SciOPS + ENP February 2022
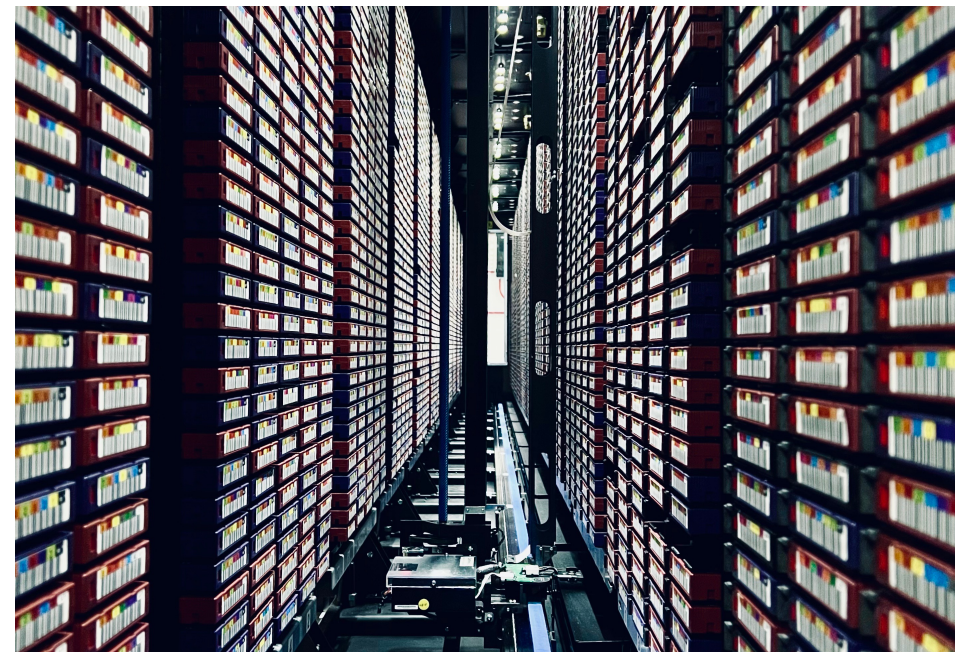
Topics

- Recent Changes

- OSG

- Swif

- XRootD

Bryan Hess
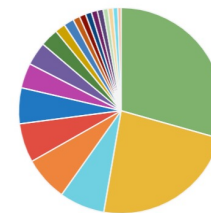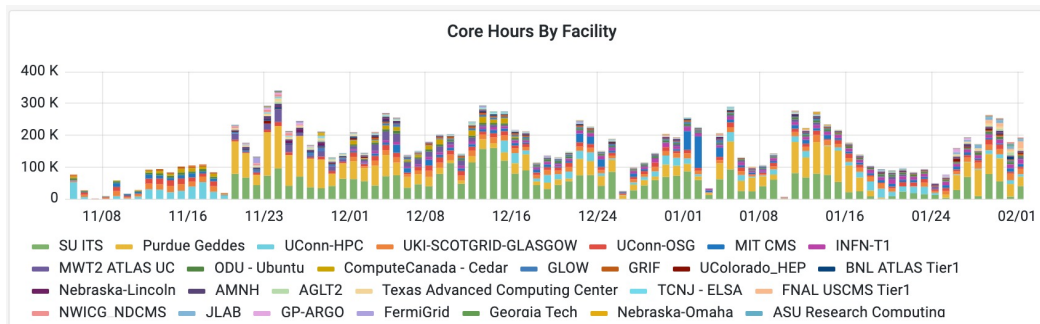
Tuesday, February 1, 2022

Jefferson Lab

Changes in the Past Month

- Slurm on the farm was upgraded to Version 20
  - This was done to improve GPU support, to keep up with security updates, and to provide new management capabilities.
  - There is behind the scenes follow on work from this related to disk and node management. Memory and GPU management is looking solid. There is a thread management issue being looked at for a maintenance day change (for jobs nearing 4000 threads)
  - Current focus is on local /scratch disk management, which was historically error prone. There is a slurm side to this to make sure we schedule and enforce limits, and there is also a swif scheduling aspect of this on the next slide
- Additional SSD buffering in place for raw data collection from halls.
- Reminder that automatic read back verification of raw data is now the norm.
- Last areas migrated to new work file servers.
- In Progress
  - Updates to Globus from v4 to v5.3
  - Long delayed data transfer nodes have shipped
  - Benchmarking for Farm Node procurement [Amitoj working with AMD]
  - Planning Feb Maintenance; Nothing that would change user workflow on the list so far.

Jefferson Lab

Open Science Grid

- We completing the upgrade from OSG 3.5 → 3.6 by Feb Maintenance Day

- The Major change (not user visible) is a transition to token-based authentication

- This is essential groundwork for token-based storage authentication authorization, which is progressing, targeted for production this summer.

- Resolved Issue: Jobs were not running at several sites (Purdue Geddes, Syracuse)

- Open Issue: OSG jobs were not launching at JLab when there were free farm cycles.
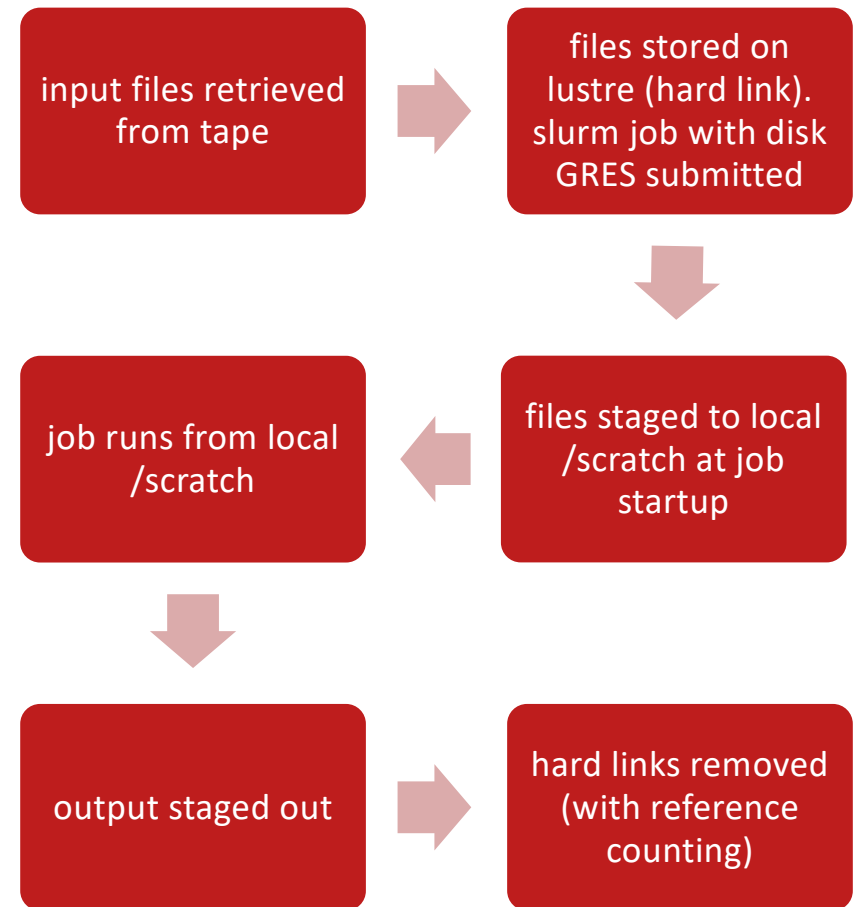

SU ITS	4 Mil
Purdue Geddes	3 Mil
UConn-HPC	974 K
UKI-SCOTGRID-GLASGOW	952 K
UConn-OSG	838 K
MIT CMS	767 K
INFN-T1	536 K
MWT2 ATLAS UC	499 K
ODU - Ubuntu	389 K

Jefferson Lab

SWIF version 2 rollout, Auger and SWIF 1 retirement

- Schedule Reminder
  - On **Tuesday, February 8th**, jobs submitted using Auger or swif1 will have reduced priority to encourage transition to Slurm and swif2.
  - On **Tuesday, March 1st**, the deprecated commands will be removed, Auger and SWIF1 services will be shut down, and documentation will be removed. The swif command will become an alias for swif2.
- Status and process
  - Uptake of swif2 has been steady and user reports constructive
  - We have regular working group meetings to look at bugfixes and issues
  - Focus is on bug fixes and clarifications now
  - Refactoring, performance improvements, new features will be gathered for maintenance day releases

Jefferson Lab

swif disk space management for worker nodes

- Local scratch disk is used by default

- If you specify a disk space requirement in swif, and it is less than the sum of the input file sizes, swif will fail the job.

- If you *do not* specify a disk space requirement in swif, it will fail. Future: warn that you have not specified a local disk size and it will allocate 1.1 * $\sum$ (input file size) in Slurm

- Slurm will respect node disk size and not over-commit /scratch

- Next step: disk use enforcement at run time

- Ask for $\sum$ (input sizes) + $\sum$ (output sizes) if using (default) local disk

- expert: If you specify absolute paths, you may read from /work, /cache, or /volatile directly

| input files retrieved from tape | → | files stored on lustre (hard link). slurm job with disk GRES submitted |

| job runs from local /scratch | ← | files staged to local /scratch at job startup |

| output staged out | → | hard links removed (with reference counting) |

Jefferson Lab

XRootD

- Current Status: We have two data transfer nodes available from off site
  - dtn-eic: off-site read-only access to /work/eic{2,3}
  - dtn1902: off-site read-only access for /work/osgpool
- XRootD Servers live on the Science DMZ network, which is outside the firewall
- When this was initially implemented, we had no ability to block sites that were "known bad actors" and so we limited exposure to the ESNet list of R&E networks
- Markus pointed out that this was limiting, but we now have cybersecurity monitoring and blackhole routing capabilities, so it is possible to open it up to the internet.
  - This received final cybersecurity and was opened 2/2/22.
- Read/Write access is under development and requires that we build out a reliable token based authorization/authentication framework.
- We are in the final negotiation with CILogon (at NCSA/ U. Illinois) to host the InCommon software stack to do this. Legal, Cybersecurity, and site office concurrence needed.
- Writeable XRootD storage is the first target for this capability.

Jefferson Lab