

# SubMIT

## Recent & Future Upgrades

Zhangqier Wang  
SubMIT Review  
May. 22<sup>nd</sup> 2024





# Outline



- TO ADD



# Accomplished Upgrades

- All the upgrade plans reported in the last review is accomplished
  - Operation system fully upgraded to AlmaLinux 9
    - All the crucial software are either upgraded according or replaced with other Alma9 packadges with the same functionality
  - Storage system migrated from Gluster to Ceph
    - Gluster disks decommissioned and used by Ceph
    - New disks put for Ceph: storage 500TB -> 1.5 PB





# New Location



- Remain compact
  - All submit machines contained within 4 racks
  - Three switches with 100Gib link
  - Compact structure

PICTURE



# Unexpected Outage



- On Saturday, March 22, a scheduled MIT router upgrade, during which the network was switched to a secondary router.
- The LNS network to the campus network is not properly connected via the secondary router, causing 6 hours down-time.
- The SubMIT had the check in person when outage happened and quickly identify the issue.
- Keep in contact the with the IS&T about the upgrade the timeline and informed users

News

## Unexpected SubMIT Outage During MIT Network Maintenance

by **David Walter** – March 24, 2025

On Saturday, March 22, the SubMIT computing analysis facility experienced an unexpected six-hour outage, between 1pm and 7pm, due to [...]

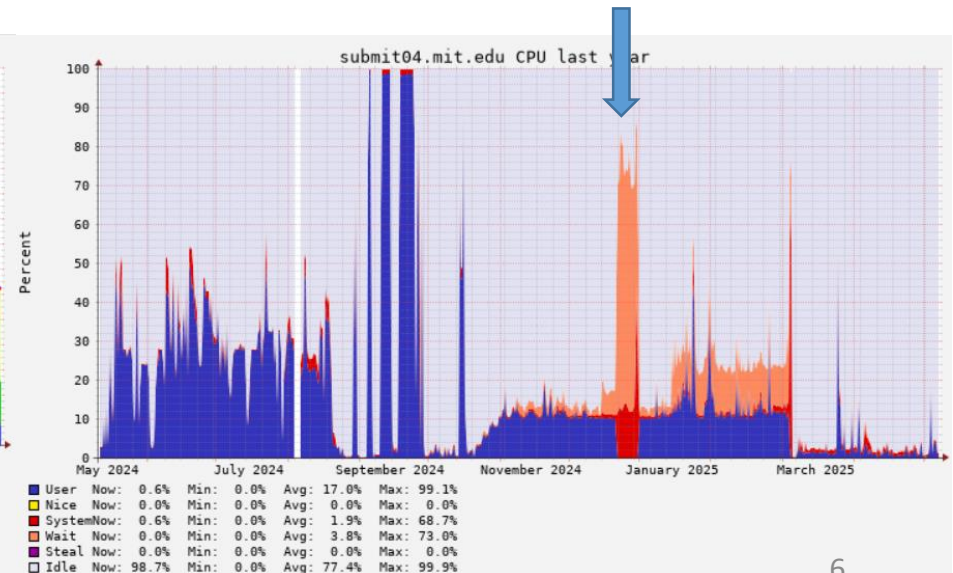
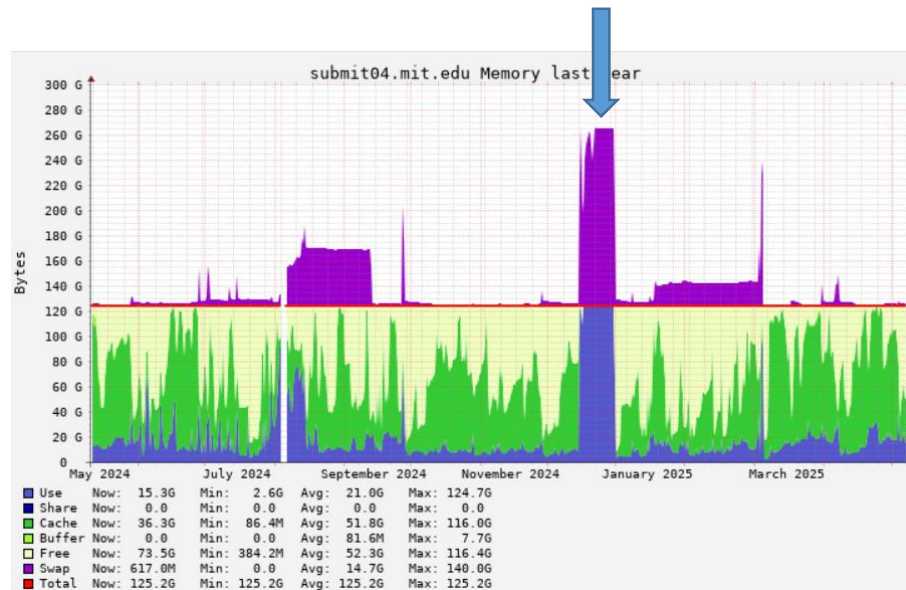
[Continue reading »](#)

- LNS router configuration adjusted to avoid such outage in the future



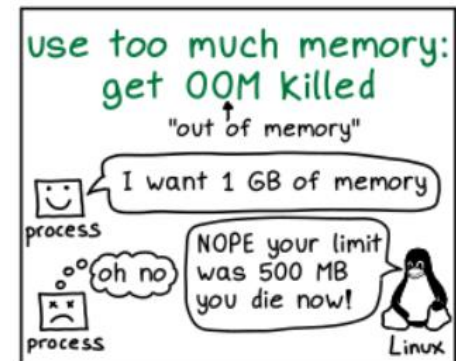
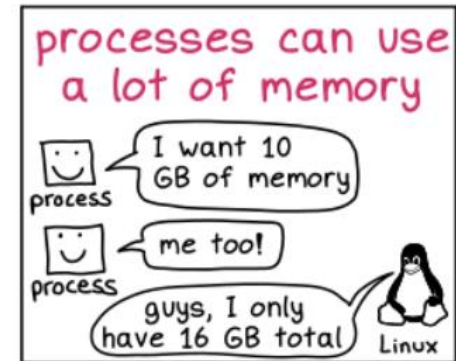
# Improper Operations

- Some user jobs unintentionally consume excessive CPU, memory, or I/O
  - Infinite loops or runaway processes
  - Excessive parallelization without limits
  - Poorly optimized code
- A single mistake brought down the entire server, along with the crucial service on it.
  - Sometimes we have to reboot the machine to bring it back





- Implement Control Groups (cgroup)
  - accounts for, and isolates the resource usage
- The limits set up is far beyond an ordinary WFs
  - 70% memory of the machine
  - 90% of the CPU of the machine
  - Reserve certain resources for system service
  - Ensure that we don't constrain the productivity
- A safe guard to all the users for a stable system
- More resources freed under this protection
  - 10 Ceph servers put in slurm pool under cgroup protection



TO WORK

- Cgroup partially implemented for basic protection
  - Limits applied on each machine to avoid overload and jammed the crucial service
- A more sophisticated cgroup set up on the way.....
  - Cgroup v2 is implemented in the last upgrade,
  - Enable cgroup in slurm, using cgroupv2.
  - Improve cgroup setting on each server
- So that we have
  - Predictable performance for users
  - Reduced node crashes
  - Simplified cluster management
- Ensure a stable environment for all users





- Keep AlmaLinux version upgraded (current version 9.5)
- **Improvements:**
  - Security patches and kernel stability fixes
  - Enhanced **cgroup v2** support for better resource control
  - Faster systemd boot and updated core packages
    - **CephFS, NVIDIA Driver, Slurm, etc.**



News

## System Maintenance Completed

by Zhanqier Wang – April 18, 2025

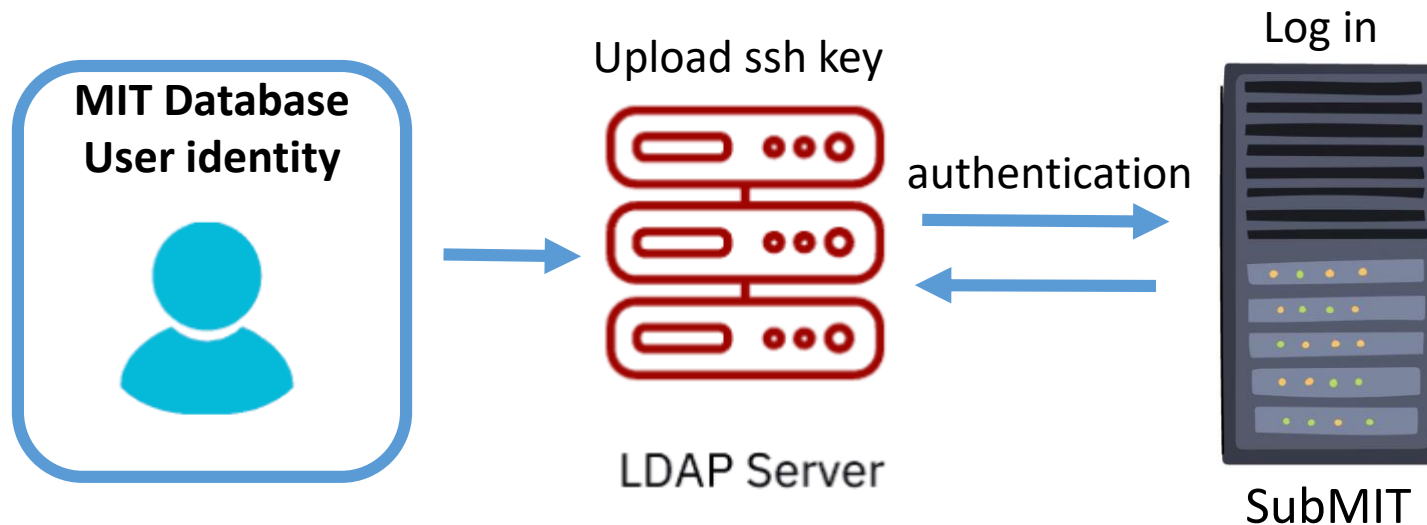
On Thursday afternoon, April 17, we performed scheduled system maintenance to apply a series of important upgrades. This required temporarily [...]

## Upgrade Policy

- Security, compatibility, performance improvements
- 30 days after the new release of Alma9 version
- Includes crucial softwares
  - Batch: htcondor, slurmd
  - Storage and access: Ceph, xrootd
  - others: GPU drivers, wordpress, Julia, ...
- Software upon request

**PLAN**

- The SubMIT user information is managed by LDAP, which is connected to the database of Athena to identify users.



- Take over LDAP server
  - reduce turn around time to manage research groups (GID) and manual operations
  - Add more detailed user metric – division, lab, etc.
- Streamlined account creation





# Other Plans



- Self cleaning mechanism
  - Orphan processes occur and left running for weeks without intervention
  - Consume extra resources
  - Automatically eliminate the orphan processes after XXX running
- Remote network management
  - Kickstart via pixel boot
    - Installation configuration ready
    - kickstart server in preparation
    - A consistent installation configuration



# Summary



- TO ADD



The End