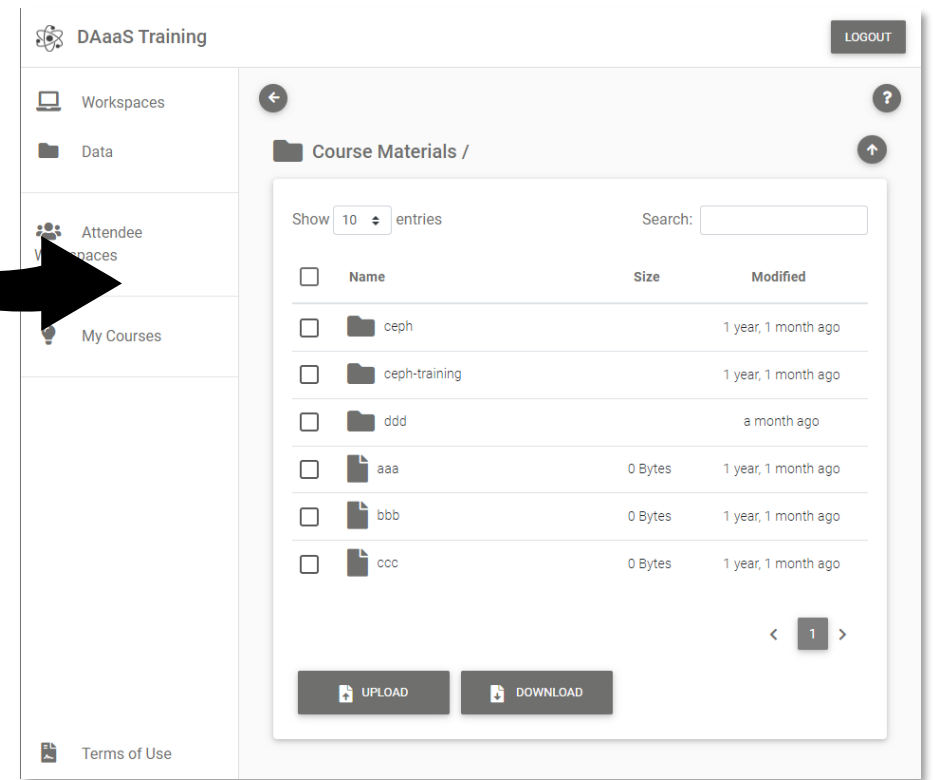
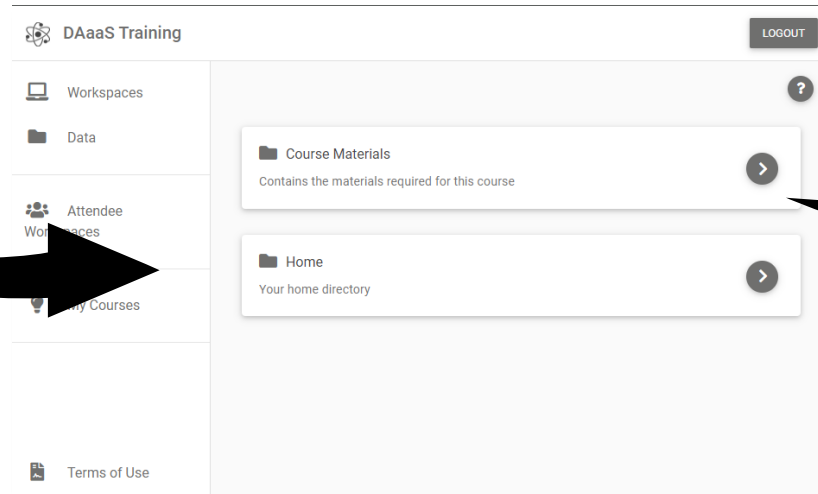
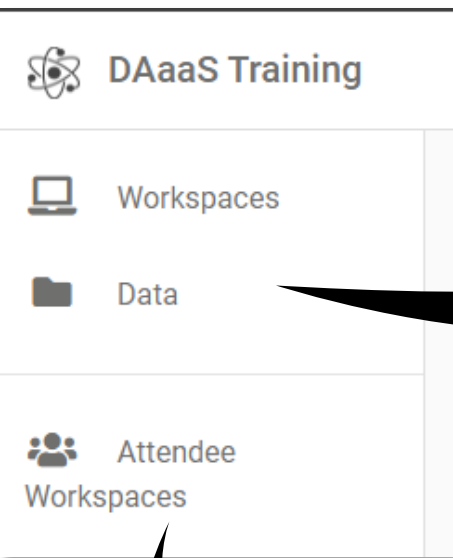
I have read and accept the terms and conditions.

Terms and Conditions

For general help with this site, find the  in the top right (these pages have specific information for each area of our site).

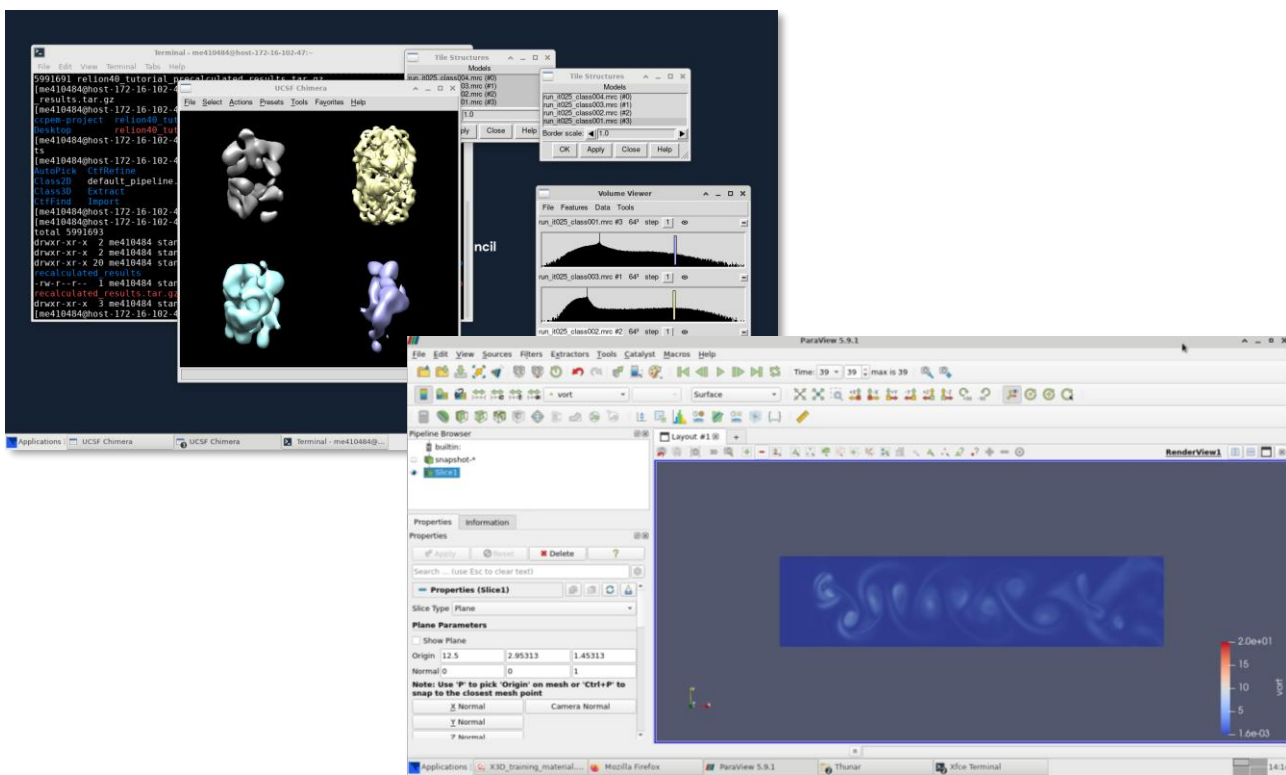


User Workflow for a course running Doppio (Web based GUI for CCP-EM)



- **Data**
 - Home
 - One per user
 - Course Materials
 - Writeable by course tutors
 - Read only for attendees
- **Attendee Workspaces (course tutors only)**
 - Preview all users' workspaces
 - Connect and provide help

CoSeC/CCP Training



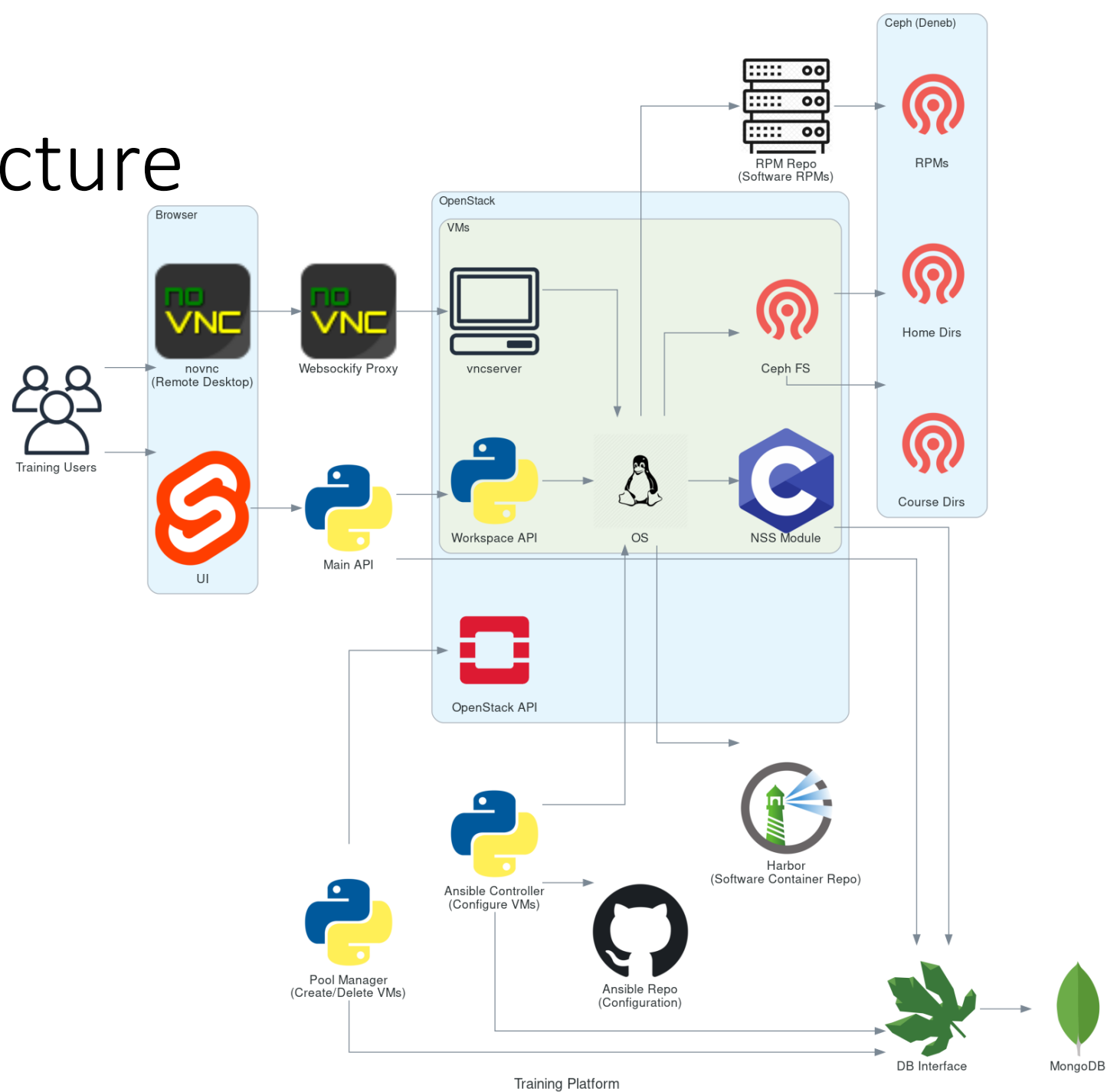
- **CCP-WSI OpenFOAM Parallel Performance Training**
 - June 2023
 - Training for applications scientists to simulate Wave Structure Interaction type problems
 - 35 users
- **CECAM School for CFD and HPC**
 - October 2023
 - CFD with Code_Saturne and HPC
 - 20 users
- **IBSim 4i 2023**
 - October 2023
 - Presenting software for Image Based Simulation
 - 30 users
- **Spring8**
 - November 2023
 - Teaching CCP-EM tools for EM model building
 - 46 users
- **CCP-EM EMBL Course**
 - February 2024
 - Teach new users at EMBL how to use our Doppio software to process cryoEM data
 - 20 users
- **Incompact3D Training Event**
 - April 2024
 - Training event for the UKTC community with Incompact3D
 - 15 users
- **CCP-EM NEMI NL CryoEM School**
 - June 2024
 - Teach users from Dutch CryoEM community how to use CCP-EM tools
 - 25 users

Challenges

- Resourcing hardware
- Particularly graphics cards
- One card is needed for each workspace, so one per user
- For these short-term courses, this creates large spikes in demand



Architecture



Systems capabilities

Storage and Databases

- Storage capabilities for petabytes of data on:
 - Spinning disk
 - Tape
- Centralised database hosting

Cloud

- Over 100,000 CPU cores
- Over 250TB RAM
- GPU/Graphics cards also available

Hardware Operations

- Installing and managing physical hardware

19^c ORACLE[®]
Database



openstack[®]

