



User Workflows on SubMIT

Jan Eysermans on behalf of the SubMIT team
Annual Review
June 3 2026

Overview



- **Account creation and login**
- **Accessing SubMIT resources**
- **Storage resources**
- **Computing resources**
- **Software resources**
- **Other tools**

Creation of new accounts



Streamlined account creation based on a request by the new user



Typical information we ask to include in the request

- Short motivation why and project context
- Role (student, PhD, postdoc, researcher, professor)

After approval, the welcome email directs the user to setup the necessary login via TouchStone

- Gives access to setup ssh keys to login into SubMIT system

For non-MIT users, a MIT guest account has to be created beforehand

- Full name, email address, date of birth, institution



How to access SubMIT?

Several ways of accessing SubMIT

- Classic access via terminal – requires to setup the keys
- A more modern way via a web-browser using JupyterHub
- VSCode terminal like
- Interactive GUI via X2Go (Linux) or X-Win32 (X11 server on Windows)

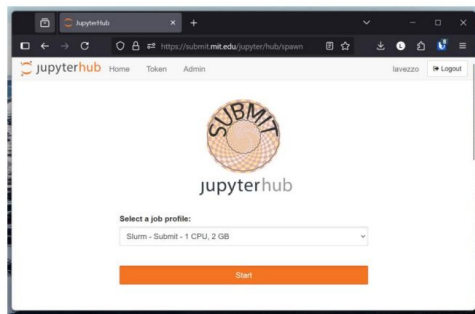
Secured authentication via Touchstone and ssh keys

All options provide full access to the SubMIT resources and functionalities

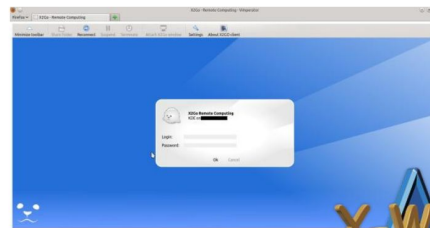
Terminal

```
lavezzo@submit05~$  
Microsoft Windows [Version 10.0.26100.3775]  
(c) Microsoft Corporation. All rights reserved.  
  
C:\Users\\lave>ssh lavezzo@submit.mit.edu  
Last login: Wed May 14 14:01:57 2025 from 18.4.134.206  
(base) [lavezzo@submit05 ~]$
```

Web-browser



X-Win32



What can users get on SubMIT?



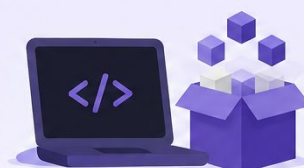
storage

- Home space for code, scripts and small outputs
- Work space for software, libraries and user environments
- Large-scale storage on CephFS
- High-speed access across the cluster



computing

- Interactive login and analysis nodes for development and exploration
- HTCondor for large-scale batch processing
- Slurm for workload management and dedicated workflows
- High-memory and large-core-count machines for demanding jobs



software

- Extensive centrally managed scientific software stack
- Support for conda environments
- Apptainer (Singularity) containers
- Docker support where appropriate



support





- User support via ticket system
- Dedicated team ready to help
- Help with software, storage, workflows and troubleshooting

See Jacob's talk



Storage resources

What do users get on SubMIT by default?

Default storage spaces (per user)			Quota
	/home/submit/user Code, plots, small outputs	<ul style="list-style-type: none">Fast access (100 Gb/s)Backed up daily	10 GB
	/work/submit/user Software & libraries	<ul style="list-style-type: none">Fast access (100 Gb/s)	100 GB
	/ceph/submit/user Large data storage	<ul style="list-style-type: none">CephFSFast access (100 Gb/s)	1 TB
Group area storage (shared by groups)			
	/ceph/submit/group Large data storage	<ul style="list-style-type: none">CephFSFast access (100 Gb/s)	X TB

} **Doubled the quota**

Other resources and tools

- Fast mount space (/scratch), can go up to 120 Gb/s (storage space upon request)
- MIT Tier-2 site storage (storage upon request)
- XrootD available for easy and efficient file transfers (internally and externally)
 - Requires authentication via grid certificate



Computing resources

What do users get on SubMIT?

Type

Purpose

Available resources



Interactive nodes

SubMIT login nodes
Development and debug workflows interactively

6 machines



SubMIT worker nodes dedicated for compute within SubMIT computing environment

~1500 CPUs
~ 66 GPUs

Access: submit, submit-gpu, submit-gpu-express

GPU types: 4 NVidia GTX 1080, 12 V100, 8 A30, 2 RTX6000



Large-scale computing, access to external resources

Several thousands CPUs (MIT T2, T3, OSG, CMS)

Increasing computational need and complexity





Computing resources

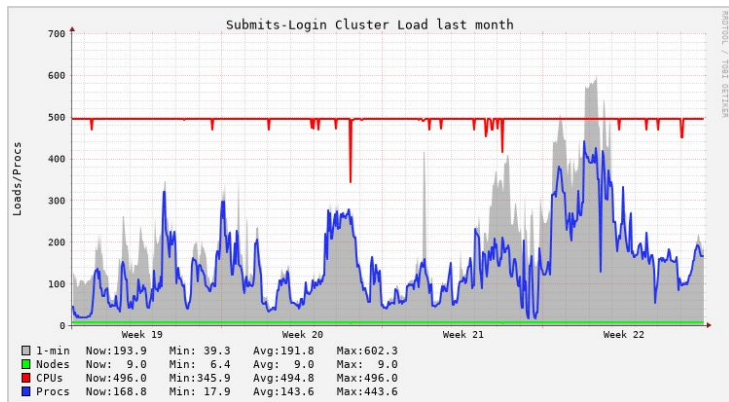
Monitoring in place for both users and admins

- Fair-share on interactive nodes maintained by users
- Scheduling on Condor and Slurm resources

Good usage and pressure of Condor and Slurm resources

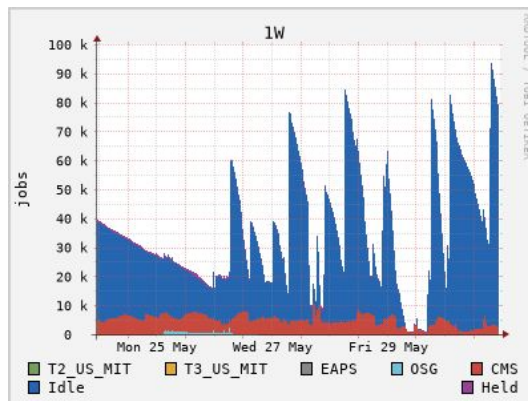
Diverse types of workflows continuously being executed at SubMIT

Load on interactive nodes



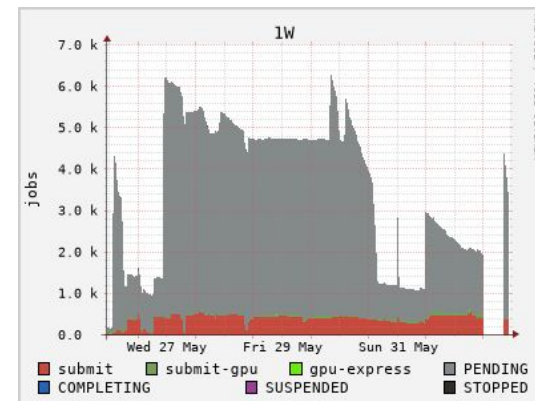
Well balanced load

Condor schedd



Up to 5-10k jobs running, often in saturation, mainly CMS

Slurm



Up to 500 jobs running, often in saturation



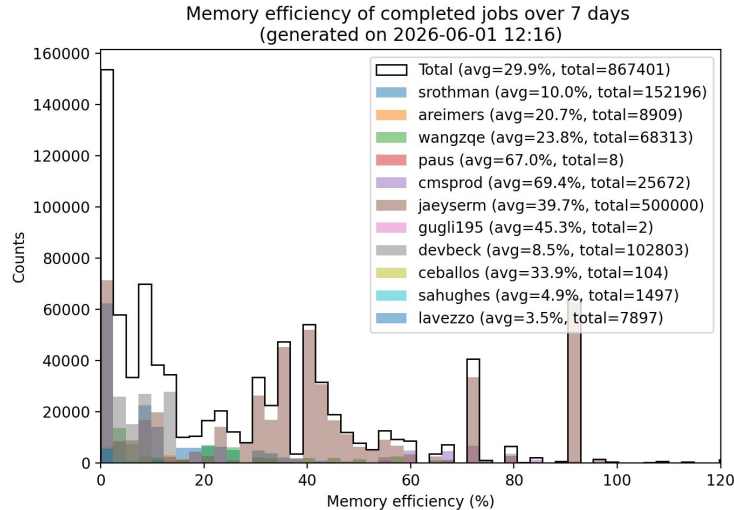
Computing resources – user job monitoring

New monitoring in place for memory and CPU efficiency

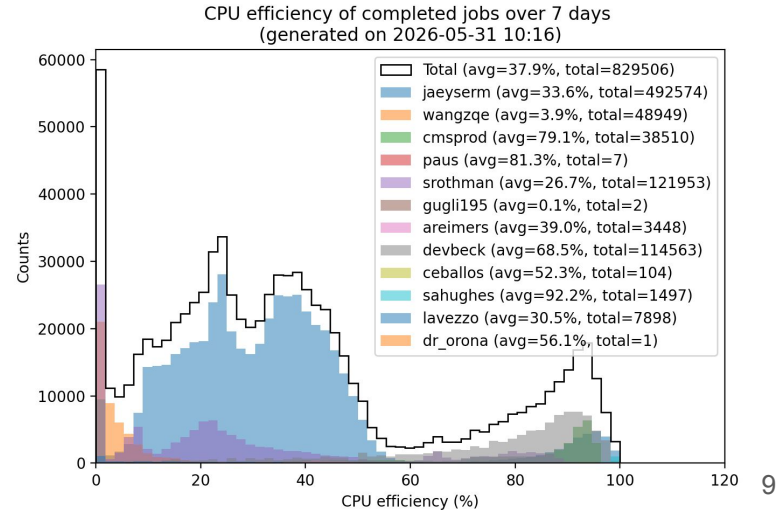
- From admin/schedd point of view, need to optimize the resource requests: the request should match the demand
- Monitoring for all finished Condor jobs per user in intervals of 7 days, 1 month and 3 months
- Users are notified manually if the job resource request largely exceeds the job demands

Similar monitoring being set up for Slurm jobs

Condor memory efficiency



Condor CPU efficiency



Software resources

Users have many tools available to ease their workflows

- Try to maintain common tools centrally
- Otherwise, users make their own environment via package managers and containers



Centrally managed standard libraries and software available natively



Users have access to conda as an environment and package manager and PIXI packager (new!)



Containers supported via podman/docker and singularity



Specialized software supported: OpenMPI, Dask, CernVM-FS, XRootD, ... as requested by users!

Others tools

Many other tools available at SubMIT for the users to get their work done *efficiently*

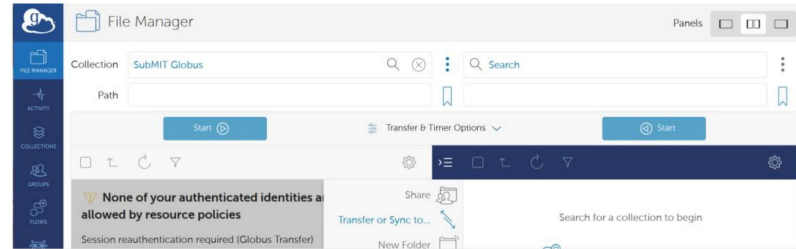
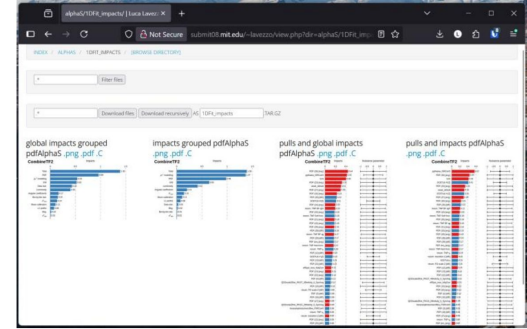


Parallel computing software
Dask Gateway and OpenMPI
integrated with Slurm



Support and guidance for
VSCode and AI-integrated
applications (Claude, Codex)

Personal and group
websites to share files
and display plots



Globus endpoint for data transfer and sharing

Future plans



SubMIT keeps being a mature and performance computing cluster

- OSG authentication via tokens ongoing
- XrootD authentication via tokens, so members without a certificate (typically non-LHC users) can also use XrootD for larger data transfer
- Support of Slurm user monitoring, similarly as implemented for Condor

Summary



SubMIT keeps being a mature and performance computing cluster

- Easy to join for members of the Department of Physics or external users
- Straightforward to interact with via terminal or web-browser (among others)
- Large, fast storage space available with a modern and efficient filesystem
- Suite of software tools centrally maintained or supported for users to access any software they need
- Access to large compute pools via batch-submission using Condor and Slurm

We continuously maintain the core tools, improve monitoring and expand on the user's needs

based