Fermilab U.S. DEPARTMENT OF ENERGY Office of Science

HEPCloud – Computing Facility Evolution for HEP

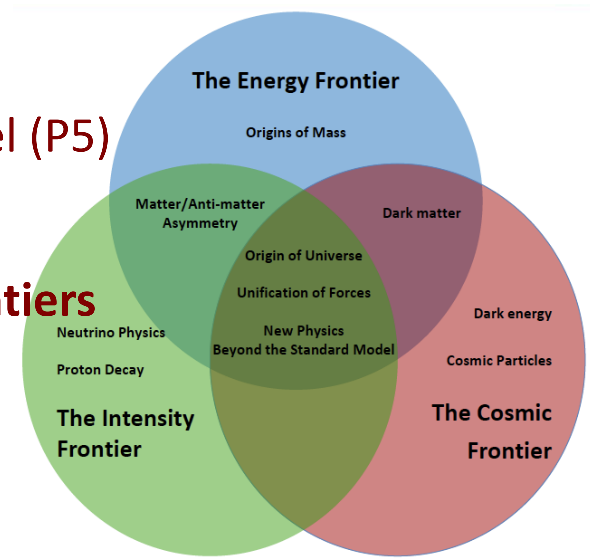
Panagiotis Spentzouris, Fermilab

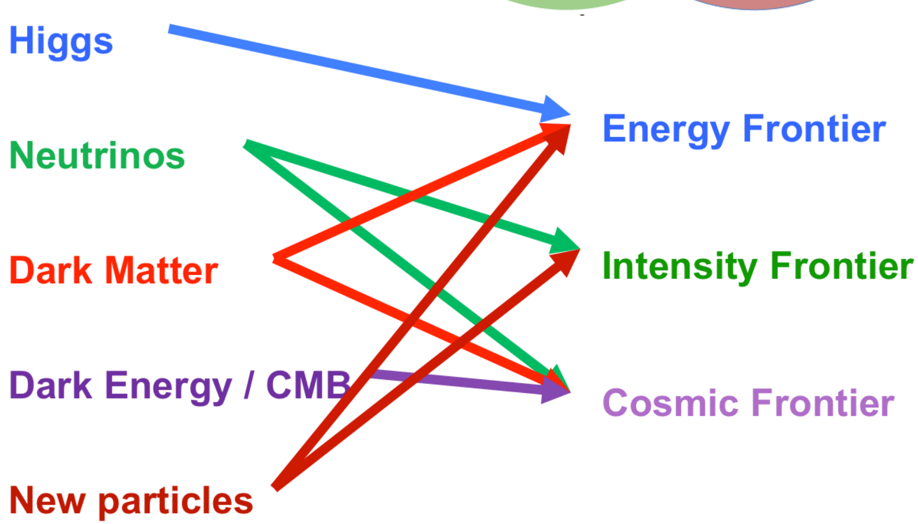
SC2015

17 November, 2015

High Energy Physics Science Drivers

The 2014 Particle Physics Project Prioritization Panel (P5) report identified
5 Science Drivers
Mapped to 3 Frontiers

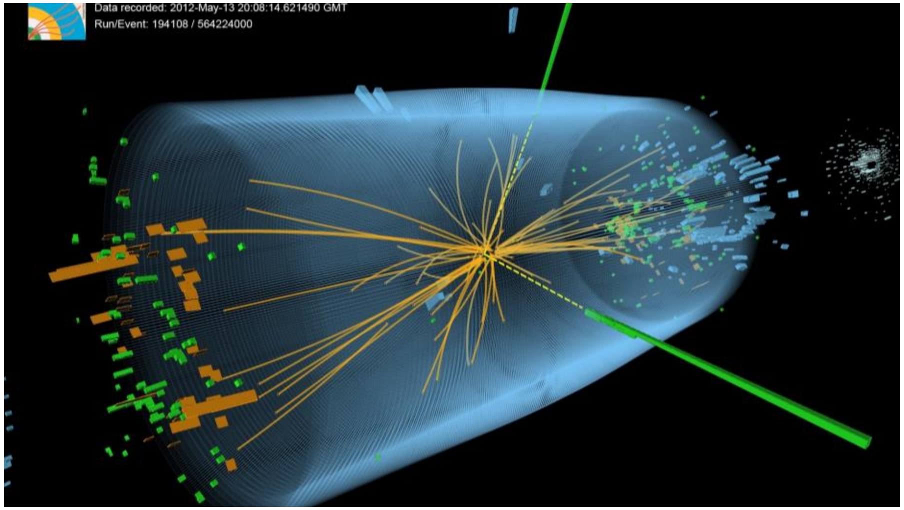


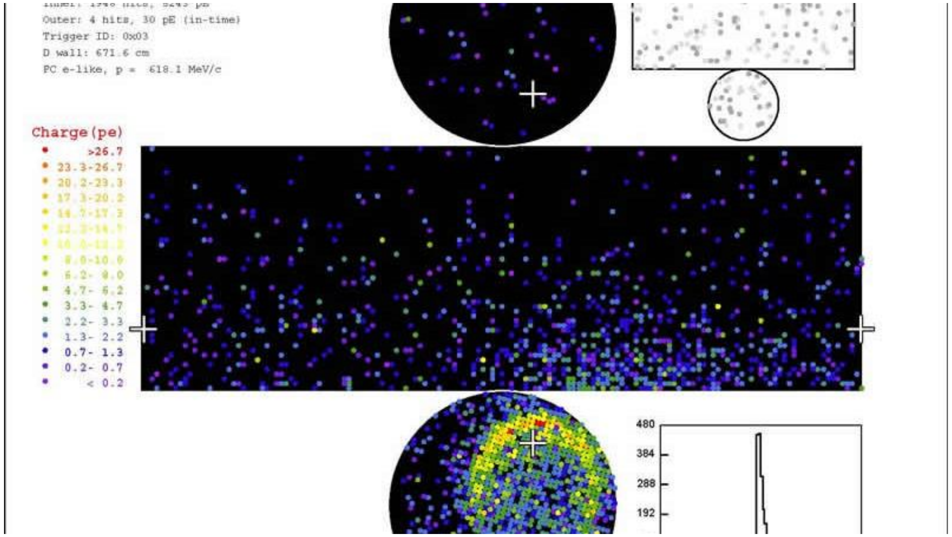


- At the **Energy Frontier**, high-energy particle beam collisions seek to uncover new phenomena
  - the origin of mass, the nature of dark matter, extra dimensions of space.
- At the **Intensity Frontier**, high-flux beams enable exploration of
  - **neutrino interactions**, to answer questions about the origins of the universe, matter-antimatter asymmetry, force unification.
  - **rare processes**, to open a doorway to realms to ultra-high energies, close to the unification scale
- At the **cosmic frontier** we seek to understand the nature of the contents of the universe: ordinary matter, dark mater and dark energy.

Where we are today

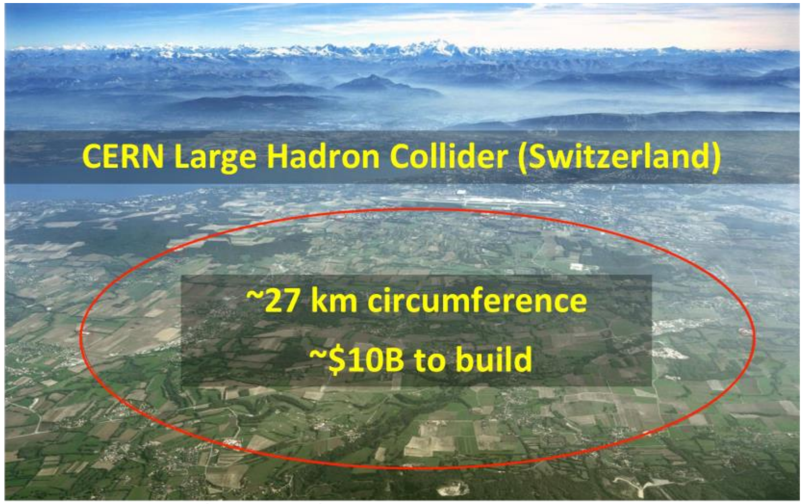
- **Discovery of the Higgs particle at the CERN Large Hadron Collider (LHC), responsible for electroweak symmetry breaking and the mass of elementary particles**
  - No physics beyond the "Standard Model" of HEP has been observed
- **Neutrinos oscillate, thus have mass**
  - No answers on mass hierarchy or symmetry properties
    - Potential explanation for matter anti-matter asymmetry observed in the universe!



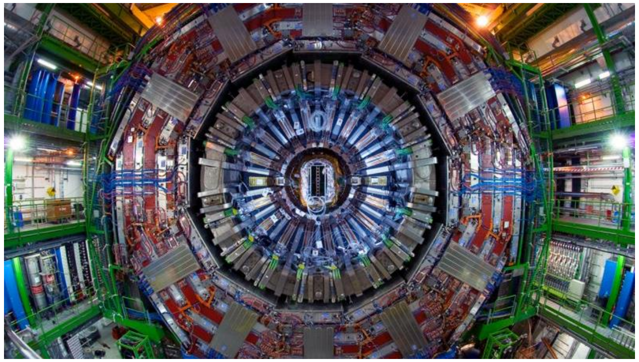


Next Steps (Energy Frontier)

- More **powerful** detectors and accelerators to facilitate discovery
  - Major upgrades for the **LHC,** for higher energy and luminosity, and for the **CMS & ATLAS** detectors
- **Computing** essential for the success of the program, faces many challenges
  - highly distributed environment, thousands of collaborators,
  - as the program evolves, higher data rates will require utilization of new techniques and technologies

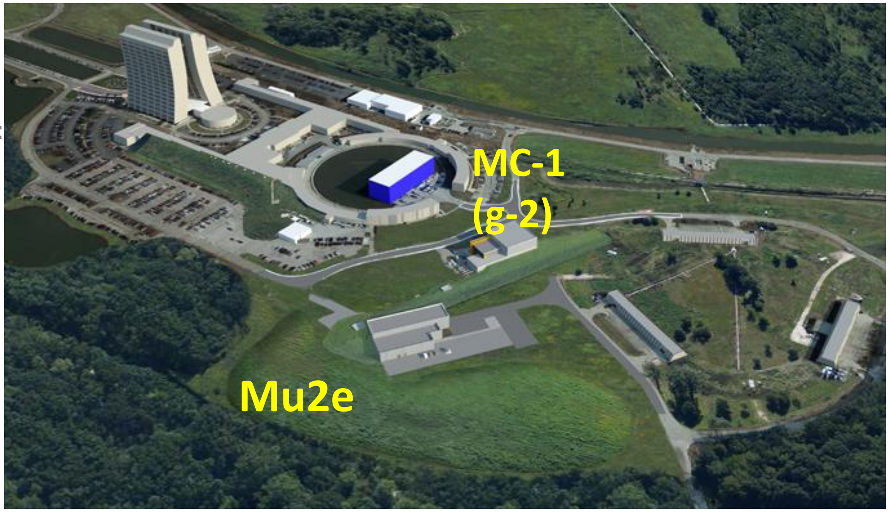


Courtesy S. Myers (IPAC 2012)


CMS detector at the LHC

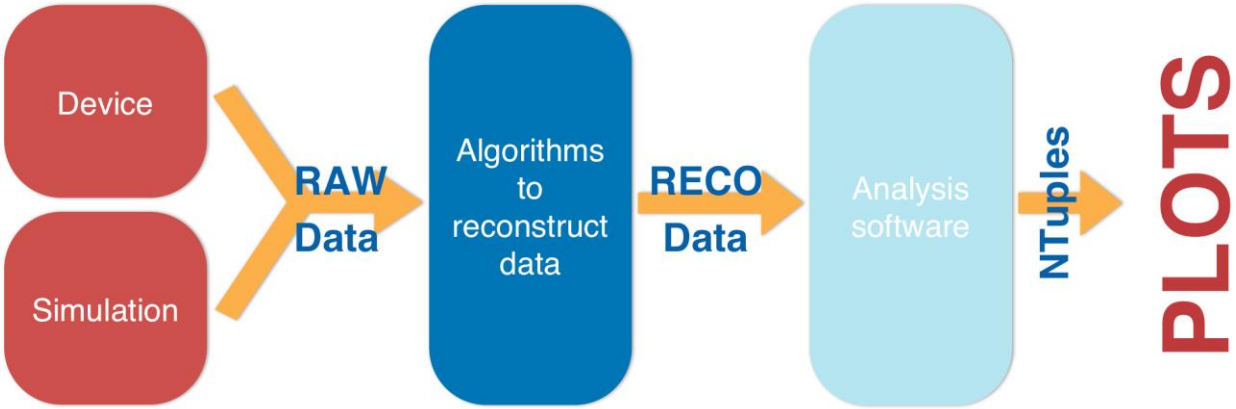
Next Steps (Intensity Frontier)

- A **high-intensity proton** accelerator for
  - **neutrino oscillation** experiments
    - Mass hierarchy, matter-antimatter asymmetry
  - rare process experiments
    - New particles and interactions
- Staged approach at Fermilab
  - Major complex improvements
  - Short and long baseline neutrino oscillation experiments
  - Rare decay experiments
- **Computing** essential for the success of the IF program and its evolution, new challenges to meet
  - Many experiments with different timelines and limited resources that need a fully supported computing ecosystem



HEP computing model: distributed facilities

- Most HEP experiments utilize distributed High Throughput Computing (HTC), with global workflow, scheduling, and data management, enabled by high-performance networks
- Every stage of the experiment requires massive computing resources

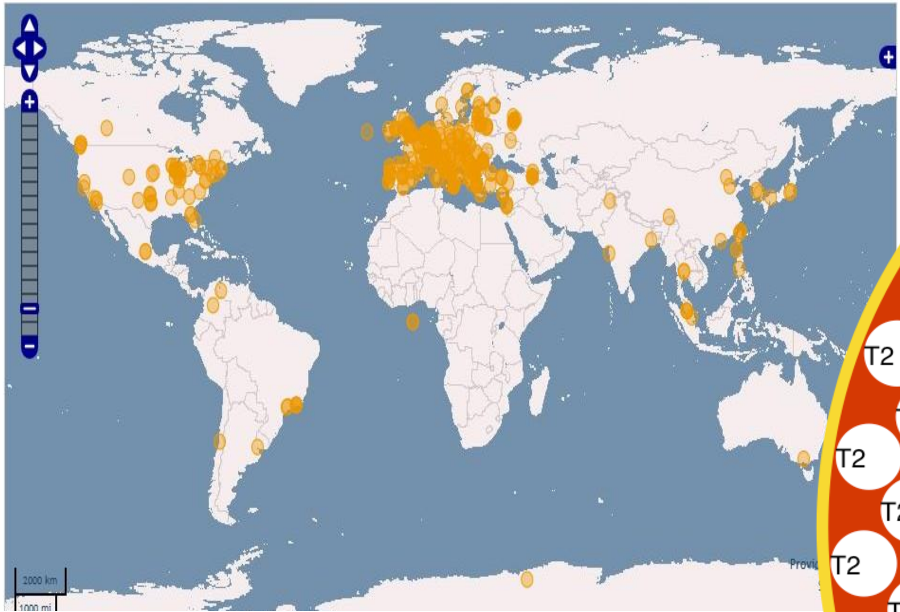


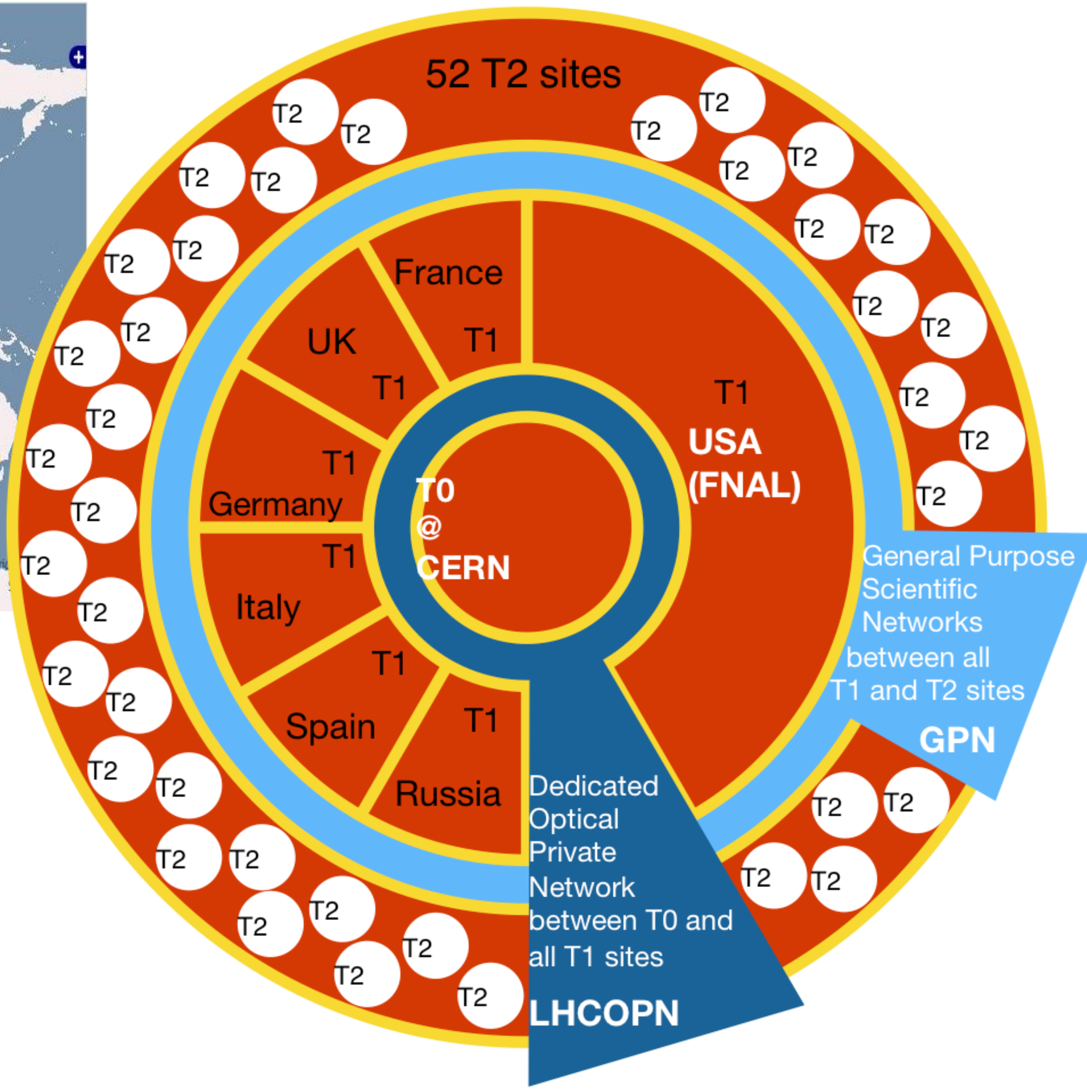
- Computing resources are either experiment owned (e.g. LHC) or deployed for a specific program (e.g. IF at Fermilab)

HEP HTC Facility Services

- Distributed compute resources presented as **a coherent job execution environment** optimized for maximum throughput
- Support of workflow management aspects required for the fidelity of the scientific results
  - book-keeping, provenance tracking
  - schedulers and workflow systems, information systems, data bases, trust relationships and security protocols, monitoring and problem resolution
- **Data Management and access essential** for extracting science from massive data stores in a distributed environment
  - support federation of storage systems, high performance networks
  - services across many data centers for local and remote access to data

CMS distributed computing model





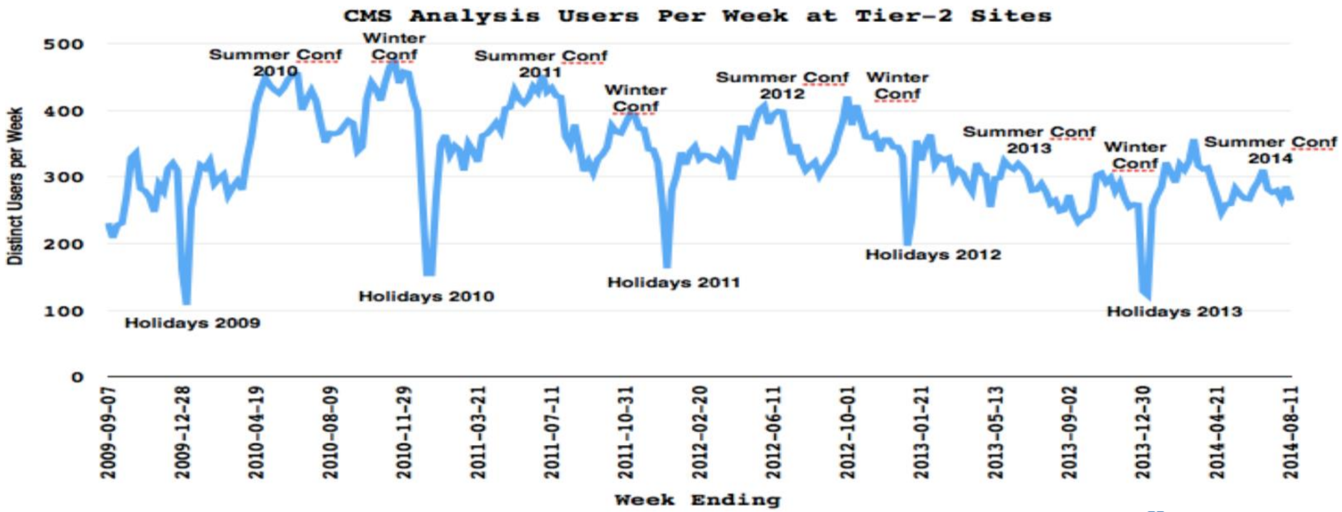
Distributed computing model empowering scientists from around the world to perform data analysis!

- ~ 200 institutions from ~50 countries

Current size of computing infrastructure

- World-wide CMS experiment: ~**100K cores**, including Tier-1 and Tier-2 facilities
  - US CMS: ~**15K** cores at Fermilab, ~**25K** cores at Tier-2 and Tier-3 sites
- General purpose (non-CMS) at Fermilab: ~**12K cores**
- Similar values for world-wide and US ATLAS, and for the ATLAS Tier-1 center (at BNL) and Tier-2 facilities

Resources are utilized in "burst" mode



Evolution of HEP experimental program

2015-2020 (large, mid-size programs)

FNAL MI-LB neutrinos (IF)

- MINOS+, MINERvA, NOvA

FNAL Booster-SB neutrinos

- uBooNE, SBND, ICARUS (IF)

FNAL Recycler-muons (IF)

- g-2

KEK-heavy flavors (IF)

- Belle II

LHC beams: Run 2

- ATLAS, CMS (EF)
- LHCb (IF)

2020-2025

…

Recycler-muons (IF)

- Mu2e

LHC Run 3 (phase 1 upgrade)

…

2025-…

Long Baseline Neutrino Facility (LBNF)

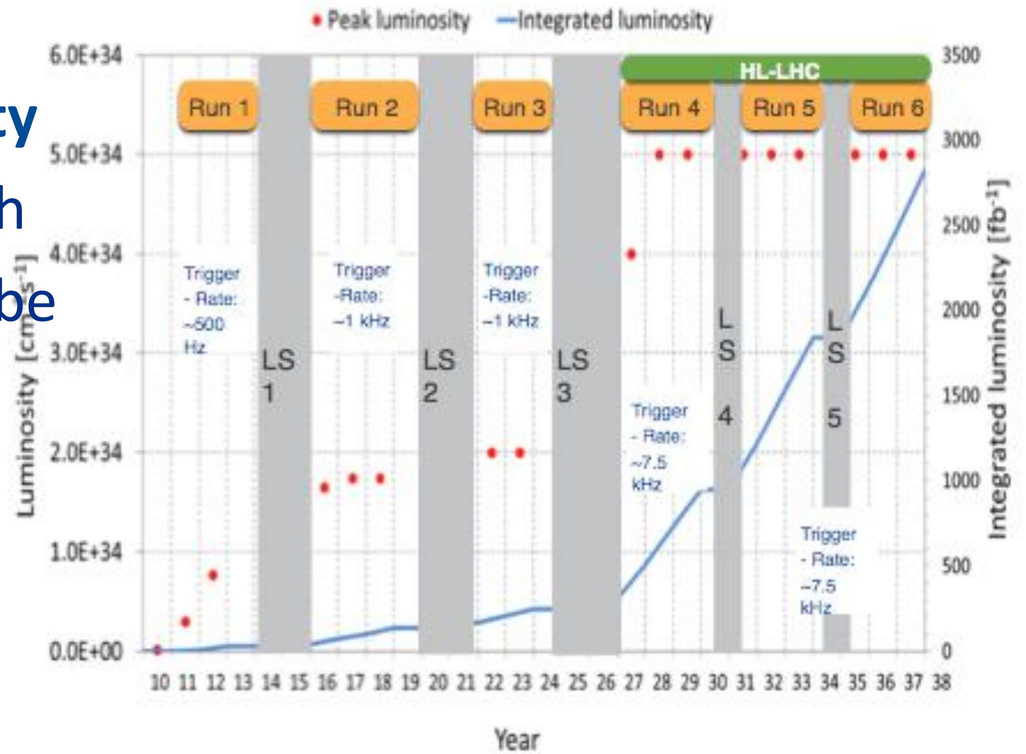
- **DUNE (IF)**

LHC Run4 (HL-LHC)

…

Future computing requirements: Compute and Data Needs

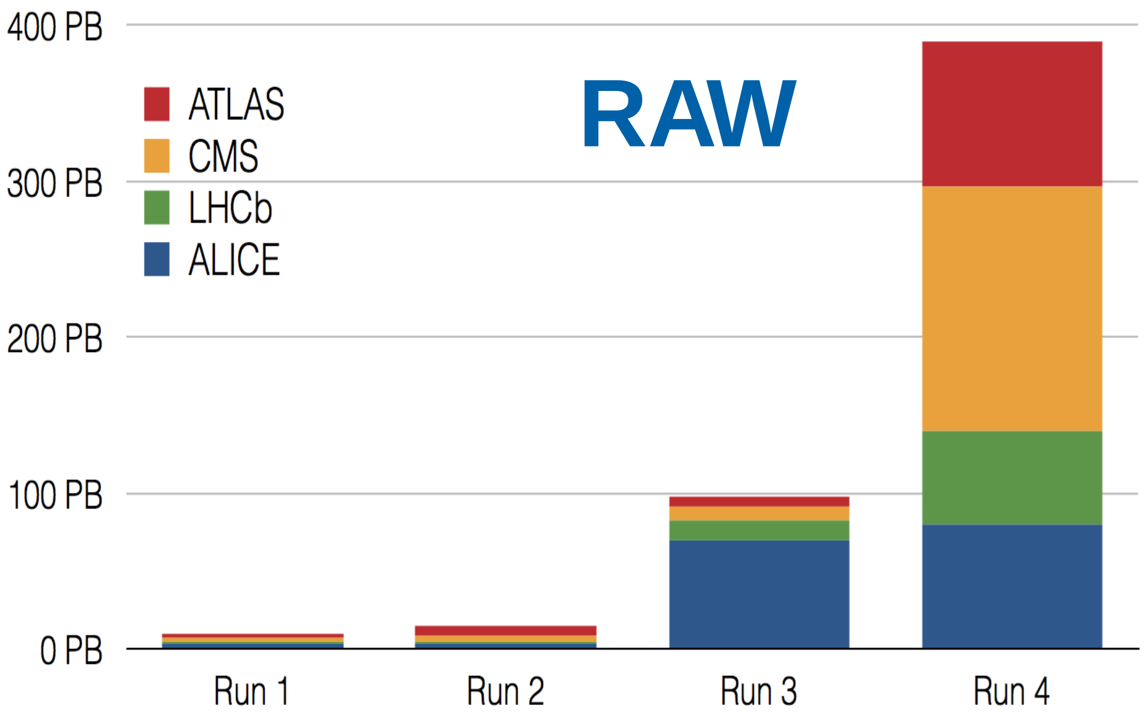
- Two new programs coming online (**DUNE**, **High-Luminosity LHC)**, while new physics search programs (**Mu2e, Belle2**) will be operating
- Increased precision & event complexity, higher luminosity, will push computing needs to **~10X-100X of current HEP capabilities**
  - *Lower value assumes optimized algorithms and new approaches*



Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Tape [PB]	2.8	2.8	2.8	2.8	19.24	54.43	103.55	153.89	204.64	255.39
Disk [PB]	4.00	4.00	5.00	8.00	27.98	79.17	115.68	153.10	190.82	228.55
CPU [kHepSPEC]	45.00	45.00	50.00	55.00	328.31	568.98	567.54	609.45	643.14	672.60

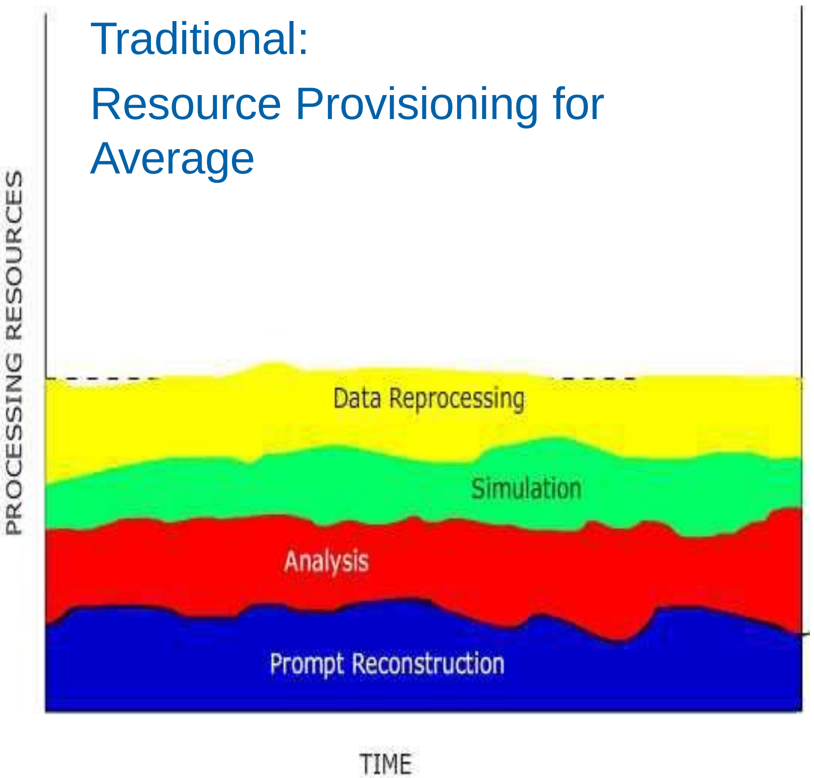
Belle2 computing requirements

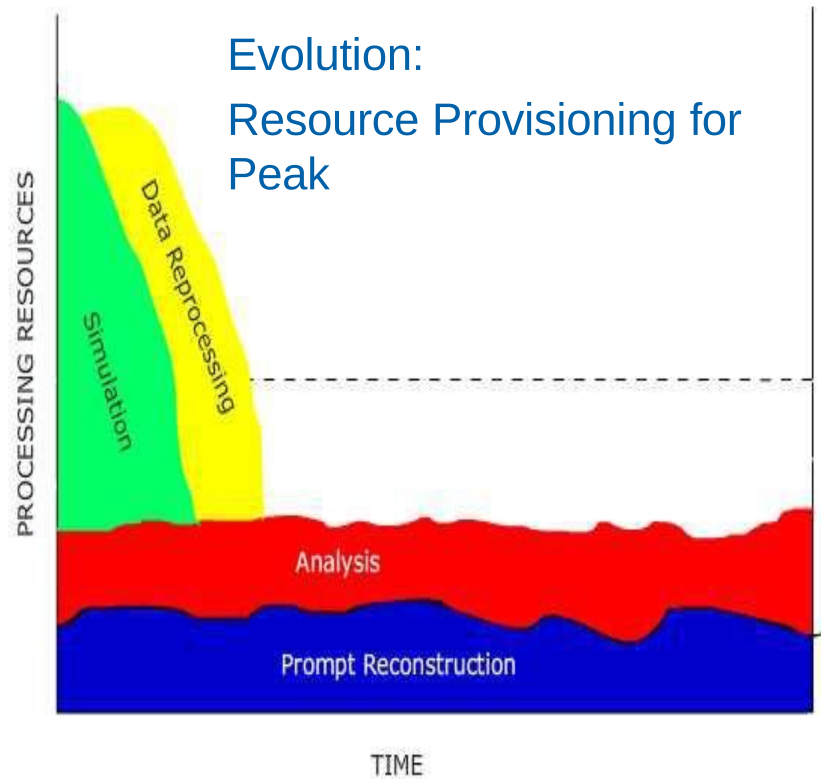
LHC expected data volumes



- **LHC Run 4 will start the exabyte era for HEP!**
  - Derived data (reco, simulation) x8 of RAW…
- **How do we process and analyze all these data?**

Evolving provisioning model (for cost-effectiveness)





- Experiments don’t need all the resources all the time
  - provisioning needs to be adaptable, providing facility “elasticity” to handle “burst usage”

➔ ***incorporate and manage*** “rental” resources

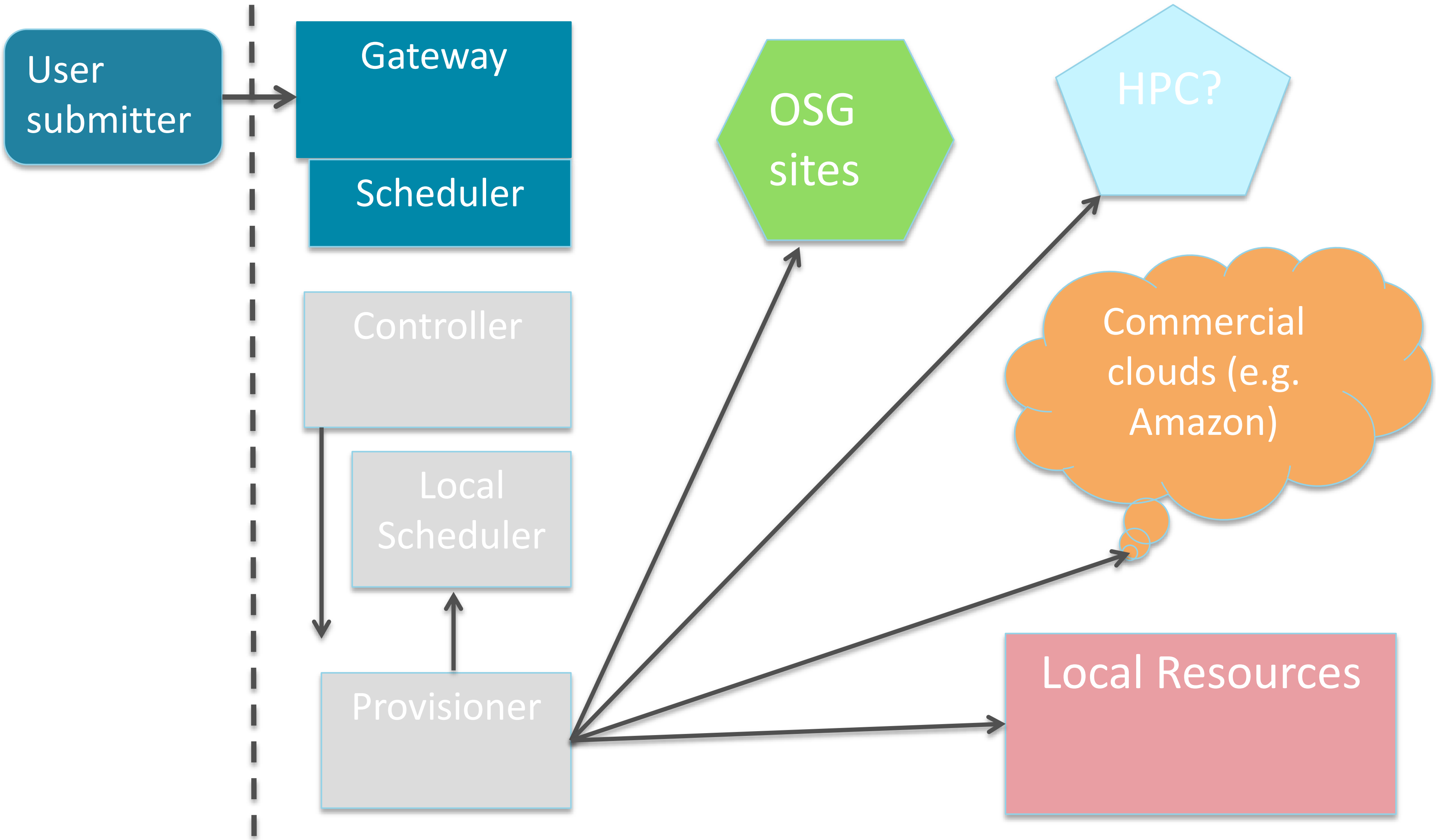
Evolution of the HEP computing paradigm

- Currently, US HEP facilities support dedicated and shared resources, for data and compute intensive workflows
- Industry trend is to use Cloud services: Infrastructure, Platform & Software as a Service, to
  - Reduce cost of provisioning and operating; provide redundancy or failover; rapidly expand and contract resources; pay only for the resources needed/used.
- Similarly, US-HEP facilities could ***incorporate and manage*** "rental" resources, achieving "elasticity" that satisfies demand peaks without overprovisioning local resources.
  - Also, separating compute from storage services will allow them to scale independently (additional cost and service optimization)

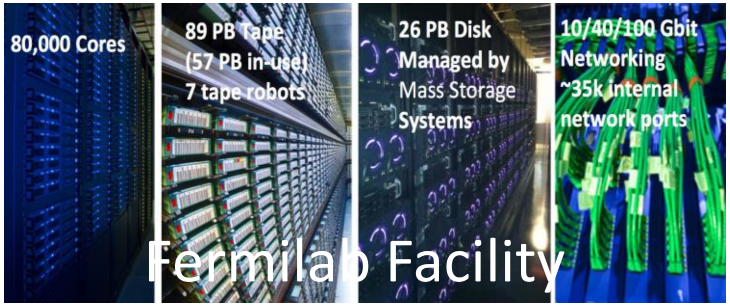
HEPCloud Facility vision

- The HEPCloud Facility is a portal to an ecosystem of computing resources, commercial or academic.
- Provides “complete solutions” to all users, with agreed upon levels of service
  - The Facility routes to local or “rental” resources based on efficiency, cost, workflow requirements and target compute engine policies.
  - Manages allocations of users to target compute engines
- Provides storage services appropriate to the system that the workflow is routed

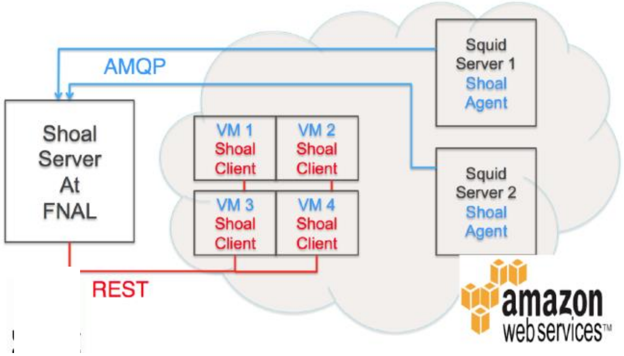
HEPCloud Facility concept



New paradigm changes role of HEP Facilities



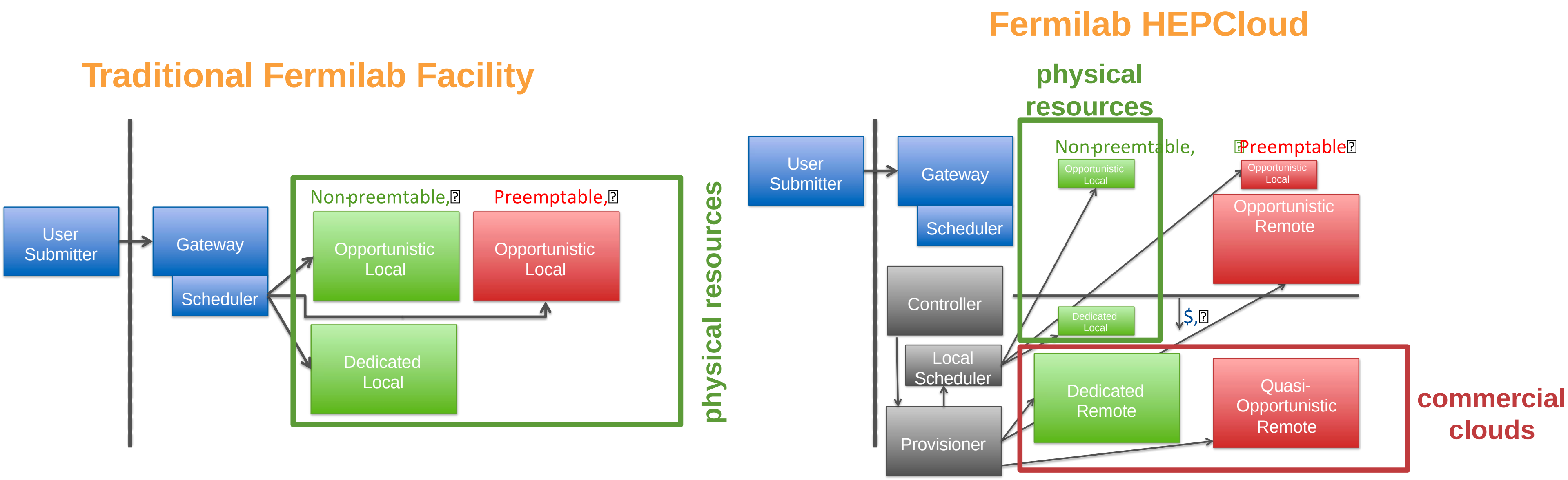
- Moving away from stand-alone, stove-piped facility solutions
- Labs could provide leadership (and shared resources) in the ecosystem
  - a large pool for offerings, including archival capabilities, compute, data management, …
    - ➢ potentially linking all US-HEP computing
  - Provide complete solutions to users
- Users don't have to decide on target compute engines or manage resources
  - Sites decide based on efficiency, cost and workflow requirements for "owned" or "rented" resources



Fermilab HEPCloud Facility project

- The goal is to integrate “rental” resources into the current Fermilab computing facility in a manner transparent to the user. Objectives include
  - A seamless user environment for all resource types, including necessary tools and infrastructure
  - The architecture, including network, needed to support required data rates.
  - The policies and middleware for efficiently using and prioritizing the use of different resources
  - The information security policies, procedures and monitoring.
- Partnership with resource providers necessary to understand requirements, architecture and software development needs

Fermilab HEPCloud project



FNAL HEPCloud Project - Demonstrators

NoVA Processing

Processing the 2014/2015 dataset
16 4-day “campaigns” over one year
Demonstrates stability, availability, cost-effectiveness
Received $30,000 academic grant from Amazon Web services

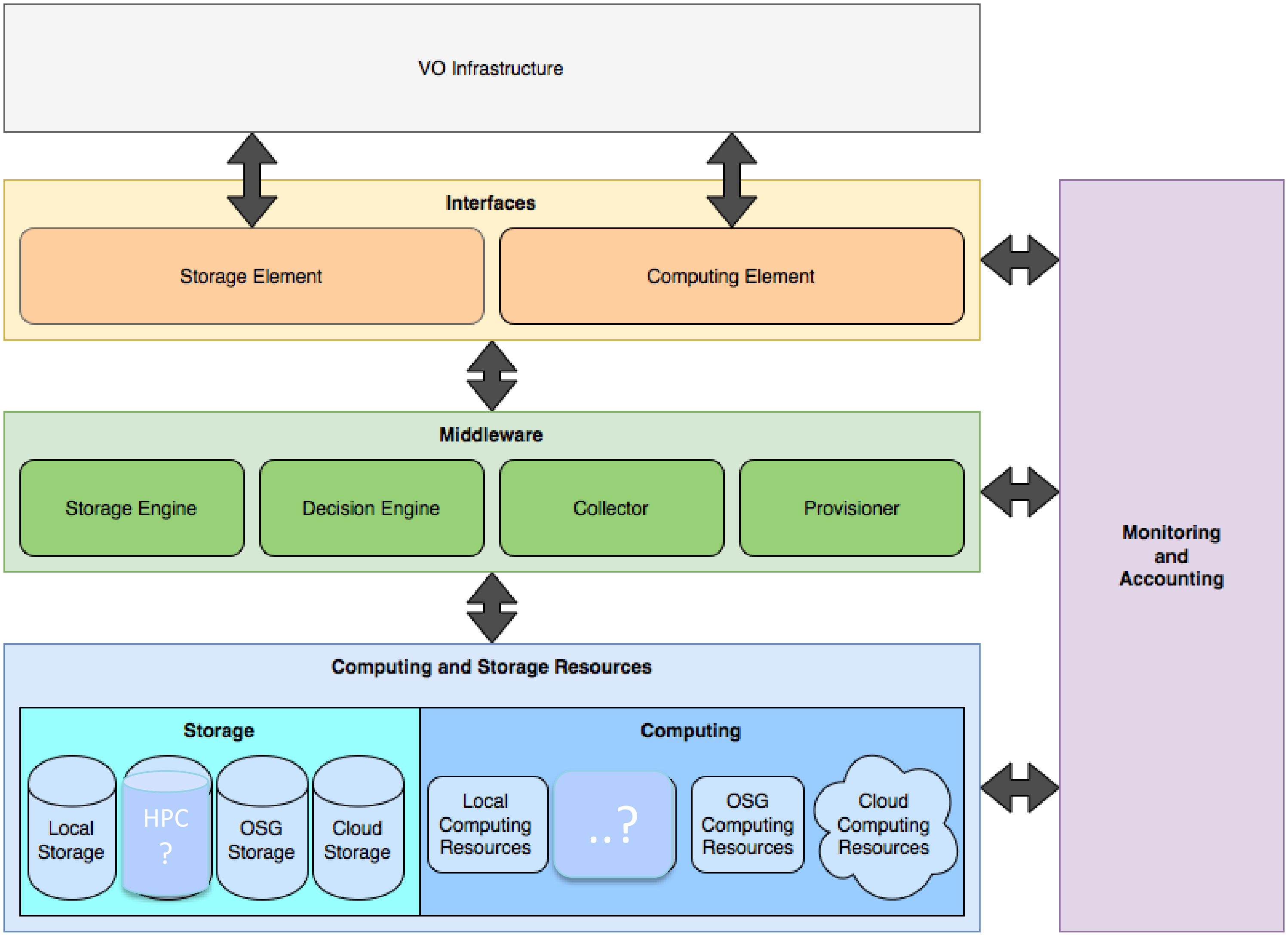
Dark Energy Survey - Gravitational Waves

Search for optical counterpart of events detected by LIGO/VIRGO gravitational wave detectors (FNAL LDRD)
Modest CPU needs, but want 5-10 hour turnaround
Burst activity driven entirely by physical phenomena (gravitational wave events are transient)
Demonstrates provisioning to peak

CMS Monte Carlo Simulation

Generation (and detector simulation, digitization, reconstruction) of simulated events for Run 2 after beam spot determination
56000 compute cores for 1 month, steady-state
Demonstrates scalability
Received academic grant covering 90% of incurred costs

HEPCloud Architecture

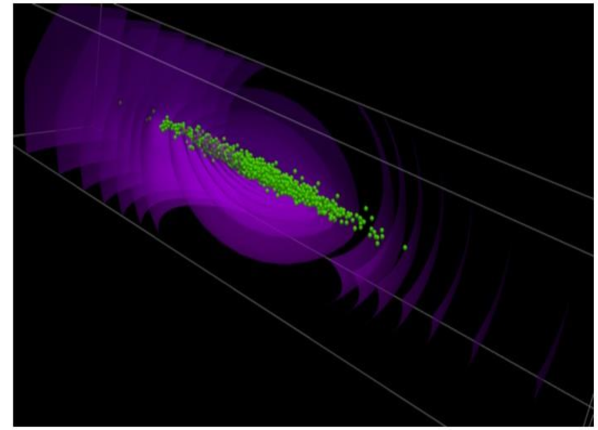


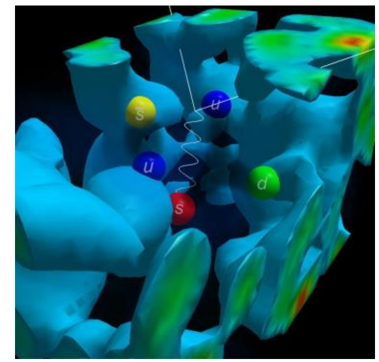
HEPCloud and HPC

- A very appealing possibility, as we are approaching the exascale era, is to consider HPC facilities as a potential compute resource for HEPCloud
  - and, in the other direction, consider HEPCloud facility services (e.g. storage) as a potential resource for HPC facilities
- Investigate use cases with workflows that will allow such utilization within the constraints of allocation, security and access policy of HPC facilities.
- Initiate work with HPC facilities to fully understand constraints and requirements that will enable us to develop the HEPCloud process, policies and tools necessary for access of HPC resources

HPC and HEP

- Long tradition of successful HEP projects using HPC resources
  - Important scientific applications from LQCD, Cosmology, Accelerator Modeling utilizing ASCR HPC facilities
- Current efforts: enable traditional HEP workflow applications to run on HPC resources
  - E.g generating Atlas LHC events at ALCF



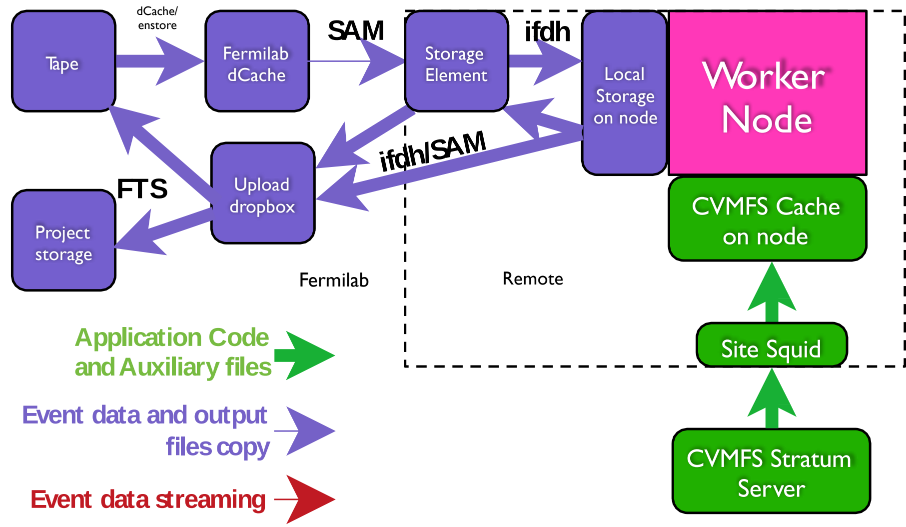


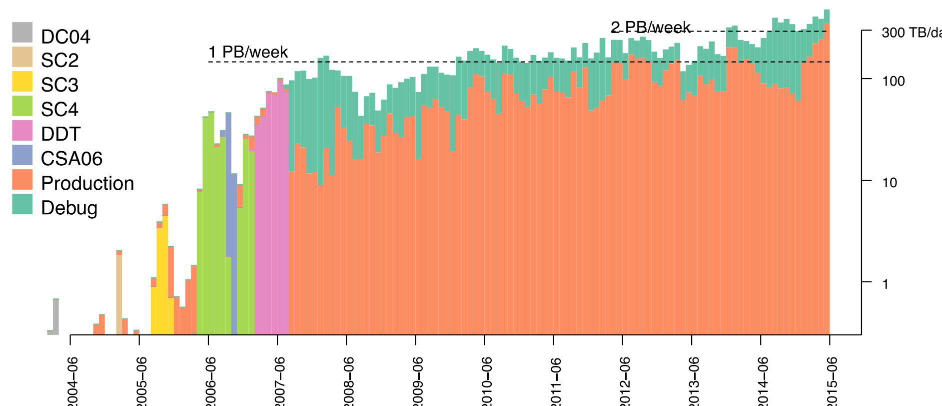
Potential use cases for HPC through HEPCloud

- Build on current successful efforts: Event Generation in fully automated production mode
  - minimum data ingress, simpler workflow
  - a good use case candidate to investigate with ASCR HPC facilities and test with the Fermilab HEPCloud project
- Could consider an HEP-ASCR “end station” concept as the entry point to HPC resources through HEPCloud

Data Management and Archival

- HEP uses various solutions to provide distributed access to data:
  - Experiment specific for Atlas and CMS or shared for Neutrino and Muon experiments
- HEP has tools and experience for distributed data management at the exabyte scale
- Make available through HEPCloud, if HPC connection successful, potentially beneficial to HPC facilities.

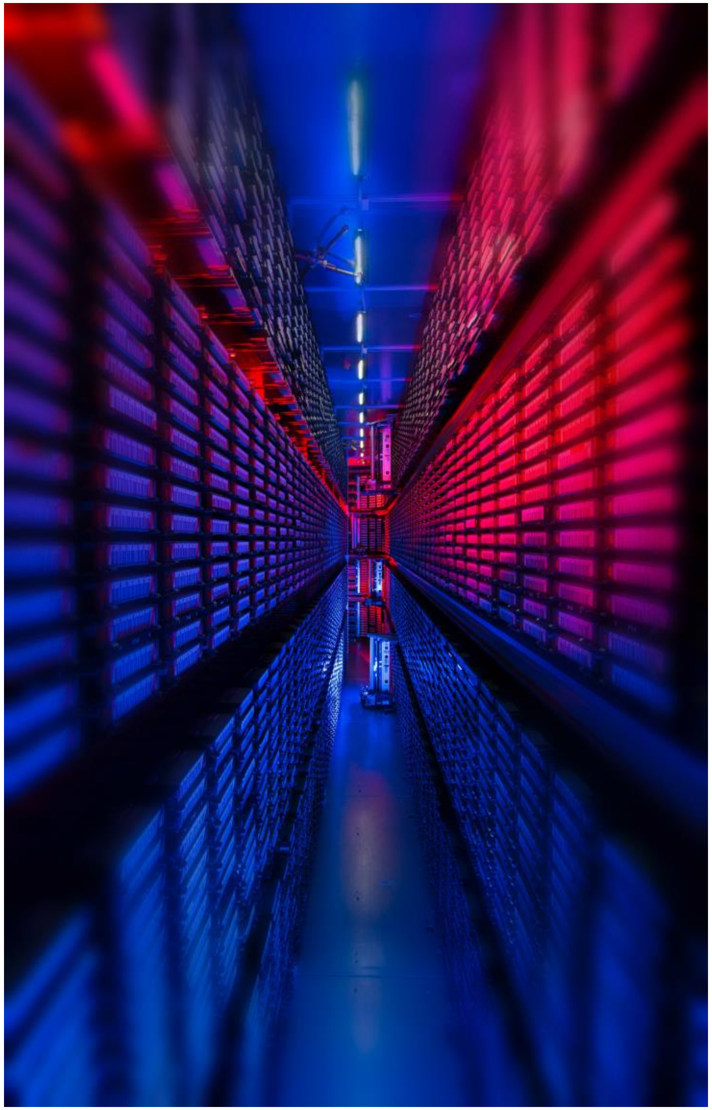




CMS transfers: more than 2 PB per week

Active Archival Facility

- Working with science communities outside HEP
  - To enable management, distribution and access of data globally
- Example: Fermilab's Active Archival Facility (AAF)
  - Provide services to preserve integrity and make available important scientific data
  - Projects:
    - Genomic research community is archiving datasets at the AAF which provides access to ~300 researchers world wide
    - University of Nebraska and University of Wisconsin are using AAF for long term archival



Summary

- Large increase in demands for HEP computing changes requirements/expectations for computing facilities
- HEPCloud, a promising new paradigm to provide efficient, "elastic", cost effective solutions
  - A portal integrating "rental" services, potentially linking US-HEP computing resources, and potentially utilizing appropriate HPC facility offerings
- Project started at FNAL to evaluate approach. Connection to HPC very important for success.

Questions?

1. User submits set of jobs to VO-specific job management agent (In our example the VO is the CMS Experiment at the LHC).