

subMIT Overview/Status

Josh Bendavid

Basic Computing Services (subMIT) Users Meeting
Oct. 22, 2024

Introduction

- subMIT system provides an interactive login pool + scale-out to batch resources
 - Home directories
 - Convenient software environment (Alma Linux 9 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
 - 800TB of spinning disks
 - Local storage (1TB/user), 10's of TB for larger group datasets
 - 40TB 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
- System is located in B24 basement, with 100gbps uplink

Introduction: subMIT Website



subMIT

Getting physics things done at MIT

[Overview](#) [News](#) [People](#) [Contact](#) [About](#) [Users Guide](#) [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
- 100s of cores and 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

Table of Contents

User's Guide - subMIT

Tutorials - subMIT

Future Work - subMIT

Indices and tables

Next topic

This Page

Show Source

Quick search

Go

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)

Planned Upgrades:

- Move to AlmaLinux

Indices and tables

- Index
- Module Index
- Search Page

Introduction: Project Organization

- Formally the project is organized as ***Basic Computing Services*** in the Physics Department
 - **Project Team:** Implementation/Operations/Maintenance of the system
 - **Users Group:** Contact point between the user community and the project team, forum for user feedback, requests, information flow to and from users
 - **Steering Committee:** Faculty oversight, funding, etc
 - See https://submit.mit.edu/?page_id=6

Users Group In Practice

- Regular meetings (every few months)
 - 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

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)
 - /ceph/submit/data/user/<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
 - Group space in /ceph/submit/data/group/<groupname>

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.

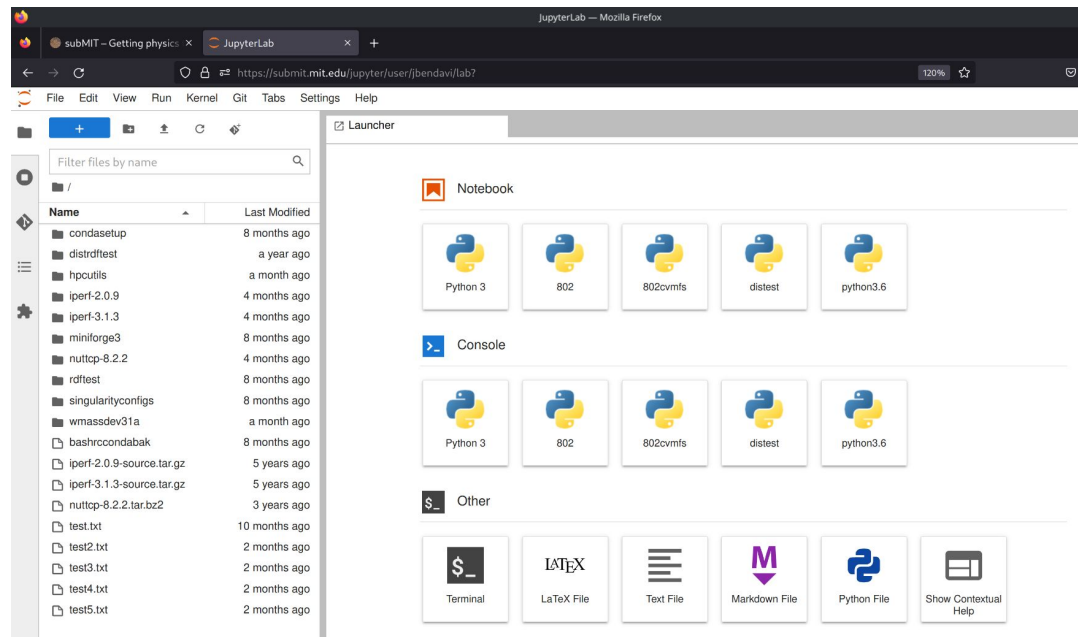
- **Conda:** your conda environments should be automatically loaded as kernels by Jupyterhub, and can be used in notebooks. See User Guide for more info.

- **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.



- 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
- Primary usage is research, but also being used for classroom exercises

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 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
 - February 2024 workshop: <https://indico.mit.edu/event/956/>
- 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

Linux Distribution Upgrade

- Previous CentOS 7 distribution reached EOL for maintenance updates in June 2024
- Decision by Red Hat to reorganize CentOS project and releases disrupted the logical upgrade path from CentOS 7->8
- Decision was 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 included Users Group and broader community
- System has been fully upgraded, with user facing services (ssh, Jupyter, batch queues) switched to Alma 9 by default in April
- Dedicated documentation to ease user transition
 - <https://submit.mit.edu/submit-users-guide/future/alma.html>
- Well-supported and documented use of containers to keep older software environments available where needed

Mass Storage Upgrade

- Previous mass storage system (/data/submit) was 500TB of spinning disks in a Gluster distributed filesystem
- **Users experienced throughput bottlenecks for high performance analysis**
 - unable to effectively leverage the throughput of a large number of disks in parallel
- **Some other performance and operational issues related to user access patterns (system responsiveness/reliability problems for users)**
 - Large number of small files, large number of files in a single directory
- A number of limitations in maintenance and flexibility, plus suboptimal failure modes (files appear to have vanished when they are only temporarily offline)
- These issues also drive extra demand for scarce/expensive NVMe storage (/scratch area)
- **Mass storage has been migrated to higher performance and more robust CephFS**
 - Migration completed in September
 - Additional disks added and more efficient erasure coding -> 800TB of usable space

Mass Storage Upgrade

- Mass storage fully migrated to CephFS
 - More space
 - More robust
 - More responsive
 - Better monitoring for project team

