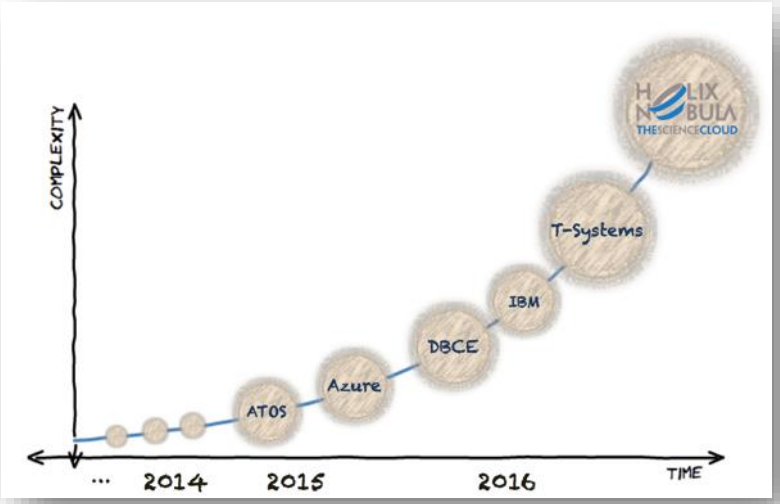
Building access layers to Cloud and HPC from the facility

Gavin McCance
CERN IT Compute and Monitoring Group

Burt Holzman
Fermilab Scientific Computing Division

CERN private cloud

- CERN has large private Openstack cloud
- Batch service almost fully migrated to HTCondor now



Exploring hybrid cloud since 2013 through various projects

HELIX NEBULA THESCIENCECLOUD

2015

2nd Mar. Atos
- *End: 31st of March 2015*
- ATLAS simulation jobs
- Single core VMs
- Up to 3k VMs for 45 days
- 1st Cloud Procurement

6th Nov. DEUTSCHE BÖRSE CLOUD EXCHANGE
- *End: 18th of Dec. 2015*
- Target all VOs, simulation jobs
- 4-core VMs, O(1000) instances
- 2nd Cloud Procurement

1st Aug. T-Systems
- *End: 30th of Nov. 2016*
- Provided by OTC IaaS
- 4-core VMs, O(1000) instances
- 500TB of central storage (DPM)
- 1k public IPs through GÉANT
- 3rd Cloud Procurement

23rd Mar. Microsoft Azure
- Sponsored Account
- "evaluation of Azure as an IaaS"
- Any VO, any workload
- Targeting multiple DCs:
  - Iowa, Dublin and Amsterdam
- *End: 30th of Nov. 2015*

2016

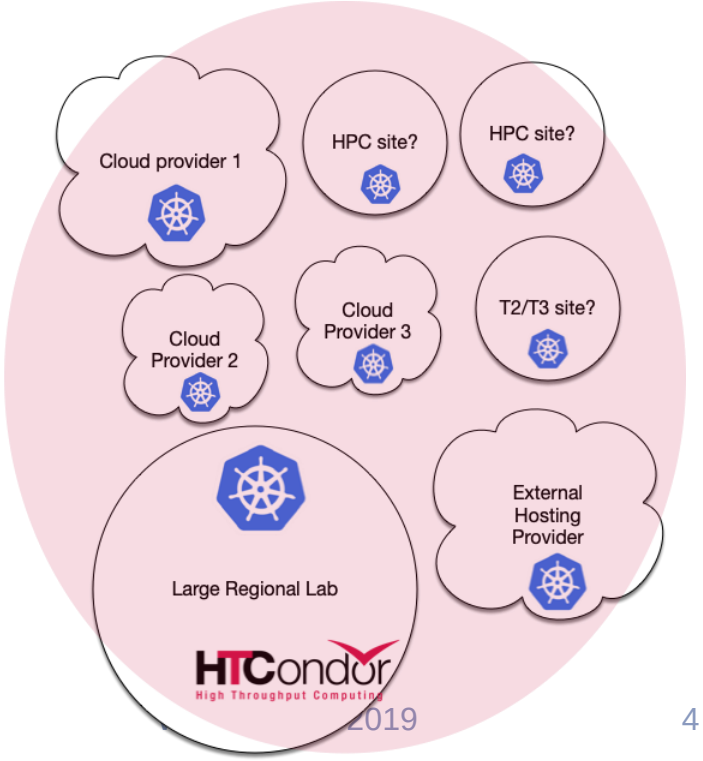
20th Nov. SOFTLAYER an IBM Company
- Agreement between IBM and CERN
- CERN PoC to evaluate:
  - Resource provisioning
  - Network configurations
  - Compute performance
- Transparent extension of CERN's T0
- *End: 13th of May 2016*

Have partnered with many EU-based vendors

CERN Hybrid cloud

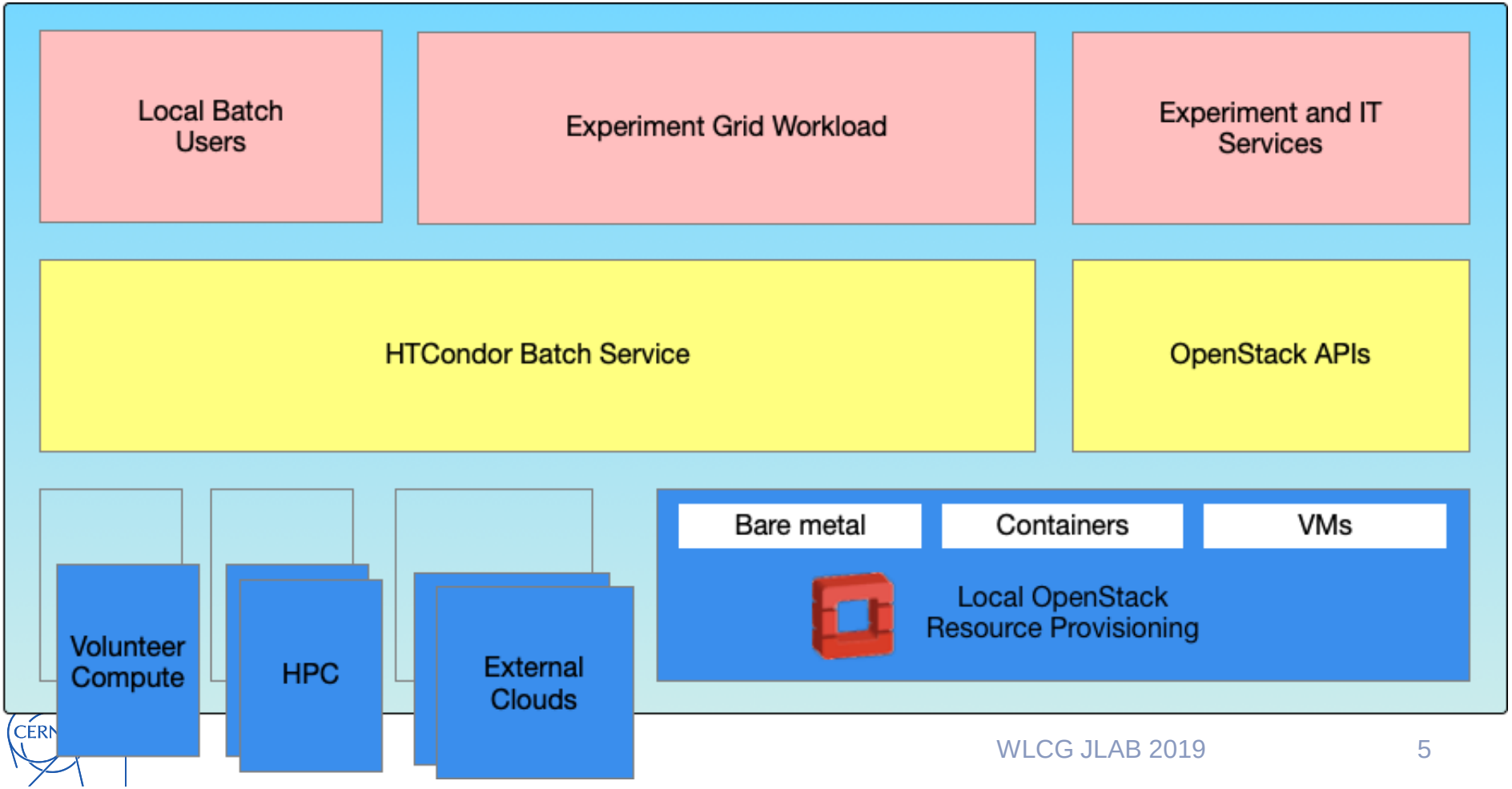
- Focus on external resources has been cloud rather than HPC
- Initial efforts focused on understanding workloads rather than operations
- Since 2017, efforts have been focused on bringing hybrid cloud into our standard site operations and on understanding the procurement model
- Primarily procuring flat IaaS capacity within scope of these projects
  - Both as VMs and federated Kubernetes
- Recently ~O(10-15k) cores, about 5% of our capacity

- HTCondor makes integration easy!
  - As per CMS global pool, very easy to use integrate external resources dynamically



Overall approach

- Single logical HTCondor pool, routing jobs to all underlying resources, internal and external



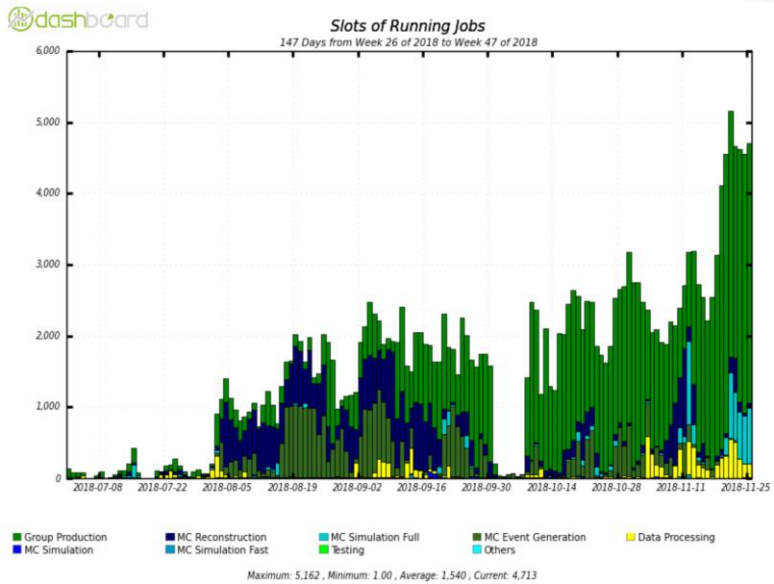
Standardised Operations

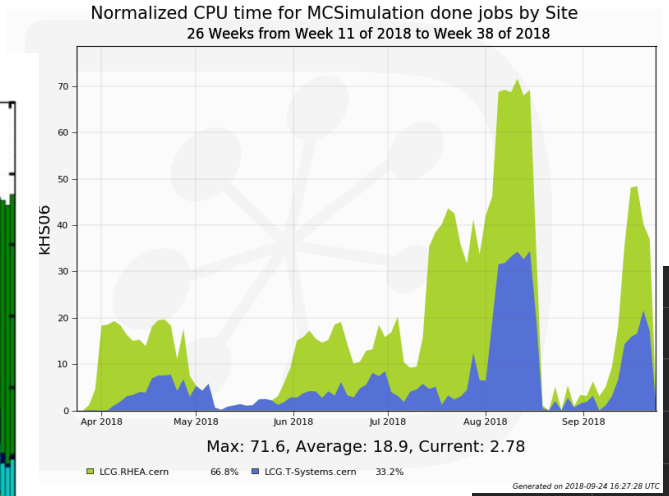
- To reduce costs, as much as possible, we try to use the same ops tooling internally and externally
  - though devil in the details sometimes e.g. NAT setups, etc
  - generally a one-off standing cost to integrating a new cloud

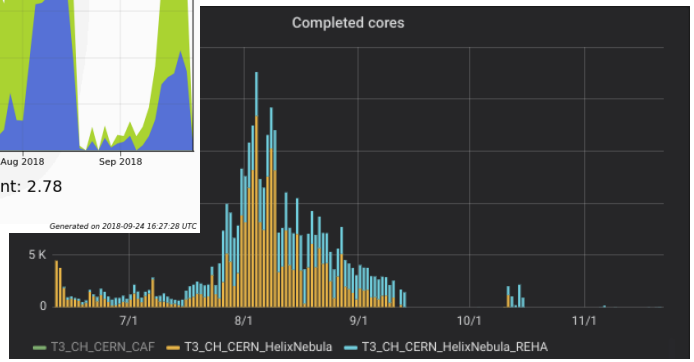
Platform	HTCondor High Throughput Computing
Configuration & Monitoring	puppet labs, collectd, Grafana
Personalization	cloud-init
Provisioning	HashiCorp Terraform

Workloads

- Various experiment workloads tried
  - Monte Carlo, reconstruction from data @CERN, analysis; some GPU
  - Monte Carlo still most common workload
- Challenges on how to identify appropriate workload automatically, and route
  - Routing it manually via ”the special CE”, per project, per experiment, is not very efficient

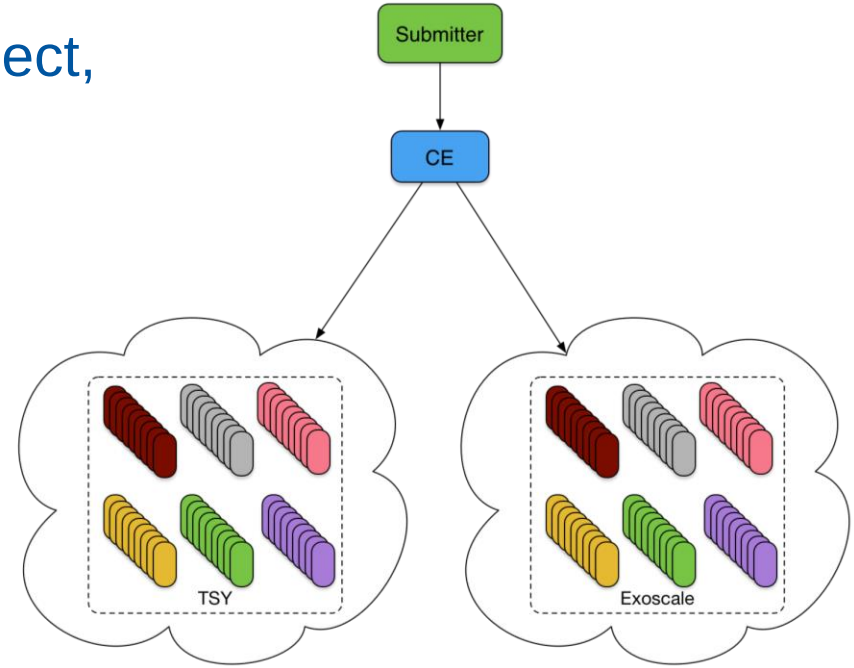






Consolidated multi-cloud

- Within recent HNSciCloud project, we had 8 WLCG sites with grants in the “buyer’s group"
  - Chose to consolidate these in a ”single WLCG” queue managed @CERN via our HTCondor
- Ops economy of scale
- Single entry point for experiments
- Successful, though some challenges
  - Managing fair-share across multi-clouds / grants
    - For HNSci project, we just used the CERN default batch shares
  - Accounting back to the buyer / grant-holder



Summary

- Strategy of hiding complexity and offering HTCondor as the platform over multiple clouds works well
  - One-off onboarding cost per cloud (sometimes significant) but handled by one ops team
  - Fits well with public procurement model

- Smaller pools of grants can be operationally consolidated
  - Economy of scale and simpler for experiments
  - Need to meet needs of grant holder and show-back that you really have met those needs; solid accounting

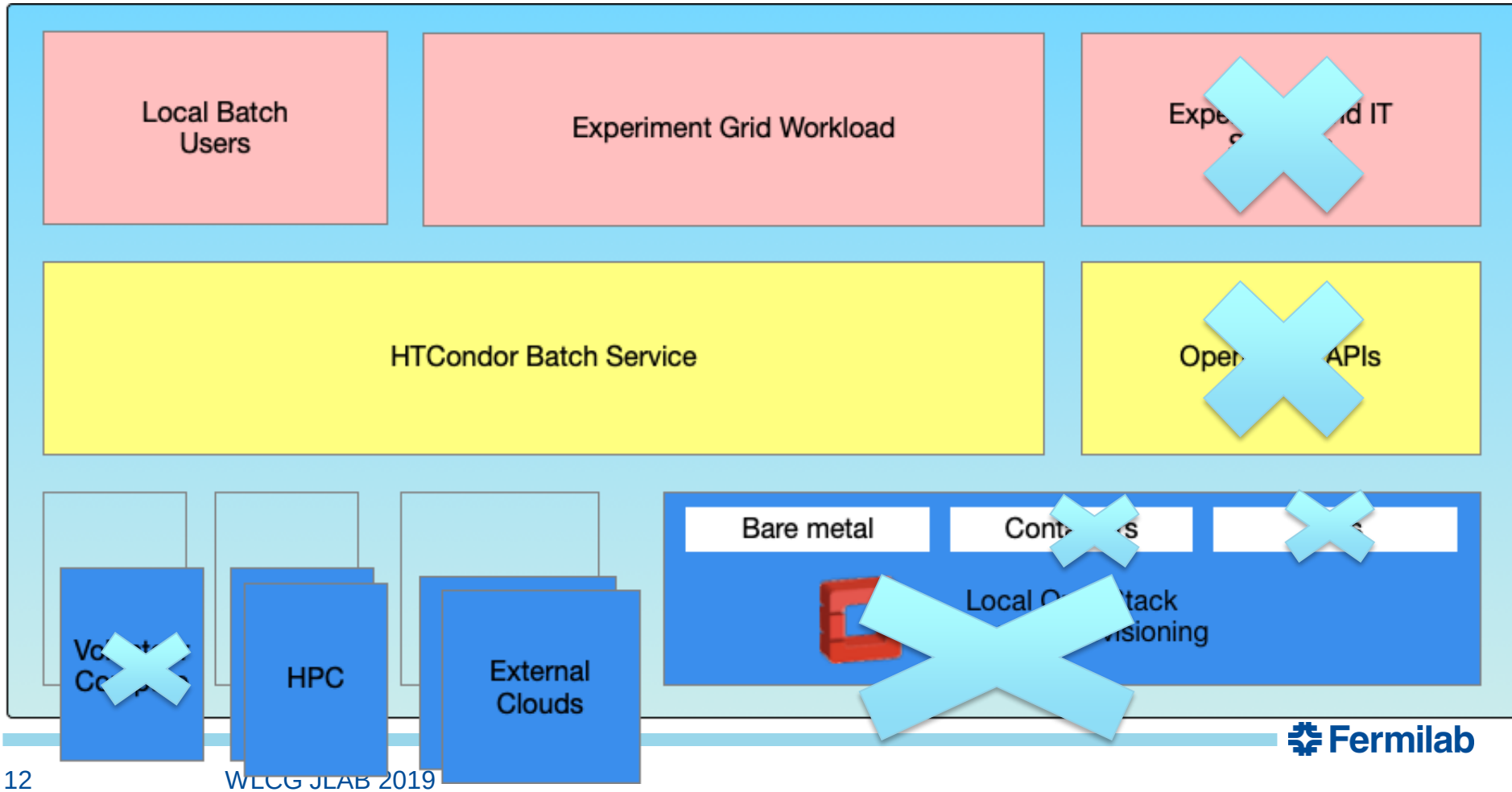
Cloud Challenges

- Public procurement with suitable discount
  - TCO advantage is still not obvious
  - Not all clouds are the same level of maturity
- Challenges in making business-as-usual
  - Project-driven "specials” e.g. extra monitoring and reporting are expensive on sites and experiments
- Cost control is critical
  - Tagging appropriate payloads and identifying which are the expensive ones e.g. data costs
  - Accounting back to the users

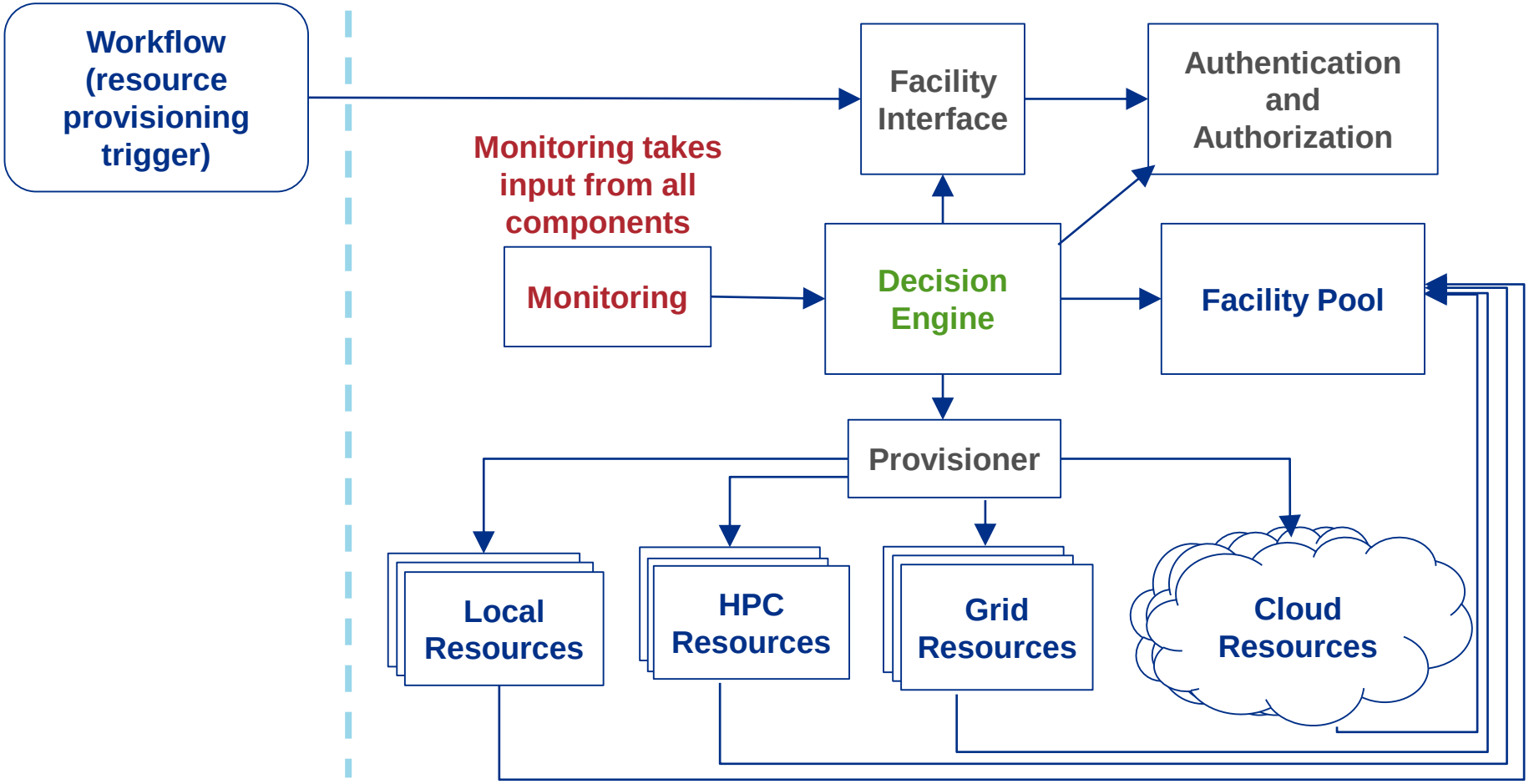

GLOBE OF SCIENCE AND INNOVATION / GLOBE DE LA SCIENCE ET DE L'INNOVATION

Common CERN/FNAL Approach

- **Expand the facility** to include off-premises resources



Fermilab Architecture (HEPCloud)

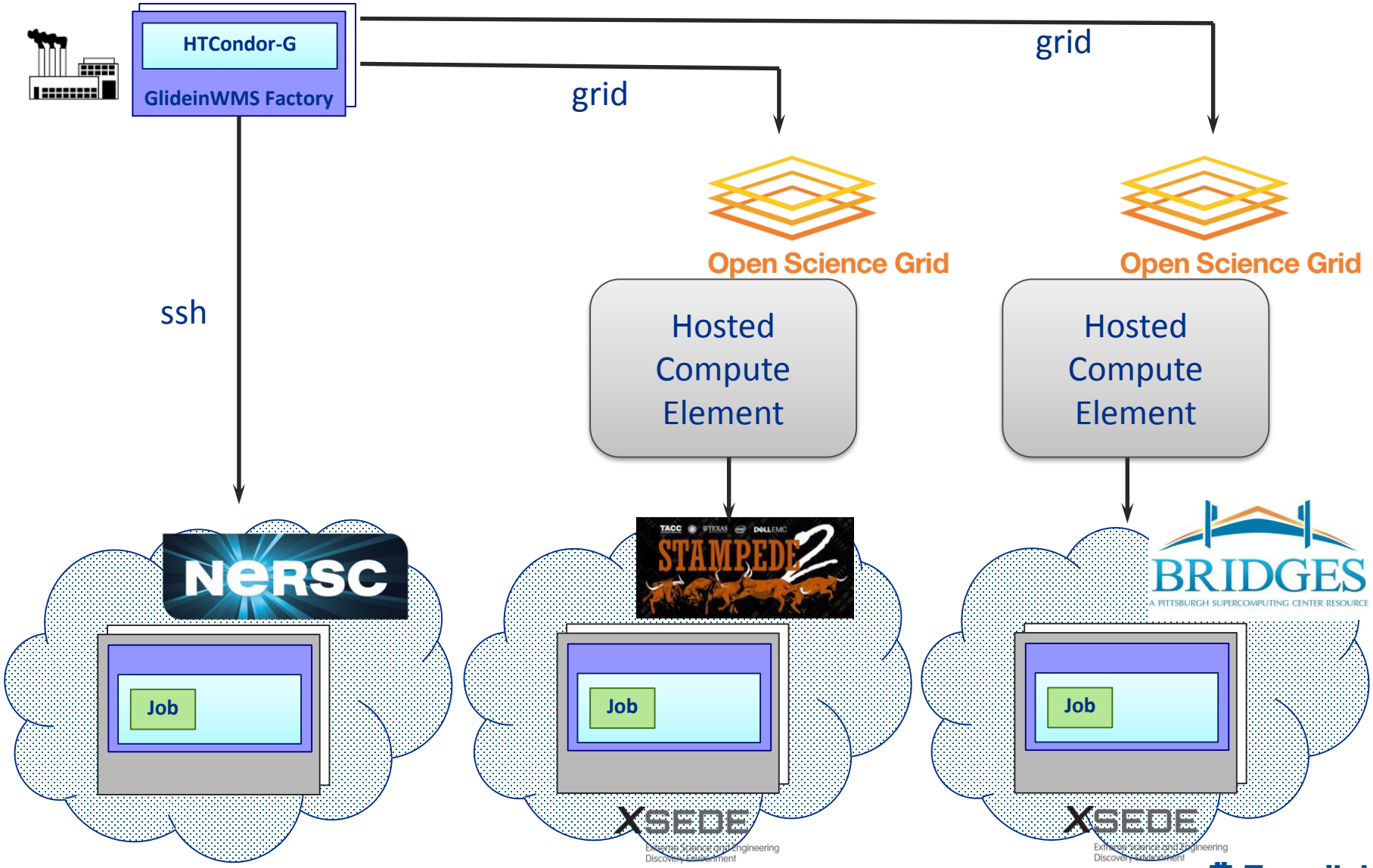


Standardized Operations

- Different tools when necessary for external vs. internal resources

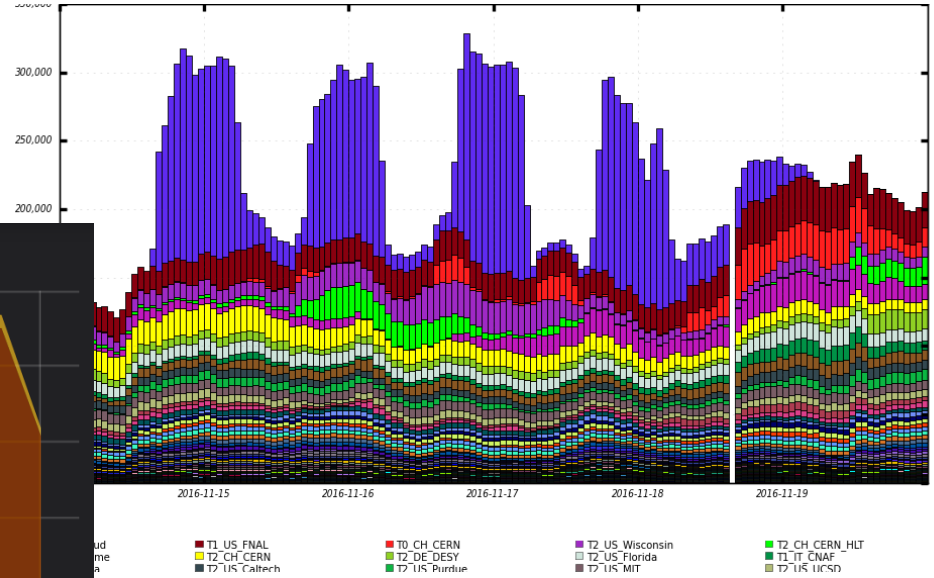
	On-premises	Off-premises
Platform	HTCondor High Throughput Computing	
Configuration & Monitoring	puppet labs, Prometheus, CHECK_MK, Grafana	
Personalization	puppet labs	SHIFTER, S, cloud-init
Provisioning	okd, ??,	HTCondor High Throughput Computing

Provisioning HPC with HTCondor

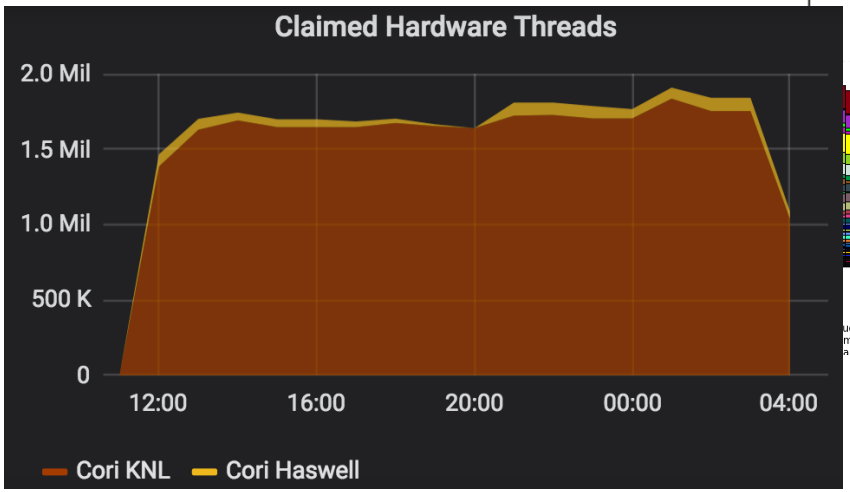


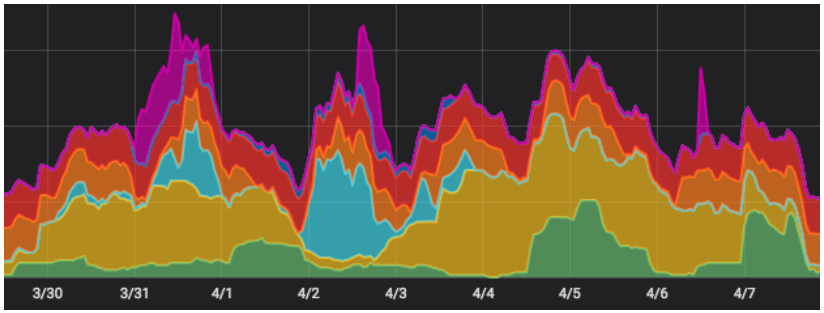
Does it scale?

160kcores on Google

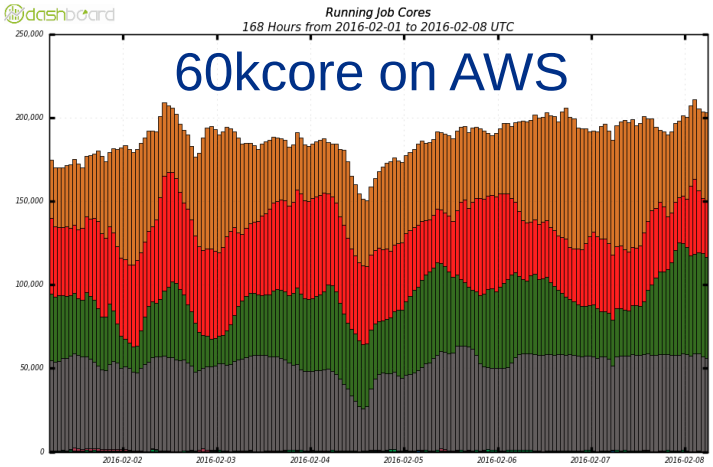


2M threads @ NERSC





Different colors correspond to
NERSC Cori-1, Cori-2, Edison; PSC Bridges;
TACC Stampede2



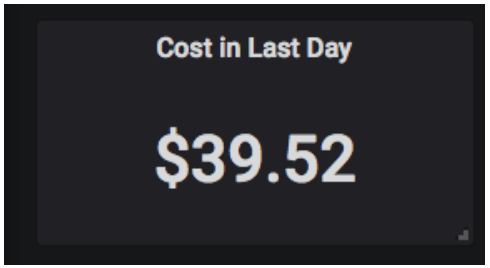
60kcore on AWS

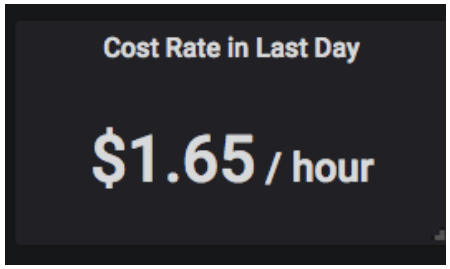
Summary

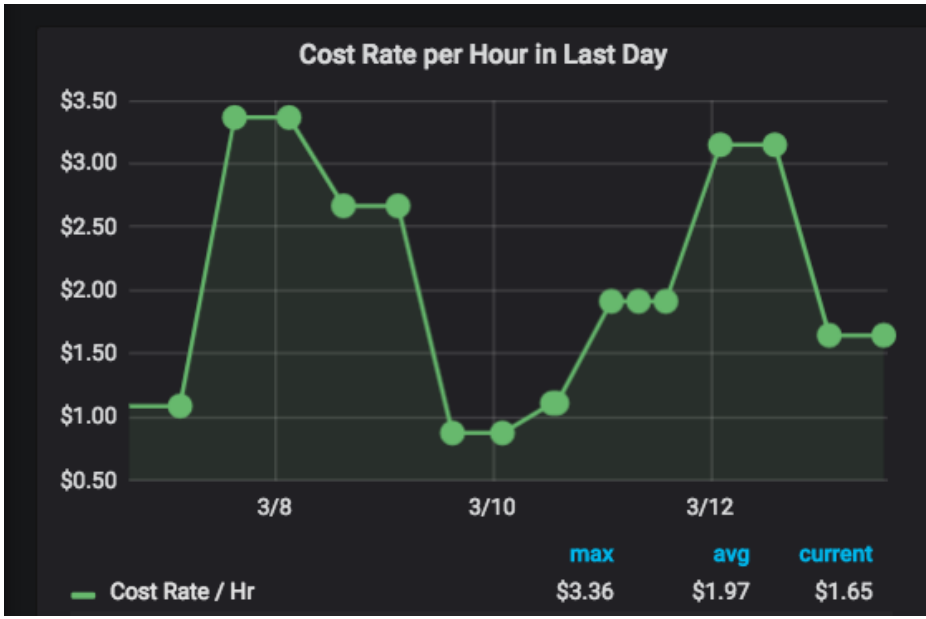
- Strategy of hiding complexity and offering HTCondor as the platform over multiple clouds works well
  - One-off onboarding cost per cloud (sometimes significant)
  - Fits well with public procurement model

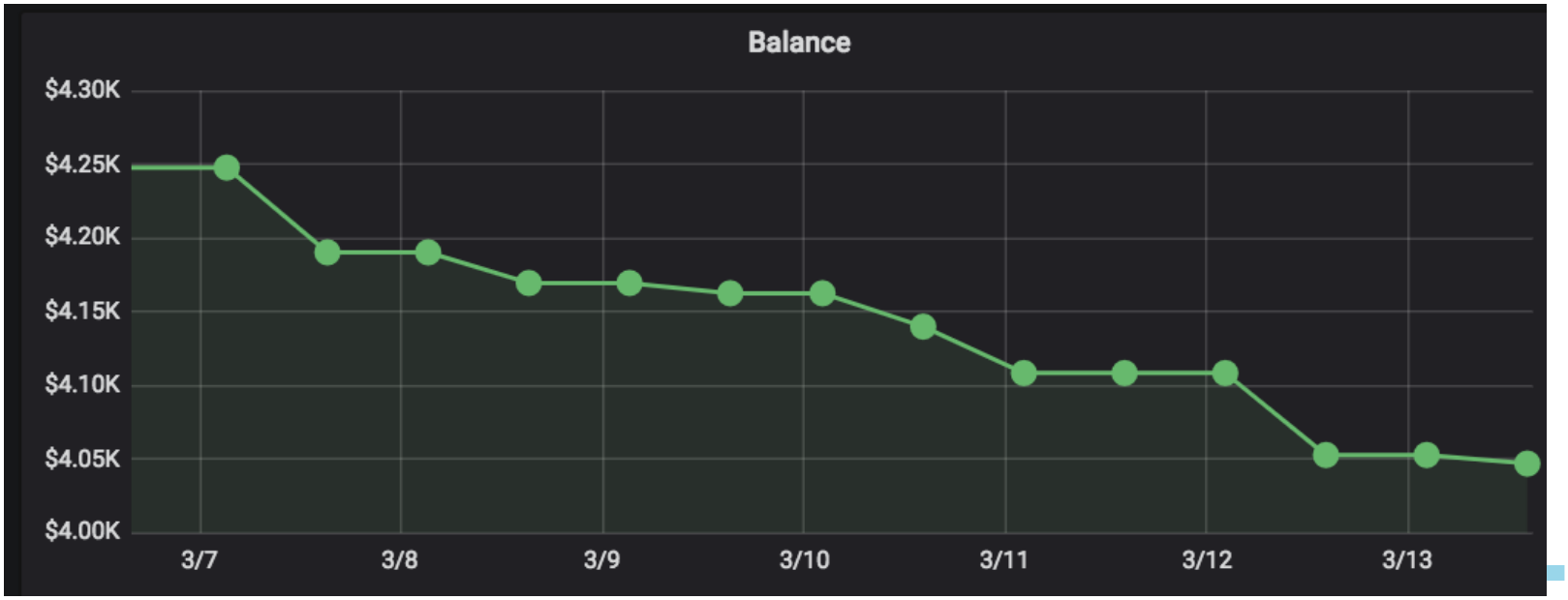
- Smaller pools of grants can be operationally consolidated
  - Economy of scale and simpler for experiments
  - Need to meet needs of grant holder and show-back that you really have met those needs: solid accounting

Accounting @ Cloud

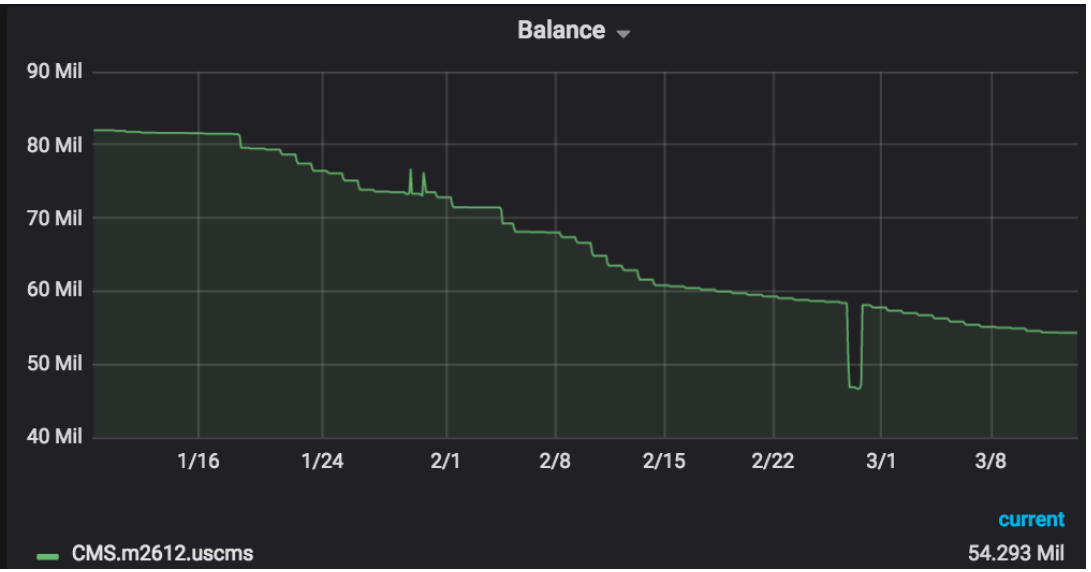


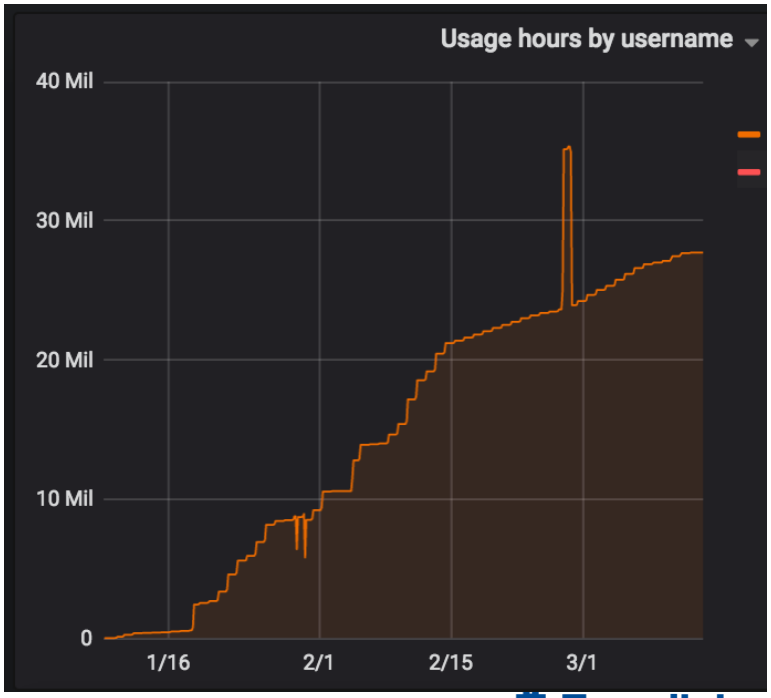






Accounting @ HPC (NERSC)



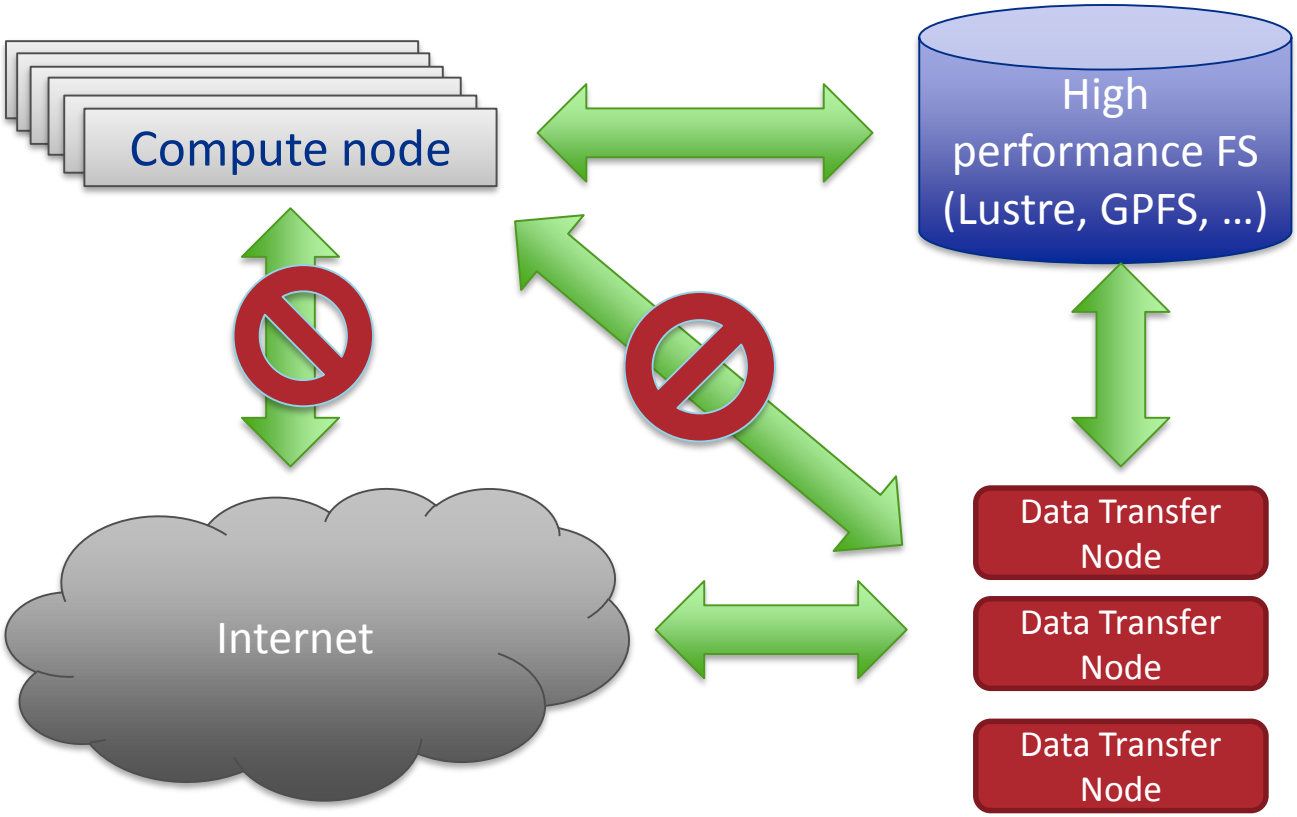


~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model

HPC Challenges - Data

- HPC Centers are built for large IOPS between local compute and storage, not over the WAN



NERSC provided 500 TB of space for CMS input data

~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model
  - Work with sites to meet experimental needs (high performance streaming over WAN?)
  - Use common data management/handling frameworks (e.g. Rucio) to prestage data

~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model
  - Work with sites to meet experimental needs (high performance streaming over WAN?)
  - Use common data management/handling frameworks (e.g. Rucio) to prestage data

- HPC sites prefer large parallel workflows

Leadership sites prefer Leadership jobs

Argonne Leadership Computing Facility

an Office of Science user facility

Production Queues

- Priority is given to jobs using at least 20% of Theta (>=802 nodes) (previously >=648 nodes)
- There is a global limit of ten (10) jobs running per user
- There is a global limit of twenty (20) jobs in queue per user
- There is a minimum job time of thirty (00:30:00) minutes for the default queue
- Wall-clock limits are a step-wise function designed to encourage scaling:
  - node count >= 128 nodes (minimum allocation): maximum 3:00:00 hours
  - node count >= 256 nodes : maximum 6:00:00 hours
  - node count >= 384 nodes : maximum 9:00:00 hours
  - node count >= 640 nodes : maximum 12:00:00 hours
  - node count >= 802 nodes : maximum 24:00:00 hours

~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model
  - Work with sites to meet experimental needs (high performance streaming over WAN?)
  - Use common data management/handling frameworks (e.g. Rucio) to prestage data

- HPC sites prefer large parallel workflows
  - Can use late-binding scheduling to aggregate smaller workflows

~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model
  - Work with sites to meet experimental needs (high performance streaming over WAN?)
  - Use common data management/handling frameworks (e.g. Rucio) to prestage data
- HPC sites prefer large parallel workflows
  - Can use late-binding scheduling to aggregate smaller workflows
- HPC allocations – most are annual proposal awards, targeting solving specific scientific problems

HPC Awards - Example

ASCR Leadership Computing Challenge (ALCC)

Current Awards

- **Eric Lancon** (Brookhaven National Laboratory), **Taylor Childers** (Argonne National Laboratory), **Paolo Calafiura** (Lawrence Berkeley National Laboratory) received 80,000,000 core-hours on Theta and 80,000,000 core-hours on Titan for "Scaling LHC proton-proton collision simulations in the ATLAS detector."

~~Cloud~~ HPC Challenges

- HPC facilities generally are built for a different compute/data model
  - Work with sites to meet experimental needs (high performance streaming over WAN?)
  - Use common data management/handling frameworks (e.g. Rucio) to prestage data
- HPC sites prefer large parallel workflows
  - Can use late-binding scheduling to aggregate smaller workflows
- HPC allocations – most are annual proposal awards, targeting solving specific scientific problems
  - Successes in this program – but we should work with agencies and facilities to establish a better fitting path for HEP