



User Support and Slurm User Accomodation for SubMIT

Chad Freer on behalf of the SubMIT team



SubMIT



An MIT Physics Department analysis facility.

→ provide ecosystems to many research areas

<https://submit.mit.edu>

subMIT system provides an interactive login pool + scale-out to batch resources

- Home and Work directories
- SSH or Jupyterhub access
- Convenient software environment (CentOS7 native (moving to Alma Linux 9), docker/singularity images, conda)
- Local batch system with O(1000) cores, >50 GPU's 8 A30s and ~45 1080s (more being integrated)
- Local storage (1TB/user), 10's of TB for larger group datasets (gluster distributed file system with ~400TB of spinning disks)
- Fast networking: 100 Gbps ethernet
- Convenient access to larger external resources (OSG, CMS Tier-2 and Tier-3, LQCD Cluster, EAPS)
- Strong focus on user support (ticketing system, AI-based chatbot in development)

SubMIT Account Setup and Data Storage



Creating Accounts:

1. Anyone with MIT ID can set up an account by simply uploading ssh keys
 - a. MIT accounts possible for external people collaborating with MIT groups
2. Account gets automatically created with all of the user spaces and permissions.
3. Immediate access to all SubMIT machines as well as batch systems

User space:

1. 5 GB for User's Home → /home/submit/<user>
2. 10 GB for User's Work → /work/submit/<user>
3. 1 TB for User's Data → /data/submit/<user>

NVME and additional space:

1. 30 TB of NVME scratch space for use by groups.
2. Additional Group space in /data as needed

T2 and Xrootd:

1. MIT T2 is mounted
2. XRootd access to typical CMS resources as well as group space on /data and T2
3. Boston XCache is available

A Word More on Data Storage



Additional Storage:

1. If a user needs additional space, the basic quotas can be expanded to match needs
2. Hadoop storage through the MIT T2 can also be provided with xrootd access

Group Storage:

1. Provide storage with common access for all users within groups
2. Can be provided through /data with amount dependent on group's need
3. If fast access is required, access to scratch storage(NVME) can also be provided

Xrootd:

1. /data has xrootd access
2. Allows access to data in Condor jobs where the mount is not available

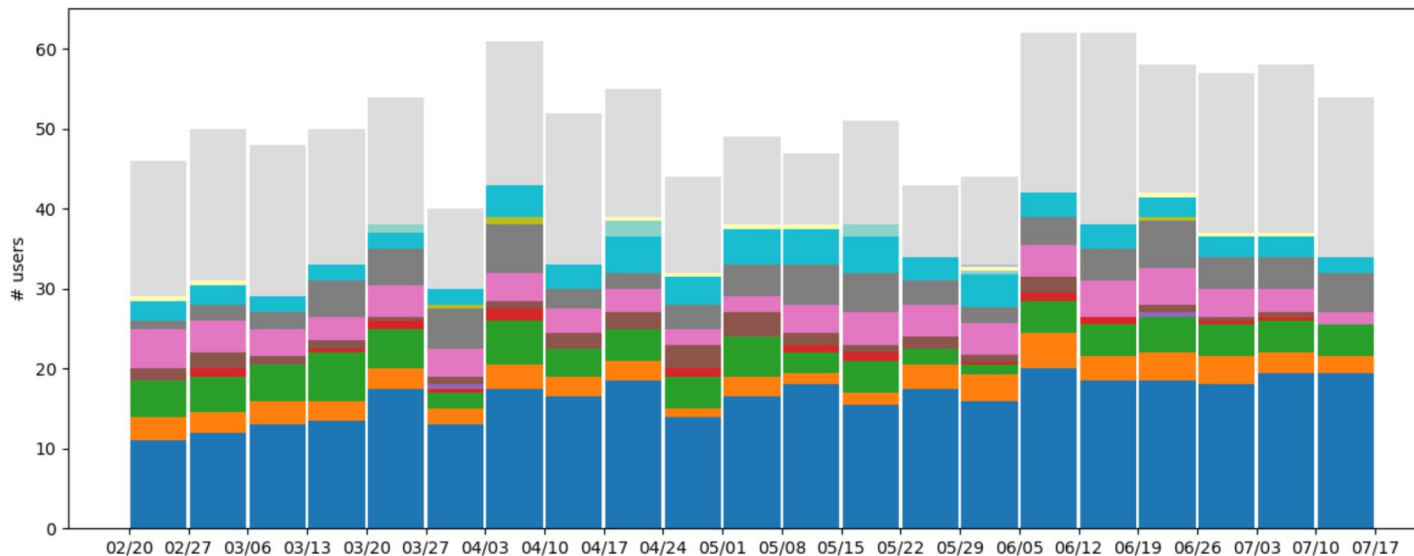
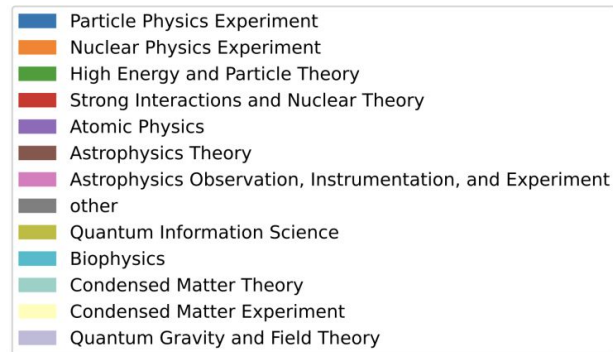
SubMIT User Usage Metrics



SubMIT now has over 300 users!

<https://submit.mit.edu/?p=305>

50-60 Unique users log onto SubMIT in a given week (Mostly from Physics)



SubMIT User Usage Metrics Continued

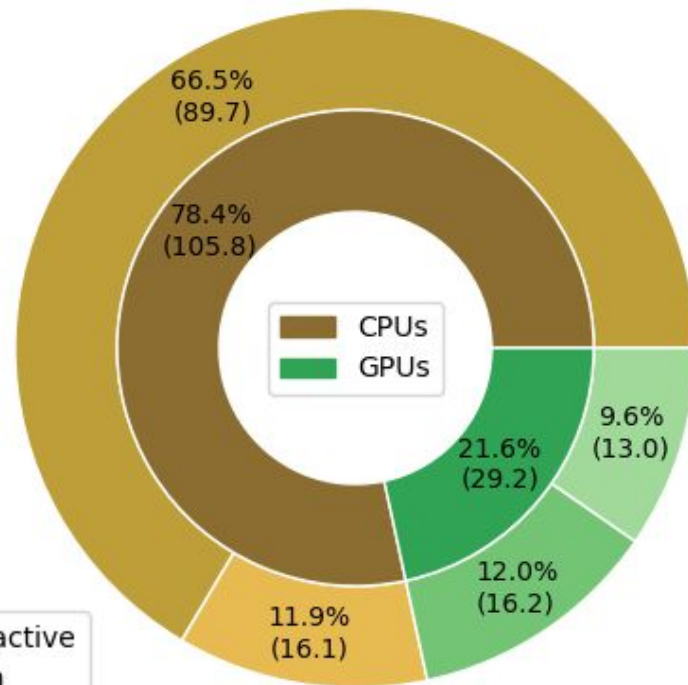
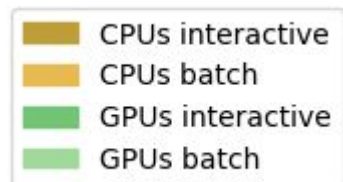


Usage for SubMIT resources:

1. Users' tend to use CPU resources but there is still a healthy number of users interested in GPUs
2. Most users interested in CPUs are just doing interactive session work
3. Significant interest in using GPUs through batch computing systems!

Takeaways:

1. Users are taking advantage of multiple methods for working!
2. Significant work being done on GPUs
3. Need to maintain and support all options.



Batch Computing Systems



Batch Computing:

1. Computerized batch processing is a method of running software programs in batches automatically.
2. Users are required to submit the jobs, but no other interaction is required

Slurm vs HTCondor:

1. High Throughput or HTCondor:

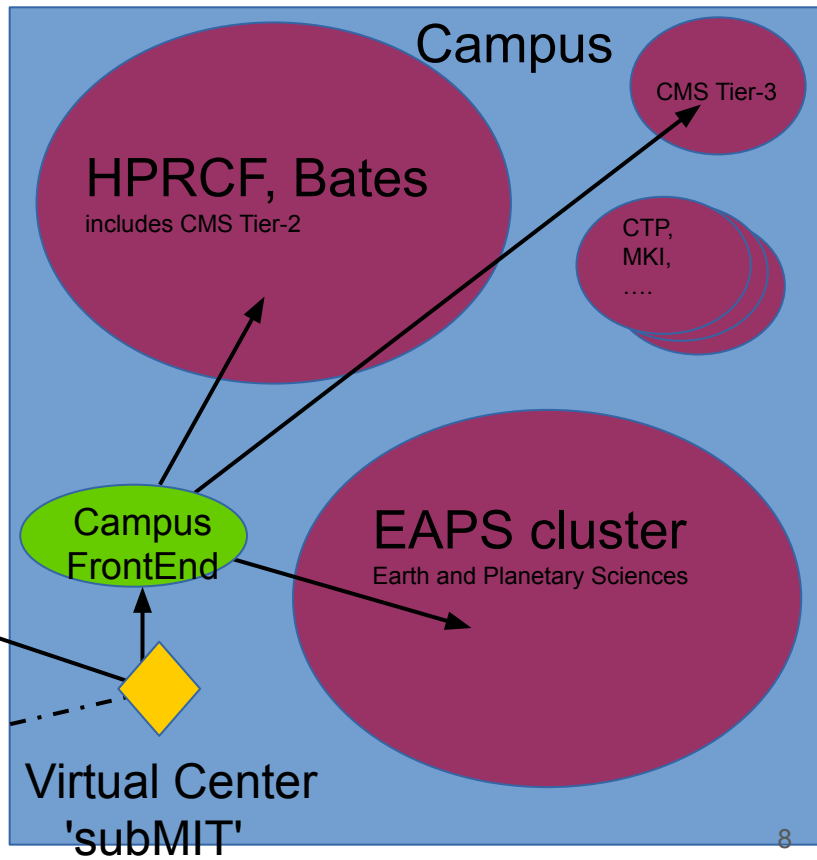
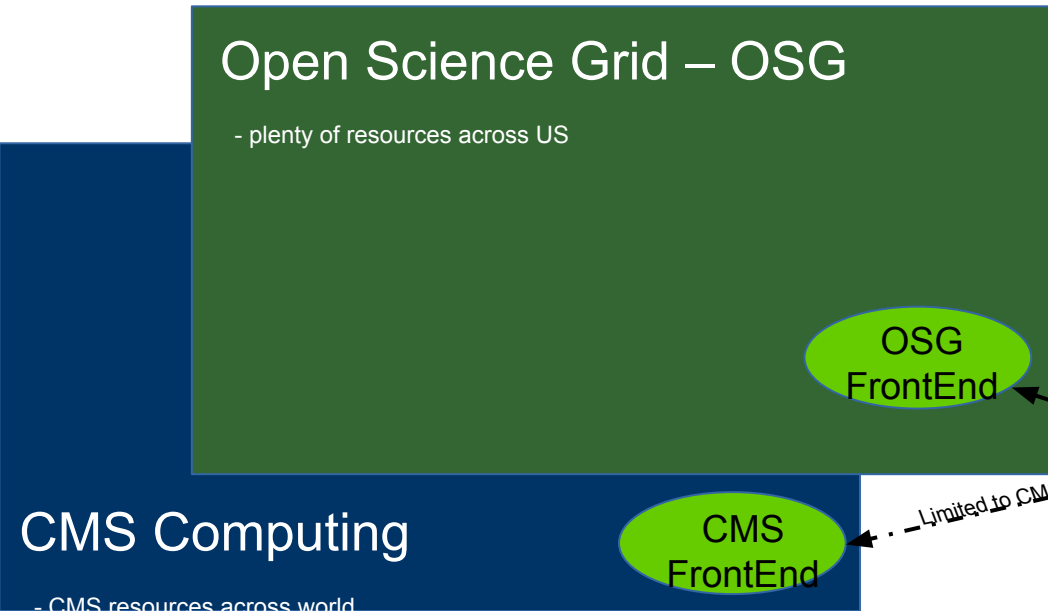
- a. high-throughput computing software framework for coarse-grained distributed parallelization of computationally intensive tasks.
- b. It can be used to manage workload on a dedicated cluster of computers (No Common Mount points/filesystems)

2. Simple Linux Utility for Resource Management or Slurm:

- a. allocating access to resources (computer nodes) to users for some duration of time so they can perform work
- b. framework for parallel jobs such as Message Passing Interface (MPI) on a set of allocated nodes
- c. Home and other mount point available on all nodes

HTCondor on SubMIT

Connected to all resources on campus



Batch computing and GPU access



HTCondor:

1. Connections to MIT T2 (~25k cores) and T3 (~700 cores)
2. Additional connection to MIT Earth and Planetary Sciences (EAPS) cluster
3. External to MIT, we have connection to the CMS Global Pool and OSG

Slurm:

1. O(2000) cores on SubMIT available for slurm jobs (more available through connection to CTP)
2. Dask setup through slurm cluster
 - a. Can be launched through terminal or jupyter
 - b. Conda environments can be used with Dask
3. O(50) GPUs made available through slurm
 - a. Can be used for batch jobs or for interactively through salloc
4. Jupyterhub uses slurmspawner so GPUs are available for Jupyterhub users
5. Includes a test partition for AlmaLinux9



Slurm Partitions on SubMIT

1. Main partition (“submit”):
 - a. Main partition for SubMIT
 - b. O(2000) CPU cores available to users spread across 19 machines

2. GPU partition (“submit-gpu”):
 - a. Main submit GPU partition
 - b. 4 servers with 48 CPUs and 2 A30 GPUs each

3. Additional GPU resources (“submit-gpu1080”):
 - a. Secondary GPU partition
 - b. 12 servers with 24 CPUs and 4 1080 GPUs each

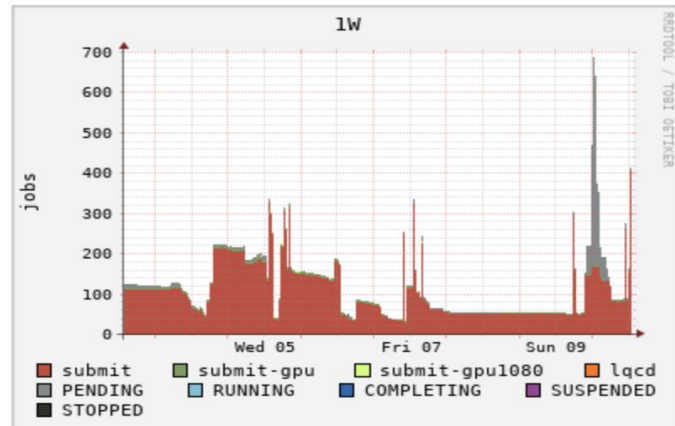
4. AlmaLinux test partition (“submit-alma9”):
 - a. Test partition for AlmaLinux
 - b. Currently only 2 new servers but will be expanded as we upgrade servers
 - i. Will continue to provide support for CentOS through singularity

Monitoring on SubMIT



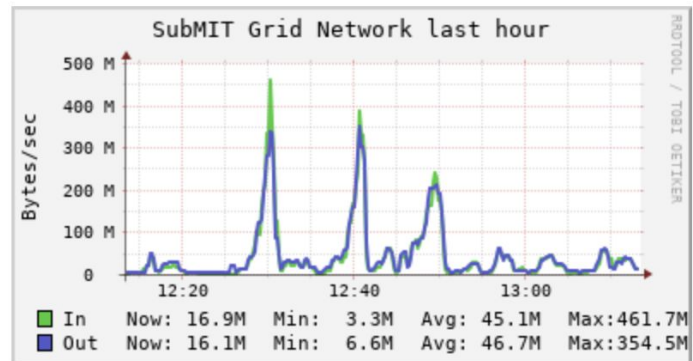
1. CondorMon and SlurmMon:

- Simple website showing a summary of the slurm jobs
- Shows total and individual user jobs
- rrd tools showing plot summaries for 2 hours, 1 day and 1 week
- Shows the usage split by computing center or slurm partition used



2. Ganglia:

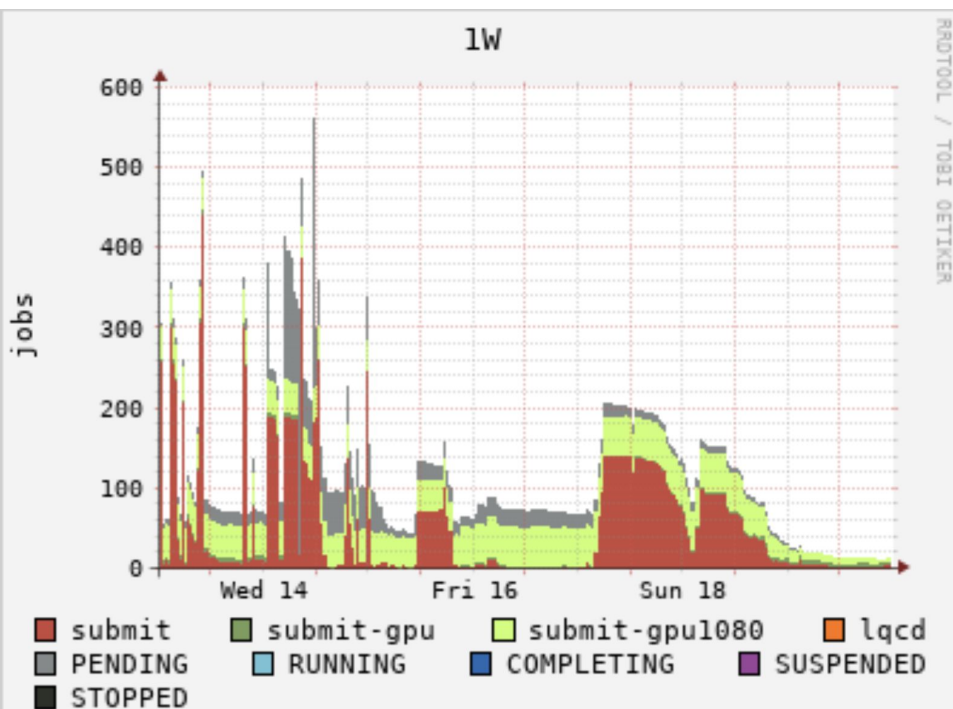
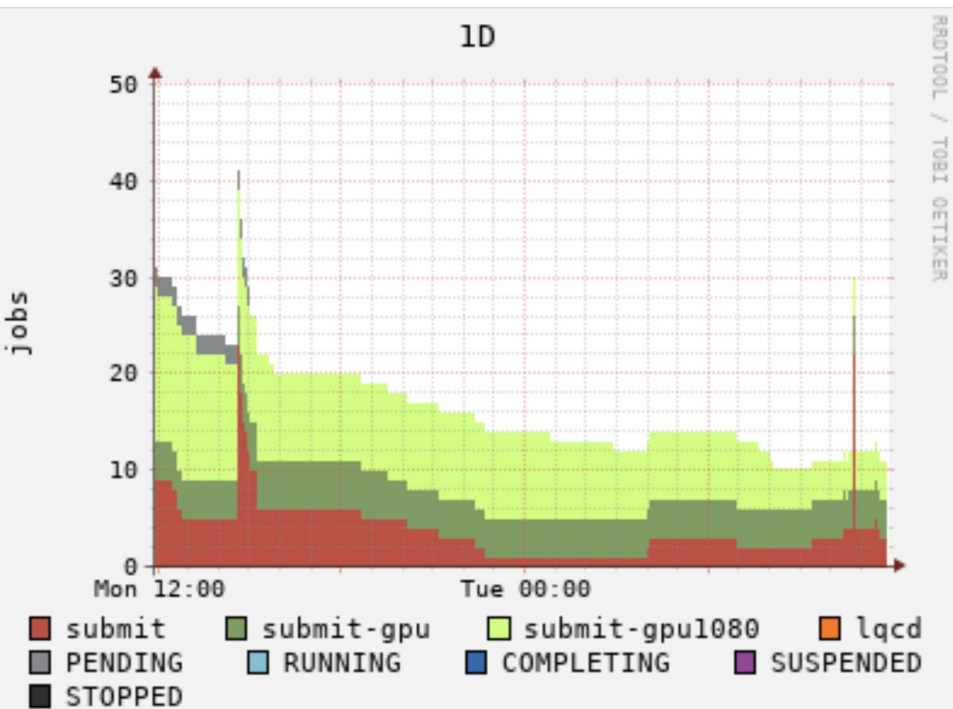
- Monitor for all server machines in the SubMIT system
- Can Monitor CPU usage, Network, Memory, etc
- Useful for checking load on individual servers



More on GPUs

GPU availability:

- Available only through slurm to avoid conflicts
- Can also be accessed for interactive use through salloc or through jupyterhub



Software Availability



Conda:

1. User encouraged to use conda
2. Conda environments will be available on all slurm nodes

CVMFS:

1. CVMFS is available
2. Additional MIT hosted CVMFS instance (available on SubMIT, T2 and T3)
3. Any package a user needs can be made available through this CVMFS instance
4. Singularity images hosted on cvmfs can be run directly
 - a. `cmssw-slc6/cmssw-cc7/cmssw-el8/cmssw-el9` for CMS-maintained images suitable for CMSSW development
 - b. `/cvmfs/unpacked.cern.ch` is available for a range of other images
 - c. Additional images can be stored on MIT-hosted CVMFS instance

Docker:

1. Users have direct access to docker

Jupyterhub and Environments



Jupyterhub Options and GPUs:

1. Jupyterhub uses Slurmspawnner
2. Access to GPUs
3. Can use slurm reservations for large groups



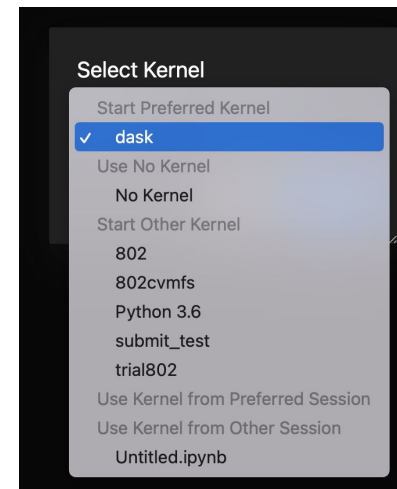
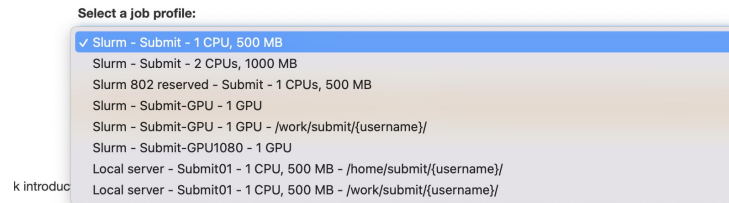
jupyterhub

Connection with User's Conda environments:

1. Jupyterhub loads the user's conda environments!
2. Central conda environments are also available
3. Provides a lot of flexibility for workflows

Dask:

1. Dask can be run through jupyterhub
2. Can access conda environment



SubMIT User Support



User's Guide:

1. Introduction to the SubMIT system and how to get your account set up
2. Guides for setting up environments and workflows
3. Several tutorials with examples on how to use batch computing, GPUs, Docker.

submit-help@mit.edu:

1. Email help-desk that connects you to the SubMIT team directly
2. Quick response time (Usually well within a few hours)
3. Willing to accommodate specific workflows/requests

Monitoring:

1. Monitoring tools available for all users
2. Monitors for the batch computing tools
3. Ganglia for server/network usage

SubMIT User's Guide



Contents of the User's Guide:

1. Main section of the User's Guide designed to help users with getting started.
2. Describes resources (storage, software, batch systems, etc) available
3. Simple commands/scripts to help start work

Tutorials:

1. A few examples of workflows
2. We are willing to support more tutorials if requested!
3. More examples on github:
<https://github.com/mit-submit/submit-examples>

Additional Sections :

1. Future work: AlmaLinux testing partition
2. We are moving to AlmaLinux and have provided a slurm partition for users to test with

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

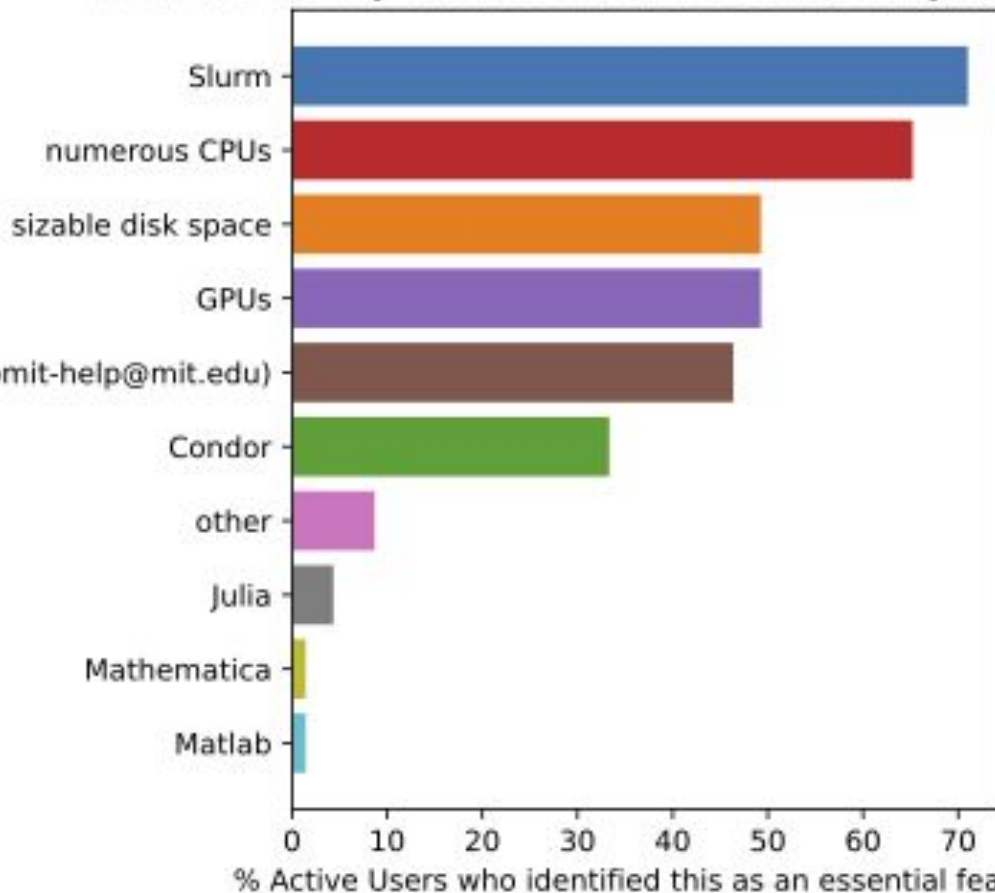
Tutorials:

- 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)

SubMIT User Metrics Continued



What are the key features that are essential for you on subMIT?



Quick Takeaways:

Widespread needs

No one answer meets all

Slurm is essential for a majority

User Support is essential!

Need a consistent team to cover questions

User support takes a **great deal of time!**

Last Month:

~20 separate threads answered by 6 SubMIT team members

SubMIT Support Examples



Request to add sage math to SubMIT (Wati Taylor) :

1. Requested an installation of sage math to be made available to use on a batch system:
2. <https://hub.docker.com/r/sagemath/sagemath>
3. Installed through docker and consistently had several hundred jobs running on slurm for weeks

Requested an installation of GAP (Xiao-Gang) :

1. Requested an installation of GAP:
2. <https://hub.docker.com/r/gapsystem/gap-docker/>
3. Installed through docker and provided centrally through cvmfs.

Requested continuous access to a large number of GPUs (Simon Rothman) :

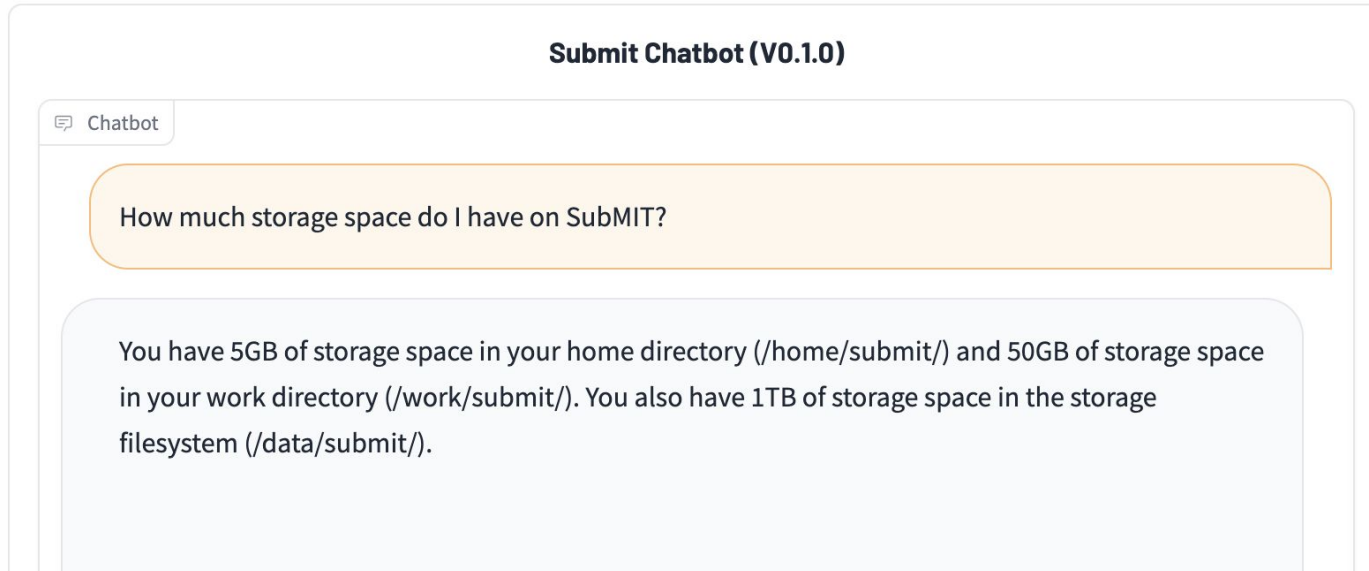
1. Requested access to GPUs
2. Ran on ~50 GPUs through slurm for weeks
3. Simon is one of our consistent users and has given great feedback on the system!

SubMIT Chat Bot



SubMIT team is currently working on a chat bot:

1. Supplement the SubMIT help team (User support is time consuming)
2. Trained on the User's Guide and SubMIT documentation
3. Currently in the development stages! Will announce a prototype in the coming months



Summary

SubMIT covers the entire MIT Physics Department and some others

1. Over 300 users with ~60 active users per week
2. Users have a wide variety of backgrounds and needs
3. Multiple active analyses that are closing in on publication as well as work for conferences

Access to resources through batch computing

1. Access to both Slurm and HTCondor (Slurm being the more popular option)
2. Access to various software through conda, CVMFS, docker, etc
3. Jupyterhub with access to your own conda environments

User feedback has been extremely helpful

1. help-desk is both essential and time consuming!
2. We have formed a users-group at SubMIT with monthly meetings and have plans for chatbot
3. No single answer for everyone

SubMIT Support 802



Conda environment made centrally:

1. Specific environment requested for 802 class
2. Environment made available to users through jupyterhub (Students do not ever need ssh!)
3. Willing to reserve machines through slurm to support a large number of users
4. Can upload notebooks directly from github provided by 802 lecturers