

subMIT Status and Plans

Josh Bendavid
subMIT Users Group Meeting
July 25, 2023

Introduction

- subMIT system provides an interactive login pool + scale-out to batch resources
 - Home directories
 - Convenient software environment (CentOS7 native, docker/singularity images, conda)
 - SSH or Jupyterhub access
 - Local batch system with $O(1000)$ cores, >50 GPU's
 - Additional storage for software installation/development, large datasets
 - Convenient access to larger external resources (OSG, CMS Tier-2 and Tier-3, LQCD Cluster, EAPS)
- User support is a key feature of the system
 - Beyond basic troubleshooting
 - Help users make optimal use of the available resources
 - Expert advice on designing/improving workflows
 - Customize and evolve system configuration to accommodate user needs as appropriate

Introduction

- Storage and networking
 - Local storage (1TB/user), 10's of TB for larger group datasets
 - 30TB of ultra-fast NVME storage with room for future expansion
 - Fast networking: 100 Gbps ethernet
 - RoCE (RDMA over Converged Ethernet) has been partially tested/commissioned, should be possible for MPI applications
 - Full IPv6 support (dual stack with IPv4)
- Additional resources recently or currently being integrated
 - More disk storage (100TB contributed from ABRACADABRA)
 - Integration of existing computing resources from research groups
 - Purchase of several large core count/high memory machines by research groups for additional computing resources and to support specialized workflows and/or R&D where large single node scaling is useful
 - Current “high density” template, Dual AMD EPYC 192 core/384 thread with 0.75-1.5TB of memory
 - Physical location of servers in the process of being consolidated in B24

Introduction: subMIT Website



subMIT

Getting physics things done at MIT

[Overview](#) [News](#) [People](#) [Contact](#) [About](#) [Users Guide](#) [A2rchi](#) [JupyterHub](#)

Overview

The subMIT login pool is designed to let users login safely, prepare and test their research, and submit their jobs to the large computing resource of their choice. There are for now a limited number of resources connected but we are working on quickly expanding them.

[Get your account on SubMIT Portal](#)

Access

ssh <user>@submit.mit.edu

jupyterhub

Status

Servers

Slurm queue

Condor queue

Expert

Resources

- >1 TB of free storage per user
- 1000s of cores and >50 GPUs available interactively and through Slurm
- Access to OSG, CMS T3 and T2, LQCD Cluster, and EAPS

Software

- Python, anaconda, Julia, Matlab, singularity, and much more!

- Website (with User's Guide/Instructions):

<https://submit.mit.edu/>

- Overview and general information
- Direct JupyterHub access
- User's Guide:

<https://submit.mit.edu/submit-users-guide/>

User's Guide - subMIT

Contents:

- User's guide - subMIT login pool
- Getting started
- Things that work and things that do not
- Available software
- Running interactively and batch jobs
- User quota and storage at submit
- Monitoring at submit
- GPU resources
- Data backup

Tutorials - subMIT

Tutorials:

- Tutorial 1: Native System (python, Julia, matlab)
- Tutorial 2: Batch Job (HTCondor and Slurm)
- Tutorial 3: Containers (Docker and Singularity)
- Tutorial 4: Package Manager (Conda and Jupyterhub)
- Tutorial 5: GPU Example (submit-gpu and GPU batch options)

Future Work - subMIT

Planned Upgrades:

- Move to AlmaLinux

Indices and tables

- Index
- Module Index
- Search Page

Users Group In Practice

- Monthly meetings
 - Advertised and open to the broader community
 - Topical presentations from project team, Users Group representatives, or other users or community members
 - Forum for feedback and information flow between the user community and the project team
 - Regular timeslot: Tuesday 10:00-11:00 EST
- Users Group representatives
 - Identified representatives from research groups across the department
 - Attend the monthly meetings
 - Provide feedback from your groups/community
 - Distribute information/news from the project team

Users Group Representatives

- Users group has been formed (JB as coordinator)
- Current Users Group representative (associated faculty/group)
 - Yin Lin (Phiala Shanahan)
 - Siddharth Mishra-Sharma (Jesse Thaler)
 - Prajwal Mohan Murthy (Bob Redwine)
 - Kaliroë Pappas (LNS Neutrino/Dark Matter)
 - Sunghan Ro (Julien Tailleur)
 - Yitian Sun (Tracy Slatyer)
 - Molly Taylor (LNS Heavy Ion Group)

Storage breakdown

- Several different storage areas are available covering different use cases
 - /home/submit/<username>
 - Home directories (nfs server), redundant disk array with backups
 - 5GB quota
 - Use for software development and (small) critical data
 - /work/submit/<username>
 - Work directory (nfs server), no backups (but redundant disk array)
 - 50GB quota
 - Use for software installation (conda or docker/singularity images)
 - /data/submit/<username>
 - Large distributed disk system, no backups, but redundancy against disk failure (“erasure coding”)
 - 1TB user quota, larger quotas available in dedicated group directories
 - Store large datasets here
 - /scratch/submit/<username>
 - Fast NVMe SSD array
 - Commissioned by several groups for high performance data analysis
 - /cvmfs/
 - Read-only distributed storage for distributing software, singularity images, etc
 - Several CERN-related repositories are available
 - Local repository /cvmfs/cvmfs.cmsaf.mit.edu where additional software or data can be added if needed
- Flexible tiered storage system, can accommodate a wide range of user needs
- Larger datasets encouraged to use shared group space, but quotas can be increased when needed

Interactive Use: Terminal or JupyterHub



jupyterhub

Select a job profile:

Slurm - Submit - 1 CPU, 500 MB

Start

Quick introduction:

• Spawn server menu:

- Slurm - Submit - 1 CPU, 500 MB: spawns a server on submit slurm partition.
- Slurm - Submit - 2 CPUs, 1000MB: similar as above, with more resources allocated.
- Slurm - SubmitGPU - 1 GPU: spawns a server on submit-gpu slurm partition, requesting 1 GPU.
- Slurm - SubmitGPU1080 - 1 GPU: spawns a server on submit-gpu1080 slurm partition, requesting 1 GPU.
- Local server - Submit01 - 1 CPU, 500 MB, /home/submit({username}): spawns on submit01, in your /home/submit({username})/ directory.
- Local server - Submit01 - 1 CPU, 500 MB, /work/submit({username}): spawns on submit01, in your /work/submit({username})/ directory.
- GPUs: you can use GPU resources in your notebooks or Jupyterhub's terminal if you spawn a server on submit-gpu or submit-gpu1080, supported through Slurm.
- Conda: your conda environments should be automatically loaded as kernels by Jupyterhub, and can be used in notebooks. See User Guide for more info.
- Singularity: you can manually set up a kernel based on a singularity environment's python. See User Guide for more info.

For more information about Submit, conda, GPUs, Jupyterhub, etc., see:

User Guide

For any questions, comments, or feedback, please send an email to submit-jupyter.

Name	Last Modified
condasetup	8 months ago
distrdtest	a year ago
hpcutlis	a month ago
iperf-2.0.9	4 months ago
iperf-3.1.3	4 months ago
miniforge3	8 months ago
nuttcp-8.2.2	4 months ago
rdftest	8 months ago
singularityconfigs	8 months ago
wmassdev31a	a month ago
bashrcondabak	8 months ago
iperf-2.0.9-source.tar.gz	5 years ago
iperf-3.1.3-source.tar.gz	5 years ago
nuttcp-8.2.2.tar.bz2	3 years ago
test.txt	10 months ago
test2.txt	2 months ago
test3.txt	2 months ago
test4.txt	2 months ago
test5.txt	2 months ago

- Interactive Jupyter session available directly from website with touchstone authentication (subMIT account still required)
- SLURM is used to efficiently share resources between interactive and batch use

Communication Channels

- User support mailing list: submit-help@mit.edu
- Experimental large language model application under development for interactive user support and to augment support ticket handling
- Slack workspace: <https://mit-submit.slack.com>
 - “help-desk” channel
- Monthly Users Group Meetings
 - Open for discussion
 - Open for user contributions: full set of Users Group representatives can be contacted at submit-usersgroup@mit.edu
- Annual subMIT workshop
 - January 2023 workshop: <https://indico.mit.edu/event/647/>
- In addition to direct interaction with the subMIT project team, users are encouraged to discuss with Users Group representative from their own group or “nearby” group

A2rchi: Ai Augmented Research Chat Intelligence

- Newest project-team member
- ChatGPT-based with input from subMIT Users Guide and eventually FAQ/historical support tickets
- Will assist with User Support
 - Interactive chat on subMIT website:
https://submit.mit.edu/?page_id=288
 - Integration with submit-help@mit.edu and ticketing system to improve email support/ticket response
- Presentation from J. Heitkoetter at Physics Dept. LLM Workshop Last week
 - <https://indico.mit.edu/event/759/>



A2rchi: Interactive Support Chat



subMIT

Getting physics things done at MIT

Overview

News

People

Contact

About

Users Guide

A2rchi

JupyterHub

Chatbot

I want to submit 100 GPU jobs, all running the exact same command "python run.py." How do I do this?

You can use the Slurm job scheduler to submit multiple jobs with the same command. You can create a job script that contains the command you want to run and then submit it to the submit-gpu partition with the command `sbatch job_script.sh`. You can also use the `--array` option to submit multiple jobs with the same command. For example, if you wanted to submit 100 jobs, you could use the command `sbatch --array=1-100 job_script.sh`. This will submit 100 jobs with the same command. You can find more information about submitting jobs with Slurm here: <https://slurm.schedmd.com/sbatch.html>.

[Click here to read more](#)

Chatbot

Could you give an example?

Sure! Here is an example job_script.sh file:

```
#!/bin/bash

#

#SBATCH --job-name=test

#SBATCH --output=res_%j.txt

#SBATCH --error=err_%j.txt

#
```

Linux Distribution Upgrade

- Current CentOS 7 distribution reaches EOL for maintenance updates in June 2024
- Decision by Red Hat to reorganize CentOS project and releases has disrupted logical upgrade path from CentOS 7->8
- Decision taken to upgrade from CentOS 7 to Alma Linux 9, considering:
 - Ease of transition
 - Support lifetime
 - Functionality
 - Direction being taken at other universities and labs (CERN, Fermilab, etc)
- Discussion has included Users Group and broader community
- Performance-sensitive services already upgraded (NVME storage)
- New large 384-thread machines using Alma 9 from the start
- Recent change (June 21, 2023) by Red Hat to CentOS source code policy introduces further uncertainty to enterprise linux ecosystem and Alma/Rocky Linux Projects
 - Carefully monitoring developments, but continuing with Alma 9 migration
 - Current plan from Alma Linux project is to maintain ABI compatibility with RHEL but not 1:1 source code or bug compatibility

Linux Upgrade: Testing and User Feedback/Support

- CentOS 7 environment will continue to be available through singularity containers
- Test environment available with a few machines to start:
 - **ssh access:** submit-test.mit.edu
 - **SLURM partition:** submit-alma9
 - Jupyter test instance possible in the near future
- Please test the new environment and provide feedback
- Full transition foreseen by the fall

The screenshot shows a web page titled "Move to AlmaLinux" with a "SUBMIT" logo in the top left. The page content includes:

- Table of Contents:** Move to AlmaLinux, How we plan to upgrade submit, Testing on Alma Linux, Will CentOS still be available, Feedback on AlmaLinux.
- Previous topic:** Tutorial 5: GPU Example (submit-gpu and GPU batch options).
- This Page:** Show Source.
- Quick search:** A search bar with a "Go" button.
- Move to AlmaLinux:** In December 2020, Red Hat announced that development of CentOS, a free-of-cost downstream fork of the commercial Red Hat Enterprise Linux (RHEL), would be discontinued. As such, we will move to AlmaLinux, a community-supported, production-grade enterprise operating system that is binary-compatible with RHEL. This page covers the planned upgrade and testing of AlmaLinux on submit.
- How we plan to upgrade submit:** Because it is hard to predict how an operating system upgrade will affect everyone's work, we will upgrade the system in stages and encourage users to test their workflows on AlmaLinux. For now, the submit system uses CentOS 7 with AlmaLinux machines available only through slurm using a test partition. For the future upgrades, the submit machines will get upgraded with announcements in the submit-users email group.
- Testing on Alma Linux:** In order to start testing with AlmaLinux, we have set up a testing slurm partition which has machines with AlmaLinux. This partition is named "submit-alma9". The easiest way to start working is to start an interactive session through slurm with `salloc`.

```
salloc --partition=submit-alma9
```

You can also run a testing slurm job with the following submit script.

```
#!/bin/bash
#
#SBATCH --job-name=test_alma9
#SBATCH --output=res.txt
#
#SBATCH --time=10:00
#SBATCH --mem=1024-cpu=100
#SBATCH --partition=submit-alma9

<your script goes here>
```
- Will CentOS still be available:** If your workflow is constrained to CentOS, submit will provide a centrally available singularity of CentOS7.9. You can test this singularity below.

```
singularity shell /cvmfs/cvmfs.cmsaf.mit.edu/submit/work/submit-software/centos/centos7p9
```

If you need something specific besides this singularity, please email us at submit-help@mit.edu.
- Feedback on AlmaLinux:** If you run into any issues while running on the AlmaLinux testing partition, please report them to the submit team at submit-help@mit.edu. Additionally, please report any differences that you see between the base CentOS system and the AlmaLinux testing partitions.

© Copyright 2021, The SUBMIT Project Team. Created using Sphinx 4.5.0.

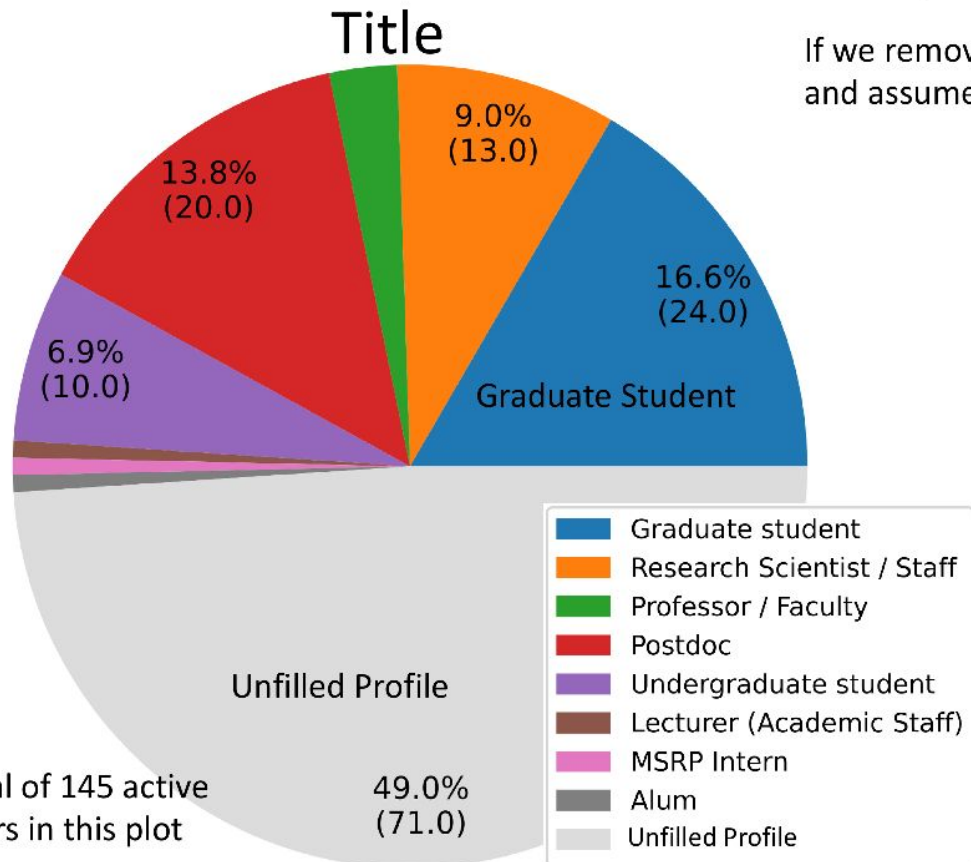
Dedicated documentation on migration:

<https://submit.mit.edu/submit-users-guide/future/alma.html>

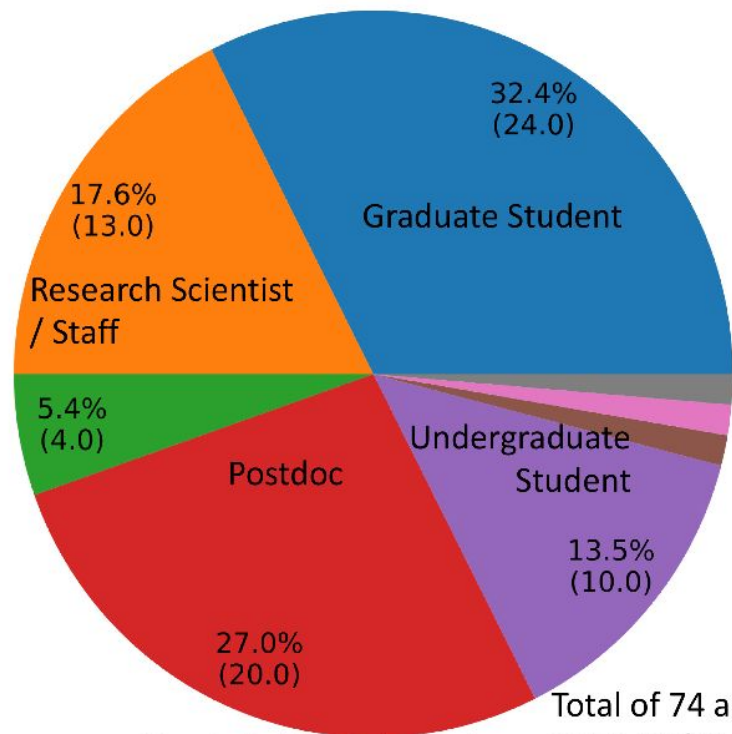
Active Users: Who are they?

"Cooooo!"

- Lecturer (Introductory Physics)



If we remove the users who have not answered ('Unfilled Profile') and assume the remainder is representative of the whole ...



- Nearly half are students
- Reasonable breakdown

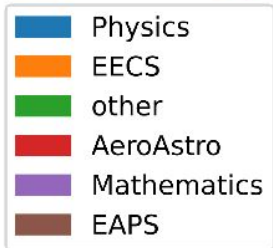
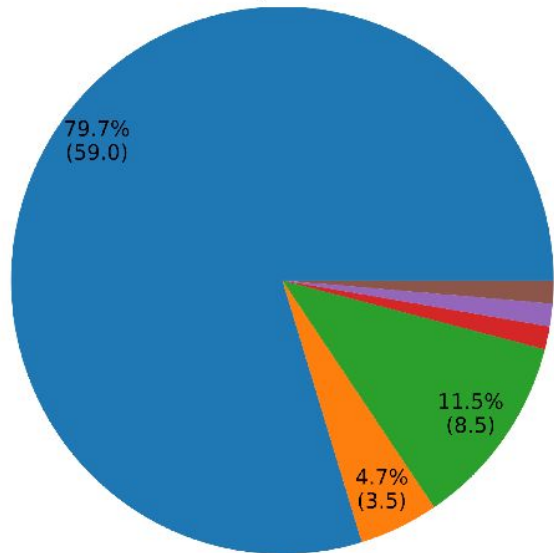
M. Heine

Note: Legend (top to bottom) is in order of wedges (counterclockwise)

Active Users: Who are they?



Department

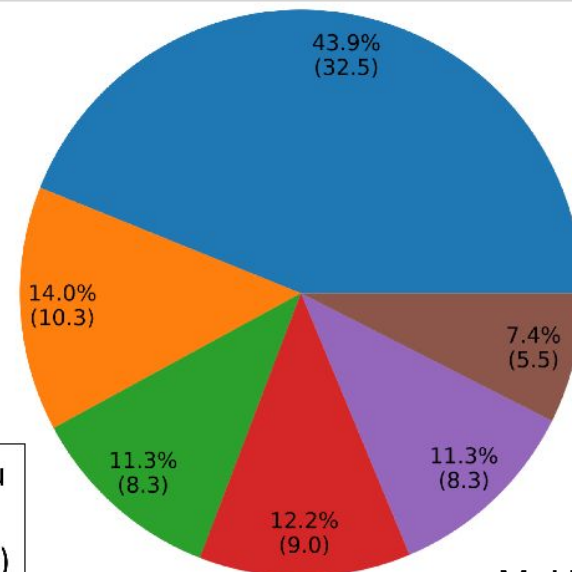
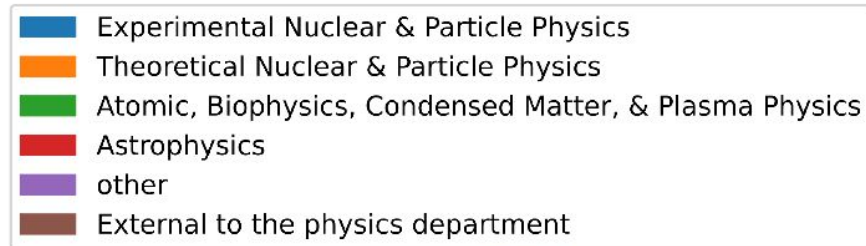


- Mostly Physics
- Experimental Nuclear & Particle Physics most represented (but not majority)
- Roughly even Division representation otherwise
- Active work to diversify (see later talk)

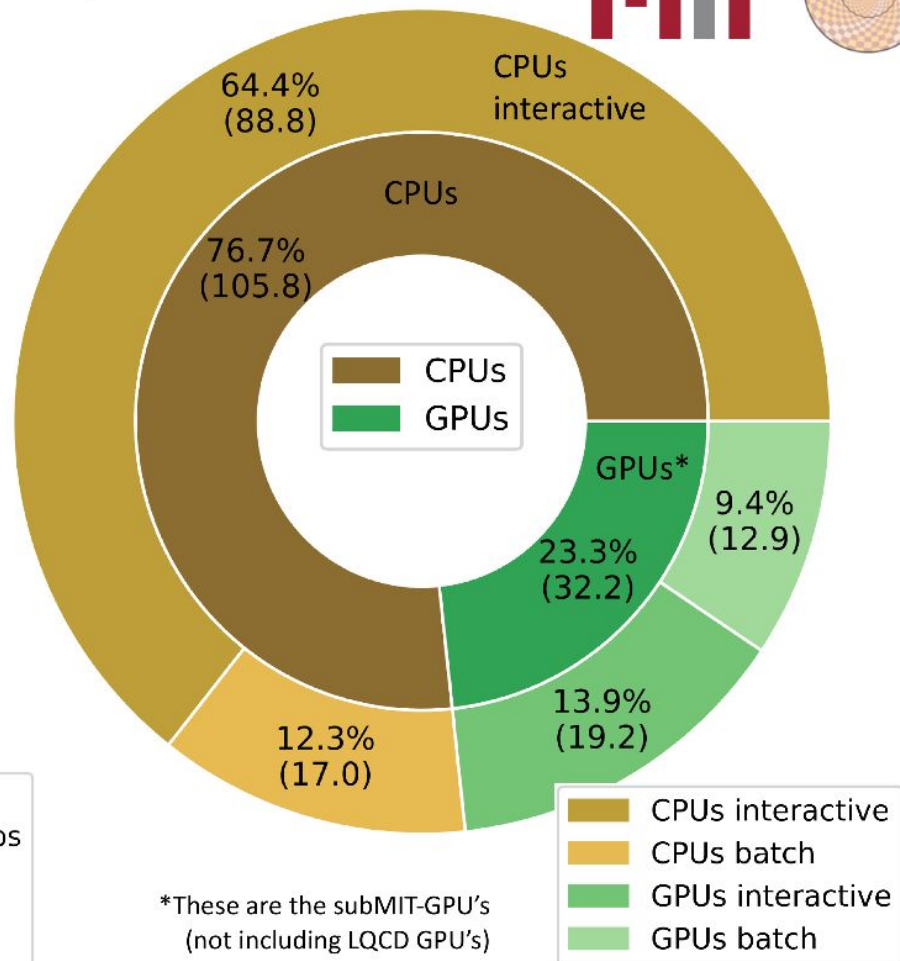
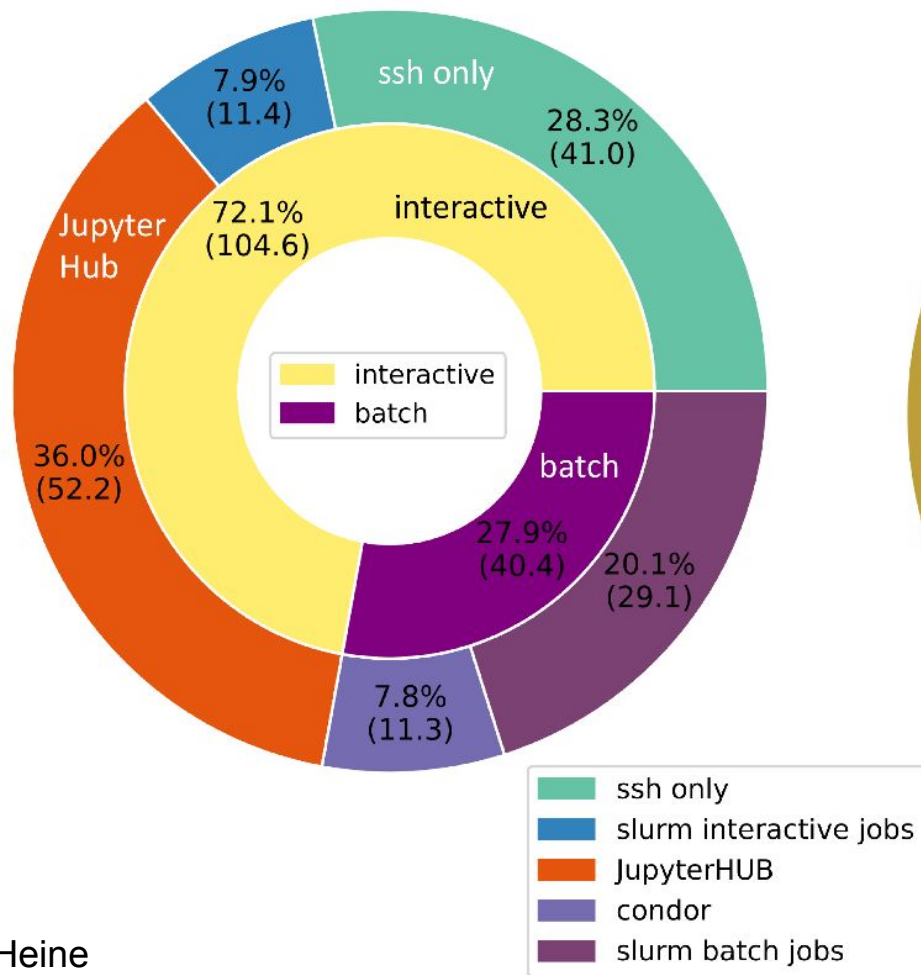
“Really appreciate the services you provide and hard work you put into this!”

- Research Scientist / Staff ; Particle Physics Experiment (LNS)

Physics Division



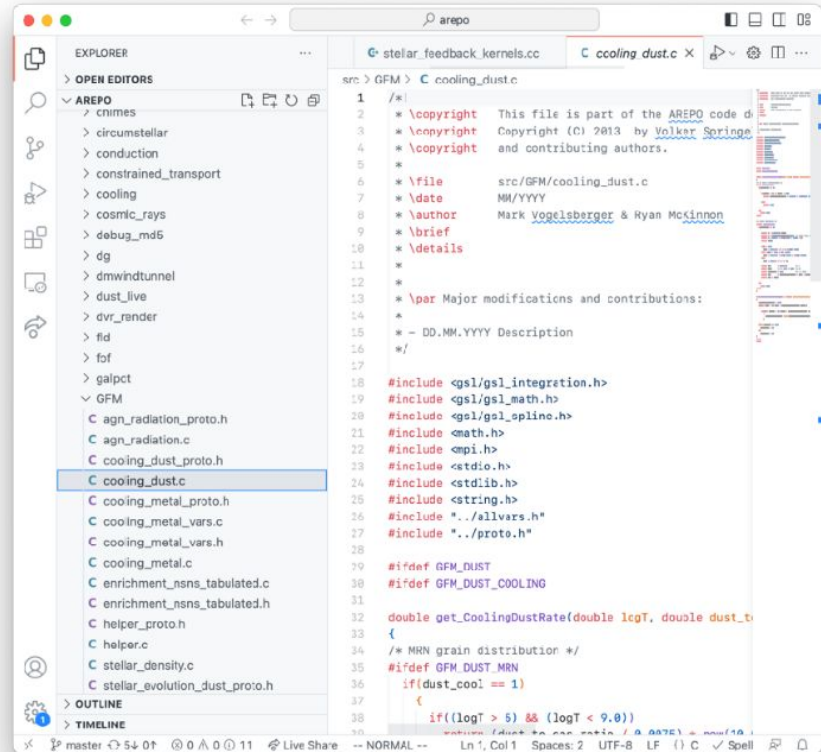
Active Users: How do they use the system?



*These are the subMIT-GPU's (not including LQCD GPU's)

Upcoming Meeting: User Training

- Based upon **user feedback**, we are developing a **user training session** for the use of **Visual Studio Code (VSC)**.
- The focus of the session will be both using **VSC locally** and **remotely through subMIT** and other MIT machines.
- **Training planned** for in **Sept/Oct** once new graduate students join.



```
1 /*
2  * \copyright This file is part of the AREPO code d
3  * \copyright Copyright (C) 2013 by Volker Springe
4  * \copyright and contributing authors.
5  *
6  * \file      src/GFM/cooling_dust.c
7  * \date      MM/YYYY
8  * \author    Mark Vogelsberger & Ryan McInnon
9  * \brief
10 * \details
11 *
12 *
13 * \par Major modifications and contributions:
14 *
15 * - DD.MM.YYYY Description
16 */
17
18 #include <gsl/gsl_integration.h>
19 #include <gsl/gsl_math.h>
20 #include <gsl/gsl_spline.h>
21 #include <math.h>
22 #include <mpi.h>
23 #include <stdio.h>
24 #include <stdlib.h>
25 #include <string.h>
26 #include "../allvars.h"
27 #include "../proto.h"
28
29 #ifdef GFM_DUST
30 #ifdef GFM_DUST_COOLING
31
32 double get_CoolingDustRate(double lcgT, double dust_t)
33 {
34     /* MRN grain distribution */
35     #ifdef GFM_DUST_MRN
36         if(dust_cool == 1)
37         {
38             if((logT > 5) && (logT < 9.0))
39                 return (dust_to_gas_ratio / A_DUST) * pow(10,
```