SciOps + ENP Meeting February 2024

Bryan Hess

Thursday, February 8, 2024

Jefferson Lab

U.S. DEPARTMENT OF ENERGY | Office of Science

JSA

Recent Operational Issue Summary

- We have had a few incidents where a user submitted jobs with very small memory requests and then rapidly allocated memory beyond the request. Normally this is caught by the CGROUP OOM killer, but due to a limitation in the CentOS 7.9 kernel, system memory was exhausted.
    - The jobs were limited, the user was contacted, and the problem has been remedied
    - The fix for the root problem is a newer kernel, which we get with Alma9.

- We had a handful of unscheduled tape library outages
    - One in January because of dropped tapes that blocked the robot path
    - An outage was avoided because one robot was able to work while the gripper on the first was replaced.
    - this morning due to a broken drive belt and failed part on both accessors [Ongoing]

- There have been several cooling irregularities
    - an air handler misconfiguration led to overcooling in the data center. We are awaiting a fix from FM&L
    - The rental chillers have tripped off simultaneously causing overheating. FM&L has a PO out for the repair of the permanent system.

Jefferson Lab

Hardware Updates

- Lustre System
  - Networking, Power, and initial software specs complete
  - A hardware limitation with the Metadata server high availability was discovered, and a replacement disk shelf has been ordered to address the issue. We should know a delivery time soon.
  - Meanwhile we are able to proceed with testing and configuration of all other components.
  - Target Delivery is May, but this may slip depending on parts availaibility.

- NVMe Storage
  - We are evaluating Weka as a flash-based complement to Lustre and successor to /work.
  - We anticipate a decision on its viability by the end of March.

- Ifarm Nodes
  - Two new ifarm nodes are in the early stages of procurement
  - Target is increased memory and local high-speed (NVMe) storage.
  - Farm18, farm19 will be retired as lifecycle replacement.

Jefferson Lab

Software Updates: Alma9

- We are going to step up the switch of Nodes to Alma9

- Ifarm9 is available for testing

- Farm16 nodes are being upgraded now.

- Pace of conversion from CentOS 7 to Alma9 will step up
  - Please let us know of any issues via ServiceNow

Jefferson Lab

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, we will default you to centos79 during this transition.
- **Alma9 Only Jobs**
  - Use constraint to specify EL9 nodes (ex. --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=centos79|el9
  - After the transition, 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
  - 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.

Jefferson Lab

Networking Updates

- ESNet/JLab Upgrade to 2 x100Gbit
  - Both circuits up.
  - Doing full bandwidth acceptance testing now
- The Infiniband Core has been updated to HDR (200Gbit)
- A long-standing bottleneck in the Ethernet Network for the Farm16 nodes has been remedied.

Jefferson Lab

Work in progress: Updates

- GitLab (code.jlab.org)
  - Ready for friendly user testing Feb 2023.
  - Accessible with federated identities from on and off site
  - Integrated CI/CD will deploy as phase 2 (Soon)
- XRootD Storage for GlueX
  - Target is streaming processing of data on OSG
  - Bringing this back to the front – auditing and upgrading data transfer nodes.
- Tape Library Outage – Completed. We now have two accessors (robots)
- Slurm Upgrade to Slurm v22 completed
- Podman support for containers on ifarm and farm in Alma9 (nearly complete)

Jefferson Lab