subMIT Overview

Josh Bendavid

Basic Computing Services (subMIT) Review

July 18, 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

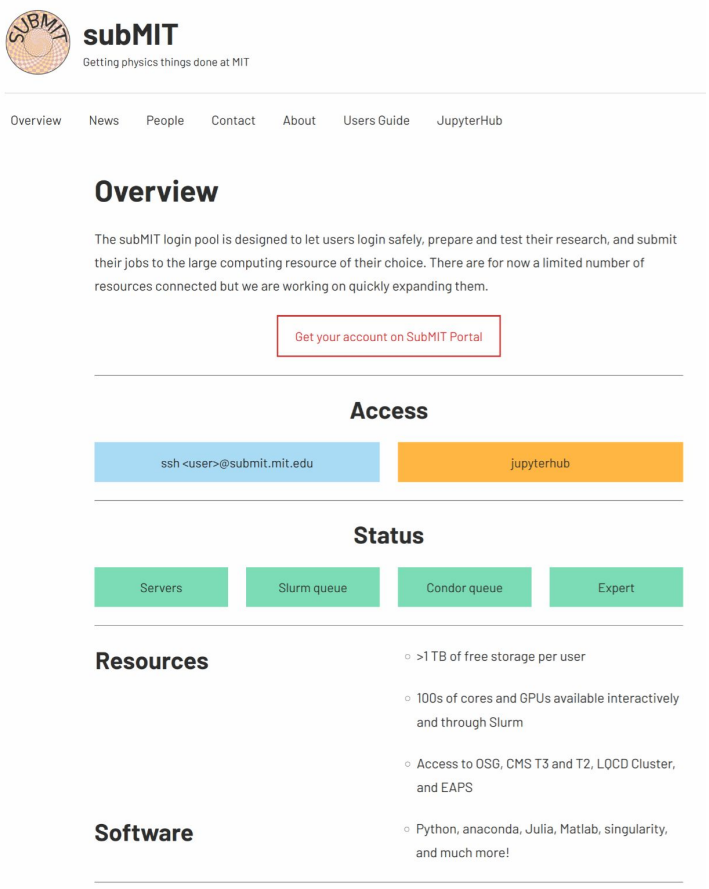
Introduction

- 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
- 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

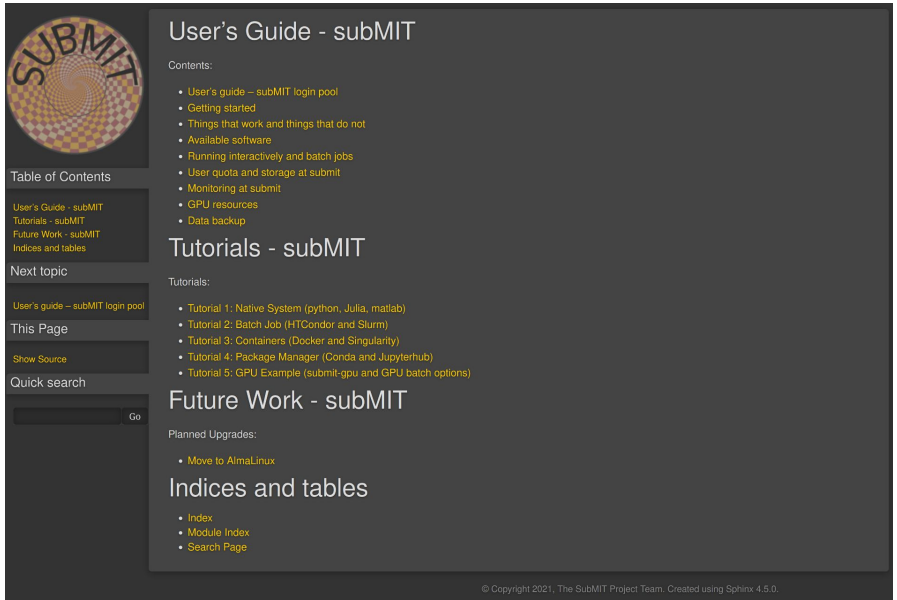
Physical Location and Infrastructure Support:

- Physical location: Machines at both Bates and Building 24 up to now
  - 100 Gbps uplink in both locations
  - Split physical location doesn't make sense for a facility of this size
    - Single location also strongly preferred for RoCE (RDMA) support
  - B24 preferred for local maintenance and support reasons
    - Network and power have been recently upgraded, current infrastructure is sufficient for a system of this size + reasonable expansion
    - Power reliability in Bates was also a concern up to this point
  - Remaining services and hardware at Bates in process of being migrated to B24
    - Expected completion in August
- Infrastructure support
  - From the fall onwards infrastructure support for subMIT is needed in B24 but not at Bates
  - LNS computing support for subMIT in B24 provided through Paul Acosta and Dave Newman
    - Infrastructure: Racks, Power, Networking
    - Network: hostnames, IP addresses, user authentication, interface with IS&T
    - Redmine server for ticketing system (also used beyond subMIT project)

Introduction: subMIT Website



- 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/



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

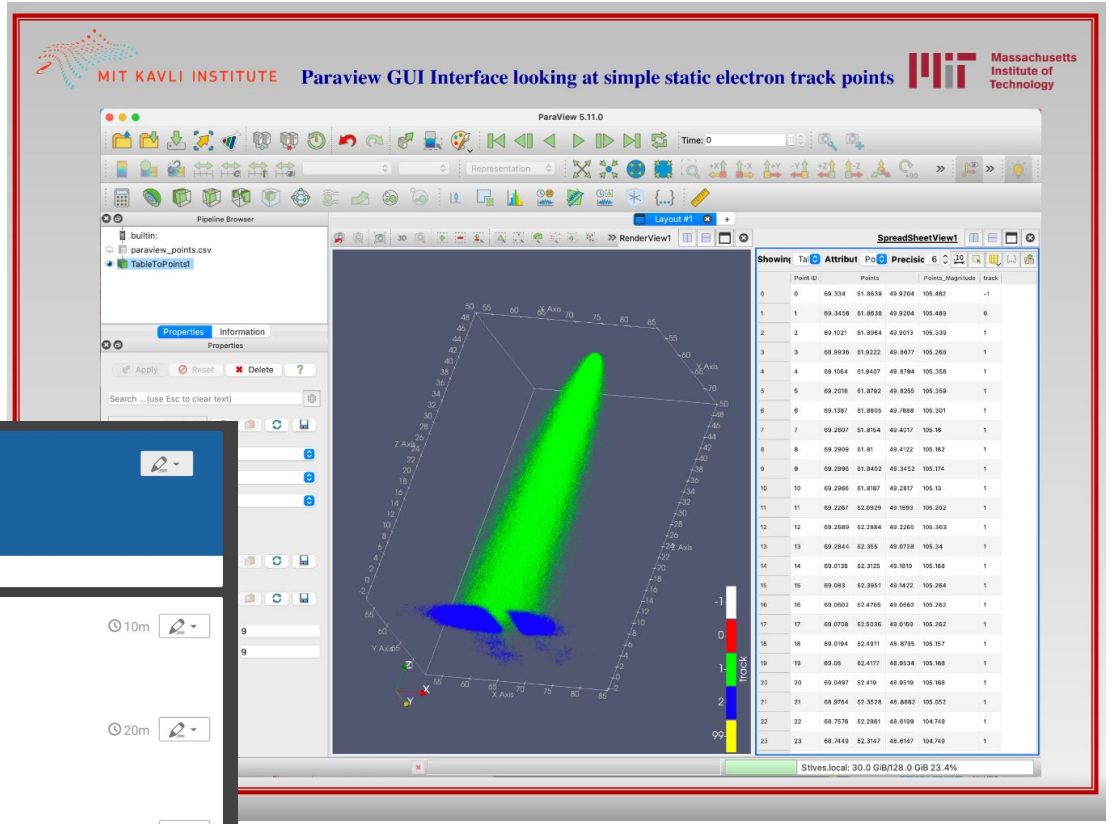
- Monthly meetings
  - Advertised and open to the broader community
  - Topical presentations from project team, Users Group representatives, or other users or community members
  - Forum for feedback and information flow between the user community and the project team
  - Regular timeslot: Tuesday 10:00-11:00 EST
  - Next meeting: Tuesday July 25: https://indico.mit.edu/event/802/
- 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)

Users Group Meetings

- E.g. presentation from April Users Group meeting on Visualization of Geant Simulations on subMIT
- https://indico.mit.edu/event/752/



Users Group Meeting

Tuesday Apr 25, 2023, 10:00 AM → 11:00 AM America/New_York

Kolker Room (26-414) (MIT)

Description https://mit.zoom.us/j/96743699673?pwd=b3h2Q3c3cVQwYW12blhMUG5SWXZCZz09

10:00 AM → 10:10 AM **News/Introduction** 10m

Speakers: Joshua Bendavid (Massachusetts Institute of Technology), Matthew Heine (staff@mit.edu)

UserGroup_04_25_2...

10:10 AM → 10:30 AM **Visualization of Geant Simulations** 20m

Speaker: Richard Foster (Massachusetts Institute of Technology)

Foster_SubMIT_Par... Foster_SubMIT_Par...

10:30 AM → 10:45 AM **Roundtable** 15m

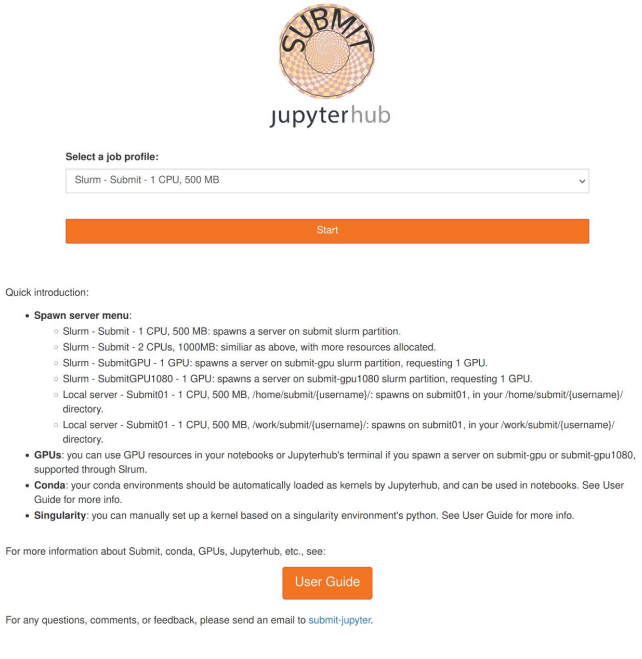
Speakers: Kaliroe Pappas (MIT laboratory for nuclear science), Molly Taylor (Massachusetts Institute of Technology), Prajwal Mohan Murthy (MIT LNS), Siddharth Mishra-Sharma (MIT), Sunghan Ro (MIT), Yin Lin (Massachusetts Institute of Technology), Yitian Sun (Massachusetts Institute of Technology)

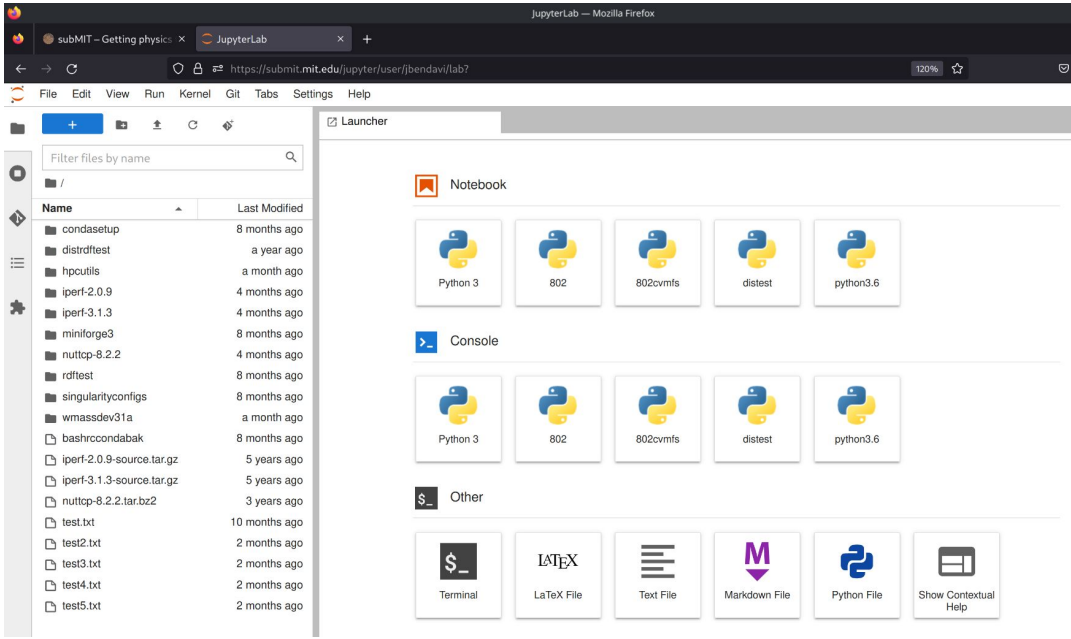
10:45 AM → 11:00 AM **Discussion** 15m

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





- 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 education applications also possible
  - Limited trial has been conducted for 802 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
  - Joint project with College of Computing, with dedicated funding
  - More discussion later + dedicated talk at LLM workshop on Friday https://indico.mit.edu/event/759/
- 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

subMIT and ORCD

- subMIT provides Basic Computing Services tailored to the physics department
  - Easier in terms of scale and scope than institute-wide solutions
  - subMIT system is already up and running and successfully supporting research across the department
- “Local” nature of support and physics expertise of project team:
  - Fast turnaround time
  - User guidance on workflow design and optimization
  - System design and evolution targeted towards physics research needs
- Longer term relationship of subMIT with ORCD and balance between institute-wide and department-level services to be defined in coming year(s)

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
- Ease transition for users through well-supported and documented use of containers
- Test instances for interactive and batch usage with Alma 9 are being set up for initial user feedback
- 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

Today's Review

- Indico page with timetable and slides:
  - https://indico.mit.edu/event/761/
- Overview of subMIT project, resources, software environment
- Discussion of User Support
- Usage Metrics
- Kolker Room + Zoom