Joe Boyd

Fermilab Computing Resources Scrutiny Group

16 February 2023

Scientific Compute Services

Department Services

- Operations – Batch/Caching/Factory/Auth (2.5 CMS and 2.5 FIFE FTE)
- Production Operations Monitoring System (1 FTE)
- Landscape monitoring (grafana/kibana/elasticsearch/kafka/Prometheus/consulting/…) (0.5 FTE)
- HepCloud (ops and dev) (0.5 CMS and 2.5 FIFE FTE)
- Jobsub (batch submission for Fermilab experiments) (1.0 FTE - heavy development past year)
- GlideinWMS development (0.65 CMS and 0.6 FIFE)
- Production support and general triage (1 FTE)
- Fermicloud (0.25 FTE)
- Continuous Integration (0.25 FTE)
- General FIFE projects – Managed tokens/Efficiency reports/Naught User policy (1.0 FTE)
- Elastic Analysis Facility (0.75 CMS and 0.75 FIFE FTE)

Resources Available Onsite

- Fermigrid Cluster
  - Reallocated some old 8 year old machines (~700 cores)
  - Adding 6656 cores now (last years money)
  - 266M core hours available (another 28M still dedicated to Rubin)
  - Need on the order of $300k per year to maintain size of Fermigrid as old nodes are retired
- Wilson Cluster
  - Omnipath for parallel processing
  - 4 x 2 nvidia Tesla V100 GPUs
  - 27 x 4 nvidia Tesla K40m GPUs
  - One power9 + 4 volta gpus (Oakridge Summit)
  - One KNL (NERSC Cori/ALCF Theta)
  - Added four A100 GPUs in one node last year
  - Will be buying LQ2 for exclusive LQCD use (probably 18 nodes with 4 NVIDIA A100s each) may get some of LQ1 for general use but not sure the number of customers for MPI type processing
- Elastic Analysis Facility (EAF)
  - Runs on OKD cluster and can expand there and into farms
  - Added 8 A100 GPUs in two nodes this past year
- We have 500k to purchase GPUs this year. Not sure what that will look like yet.
- https://landscape.fnal.gov/monitor/d/000000183/scd-summary-gpgrid?orgId=1

HPC Center Usage

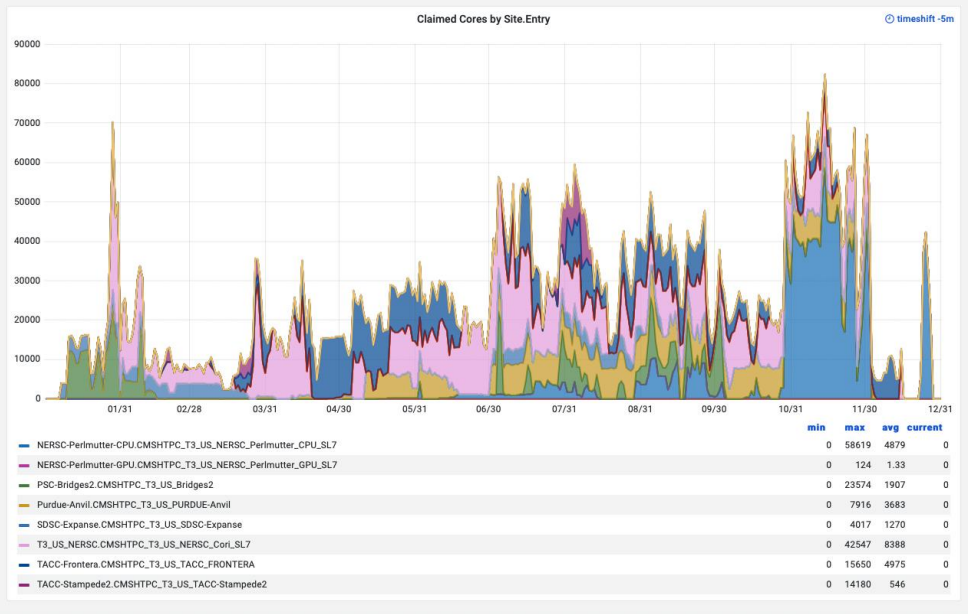
- HepCloud February Stakeholders meeting had experiment reports https://indico.fnal.gov/event/57892/
  - Mu2e – Needs big-core CPUs, no GPU, G4 multithreaded, LBL collaborators can get allocations
  - MicroBooNe – 1.8M cpu hours used
  - Icarus – planning to run some production on HPC resources
  - DUNE – 1.5M wall hours; testing Perlmutter GPUs with realistic payloads
  - G-2 - Done simulation in the past but probably no HPC need for another year
  - CMS – Using more than they've been allocated; GPU workflows by Summer 2023

NERSC Alloc	Total Hours	Actual Hours	Raw Hours	GPU Hours
FIFE	191,000	243,265	843,616	13,764
CMS	606,291	606,291	1,391,578	-
Neutrino (Perlmutter Free)	241	76,261	76,161	-
Total HEP	**797,532**	**925,817**	**2,311,355**	**13,764**

HPC Center Usage

- ACCESS current allocation runs July 2022-June 2023
  - TACC Stampede2 78830 / 75000
  - SDSC Expanse 4965298.0 / 5000000.0
  - PSC Bridges2 8440602.0 / 8440000.0
  - Purdue Anvil 36041377.3 / 36050580.0
- TACC Frontera (not part of XSEDE/ACCESS) April22-Mar23
  - 900306 / 90000
- CMS used about 25k cores DC at NERSC

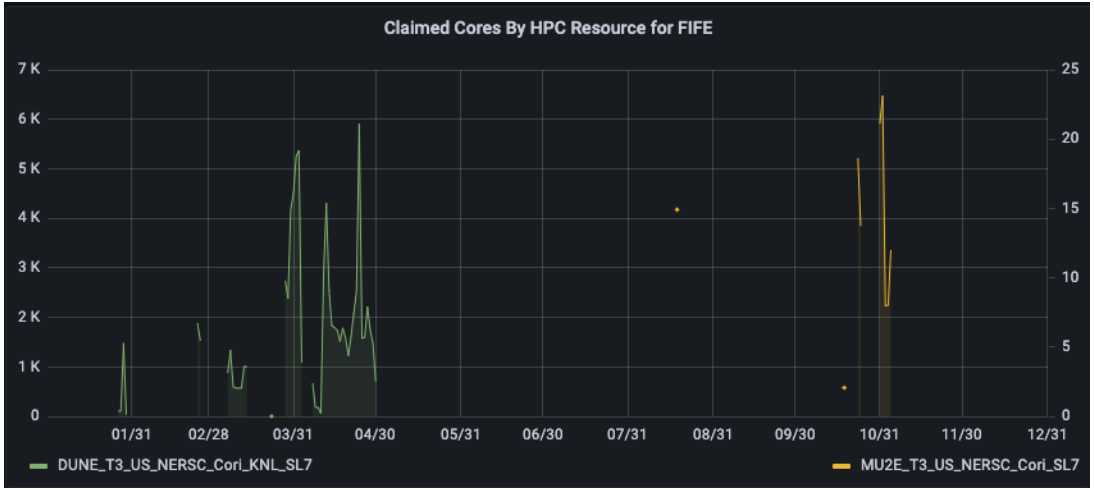
CMS NERSC Usage Constant



Fermilab

HPC Center Usage

- CMS and DUNE both working on NERSC Perlmutter GPU integration
- Plot shows DUNE usage earlier in the year and mu2e in October

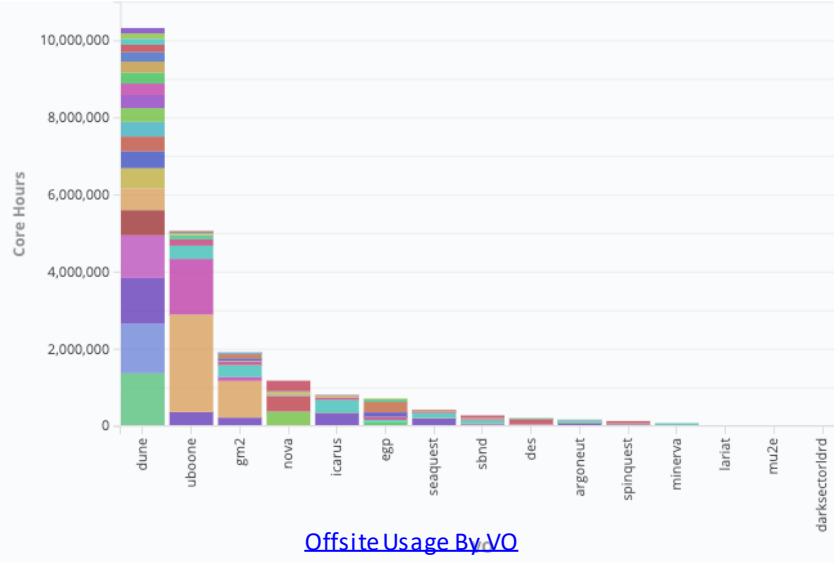


Non-CMS usage a bit more sporadic

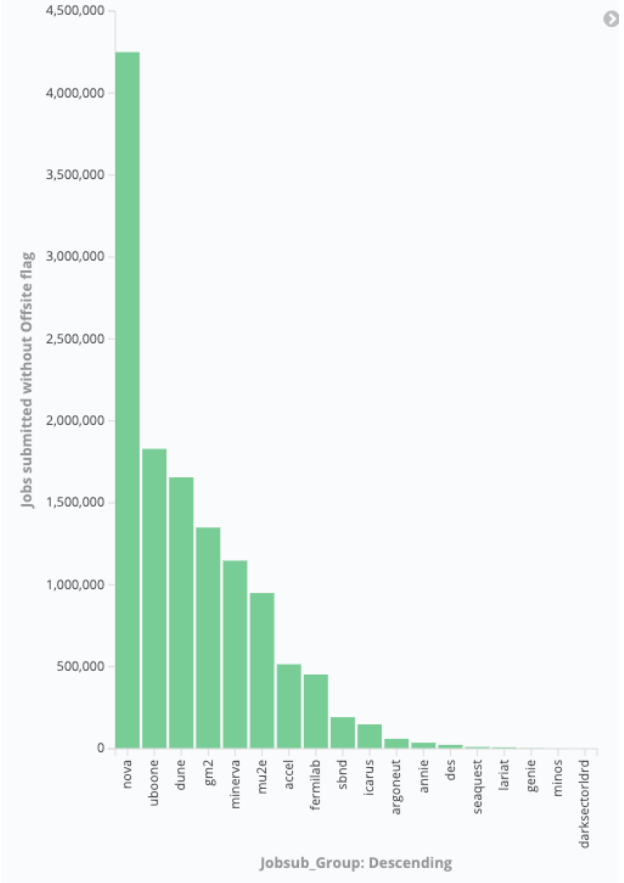
Fermilab

Offsite/Opportunistic Resources

- No significant usage of paid cloud in the last year
- Containers should limit issues at remote sites
- Left plot shows usage of OSG, right plot shows jobs submitted in last 6 months without offsite flag



Offsite Usage By VO

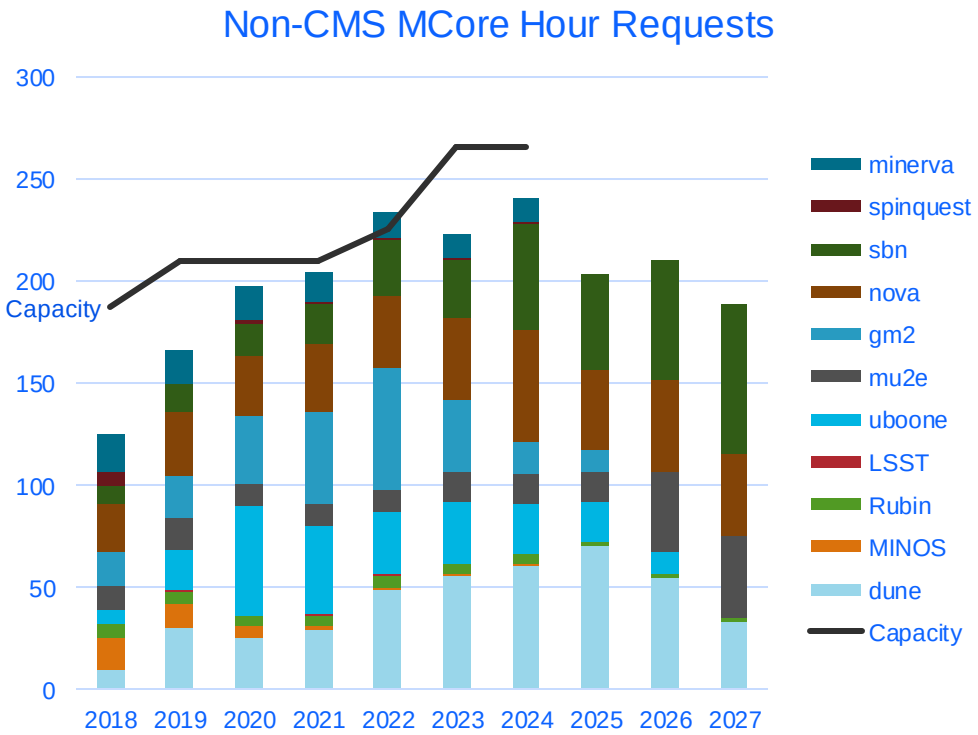


Last 6 months not submitted offsite

Fermilab

Experiment Request Summary

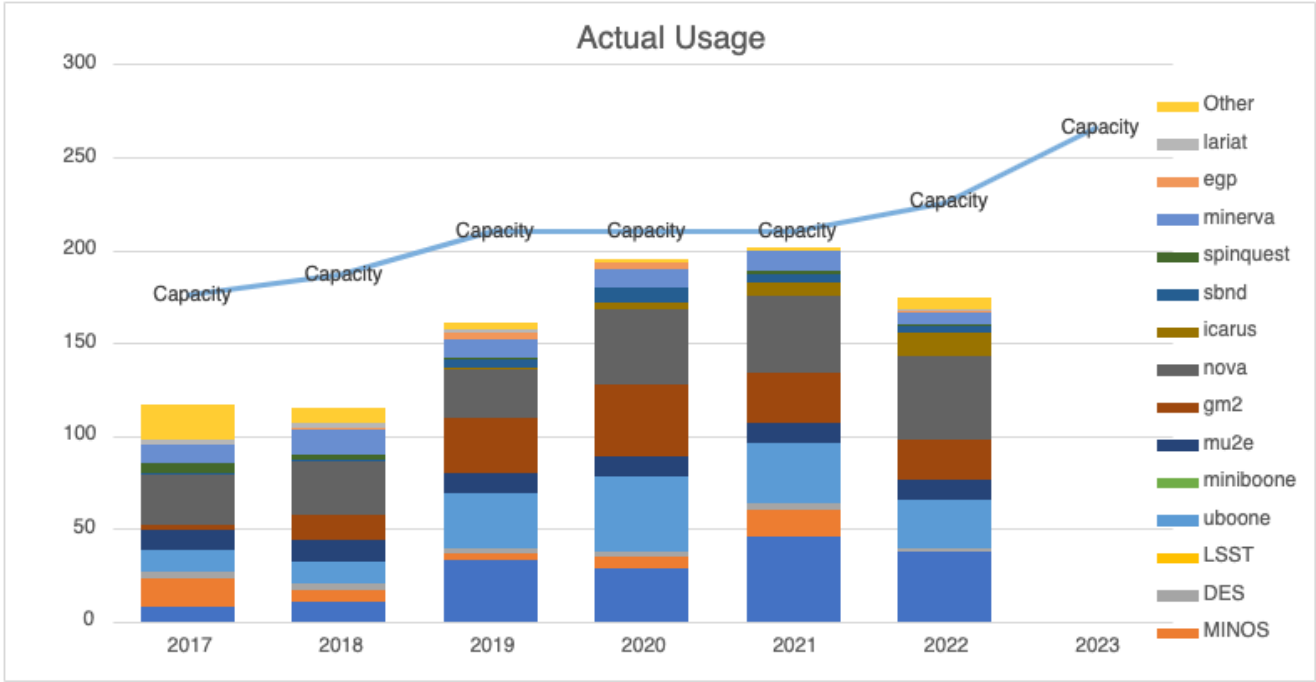
- Requests seem to be leveling off
- SBN experiments and NOVA not changing much in out years
- Mu2e increases
- DUNE decreases after 2025 along with others



Fermilab

Actual Usage

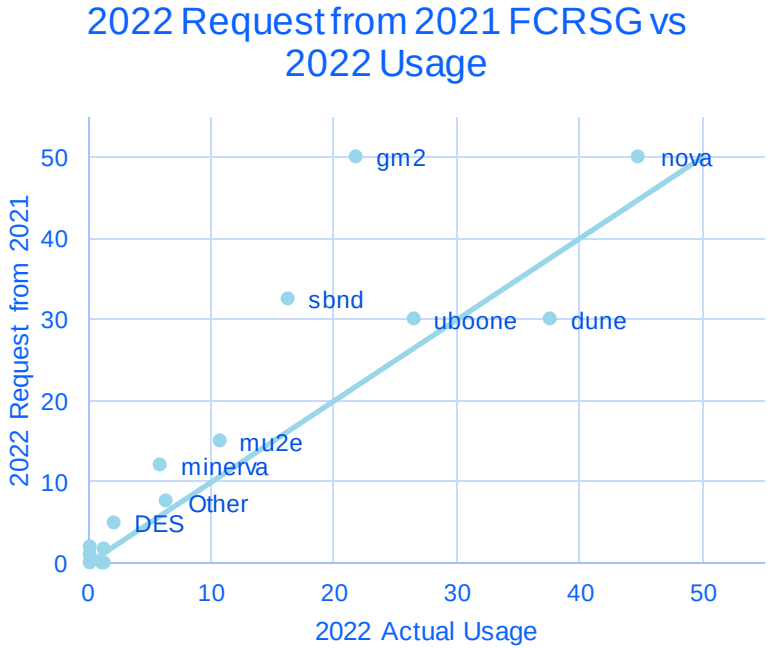
- We lose a large number of machines yearly and these capacity lines for future years do not take that into account

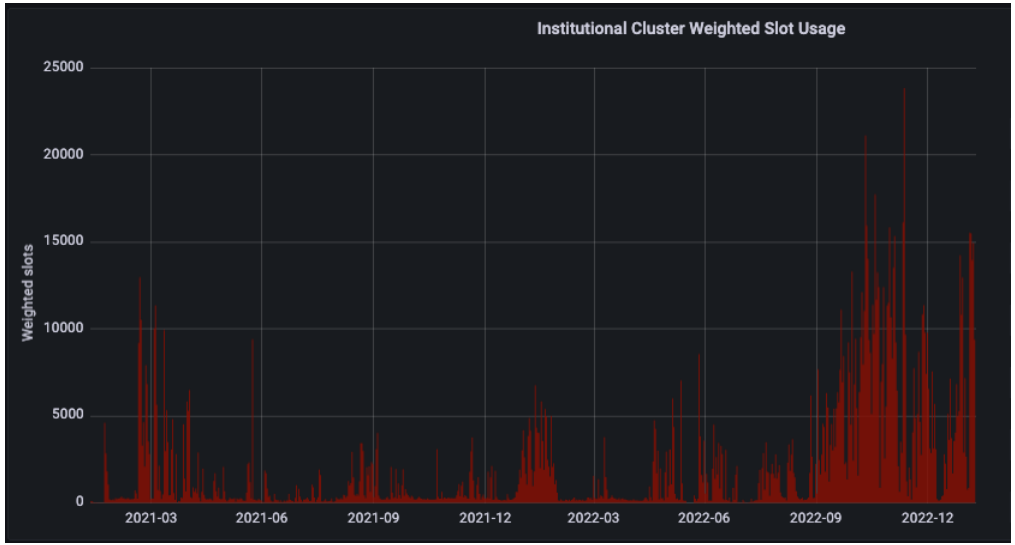


Fermilab

Actual Usage

- Only DUNE had usage higher than their request last year
- OSG ingress into Fermigrid has been very high since September





OSG usage of Fermigrid over last 2 years

Fermilab

Deliverables Coming Up

- Finish the deployment of jobsub rewrite with tokens
- Work with DUNE to deploy schedds at RAL and BNL
- Get POMS moved to github and try to get VOs to contribute development effort (DUNE and mu2e may have interest)

What Experiments Should Be Doing

- Engage with HepCloud team for any work that might be able to run on HPC resources
- Work with experiment collaborators if they have access to HPC allocations
- Submit jobs so they will run everywhere
- Consider which of their workflows may benefit from GPU access
- Consider usage of the Elastic Analysis Facility