



ECCO Summer School Cloud Computing Resources

Ou Wang, Andrew Delman, Ian Fenty

Jet Propulsion Laboratory, California Institute of Technology

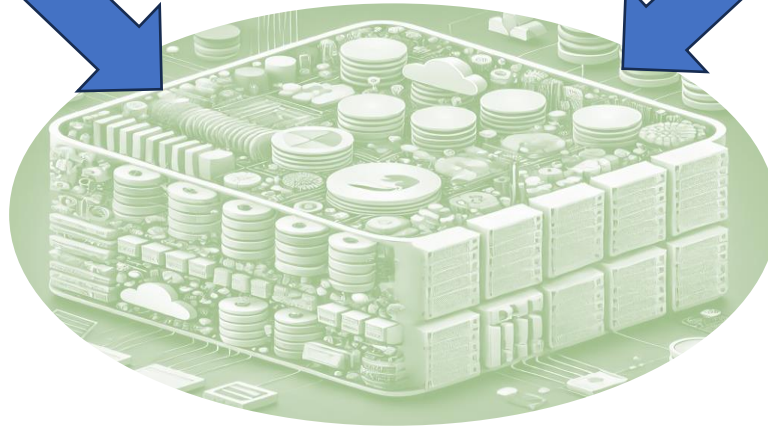
ECCO Summer School, May 19-30, 2025
Asilomar Conference Grounds, Pacific Grove, California
© 2025. All rights reserved.

Computer Resources

Parallel-Cluster (**P-Cluster**): *Parallel Computing and Analysis*



Open Science Studio (**OSS**): *Analysis*



Amazon EFS (Elastic File System; **/efs_ecco**): *Storage*

Setting up and working in the Open Science Studio

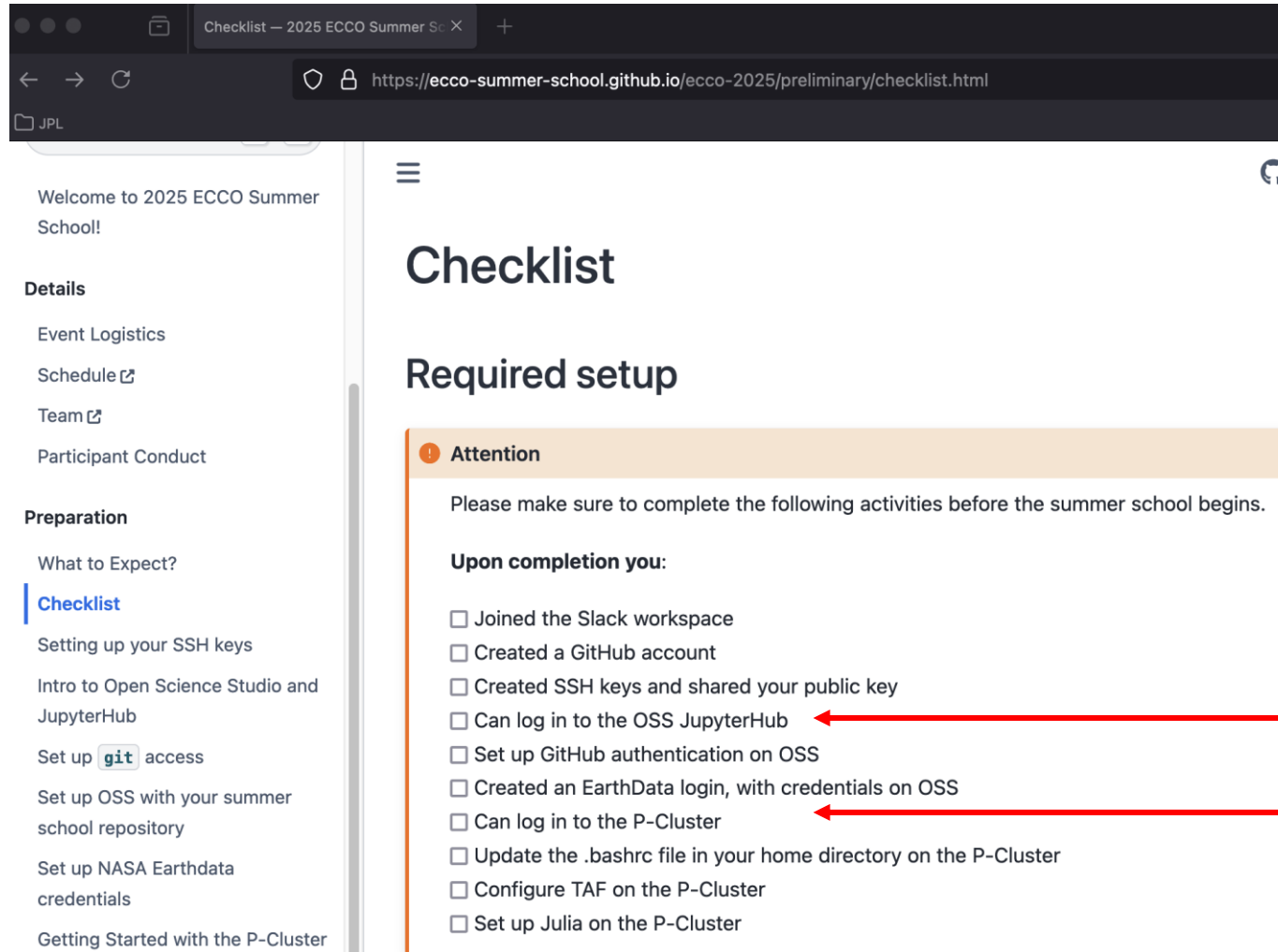
Andrew Delman – UCLA/JIFRESSE

Outline

- “Check in” with the preparation checklist
- What is Open Science Studio (OSS)?
- Logging in to OSS
- Using *git* and GitHub to collaborate on coding projects

The preparation checklist

- Don't worry if you haven't done all of these...but try them before you start your project



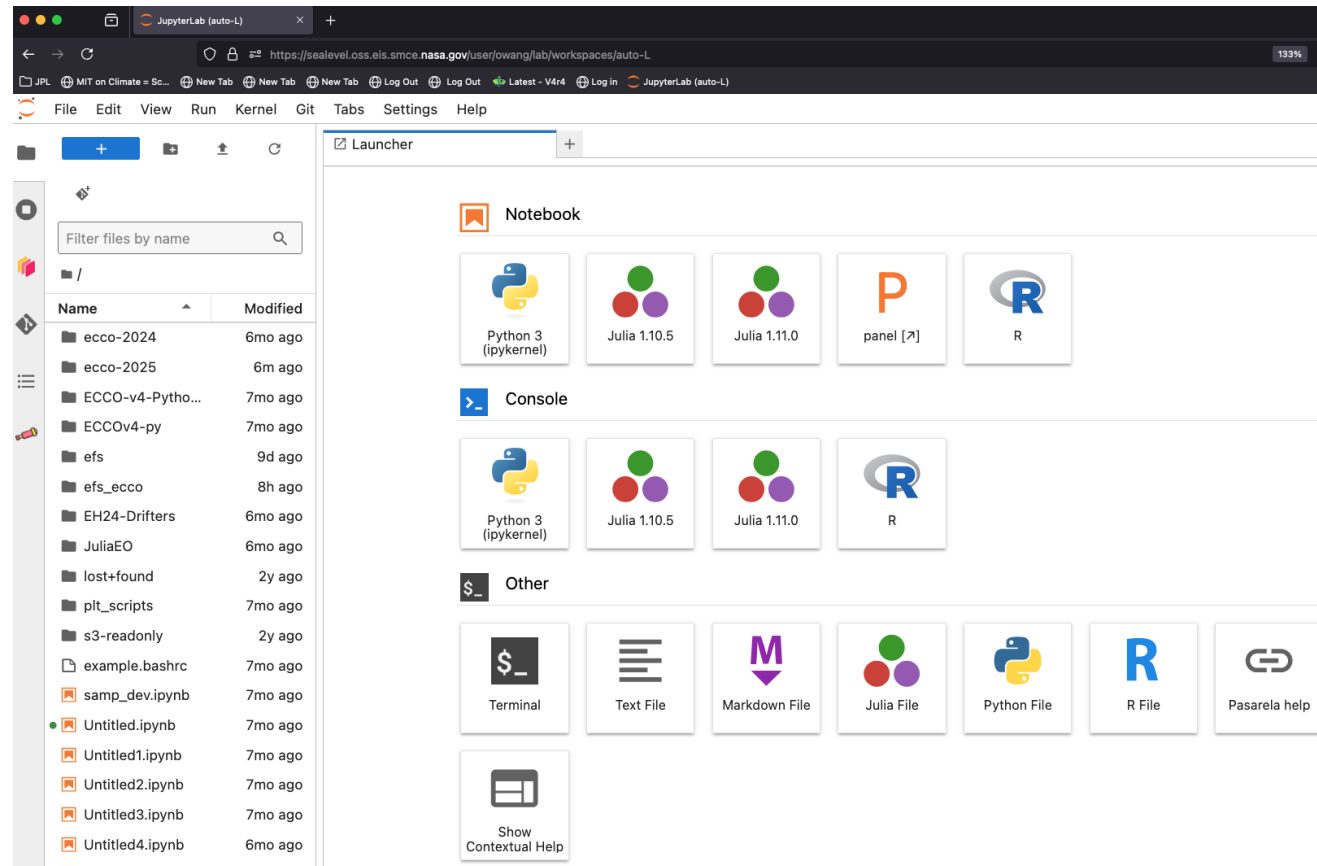
Can log in to OSS JupyterHub

Can log in to P-Cluster

- If you haven't tried logging in to OSS (<https://sealevel.oss.eis.smce.nasa.gov>), get your laptop ready to try it now!

What is Open Science Studio (OSS)?

- A platform developed by the Science Managed Cloud Environment (SMCE) group at NASA Goddard Space Flight Center (Greenbelt, MD)
- When you log into OSS (<https://sealevel.oss.eis.smce.nasa.gov>), a JupyterHub server is launched in the Amazon Web Services (AWS) Cloud, in the same region (U.S. West-2) as ECCO output and other NASA earth science data

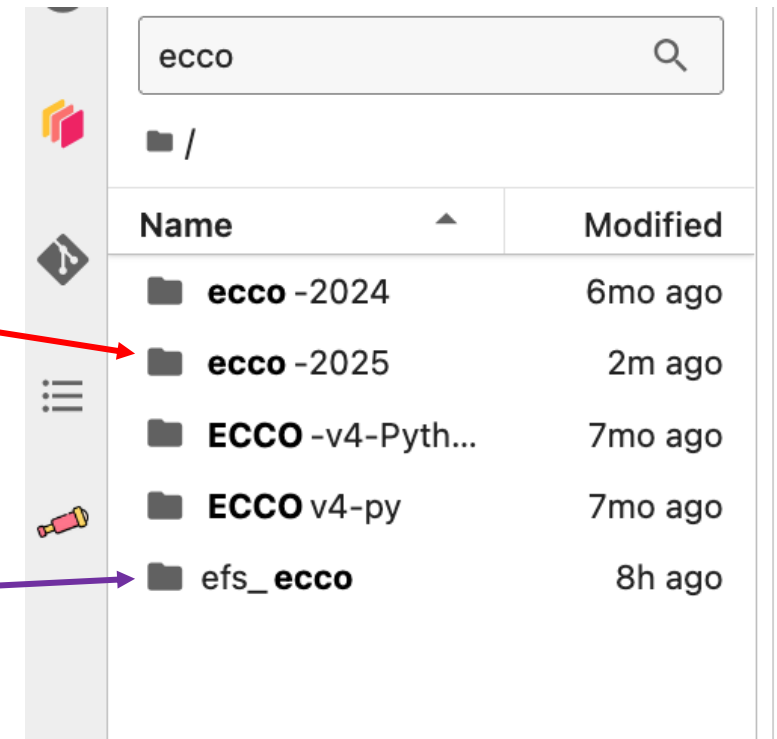


What is Open Science Studio (OSS)?

- Each user gets their own private filesystem, but also access to shared resources through `/efs_ecco`
- OSS (especially in the *ecco-2025* project directory) is the place to store your Jupyter notebooks and any associated project files such as documentation and images

***ecco-2025*: once you clone this repository from GitHub, you will find all the ECCO Summer School website info in this folder (including tutorials)**

***efs_ecco*: shared storage volume (check that you have access to this)**

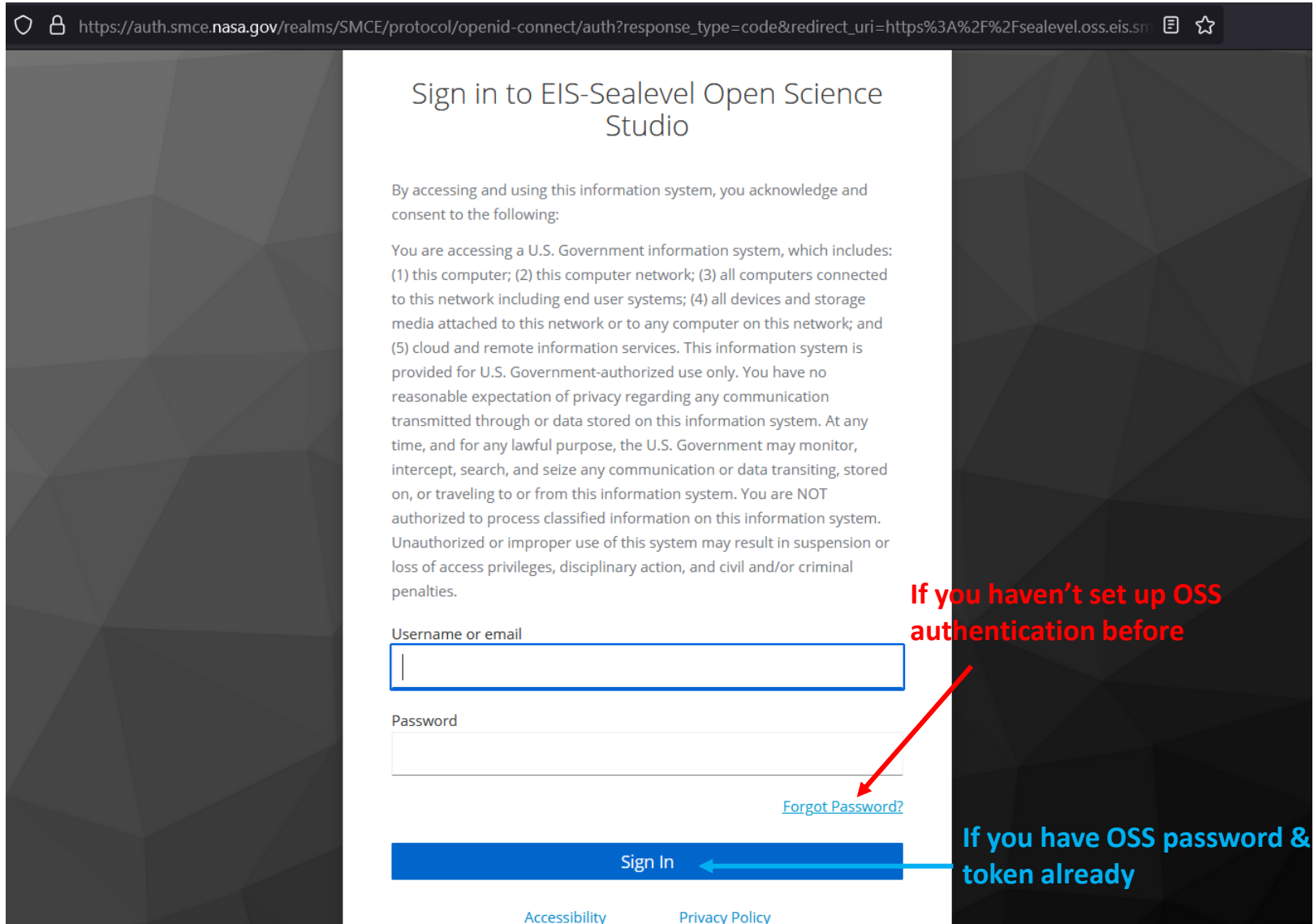


Name	Modified
 ecco -2024	6mo ago
 ecco -2025	2m ago
 ECCO -v4-Pyth...	7mo ago
 ECCO v4-py	7mo ago
 efs_ecco	8h ago

- Outside of `/efs_ecco`, **please do not store large data files on OSS**...total storage limit on OSS is 150 GB!

Logging in to OSS

- Go to <https://sealevel.oss.eis.smce.nasa.gov> – bookmark this page if you haven't yet!



The screenshot shows the login page for the EIS-Sealevel Open Science Studio. The page has a dark grey background with a geometric pattern. The main content area is white and contains the following text:

Sign in to EIS-Sealevel Open Science Studio

By accessing and using this information system, you acknowledge and consent to the following:

You are accessing a U.S. Government information system, which includes: (1) this computer; (2) this computer network; (3) all computers connected to this network including end user systems; (4) all devices and storage media attached to this network or to any computer on this network; and (5) cloud and remote information services. This information system is provided for U.S. Government-authorized use only. You have no reasonable expectation of privacy regarding any communication transmitted through or data stored on this information system. At any time, and for any lawful purpose, the U.S. Government may monitor, intercept, search, and seize any communication or data transiting, stored on, or traveling to or from this information system. You are NOT authorized to process classified information on this information system. Unauthorized or improper use of this system may result in suspension or loss of access privileges, disciplinary action, and civil and/or criminal penalties.

Username or email

Password

[Forgot Password?](#)

[Sign In](#)

[Accessibility](#) [Privacy Policy](#)

Annotations on the right side of the page:

- If you haven't set up OSS authentication before** (red text, with a red arrow pointing to the 'Forgot Password?' link)
- If you have OSS password & token already** (blue text, with a blue arrow pointing to the 'Sign In' button)

Logging in to OSS - <https://sealevel.oss.eis.smce.nasa.gov>

- If you are setting up authentication, you should get a prompt to send you an e-mail
- Then you will be asked to set up a password, and an app-based token (scan a QR code in Microsoft Authenticator, Google Authenticator, or similar app)
- After you set up two-factor authentication, when you go to the URL above, you will enter a password to sign in, and then your app-based token on the second screen

SCIENCE MANAGED CLOUD ENVIRONMENT

Forgot Your Password?

Username or email

[« Back to Login](#)

Submit

Enter your username or email address and we will send you instructions on how to create a new password.

SCIENCE MANAGED CLOUD ENVIRONMENT

adelman 

One-time code

Sign In

Logging in to OSS – server selection

- Once logged in to OSS you will see a screen to launch your personal server on the OSS JupyterHub
 - Smaller servers (especially **Large** and **Extra Large**) are sufficient for experimenting with and running most notebooks
 - Larger **ECCO Hackathon** servers could be helpful for running notebooks that use a lot of memory
 - For example, calculations with global full-depth monthly ECCOv4 output during all 26 years

Server Options

● EIS-Sealevel Open Science Studio

Images

Earth Science Server

Server Type

Standard - Up to 4GB RAM/4 CPU

Standard - Up to 4GB RAM/4 CPU

Large - Up to 8GB RAM/4 CPU

Extra Large - Up to 16GB RAM/4 CPU

ECCO Hackathon - XL - Up to 32GB RAM/4 CPU

ECCO Hackathon - 2XL - Up to 64GB RAM/8 CPU

ECCO Hackathon - 8XL - Up to 256GB RAM/32 CPU

4XL Server (On-Demand) - APPROVAL ONLY - Up to 64GB RAM/16 CPU

Logging in to OSS - <https://sealevel.oss.eis.smce.nasa.gov>

The screenshot shows the JupyterLab interface. On the left is the file browser, and on the right is the launcher. A purple arrow points to the 'efs_ecco' directory in the file browser.

File Browser:

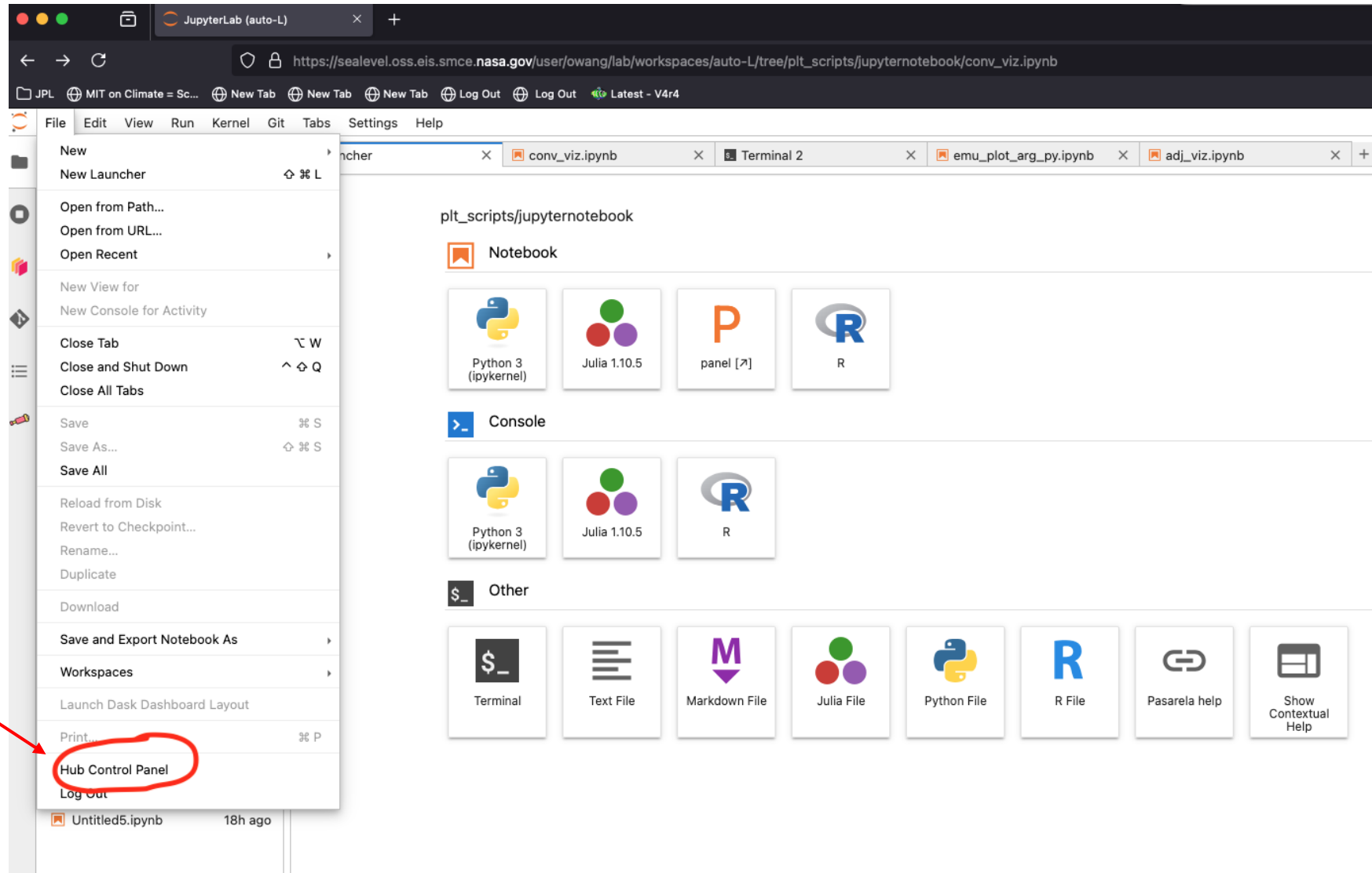
Name	Modified
ecco-2024	6mo ago
ecco-2025	6m ago
ECCO-v4-Pytho...	7mo ago
ECCOV4-py	7mo ago
efs	9d ago
efs_ecco	8h ago
EH24-Drifters	6mo ago
JuliaEO	6mo ago
lost+found	2y ago
plt_scripts	7mo ago
s3-readonly	2y ago
example.bashrc	7mo ago
samp_dev.ipynb	7mo ago
Untitled.ipynb	7mo ago
Untitled1.ipynb	7mo ago
Untitled2.ipynb	7mo ago
Untitled3.ipynb	7mo ago
Untitled4.ipynb	6mo ago

Launcher:

- Notebook:** Python 3 (ipykernel), Julia 1.10.5, Julia 1.11.0, panel [?], R
- Console:** Python 3 (ipykernel), Julia 1.10.5, Julia 1.11.0, R
- Other:** Terminal, Text File, Markdown File, Julia File, Python File, R File, Pasarela help, Show Contextual Help

Your home directory will not look like this now, but check that *efs_ecco* is mounted there

Stop OSS server - <https://sealevel.oss.eis.smce.nasa.gov>



File →
Hub Control Panel →
Stop My Server

Using *git* and GitHub

- The ecco-2025 GitHub repository (<https://github.com/ECCO-Summer-School/ecco-2025>) is the place that we will use to share notebooks, code, etc. that project teams have written
- The *git* command line interface tool allows you to manage code edits locally as well as interact with online GitHub repositories
- Note there are 3 tutorials to help you get set up with and start using *git*!
 - **Set up *git* access:** how to authenticate with GitHub
 - **Set up OSS with your summer school repository:** *git clone* the ecco-2025 repository
 - **Using *git* for version control -> Using *git* to edit and publish code:**
Get acquainted with some of the most commonly used commands
(*git clone*, *git branch*, *git pull*, *git add*, *git commit*, *git push*)

Preparation	
What to Expect	
Checklist	
Intro to Open Science Studio and JupyterHub	
Set up <i>git</i> access	
Set up OSS with your hackweek repository	
Set up NASA Earthdata credentials	
Getting Started with the P-Cluster	
Guidelines to Set Up Julia	
Tutorials	
ECCO general information	▼
ECCO data access	▼
ECCO v4 computations	▼
Julia Tutorials	▼
P-Cluster Tutorials	▼
EMU	▼
Using git for version control	^
Using <i>git</i> to edit and publish code	
Algorithmic Differentiation (AD)	▼



P-Cluster Overview

- Very much like a high-performance computing (HPC) system where you can run interactive jobs and perform parallel computing
- Log in using `ssh -i /path/to/privatekey -X USERNAME@34.210.1.198`
- Home directory `/home/USERNAME`; suitable for storing source code, configuration files, documents etc.
- **Update the `.bashrc` file in your P-Cluster home directory**
- SLURM as batch job scheduler
- Choose different partitions (queue/machine) based on job size
- Conduct run and intensive analysis, write large data in `/efs_ecco/USERNAME`
- Back up your source code on permanent places; ECCO Summer School systems go away in a couple of weeks.

Log in to P-Cluster

```
(base) [owang@penguin ~]$ ssh -i ~/.ssh/id_rsa -X owang@34.210.1.198
#####
* This is an SMCE System
* By accessing and using this information system, you acknowledge and consent to the following:
*
* You are accessing a U.S. Government information system, which includes: (1) this computer;
* (2) this computer network; (3) all computers connected to this network including end user
* systems; (4) all devices and storage media attached to this network or to any computer on
* this network; and (5) cloud and remote information services. This information system is
* provided for U.S. Government-authorized use only. You have no reasonable expectation of
* privacy regarding any communication transmitted through or data stored on this information
* system. At any time, and for any lawful purpose, the U.S. Government may monitor, intercept,
* search, and seize any communication or data transiting, stored on, or traveling to or from
* this information system. You are NOT authorized to process classified information on this
* information system. Unauthorized or improper use of this system may result in suspension or
* loss of access privileges, disciplinary action, and civil and/or criminal penalties.
*
#####
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-1083-aws x86_64)

* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:       https://ubuntu.com/pro

System information as of Mon May 19 04:51:08 UTC 2025

System load:  0.01          Processes:            177
Usage of /:   51.6% of 72.50GB Users logged in:      1
Memory usage: 18%          IPv4 address for ens5: 10.20.22.69
Swap usage:   0%

* Ubuntu Pro delivers the most comprehensive open source security and
  compliance features.

https://ubuntu.com/aws/pro

Expanded Security Maintenance for Applications is enabled.

0 updates can be applied immediately.

New release '22.04.5 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Mon May 19 04:45:30 2025 from 137.78.251.63
owang@ip-10-20-22-69:~$ pwd
/home/owang
```

- Log in using `ssh -i /path/to/privatekey -X USERNAME@34.210.1.198`

Message from the U.S. Government

Ubuntu Information

System Information

Software Update Information

Home directory: `/home/USERNAME`

SLURM: Batch Scheduler on P-Cluster

Common SLURM commands

Command	Description	Example
<code>sbatch</code>	Submits a job script to the SLURM scheduler.	<code>sbatch job_script.sh</code>
<code>scancel</code>	Cancels a pending or running job.	<code>scancel <job_id></code>
<code>squeue</code>	Displays information about jobs in the queue.	<code>squeue</code>
<code>sinfo</code>	Displays information about available SLURM nodes and partitions.	<code>sinfo</code>
<code>salloc</code>	Allocates resources for a job interactively.	<code>salloc --ntasks=2 --ntasks-per-node=2 --partition=sealevel-c5xl-demand --time=01:00:00</code>
<code>srun</code>	Submits a job or launches parallel tasks (can be used in a script or interactively).	<code>srun --ntasks=4 ./my_program</code>



Partitions (Queues/Machines): Amazon EC2 Instances

Use sinfo to check available partitions

```
[owang@ip-10-20-22-69:~]$ sinfo
PARTITION      AVAIL  TIMELIMIT  NODES  STATE NODELIST
sealevel-c5n18xl-spot*    up    infinite     50  idle~ sealevel-c5n18xl-spot-dy-c5n18xlarge-[1-50]
sealevel-c5n18xl-demand   up    infinite     50  idle~ sealevel-c5n18xl-demand-dy-c5n18xlarge-[1-50]
sealevel-c5xl-spot       up    infinite    1000  idle~ sealevel-c5xl-spot-dy-c5xlarge-[1-1000]
sealevel-c5xl-demand     up    infinite    1000  idle~ sealevel-c5xl-demand-dy-c5xlarge-[1-1000]
```

demand vs. spot:

- Differences in priority: “*demand*” instances have higher priorities than “*spot*”. AWS can terminate “*spot*” instances if there is a demand for them.
- Use “*demand*” partitions unless it is acceptable for your job to be interrupted.

c5n18xl vs. c5xl:

- Differences in computer resources. The c5n18xl instances have more CPU and memory than c5xl.
- Use c5n18xl instances for larger jobs and c5xl for smaller jobs.



Head Node, Interactive Compute Node, and Batch Compute Node

Head Node (the P-Cluster Node when SSHing to it)

Limited resources, best for editing and tiny analysis

Interactive Compute Node:

Suitable for small analysis and computing

Request an interactive node for small analysis and computing:

```
salloc --ntasks=2 --ntasks-per-node=2 --partition=sealevel-c5xl-demand --time=01:00:00  
srun --jobid=NNNN --pty /bin/bash
```

Batch Compute Node:

Running jobs in batch mode

```
sbatch --ntasks=4 --ntasks-per-node=36 --partition=sealevel-c5n18xl-demand --time=24:00:00 job.bash
```

or

```
sbatch job.bash
```

where SLURM directives are specified in `job.bash`

Try out MPI *Hello World* Examples



Run MPI Hello World Examples

This tutorial introduces examples for running a simple MPI "Hello World" Fortran code using SLURM, both as an interactive job and as a batch job.

Note

The files mentioned in this tutorial are accessible on the P-Cluster at `/efs_ecco/ECC0/misc/hello/`. Some of them are also provided as downloadable links within this tutorial.



efs_ecco

- Amazon EFS (Elastic File System); used very much like a local disk
- Bridge between P-Cluster and OSS, accessible from both systems as /efs_ecco
- Unlimited size
- Shared disk space with **unsafe** write permissions for all; **Be cautious when deleting files—double-check to avoid mistakes**
- Files will be removed after a couple of weeks

Setup Checklist

! Attention

Please make sure to complete the following activities before the summer school begins.

Upon completion you:

- ☐ Joined the Slack workspace
- ☐ Created a GitHub account
- ☐ Created SSH keys and shared your public key
- ☐ Can log in to the OSS JupyterHub
- ☐ Set up GitHub authentication on OSS
- ☐ Created an EarthData login, with credentials on OSS
- ☐ Can log in to the P-Cluster
- ☐ Update the .bashrc file in your home directory on the P-Cluster
- ☐ Configure TAF on the P-Cluster
- ☐ Set up Julia on the P-Cluster



Reminder List

- Frequently back up your source code, configuration files, documents, etc., to permanent archive locations during the Summer School and before the School ends. Systems will go away in a couple of weeks.
- Be careful when deleting files on `/efs_ecco`; the disk has shared write permissions.
- Be mindful of the disk space in your OSS directory: `/home/jovyan`. Filling up the 150 GB quota may crash the system.
- Large output files should be written to `/efs_ecco`, not your OSS or P-Cluster home directories
- Stop server on OSS and log out of P-Cluster when finished

Data Access

(/efs_ecco/ECCO/)

/efs_ecco/ECCO/V4:

- **V4r4** estimate: /efs_ecco/ECCO/V4/r4/netcdf
 - native: on native llc grid
 - latlon: on regular lat-lon grid (0.5x0.5 degree)
- **V4r5** estimate: /efs_ecco/ECCO/V4/r5/netcdf
 - native: on native llc grid
 - latlon: on regular lat-lon grid (0.5x0.5 degree)
- **V4r5 biogeochemistry** estimate:
/efs_ecco/ECCO/V4/r5_bgc/netcdf
 - native: on native llc grid
 - latlon: on regular lat-lon grid (0.5x0.5 degree)