ALICE USA Operations and Plans
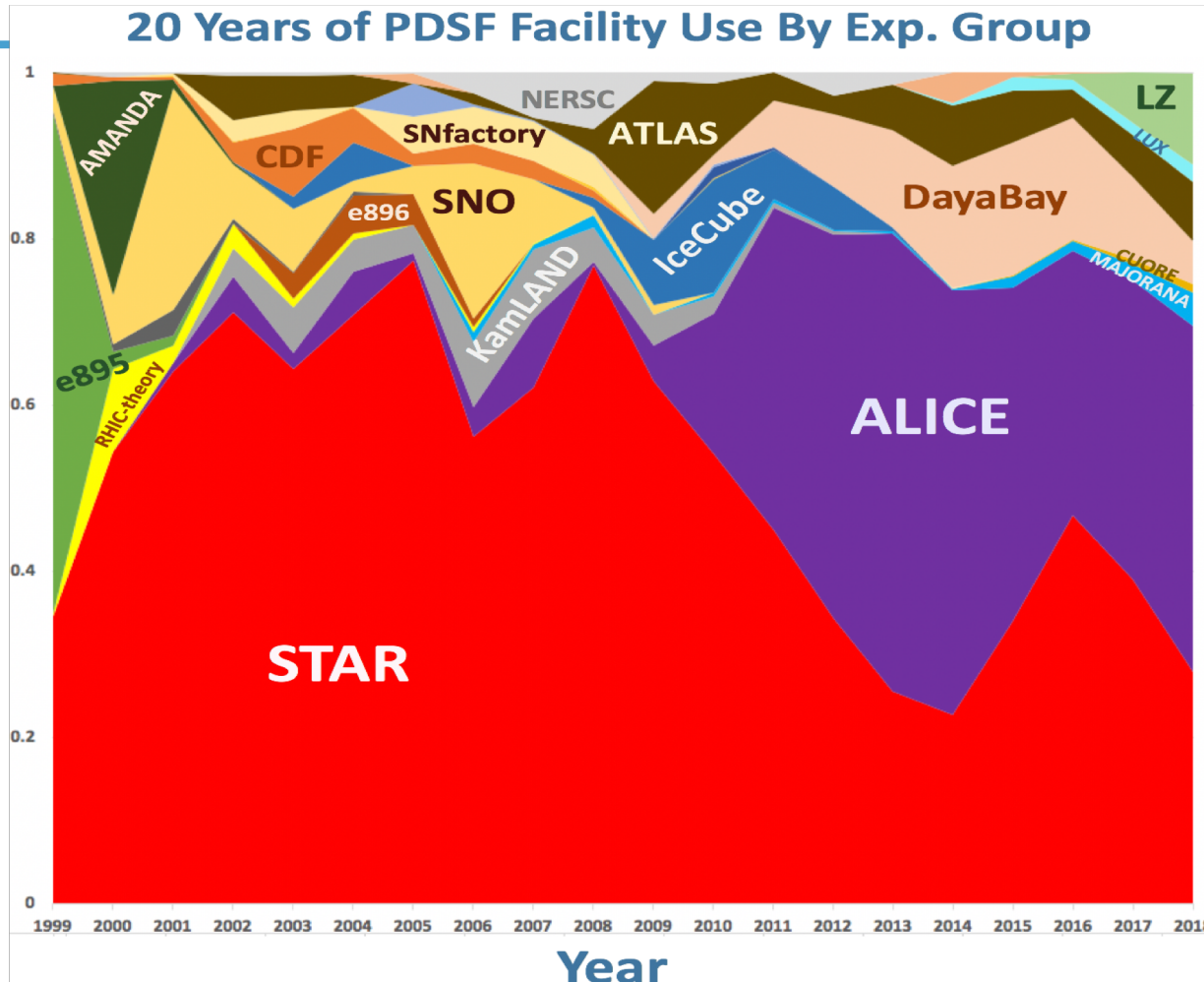
R. Jeff Porter (LBNL), Michael Galloway (ORNL), John White (LBNL)

ALICE T1/T2 Workshop
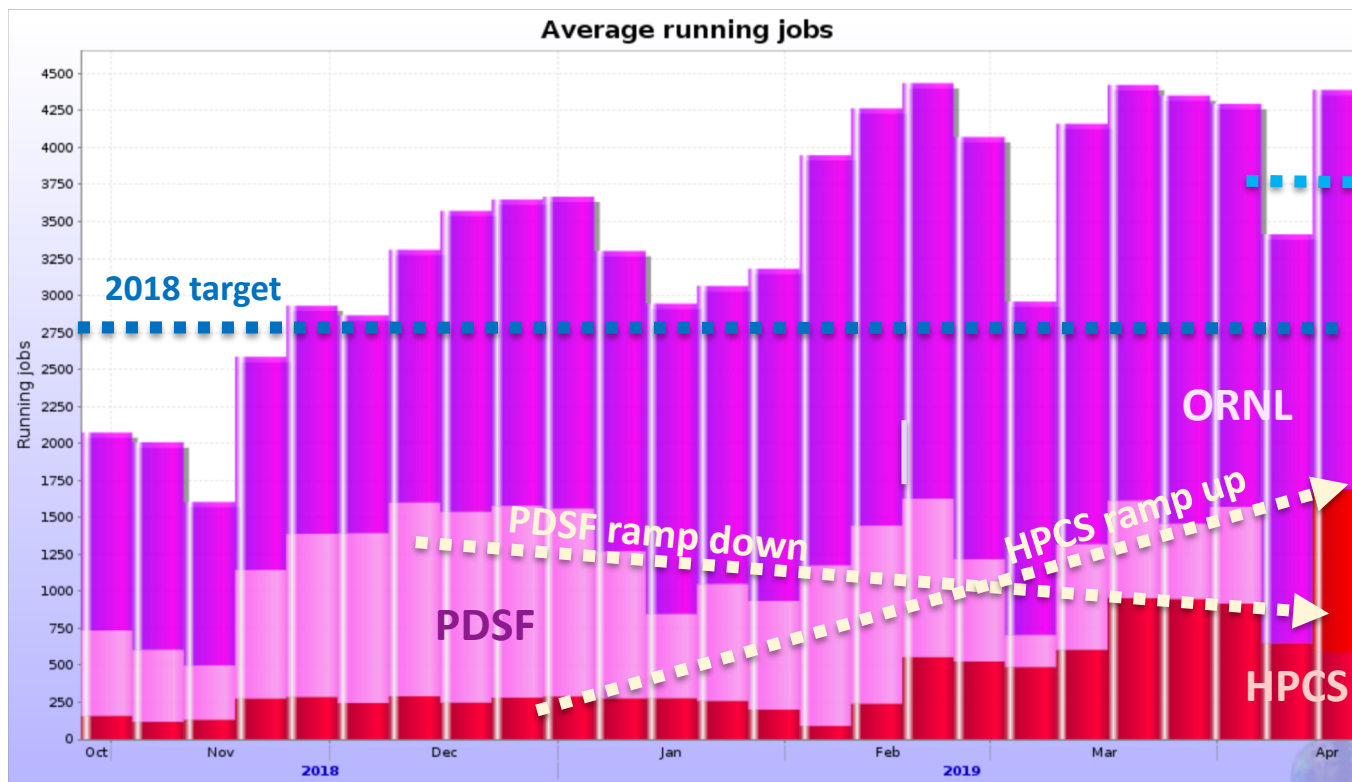
May 14, 2019

ALICE-USA Computing Project

- **Original facility in 2010:** **NERSC/PDSF @ LBNL and Livermore Computing @ LLNL**

- **Project changes in 2014**
  - LLNL withdraws from ALICE-USA – we replaced LLNL/LC with ORNL/CADES T2

- **Project issues: 2017-2018**
  - Funds held in 2017 due to US budget
  - NERSC to decommission PDSF (done in 2019)
    - Suggest we migrate to NERSC HPC systems

- **Our approach – rebuild.**
  - Build new ALICE T2 @ LBNL HPCS
  - Modest use of NERSC HPC (for tomorrow)

- **Current Effort**
  - ALICE: J. Porter, Project management, AliEn site services
  - ORNL: P. Eby, ORNL::EOS storage, M. Galloway, cluster management
  - LBNL: J. White, LBNL_HPCS::EOS storage, K. Fernsler, cluster management

Average running jobs

concurrent jobs: 2010-2018

regular (20%) annual growth

ORNL

LLNL

LBNL/PDSF

LBNL/HPCS

20 Years of PDSF Facility Use By Exp. Group

- **20+ years supporting HEP/NP Experiments.**
  - News Article: https://newscenter.lbl.gov/2019/05/01/the-little-computer-cluster-that-could

2018 Job Occupancy

Average running jobs

ORNL
New nodes in Sept & Jan.
Batch issues lowered rates
early in year

LBL PDSF.
Node removal beginning in Dec
Shutdown Apr 1st 2019

LBL HPCS
New nodes: Nov, Feb, Mar
Will take ~30 PDSF nodes
Late May, early June

ALICE at LBNL (SCG)

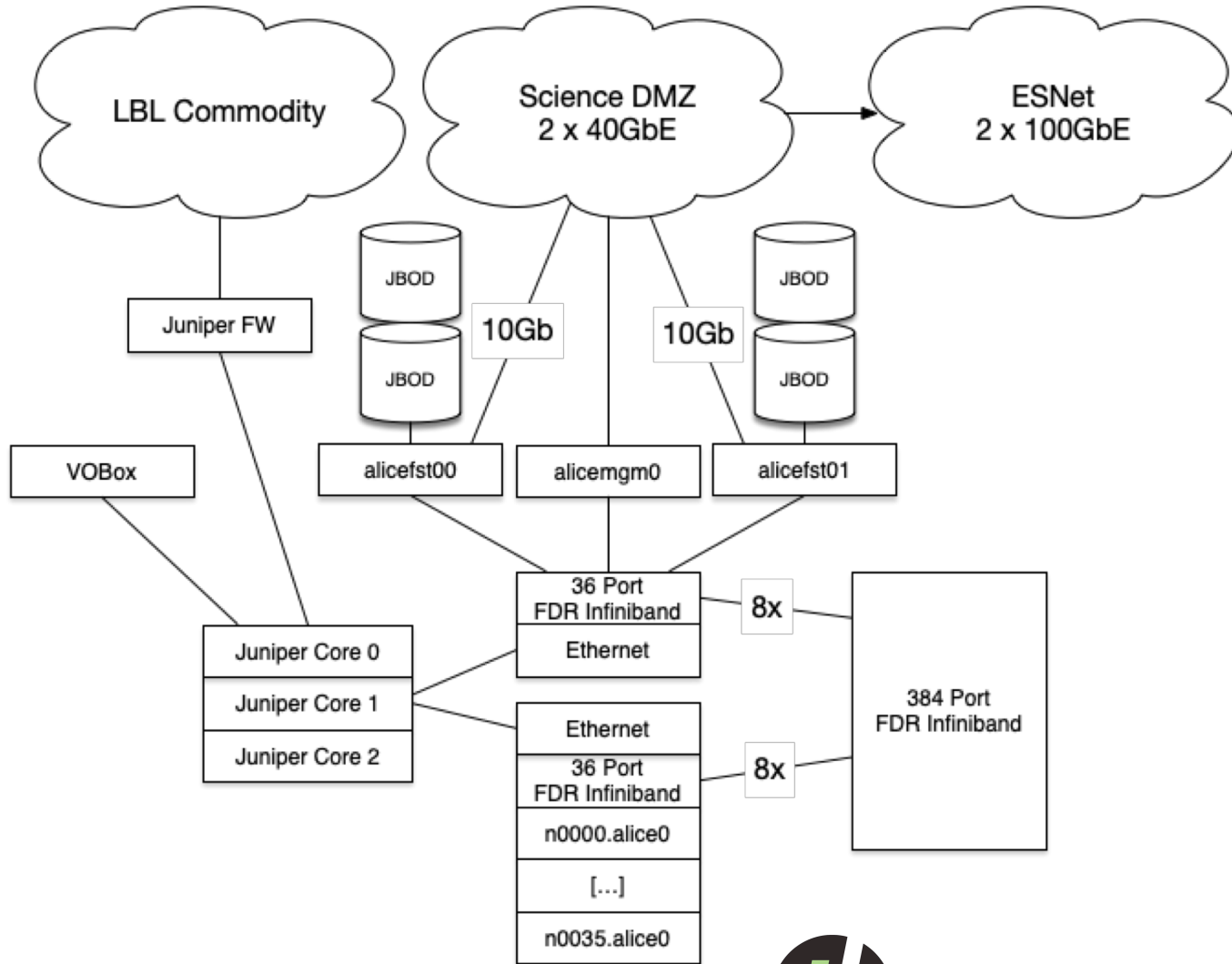## ALICE T2 Meeting 2019

John White
Lawrence Berkeley National Lab

Scientific Computing Group

- **SCG (previously HPCS)**
- **9 FTE Ranging from User Services to Engineers**
- **Core HPC Cluster business (LBNL + UC Berkeley)**
  - 2800+ nodes
- **~13PB Storage (Lustre, EOS, NAS, ZFS)**
- **2 FDR Fat Tree Fabrics (8:1), cross-linked**

SCG Network Layout

ALICE at LBNL

- **2.5PB EOS, ~1.6PB used**
  - 1 MGM, 2 FST (Stateless, Warewulf provisioned)
  - 250 Spindles, Single Disk FSID
    - Coincidentally timed deployment with ORNL ZFS performance Issues
    - Disk:FST ratio seems to be an issue
  - 10GbE external, 56GbIB internal (Per FST)
  - Recently resolved DNS trickery fixed efficiency numbers
  - Issues taking on PDSF data (3rd party transfers)

Compute Environment

- **36 CEs**
  - 4 First Gen, 32 2nd Gen
  - 1st Gen: 56core, ~2GB/core - E5-2680v4
  - 2nd Gen: 32core, ~3GB/core – Scalable Gold 6130
- **Slurm**
  - Heavy GRES usage
  - No per-job memory limits, MaxJobs per node (per generation)
  - Opportunities for LowPrio queue (scavenging cycles)

TODO

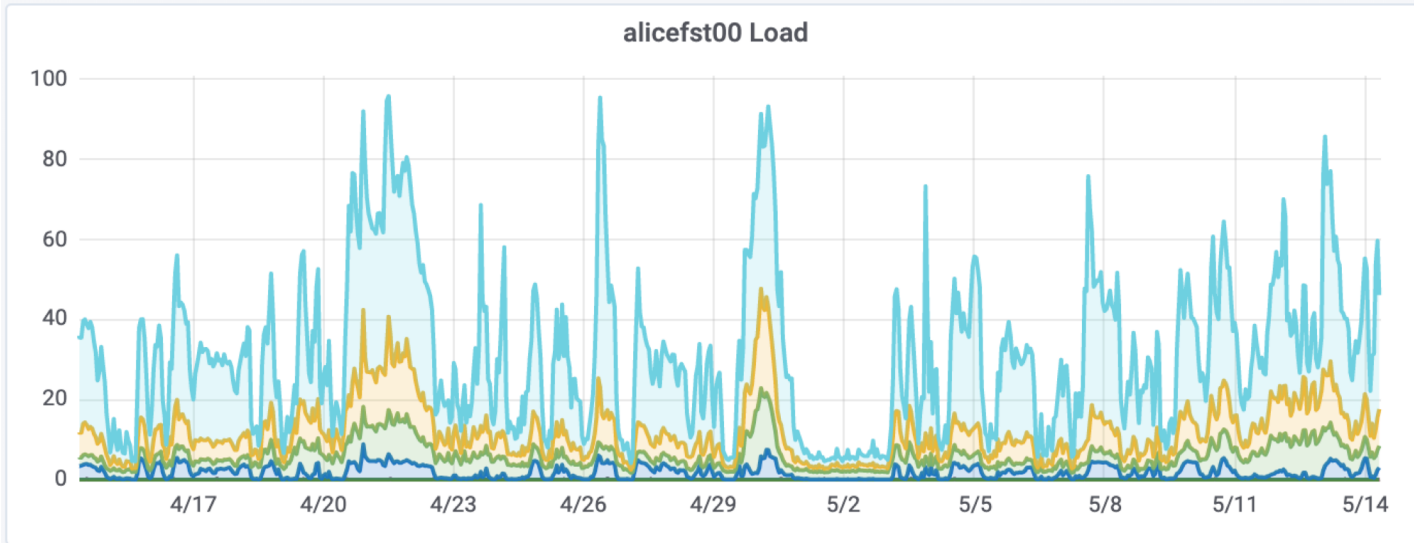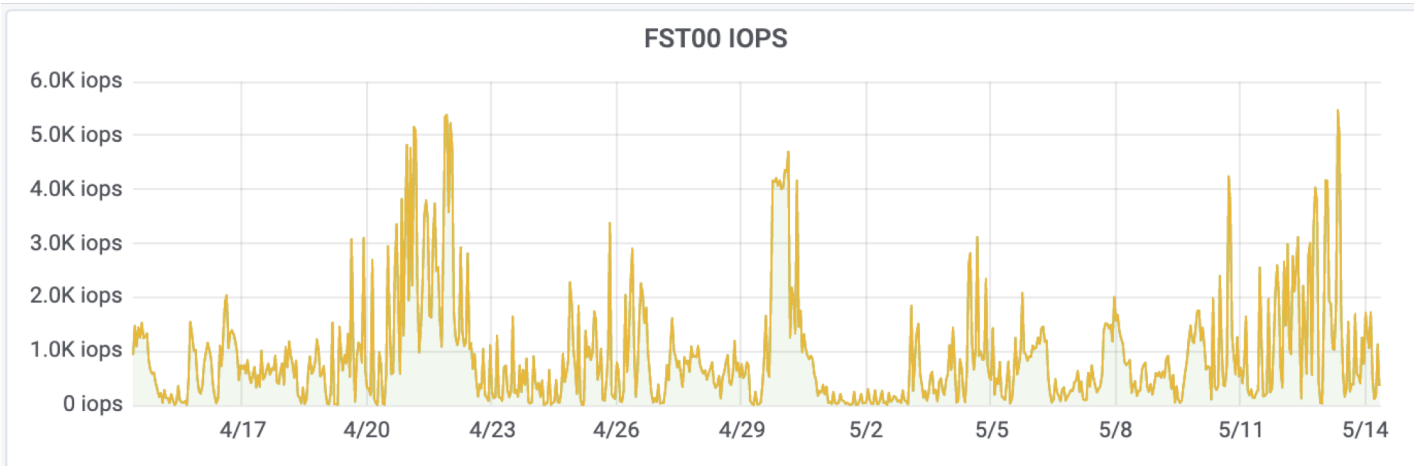- **LHCONE**
  - Waiting on FTE time from Networking group
  - Setting up a new AS from border through to EOS
  - V4/V6 would peer with ESNet, CENIC, UCB, Internal
    - 131.243.135.0/24
    - 2620:83:8000:4d00::/56
  - Blocking on several projects
    - ASA to SRX migrations
    - Zone Router Replacement
    - PBX Overhaul
    - Aruba controller upgrades

FST Load per IOPS

ORNL T2

- **Structurally Part of CADES**
  - CADES - Compute and Data Environment for Science
  - Mid Scale HPC, Storage & Cloud for ORNL R&D
    - ~20k Cores HPC (Open and Moderate PZ's)
      - Condo Environment
      - Largely CPU/Some GPU
    - 56Gb IB Fabric/10Gbe Fabric
    - 4PB Lustre
    - 2PB NFS
    - 6k VCPU's of Openstack Cloud
  - NCCS Manages OLCF (Titan/Summit/Frontier)
  - Staffed by:
    - Pete Eby/Michael Galloway/Susan Hicks/Steve Moulton

Jeff Porter LBNL

ORNL T2

- **Compute Element Current Status**
  - ALICE Dedicated
  - 2940 Slots (100 Nodes)
  - Mix of Hardware
  - Mixed 10Gbe/1Gbe Fabric
  - Torque/Maui Scheduling
    - Oddities with our Scheduling env

- **2019-2020 Planning**
  - Migration to Slurm
  - Worker Node Migration to CentOS7

Jeff Porter LBNL

ORNL T2

- **Storage Element Current Status**
  - EOS Citrine
  - 2 New FST's and 2 New JBODS (2 60x10T)
    - Single Drive FSID's
    - FST's have 20Gb Link to Fabic
  - 4 Older FST's and 5 Older (4 60xT 1 60x10T)

- **2019-2020 SE Planning**
  - Continue Migration from ZFS to Single Drive FSID's
  - EOS Services Migration
  - Prep for EOL of Original 1.5P

Jeff Porter LBNL

ORNL T2

- **Networking**
  - Connected to ORNL Border Router
    - Outside ORNL Firewall
  - Completed Migration to LHCone
  - 40Gb ESNET Uplink
    - Considering 100Gb Uplink
  - IPV6 Configured on Network
    - Still Planning Site Migration
  - Network Hardware Modernization

Jeff Porter LBNL

ORNL T2

- **Site Services**
  - Hypervisor and EOS Server Upgrades to Newer Hardware
  - VOBOX Upgrade
  - Grafana Metrics
    - Job Mix Monitor
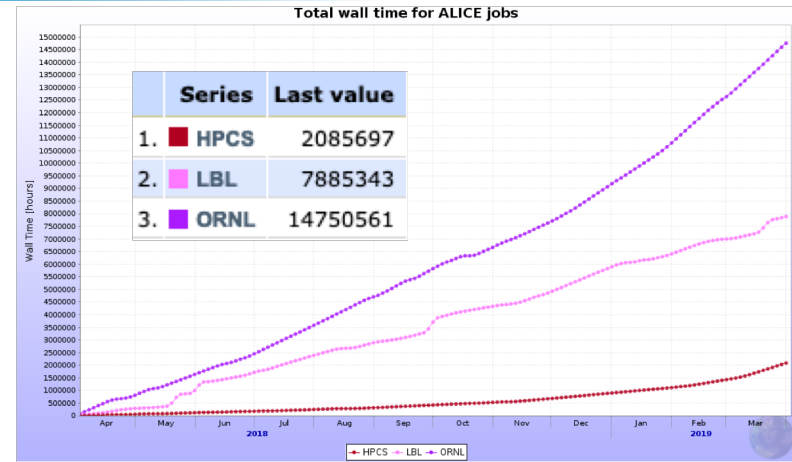 - Continue Config Management (Ansible) Development

Jeff Porter LBNL

2018 CPU Delivery Relative to Obligations

Relative to 2018 Project Plan

CPU Obligations	kHS06
ALICE-USA	43.5
LBNL T2	21.5
ORNL T2	22.0*

*ORNL cannot pledge w/o MoU

Total wall time for ALICE jobs

Series	Last value
1. HPCS	2085697
2. LBL	7885343
3. ORNL	14750561

Walltime delivered RRB-2018

USA T2 Site	Per/core CPU (HS06/slot)	Wall-time delivered (MHS06-hrs)	ALICE-USA obligation* (MHS06-hrs)	% obligation delivered
LBNL:PDSF+HPCS	18.8+16.9	148.2+35.2=183.4	178.9	102%
ORNL	12.0	177.0	183.0	97%
Total		360.4	361.9	99.6%

Issue: CPU deployment & utilization was not as planned, thus we underperformed early and overperformed late

* includes 95% uptime

2018 CPU Efficiency

- **ORNL**
  - Periods of low eff. correlated with restructuring ORNL::EOS

- **LBNL**
  - Eff. is similar to other T2s.

- **HPCS**
  - Early results track LBNL
  - Extreme low efficiency since summer correlated to very high rate of <u>analysis jobs</u> ???

PDSF = 75%
ORNL = 70%
HPCS = 55%

HPCS trend line?

Jobs efficiency (cpu time / wall time)

HPCS   LBL   ORNL

Similar ALICE T2 = 75%
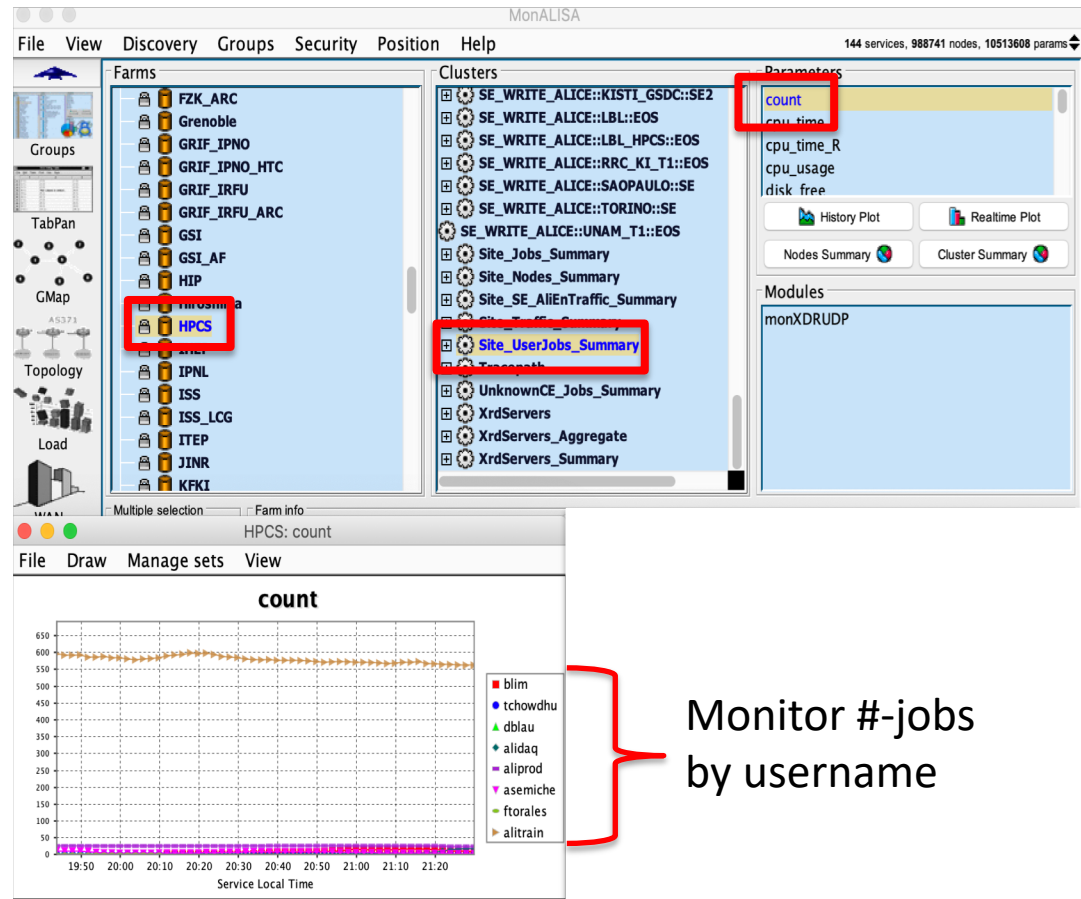
JobMix Monitor

- **Java ML Client**
  - Example client from Costin
  - Modified to sample jobs/site/user
    - Count, vmem, rss
    - Supports list of sites
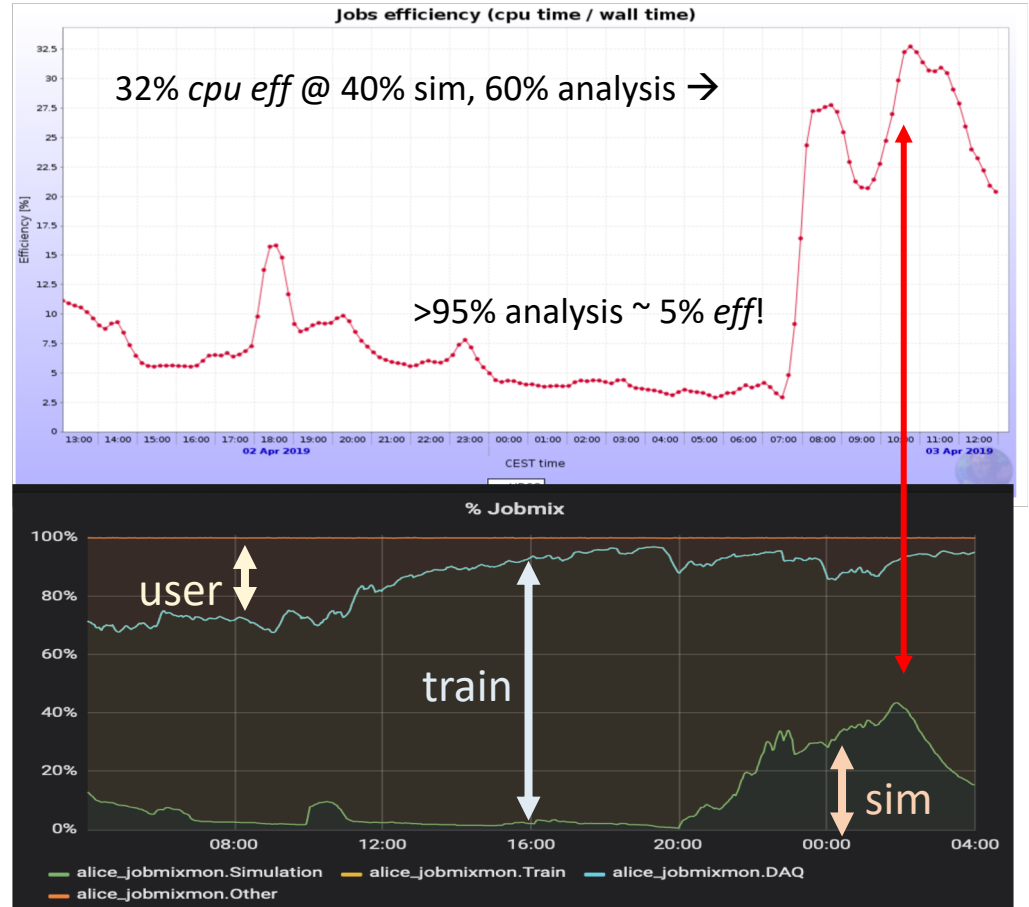 - Run as daemon with snapshots to local file

- **Python loader**
  - Reads snapshot and sums results per user type
    - Aliprod, alitrain, any-user
  - Summary per site to stdout
  - Output piped to Grafana using standard site method

Monitor #-jobs by username

JobMix monitor continued

- **Confirmed:**
  - HPCS had too many analysis jobs
  - Eff was tanking beyond expectations

- **Problem was name mismatch:**
  - Site Name = HPCS
  - EOS Name= LBNL_HPCS
  - Jobs would land and pull data from remote storages

- **Will continue to gather data to map our site expectations**

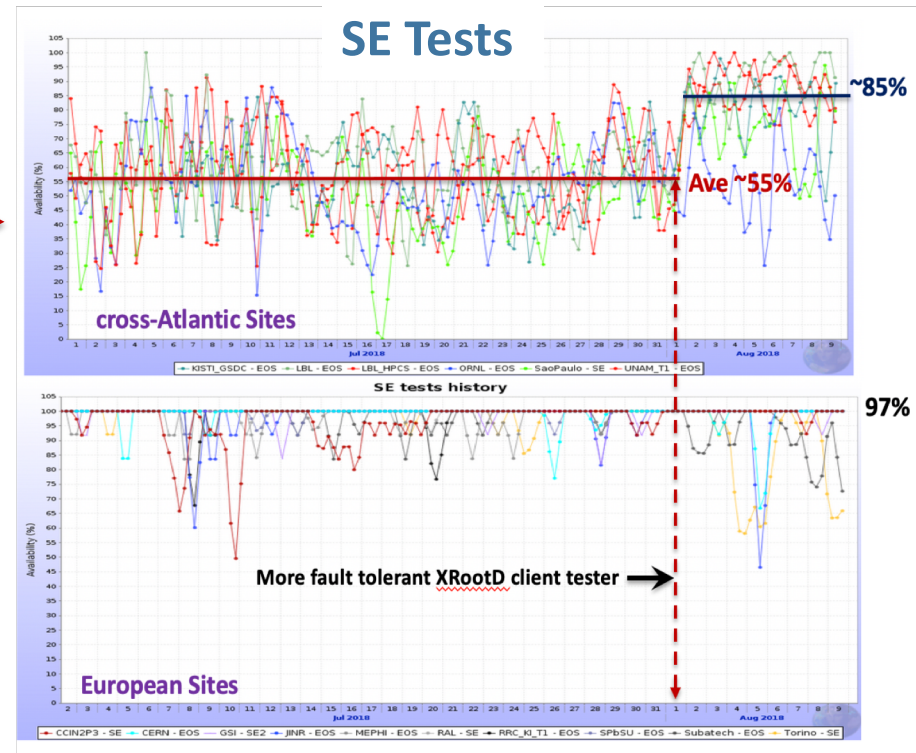- **Monitor also dumps out name & timestamp of large memory jobs**

32% *cpu eff* @ 40% sim, 60% analysis →

>95% analysis ~ 5% *eff*!

collaboration welcome!
https://github.com/alice-us-grid-sites/job-mix-monitor

- 20 -

Jeff Porter LBNL

Network monitoring, with perfsonar?

- **Three troubling episodes that suggested network problems**
  - Difficulty emptying ORNL::EOS
  - Cross-Atlantic SE test failures
  - Odd data transfers to HPCS

We still don't understand sudden drop

1 GB/s

HPCS

CCIN2P3

CERN

SE Tests

~85%

cross-Atlantic Sites

Ave ~55%

SE tests history

97%

More fault tolerant XRootD client tester →

European Sites

One of 2 routes from CERN campus dropped packets

- We've added perfsonar instance at ORNL T2
- We will add one at LBNL T2
- OSG/WLCG can provide ALICE Dashboard
- Shall we (ALICE) pursue this more generally?

Jeff Porter LBNL
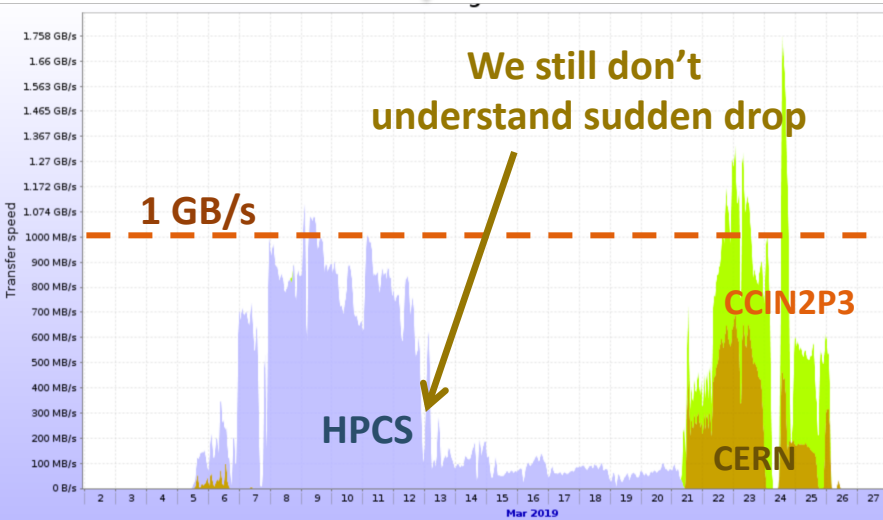
Resource Requirements & Procurement Planning

ALICE Computing Requirements vetted by the WLCG

ALICE		2018		2019				2020	
		CRSG recomm.	Pledged	Spring Request	Fall Request	2019 Fall req. /2018 CRSG	C-RSG recomm.	Fall Request	2020 req. / 2019 CRSG
CPU	Tier-0	350	350	430	350	100%	430	430	100%
	Tier-1	307	280	365	365	119%	365	365	100%
	Tier-2	313	313	376	376	120%	376	376	100%
	HLT					n/a	n/a	n/a	n/a
	Total	970	942	1171	1091	112%	1171	1171	100%
	Others								
Disk	Tier-0	26.2	26.2	34.3	31.2	119%	34.3	37.4	109%
	Tier-1	30.5	30.4	37.9	41.0	134%	37.9	49.2	130%
	Tier-2	29.0	29.0	33.9	33.9	117%	33.9	39.0	
	Total	85.7	85.6	106.1	106.1	124%	106.1	125.6	

% ALICE-USA participation

ALICE-USA Resource Deployment Plan ➔ In annual PEAP update

Resource	Installed	FY2019	FY2020
ALICE-USA Obligations			
CPU (kHS06)	43.4	51.9	53.4
Disk (PB)	3.9	4.8	6.1
ALICE-USA Plan			
CPU (kHS06)	44.0	53.0	53.0
% CPU obligation	101%	102%	100%
Disk (PB)	4.2	4.9	6.1
% Disk obligation	107%	100%	100%

DOE requests we don't overprovision

Jeff Porter LBNL

U.S. DEPARTMENT OF ENERGY | Office of Science

Overview of 2019 Project Plan

- **Maintain facility at ORNL/CADES**
  - CPU slots & 2.x PB Storage
  - Extra CPU adds safety margin!!

- **Expand new LBNL/HPCS Facility**
  - 300 CPU slots → 1200 CPU slots
  - 1000 CPU slots coming from PDSF
  - 1.2PB → 2.5 PB storage
  - Can add another 1.2 PB storage

- **Modest use of NERSC HPC**
  - ~100 CPU slots
  - Prep for next large system: 2020
  - Discuss tomorrow

- **Submitted to DOE, Nov 9**
  - Reviewed/approved at quarterly DOE call

2019 PEAP Update

Resource	Installed	FY2019	FY2020
ALICE-USA Obligations			
CPU (kHS06)	43.4	51.9	53.4
Disk (PB)	3.9	4.8	6.1
ALICE-USA Plan			
CPU (kHS06)	44.0	53.0	53.0
% CPU obligation	101%	102%	100%
Disk (PB)	4.2	4.9	6.1
% Disk obligation	107%	100%	100%

U.S. DEPARTMENT OF ENERGY | Office of Science

Jeff Porter LBNL