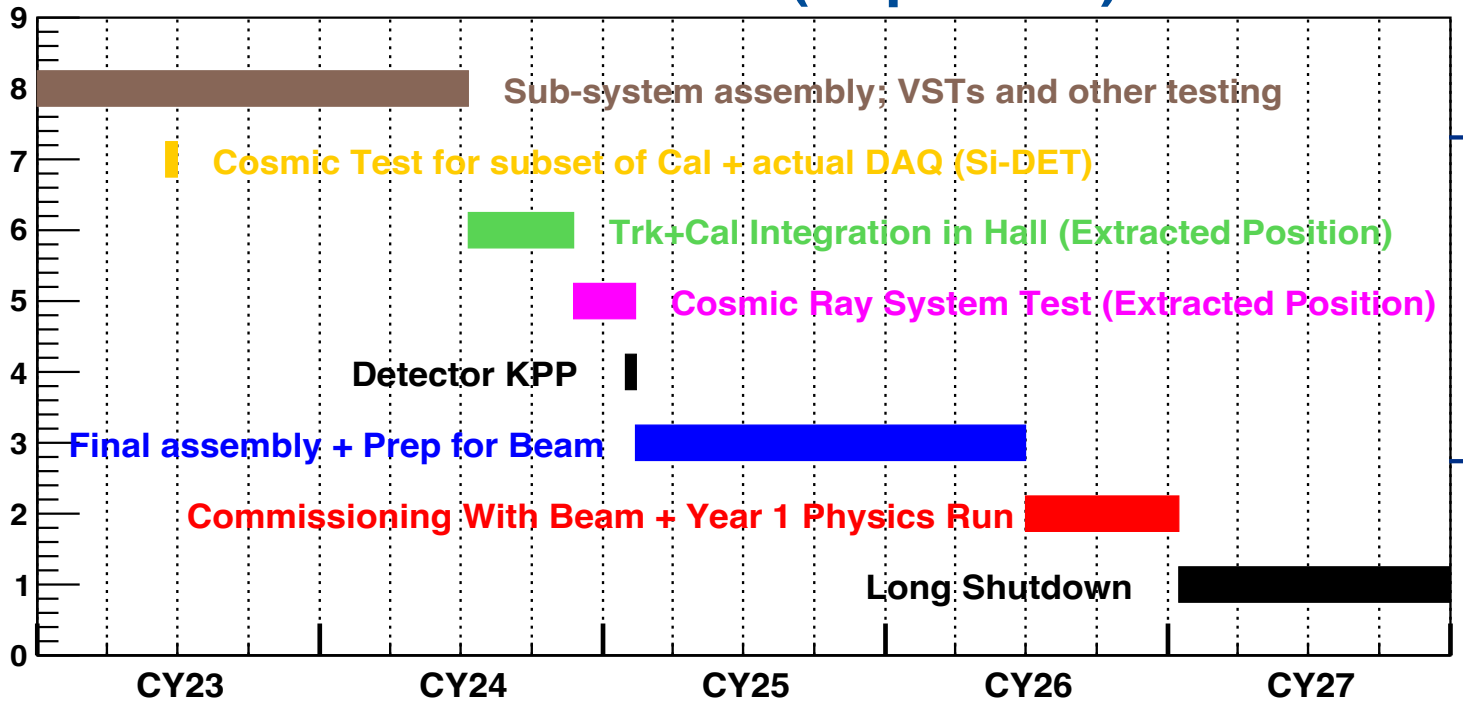
Mu2e FCRSG FY23

Mu2e-doc-44460-v3

Rob Kutschke, Fermilab DSSL

February 15, 2023

Mu2e Rebaselined Schedule (Sept 2022)

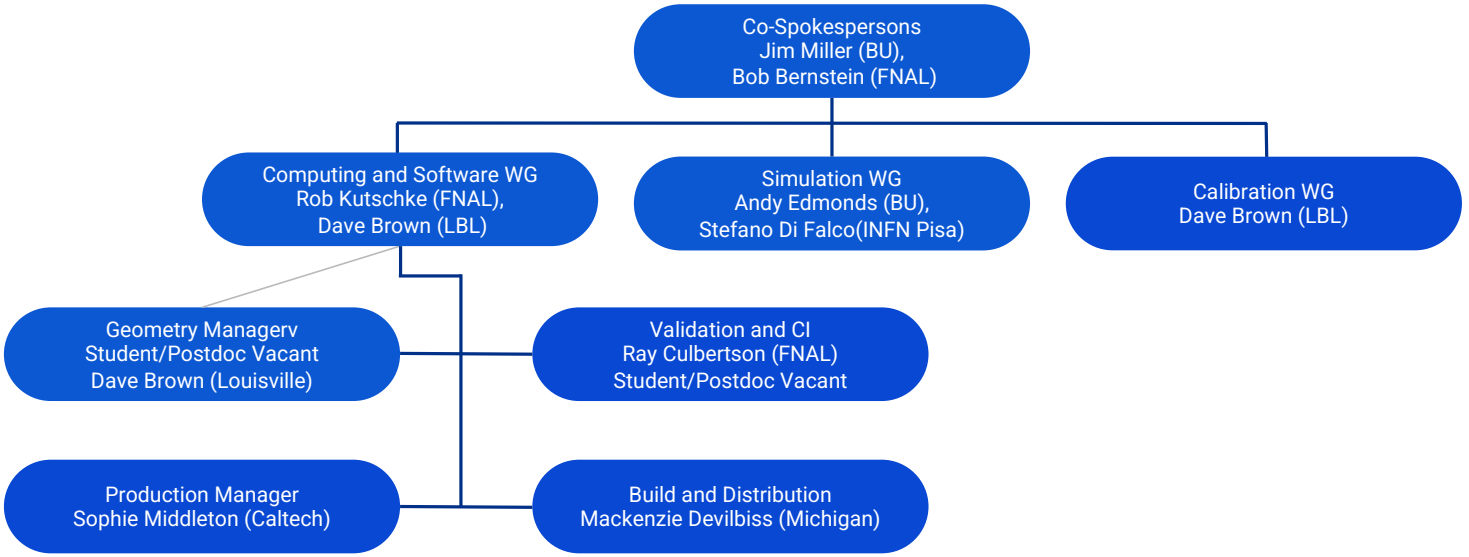


Taking cosmic ray data; sometimes intermittently, sometimes continuously

- Physics goal: enough data in Year 1 for a significant publication.
- Medium step in data rate at start of Trk+Cal Integration in Hall
- High rate data starts at Commissioning with Beam.
- Final assembly: install magnets; map B field; insert detector; install shielding; commission beam to target.

Fermilab

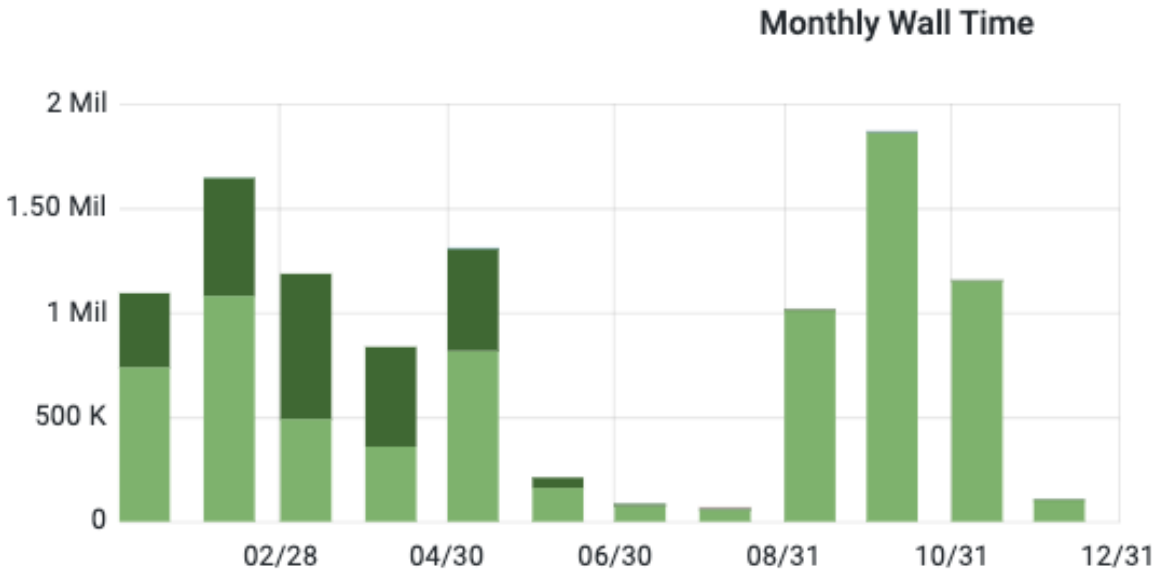
Organization Chart for Mu2e Offline Computing



- Other roles filled transiently as needed
- Above is for pre-operations phase; now planning transition to operations.
  - Ops Manager: Greg Rakness (PPD)
- Algorithm development is the responsibility of the subsystems; some people wear both hats.
- Computing workforce continues to get smaller: people diverted to on-project tasks.
  - Expect workforce increase mid CY24 when detector arrives in the hall.

Fermilab

Mu2e CPU Usage over the past year: Total 10.6 M wall clock hours



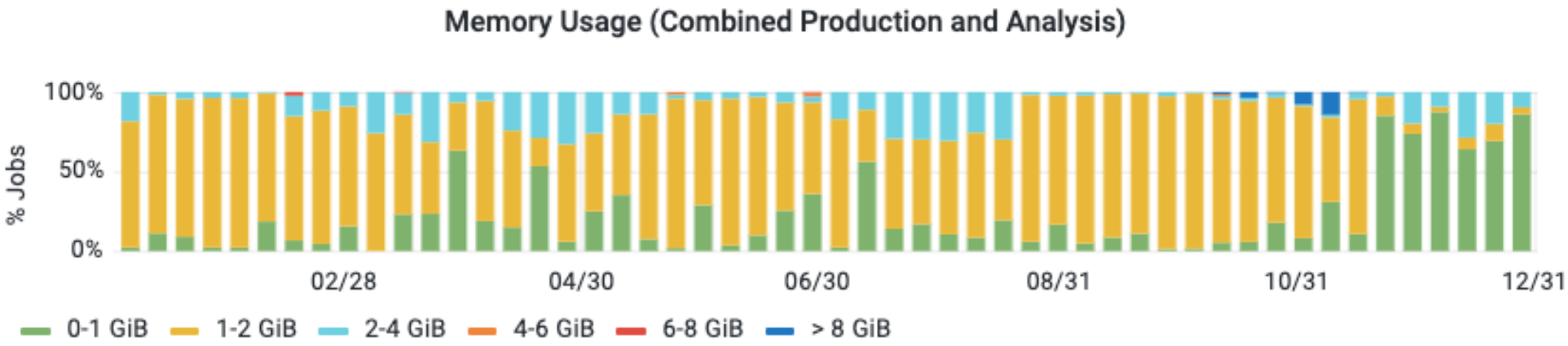
	avg	total
FermiGrid Analysis	607 K	7.89 Mil
FermiGrid Production	205 K	2.67 Mil
OSG Analysis	368	4.79 K
OSG Production	456	5.93 K
NERSC Analysis	0	0
NERSC Production	257	3.34 K

Grand total: 10.6 M

- +14M hours at ANL; using allocation of ANL collaborators; not using HEPCloud.
- Most cycles are G4 which we run in MT mode for memory mitigation.
- Production switched to POMS from in-house scripts
  - Gave up on OSG: few cycles; frequent failures requiring hand work to complete campaigns.
  - Started onboarding to NERSC last fall – slow going – long pause for holiday break – restarted recently.
- Mostly simulation campaigns so most deadlines are elastic.

Fermilab

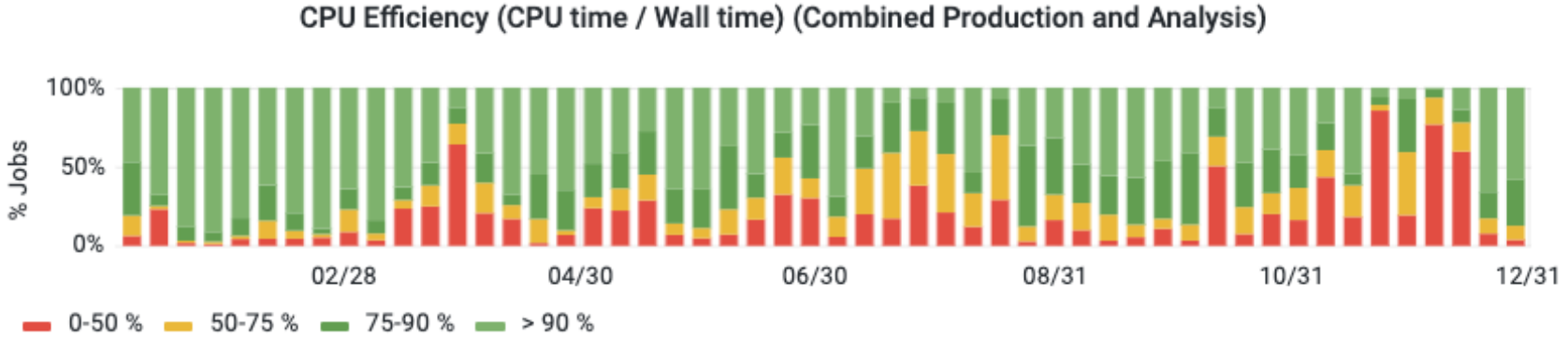
Mu2e Memory Footprint over the Past Year

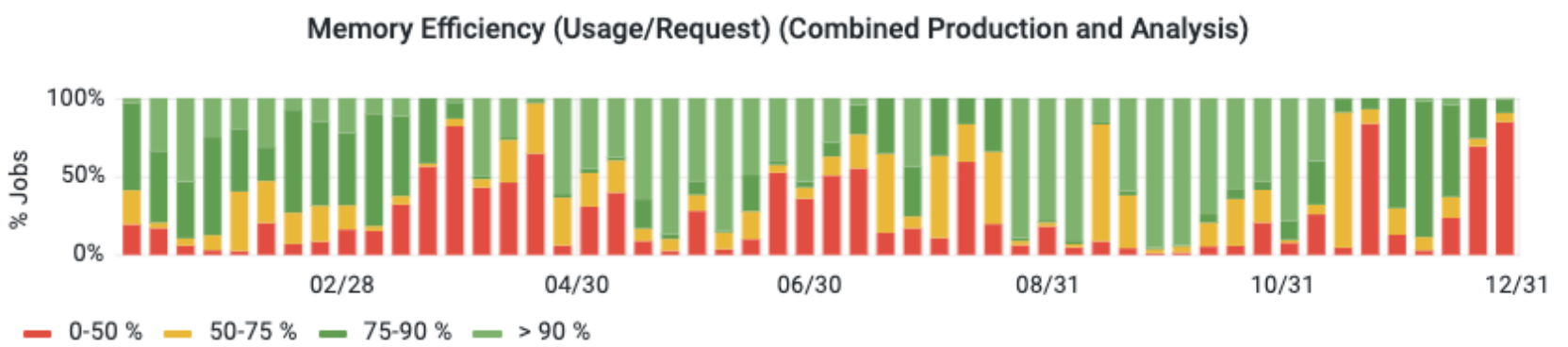


- Similar to last year
  - G4 jobs typically use 2 slots and 2 cores using ~3 GB (light blue).
  - Trigger, reco, analysis jobs typically use < 2 GB
  - High memory jobs are maximum excursion pileup and high energy tail of cosmic rays.
- Expect stable memory footprint in coming years
  - Unknown: big conditions data when we encounter experimental data?

Fermilab

Mu2e CPU and Memory Efficiency over the Past Year





- Both similar to last year; many periods of low efficiency occur when overall usage is low.
  - Some low efficiency is due to slow dcache response
  - Year averaged on-site CPU efficiency is 92% for 2022
- Solution: improved user education and monitoring: can we get an automated monthly report?

Fermilab

What do you want to achieve in computing over the next 5 years?

Goals	Where does the experiment need to contribute	Where does SCAI need to contribute
Big picture: Be ready to take and process cosmic ray data in July 2024 and beam data in July 2026. Goal to publish within a few years after end of Y1 data taking.	See below for the components of this goal	
Commission DAQ, Trigger, slow control	Trigger algorithms; detector control system; online DQM; configuration of the overall system.	Ryan Rivera's group drives the work on the systems that will be in the Mu2e Hall.
Update model of data sizes, processing times, calibration+reco workflow	Most of the work	Consulting about relative merits of some tradeoffs.
Demonstrate ability to move data from DAQ disks to long term storage at required rate, using synthetic load.	Design and execute tests. Complete by summer 2023.	RAID configuration waiting on AlmaLinux availablity. Consulting.
Develop data-logging workflow	Requirements, joint effort on design, implementation	Consulting. Add features to SCAI tools as needed.

(continued on next page)

What do you want to achieve in computing over the next 5 years?

Goals	Where does the experiment need to contribute	Where does SCAI need to contribute
Transition from SAM to to RUCIO / Metacat / Data Dispatcher	The main body of the work.	We will need lots of guidance and support. We may request new features.
Develop rapid, reliable workflows, for processing experimental data; includes support for analysis workflows. Use lessons learned and tools developed to improve simulation and general user workflows.	Requirements, joint effort on design, implementation	We may request new features in POMS and related tools
Offline DQM System	Requirements, design, implementation	DB support and support for visualization tools hosted on Mu2e website.
EVE7 Event display	Requirements, design, evaluate options, implementation	If SCAI is offering to support event displays, we are interested.
Reconstruction, calibration, alignment, DQM algorithms	Most of the work.	Consulting on design issues, best practices, better use of profiling tools, code reviews etc
Create MOU/TSW with SCAI & ITD	Joint effort.	Joint effort.
New user onboarding	Most of the work	

Campaign Schedules

	2023	2024	2025	2026	2027
NERSC	2 Mhours (CPU only)				

- Now to fall 2024:
  - Sim: cosmic ray study to be done at NERSC ASAP ~2 Mh core hours
  - Processing of experimental data will be intermittent and use few resources
- Fall 2024 to July 2026
  - Frequent but intermittent cosmic ray runs for commissioning. Small data volume (<< 1PB).
  - Continuous reprocessing throughout the entire period in order to bootstrap calibration/alignment
- July 2026 to Jan 2027
  - Commissioning with beam and year 1 run. Big step in tape and CPU needs. See tables.
- 2027
  - Processing, calibration, reprocessing, … analysis

CPU @ Fermilab Prediction Going Forward and Accuracy of Your Predictions [units of Million (1 CPU, 2GB) wall hours per CY]

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	15	11	11	11	15	15	15	40	40
Actual Used	11	10.9	10.5	10.6	N/A	N/A	N/A	N/A	N/A
Efficiency	86%	90%	81%	92%	N/A	N/A	N/A	N/A	N/A

- Numbers are the total of all HTC needs (Fermigrid + OSG)
  - Much could be run on OSG but we have only opportunistic cycles.
  - In recent years few opportunistic cycles have been available.

CPU – non-FNAL HTC Resources Going Forward and Accuracy of Your Predictions [units of Million (1 CPU, 2GB) wall hours per CY]

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	Opportunistic use only; few cycles available and high failure rates.								
Actual Used					N/A	N/A	N/A	N/A	N/A
Efficiency	%	%	%	%	N/A	N/A	N/A	N/A	N/A

- We have no dedicated off-site OSG cycles
- In recent years there have been few opportunistic cycles
  - Failure rates many times higher than on Fermigrid.
- We will submit to OSG but cannot predict what we will actually get so we do not show anything here. See previous page for our request for Fermgrid+OSG

CPU – HPC Resources Going Forward and Accuracy of Your Predictions [units of Million (1 CPU, 2GB) wall hours per CY]

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	Not used				5	5	5	5	5
Actual Used					N/A	N/A	N/A	N/A	N/A
Efficiency	%	%	%	%	N/A	N/A	N/A	N/A	N/A

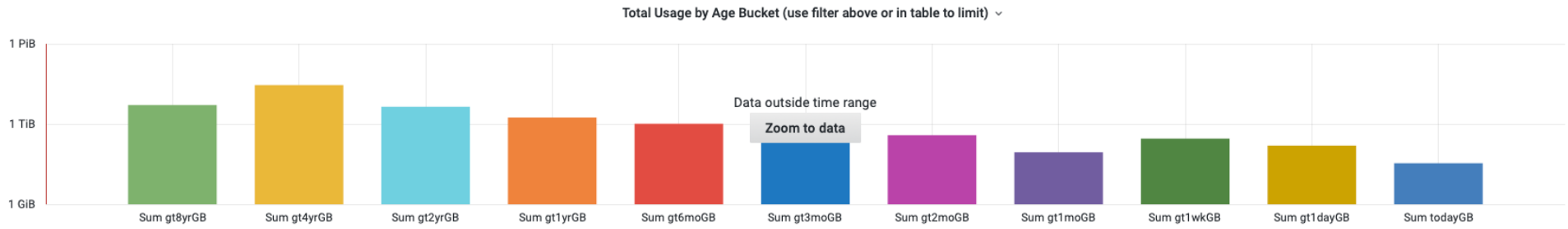
- Placeholder for one medium sized G4 campaign each year
  - No definite plans at this time.
- We also have access to 14 Mhours/year at on Theta at ANL.
  - So far no plans to use this in 2023.

CPU – GPU Resources Going Forward and Accuracy of Your Predictions [units of Million (1 CPU, 2GB) wall hours per CY]

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested									
Actual Used					N/A	N/A	N/A	N/A	N/A
Efficiency	%	%	%	%	N/A	N/A	N/A	N/A	N/A

- Mu2e currently has no significant workflows that can use GPUs
- Given our expected staffing levels and competing priorities we do not expect to have the resources to investigate GPUs until the long shutdown (2027).

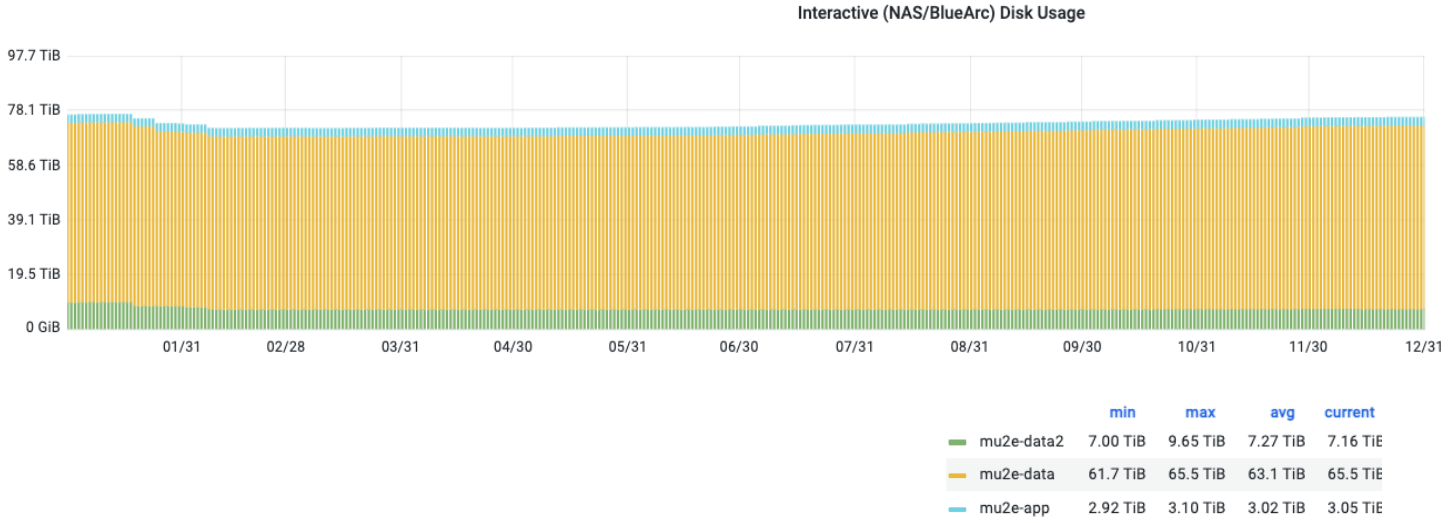
Age of files in NAS



- On Jan 1, 2023 Mu2e NAS data disks were close to full and 85% of space was > 4 years old
  - Cleanup underway
  - Now ~50% free
- NAS app disk was close to full and had 35% of space > 4 years old
  - Cleanup is underway.
  - Now ~40% free

Fermilab

NAS Usage and Projections

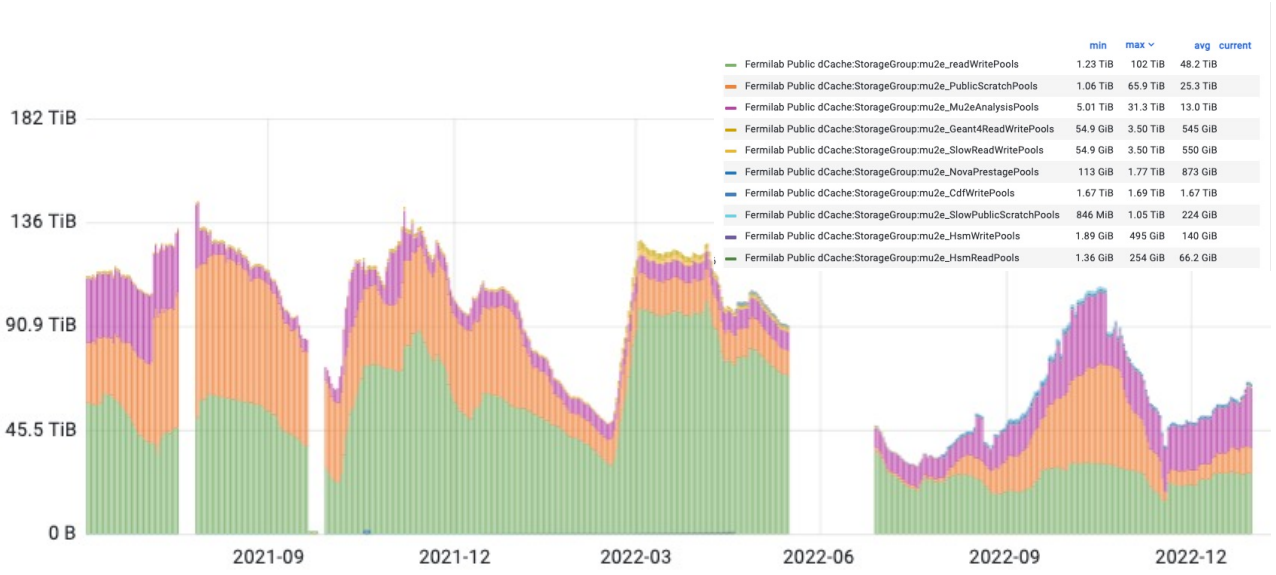


	App (TB)	Data (TB)
2022	3	80
2023	3	80
2024	4	80
2025	4	80
2026	5	80
2027	5	80

- Currently removing stale files from NAS
  - Goal: > 70% free space by end of Februrary.
- Most people use /pnfs not NAS
  - /mu2e/data is not grid visible.
- Need to do homework to respond to NAS/Ceph question.

Fermilab

dCache Usage and Predictions (in TB)

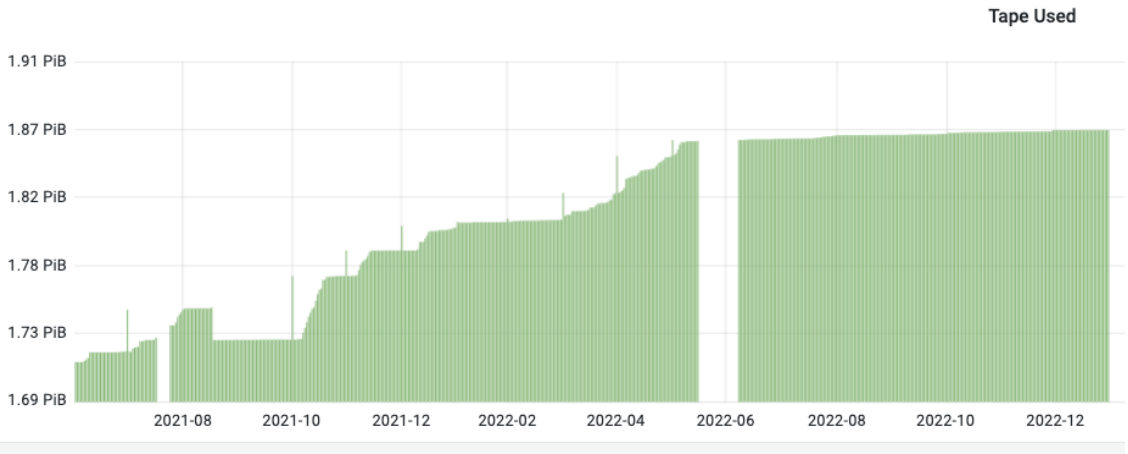


Mu2e has no significant general purpose non-FNAL storage elements

	Analysis (Persistent)	Other Dedicated (Write)
Current	TB (actual)	TB (actual)
2022	150	
2023	150	
2024	150	20
2025	300	20
2026	500	60
2027	500	60

Fermilab

Tape usage and predictions (in TB)



- Actual use only ~ 0.75 PB.
  - Above figure includes files not migrated T10K to LTO8.
  - In the process of freeing another ~0.1+ PB
- After we design the reco and calibration workflows we will revisit the file family structure.

	Total Added By End of Year
At end 2021	+TB (actual)
2022	500
2023	500
2024	1000
2025	2000
2026	15,000
2027	5,000

Fermilab

Data Lifetimes

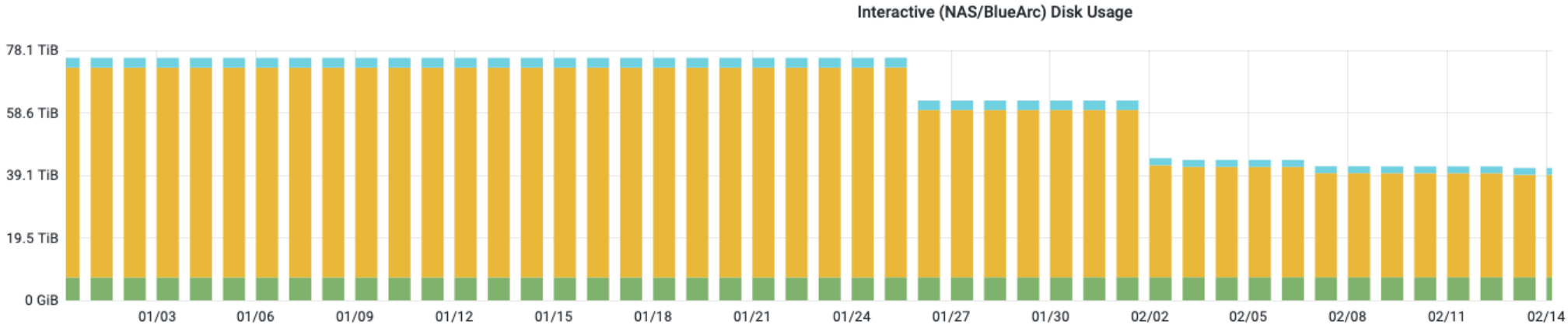
- Simulated datasets:
  - Retire when successor is established; estimate typically 3-4 years.
  - Some specific studies have no successor.
- Strawman plan for experimental data
  - Raw data: keep until N years after last paper; N to be negotiated.
  - Derived datasets: keep current generation plus previous 2.
  - Final analysis datasets for each major paper
    - These will be small and we will keep them all.
- Disk space
  - Major cleanup nearing completion
  - Plan to audit for obsolete files a once per year

Analysis Facility Use

- Mu2e does little analysis at this time
  - Analysis of simulated events
  - Growing analysis of VST and test stand data.
  - Dominantly root but python/Jupyter slowly increasing
    - Transition to python is organic – not a planned analysis model.
- A few people have looked at the Elastic Analysis Facility. Their impressions:
  - Documentation is thin; learning curve is steep
  - An inconvenience: the /mu2e disks are not mounted on EAF
  - They can get their work done well enough on mu2egpvm and their laptops
  - So they gave up.
- Next opportunity may be a group that is moving from TMVA to python based ML.
  - Hit classification and track quality classification
  - Work has just started; LBL based and may use LBL facilities

Backup Slides

NAS Usage YTD



	min	max	avg	current
mu2e-data2	7.16 TiB	7.22 TiB	7.19 TiB	7.22 TiB
mu2e-data	32 TiB	65.5 TiB	54.4 TiB	32 TiB
mu2e-app	2.15 TiB	3.05 TiB	2.79 TiB	2.15 TiB

Fermilab