

# SciOps + ENP Coordination

February 2023

Bryan Hess

Thursday, February 2, 2023

Jefferson Lab



# Farm and Processing Status

- The 24 Farm23 nodes were deployed on Jan 23, 2023
  - 256 CPU (hardware threads) per node, 512GB memory
  - 7TB local scratch space per node
- A scheduling change was made to help prevent large-job scheduling problems when the queue was long and there was contention for resources
- SWIF workflow management working well/ integrated with /cache
- Lustre is performing well

	<b>farm19</b>	<b>farm23</b>
Product Line	AMD EPYC™ 7002 Series	AMD EPYC™ 7003 Series
# of CPU Cores	32	64
# of Threads	64	128
Max. Boost Clock	Up to 3.35GHz	Up to 3.5GHz
Base Clock	2.5GHz	2.45GHz
L3 Cache	128MB	256MB
Default TDP	180W	280W
CPU Socket	SP3	SP3
Socket Count	1P/2P	1P/2P
System Memory Type	DDR4	DDR4
Memory Channels	8	8
System Memory Specification	Up to 3200MHz	Up to 3200MHz
Per Socket Mem BW	204.8 GB/s	204.8 GB/s

# Reminder: Multi-Factor Authentication for Interactive Access

---

- March Maintenance day -- March 21st, 2023
- Status
  - Login machines are up and ready for testing
  - credentials are being issued to everyone with access. 300 more to go out of ~1700.
- I will send another email reminder soon
- Tools that may help
  - <https://tmuxcheatsheet.com/quick-start/>
  - [SSH Configuration Suggestions](#)
- Questions that have come up
  - MobilePass tokens, Google Authenticator, Microsoft authenticator all work
  - Existing tokens used for hallgw access are the same

# Farm Resource Monitoring

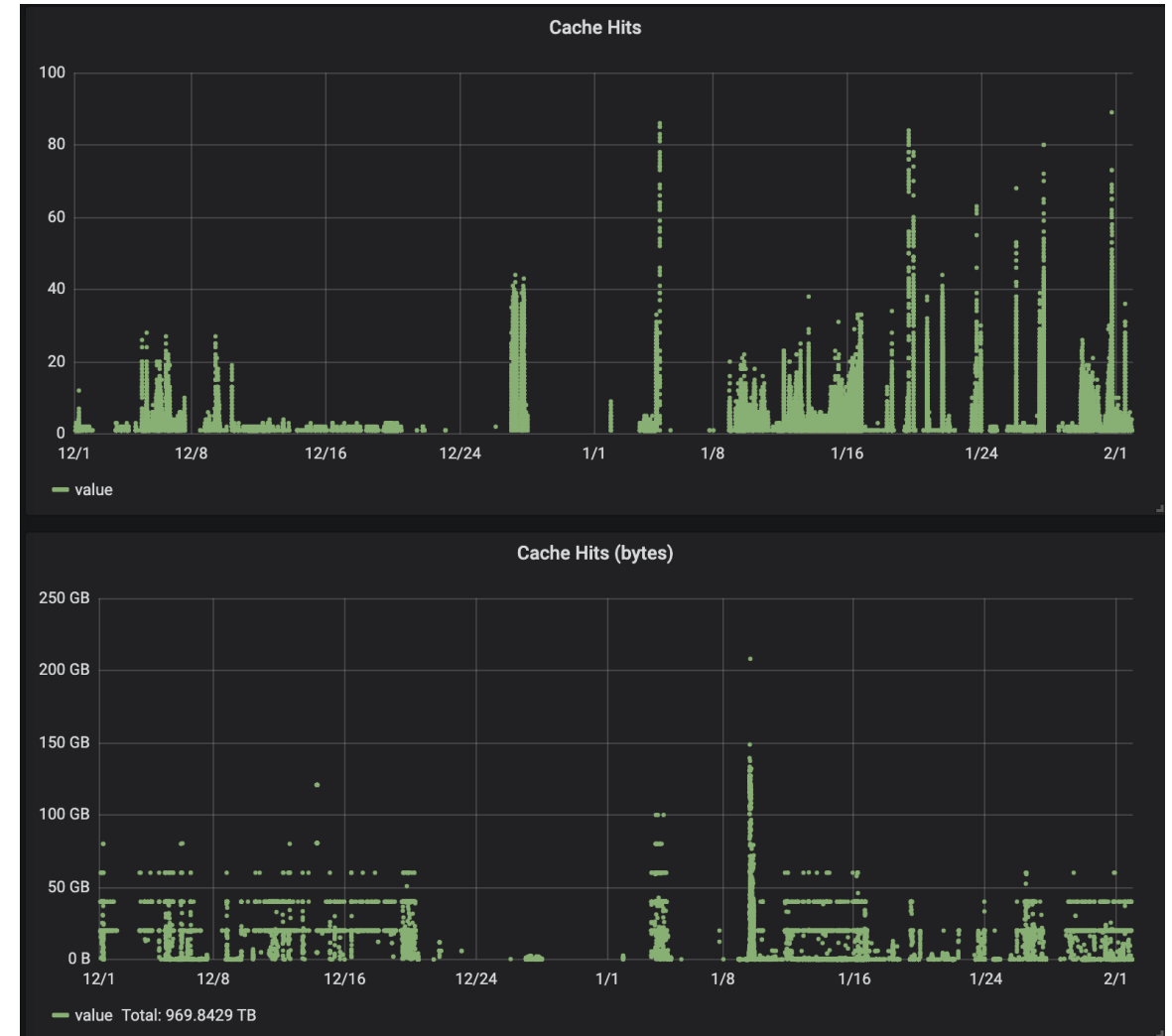
- The aspiration is to keep all CPUs busy
- Supporting this
  - Jasmine disk cache work (more on this)
  - Job Efficiency Monitoring
- We are stepping up monitoring for outliers that affect performance
  - Memory over-allocation (email and web)
  - <https://scicomp.jlab.org/scicomp/slurmJobs/memoryEff>
- This is also the focus of the LDRD, which has broader scope

User Name	Account Name	Average Walltime	Average CPU	Low Memory Efficiency Job Count ↓	Average Memory Efficiency
roberteja	clas12	29 mins	1	23,909	0.89%
keigo	halld	0 mins	1	6,380	0.02%
billlee	eic	2 mins	1	1,638	0.04%
andrsmit	halld	10 mins	5	126	0.03%
osg-eic	eic	44 mins	8	59	0.51%
erins	clas	1 mins	1	56	0.06%
mfmce	clas12	2 hrs 57 mins	3	54	2.24%
ruonanli	hallc	57 mins	1	30	0.38%
gmat	clas12	0 mins	3	22	0.02%
osg-clas12	clas12	44 mins	8	18	0.52%
nwickjlb	eic	2 hrs 20 mins	1	17	12.61%
heinricn	hallc	6 mins	1	11	0.21%

*Jobs included in this tables are finished successfully in last three days*

# Update: Jasmine Internal Disk Cache Performance

- Now fully in production
- Distinct from user visible /cache
- Aim is to avoid going to tape for current workflows
- Useful for duplicate creation, small files, "hot files" that age out of /cache.
- Current deletion strategy prioritizes high cost retrievals. This will be tuned as we learn more
- Nearly 1PB of cache hits in 2 months
- Aim: Keep the farm fed, reduce reliance on tape for current experiments
- Cache size will expand as hardware is available to down-cycle.



# NUMA Performance Considerations

---

- The farm Slurm configuration uses the `task/affinity` plugin. This binds threads to cpus.
- Further optimization that helps to avoid traversing memory cache domains. Keeping all the jobs on the the same socket will be a performance advantage.
  - For farm19, there are 32 cores per socket (64 threads)
  - For farm23, there are 64 cores per socket (128 threads)
  - Slurm CPU== hardware thread
  - A 128 CPU job on farm23 Milan can be bound to one socket
- Slurm allows for “hints” and explicit configuration that may help.
  - [https://slurm.schedmd.com/mc\\_support.html](https://slurm.schedmd.com/mc_support.html)
  - Low Level: `--cpu-bind`
  - Hints: `--hint=compute_bound`, `--hint=memory_bound`, `--hint=[no]multithread`
  - High Level: `--threads-per-core`, `--cores-per-socket`, `--sockets-per-node`

# Short Updates

- Rucio – first test system built
  - Security
  - Storage investigation
  - Aligns with our Jasmine/XRootD
- XRootD writeable storage
  - EIC test caes next
- Do we need a farm23 ifarm machine?
  - If so, how should it be provisioned? (memory, disk)
- Lustre23 purchase for farm
  - Planning is starting.
- Back to 4 halls!

