SciOps+ENP Monthly Coordination Meeting

May 13, 2024

Bryan Hess

Monday, May 13, 2024

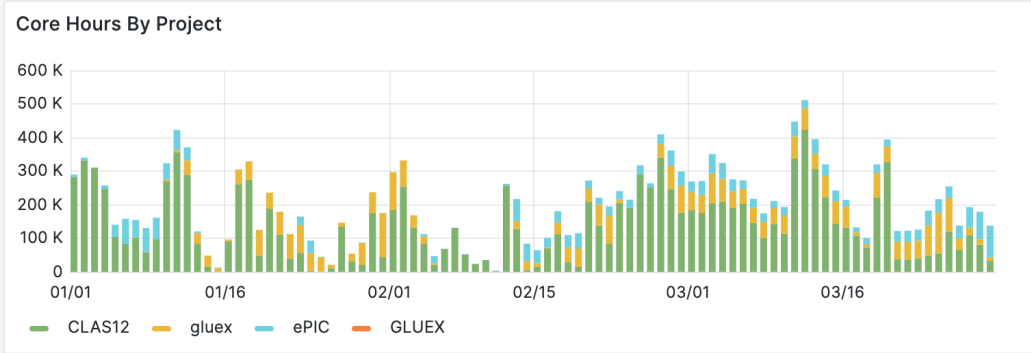
Jefferson Lab

U.S. DEPARTMENT OF ENERGY | Office of Science

JSA

Performance Overview – Storage, OSG use, Data Transport

- Data Rates from halls to Data Center are increasing.
- Both 100Gbit/sec internet connections are up and have been demonstrated to run at capacity
- Planning for 2x400Gbit is underway; This is a long-lead, multi-year project that will kick off in FY25
- OSG Core Hours usage shown below



Q2 2024 Tape Usage	Tape Writes	Tape Reads
Raw Data from Halls	1.8TB	5.8PB
All Other Data	3.3TB	3.2TB
Total	**5.1PB**	**9PB**

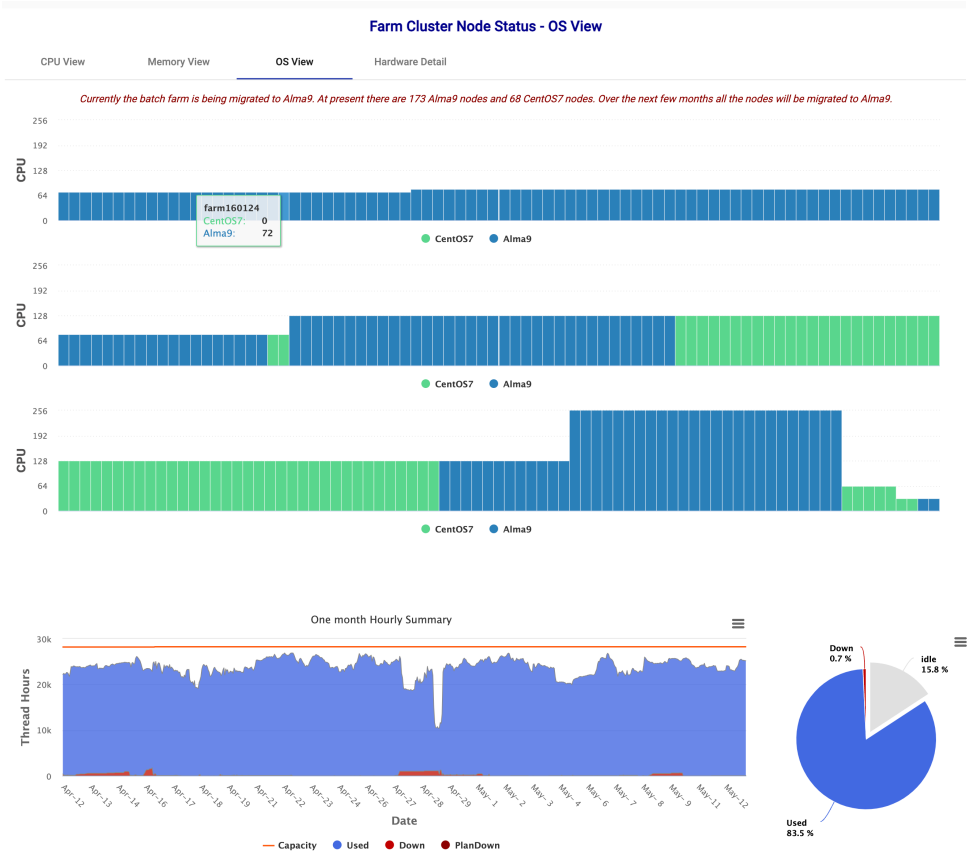
DATA – JANUARY 2024	Peak	95% traffic	TOTAL
Halls A/B/C/D to Data Center	11.78 Gbps	6.98 Gbps	836.11 TB
Internet – Inbound	5.43 Gbps	2.83 Gbps	201.4 TB
Internet – Outbound	10.13 Gbps	5.73 Gbps	595.13 TB

DATA – FEBRUARY 2024	Peak	95% traffic	TOTAL
Halls A/B/C/D to Data Center	11.91 Gbps	5.77 Gbps	1.11 PB
Internet – Inbound	50.21 Gbps	1.48 Gbps	221.13 TB
Internet – Outbound	51.21 Gbps	7.31 Gbps	832.11 TB

DATA – MARCH 2024	Peak	95% traffic	TOTAL
Halls A/B/C/D to Data Center	5.46 Gbps	4.60 Gbps	517.16 TB
Internet – Inbound	10.91 Gbps	4.27 Gbps	515.70 TB
Internet – Outbound	86.70 Gbps	32.94 Gbps	2.47 PB

AlmaLinux 9 farm deployment – Proposed Completion Plan

- Most of the farm is now running AlmaLinux9 (over 80% of CPU)
- Both the EL7 and OL9 hosts are consistently busy. Usage is high.
- On Maintenance day (May 21, 2024) we will change the default from EL7 to EL9.
  - This is a potentially breaking change for some EL7 code.
  - Prior to maintenance day and email announcement will be sent
  - Targeted emails will be sent to users with large EL7 workloads
- The last 10% of CPU will remain on EL7 while the last jobs are migrated.
- Once the final EL7 nodes are upgraded
  - the default will be removed
  - The ifarm 7 nodes will be retired
  - target: summer



Reminder: Slurm Production partition feature changes for Alma9

- **CentOS 7 Only Jobs**
  - No *current* change is required (see deprecation notice*).
  - If no constraints is specified, jobs run on centos79 until 5/21/2024.
- **Alma9 Jobs**
  - Specify a constraint to run on EL9 (Alma9) nodes: `--constraint=el9`
  - Deprecated features will not be allowed (see deprecation notice*).
  - After transition, no change required.
- **OS agnostic jobs (example: container-based workflows)**
  - During the transition, to use all farm nodes set `--constraint=el7|el9`
  - After the transition, el7 (or centos79) **MUST** be removed (see deprecation notice*).
- ***Deprecation Notice** - As nodes transition to Alma9, we will deprecate the following *features*:
  - general (same as not specifying any constraints)
  - Centos79, el7
  - amd, xeon (unused)
  - gpu, TitanRTX, T4, A100 (not used on production partition and handled by gres as needed)
  - Attempting to use deprecated features on Alma9 nodes will result in a job submission error.

Lustre24 schedule

- Testing is complete.
- Lustre24 system will be mounted on farm/ifarm systems on Maintenance day, May 21st.
- Testing of large-scale data sync will continue that week
- Next Steps
  - Determine which areas to move first
  - Continue though projects on /cache and /volatile until sync is complete
- Decommission Lustre19 system
- Augment Lustre24 system with FY24 purchase beyond 10PB.
  - This can be done without outage

Code.jlab.org status (gitlab)

- Progress Continues to build out capabilities
- Coordinating a date with RedHat to install Open Shift
  - Expected June
  - Needed for CI/CD functionality
  - Needed for expanded object store capability
- Streamlining the onboarding process for new users

Procurements and Hardware Changes

- Two new GPU nodes with A800 GPUs, sciml240[12]
  - Installed. In production.
  - See summary at https://scicomp.jlab.org/scicomp/nodeStatus/hardware
- Two new ifarm nodes for Alma9 – PO awarded. Awaiting Delivery
- Tape Storage
  - Purchasing 4 more tape drives to support Hall A load (online and offline)
  - Purchasing data mover systems for these drives
  - Lifecycle replacement of two write buffer nodes for data from halls and to augment write speed for 3 hall high-rate running.
- Networking: still waiting for network switches for farm Ethernet side (1 year lead time)
- Lifecycle of systems
  - Non-farm nodes from FY18 or earlier have been retired
  - Farm18 nodes will continue to run; low risk to operations from failures
- FY24: Storage Augmentation of Lustre24 – procurement process started
- FY25: CPU for farm (specification to be determined)
  - FY19 – now has been AMD EPYC based on cost for CPU.
  - Intel may be competitive again (e.g. Sapphire Rapids)

Projects in Progress and Status

- Power and Cooling capacity in the data center
  - Project underway with FM&L to look at a 5 year plan for power, cooling, and floor space
  - We have developed a capacity planning model for data center management
  - Adding two more PDUs, looking at HACS cooling capacity
  - Newer servers are power dense and loud. Data center is now a high noise area
- Rucio – progress continues on tape storage integration and off-site data replication
- NVMe Storage for project areas – no new updates
- Farm-local home directory – no new updates