Status & Future Computing for Intensity Frontier Experiments

Adam Lyon (FNAL/SCD)

Potential Fermilab Muon Campus & Storage Ring Experiments Workshop

May 27, 2021

Drivers (from February 2021 ICAC meeting – Liz S-K)

The main drivers behind future computing plans at Fermilab are

- CMS in the High Luminosity LHC era
- DUNE
- US Department of Energy investments in Exascale High Performance Computing (HPC)
- Artificial Intelligence/Machine Learning development

Main (evolving) paradigms

- Evolving storage systems (Data Lakes, Expanded Quality of Service [QOS] tiers)
- Use of HPC systems (storage at Fermilab, compute elsewhere)
- Dedicated computing and storage systems optimized for analysis at Fermilab
- AI/ML workflows and using AI/ML to optimize computing and storage

Caveats

It is unclear (at least to me) how the evolution described here would affect small experiments starting before 2030

You may come along for the ride for facility improvements and new systems

You may have no choice but to change paradigms (e.g. if more compute happens at HPC centers)

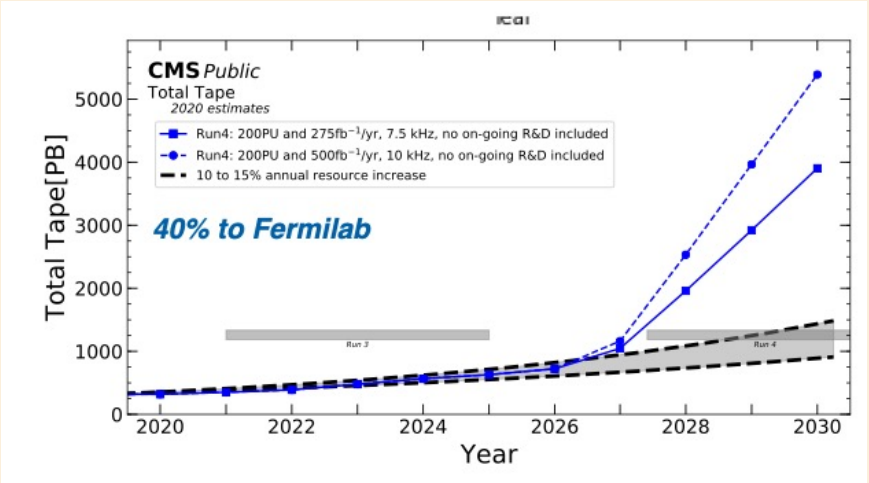
Short and small experiments may(?) be able to hang on to legacy systems and software

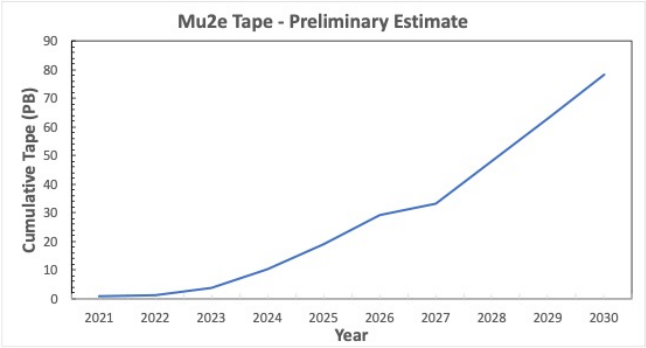
An extended Muon g-2 may be different than a new experiment

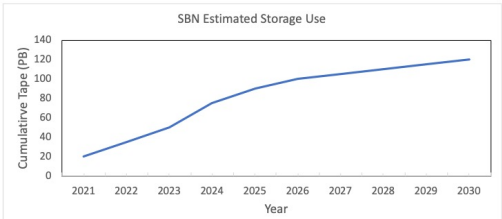
Storage needs (Current volume is ~ 300 PB)

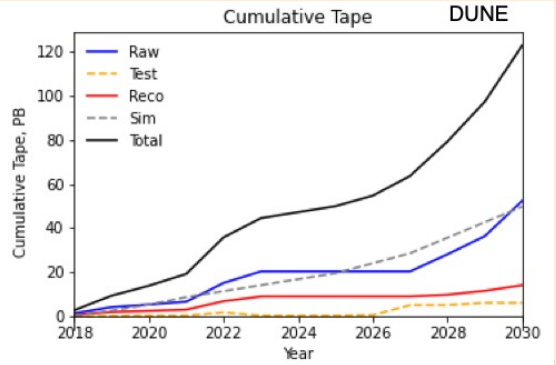
Experiment	Approx. 2030 Data Volume
CMS	1600 PB
DUNE	120 PB
SBN	120 PB
Mu2e	80 PB
Muon g-2	40 PB*

*end of Run 6









Storage at Fermilab

Fermilab wants to remain the center for data storage for experiments

- Hard to compete with HPC for computing
- HPC centers are less interested in huge scale distributed storage

Current storage architecture is an evolution of the Fermilab Tevatron Run 2 storage system

You know it as SAM

Two-tiered Quality of Service: Disk (kinda fast) & Tape (slow)

Very difficult to eliminate tape use … disk/memory more $$

Having more QOS tiers defined by capability would be helpful

e.g. latency, i/o speed, resilience, parallelism

Evolution of storage (ICAC/Bo Jayatilaka)

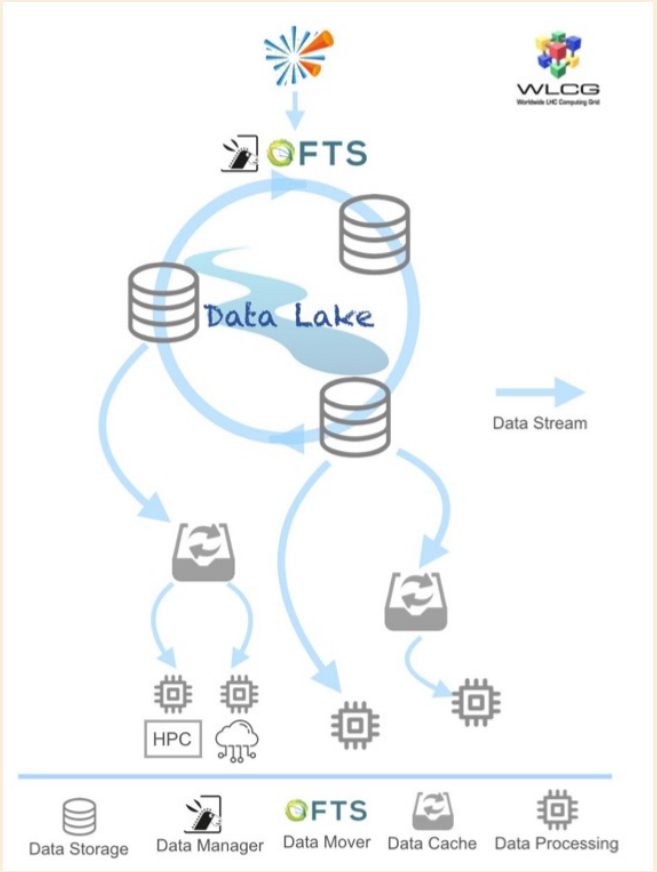
Currently, dedicated sites have associated storage

One accesses storage at that site with knowledge of site's details (e.g. endpoint, structure, limitations)

Future: Data Lakes

A storage system that looks like one site (one endpoint), but is geographically distributed among several storage centers

A compute site may have no persistent storage except for cache



Fermilab

Rucio (Robert Illingworth)

SCD would like to replace SAM

- It's based on 20 year old ideas and it's hard to scale. Difficult to deal with distributed data

RUCIO

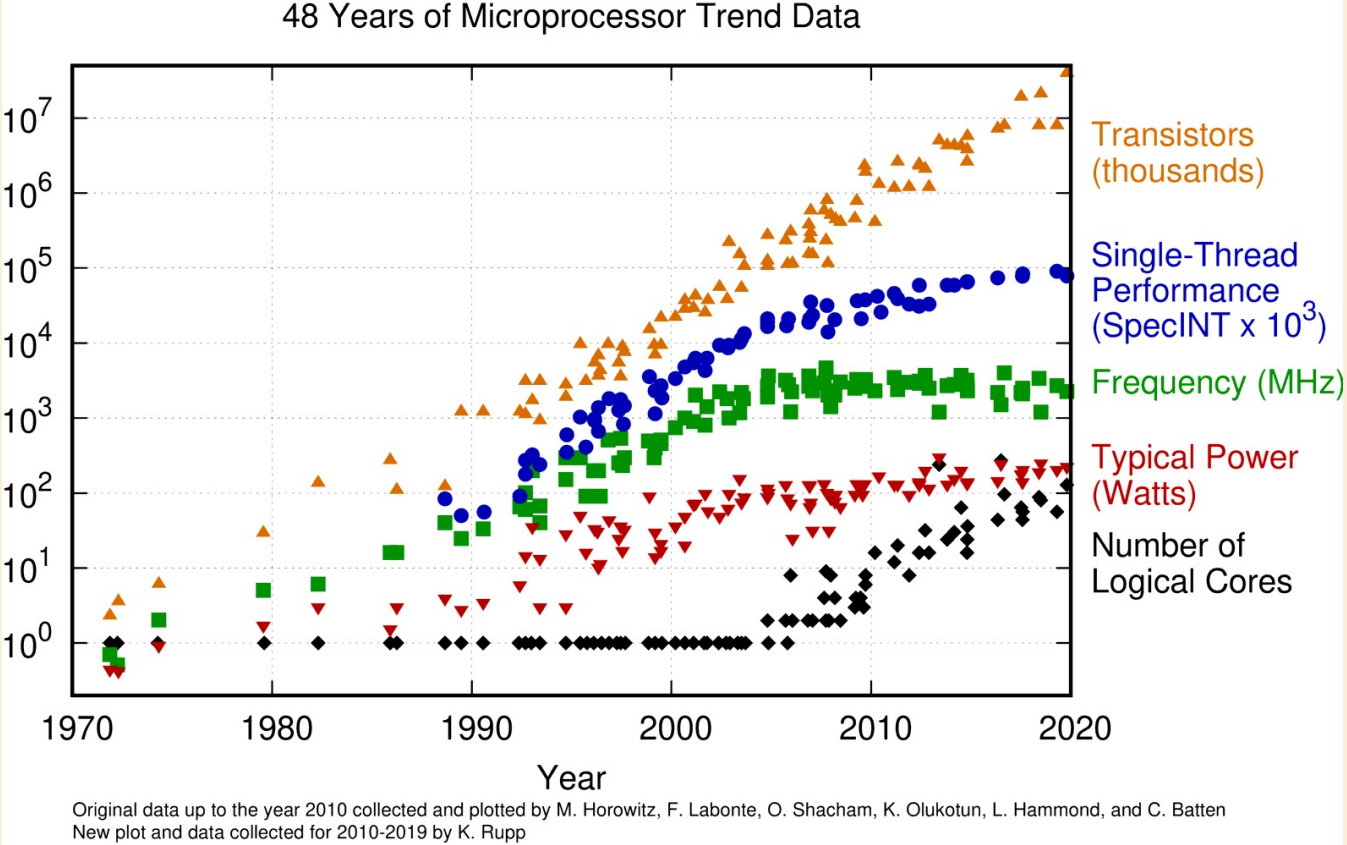
- Organized by CERN
- "manages multi-location data in a heterogeneous distributed environment"
- A modern scientific data management system
- CMS will likely adopt it; DUNE and protoDUNE are using it now
- Still File-based
- Quotas and lifecycle management
- No metadata catalog
- Working on combining SAM metadata system with RUCIO
- Eventually integrate metadata into RUCIO to completely replace SAM

Compute

Can cram more transistors on a chip (density doubles every two years – Moore's law)

But can't make them faster

Instead of making faster compute cores, give you more cores



Fermilab

Compute

DOE investing $2B in Exascale HPC – coming online soon
Scale up computing w/o scaling up power (too much)

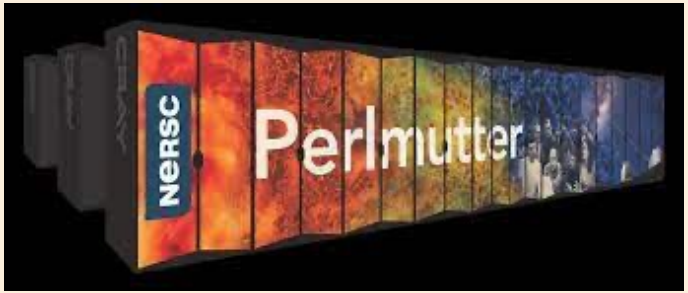
Each machine will be a combination of
CPUs and **GPUs** (NVIDIA, Intel, AMD) [no KNL]

GPUs aimed at AI/ML training and algorithm speedup

Several projects at Fermilab looking at "portable" algorithms (CPU & variants of GPUs)

Dedicated AI/ML machines and clouds (e.g. FPGA cloud)

Evolution of the Grid in this era is unclear to me

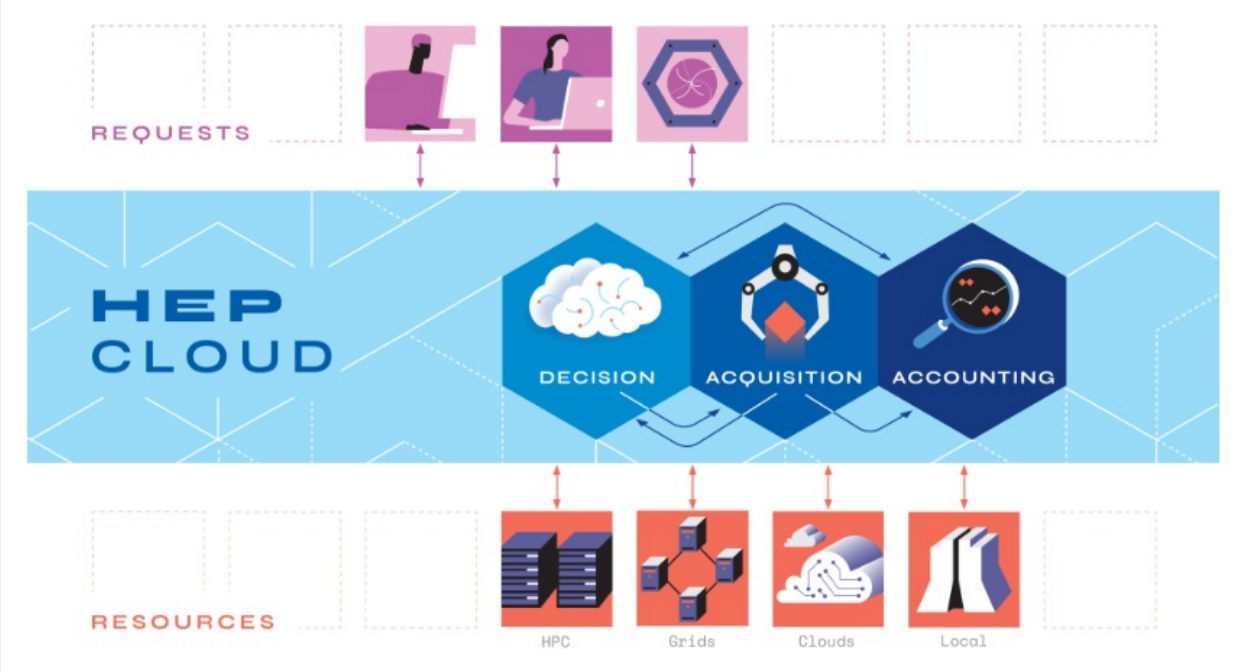


Fermilab

HEPCloud

- Already in progress
- Muon g-2 is using for simulation on HPC

- Common submission facility for sending jobs to many different types of resources



- Ultimate goal is to have the "Decision Engine" choose the most suitable and cost-effective resource for your workflow

Using HPC effectively – SciDAC & CCE projects

Scientific Discovery through Advanced Computing

DOE funded (OHEP and ASCR) Computing R&D projects

- Using HPC environment management (Spack) – ICARUS
- Speeding up algorithms with computing accelerators (GPU/KNL) for current experiments
- Portable algorithms (work on many computing backends)
- Optimized storage/Object storage
- HEP algorithms with Quantum Computing

Elastic Analysis Facility

Current Grid computing is not optimized for analysis workflows

- Traditional disk is too slow for interactive-like experience
- Batch system is not ideal for analysis work
- Not optimized for column-wise analyses (evolution of python frameworks)
- Not compatible with Jupyter notebooks

Fermilab is working on an Elastic Analysis Facility

- Containerized infrastructure and Jupyterhub deployment
- Optimized for row-wise and column-wise analyses
- Very fast low latency NVMe storage
- Mostly driven by CMS, but will be compatible with other experiments

SCD is close to a prototype!

AI/ML

Large efforts in Artificial Intelligence & Machine Learning

- for HEP experiments (classification, analysis, ...)
- for Astrophysics/Cosmology (weak lensing, galaxy classification, telescope operations, ...)
- for accelerator operations and tuning

High velocity and real time AI/ML

Use of hardware co-processors, GPU, TPU, FPGA

Other

Framework evolution

- Towards a common framework
- Easy GPU & parallelism
- DUNE has a requirement document

- Non-C++ analysis frameworks are rising (CMS coffea)

DAQ

- Fermilab continues to develop Off-the-shelf DAQ & artDAQ

Pushing away from Redmine

- GitHub?

Conclusions

- Computing is evolving slowly but surely
- Driven by CMS and DUNE, but small experiments will benefit
- Many DOE investments and R&D programs
- Questions:
  - How long can an extended experiment hang on to legacy infrastructure (will it want to)?
  - What choices will a new experiment have to make?