

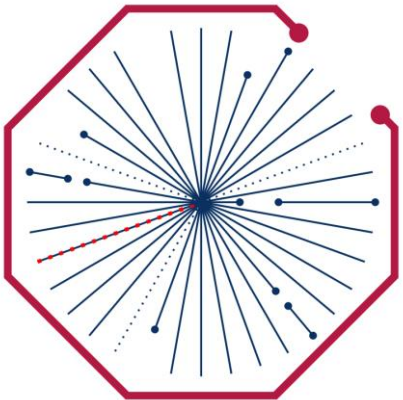


U.S. DEPARTMENT OF
ENERGY

Office of Science



ALICE-USA Computing Project Status and PEAP 2023 Update



ALICE-USA Computing Meeting @ ORNL

December 7, 2022

Irakli Chakaberia

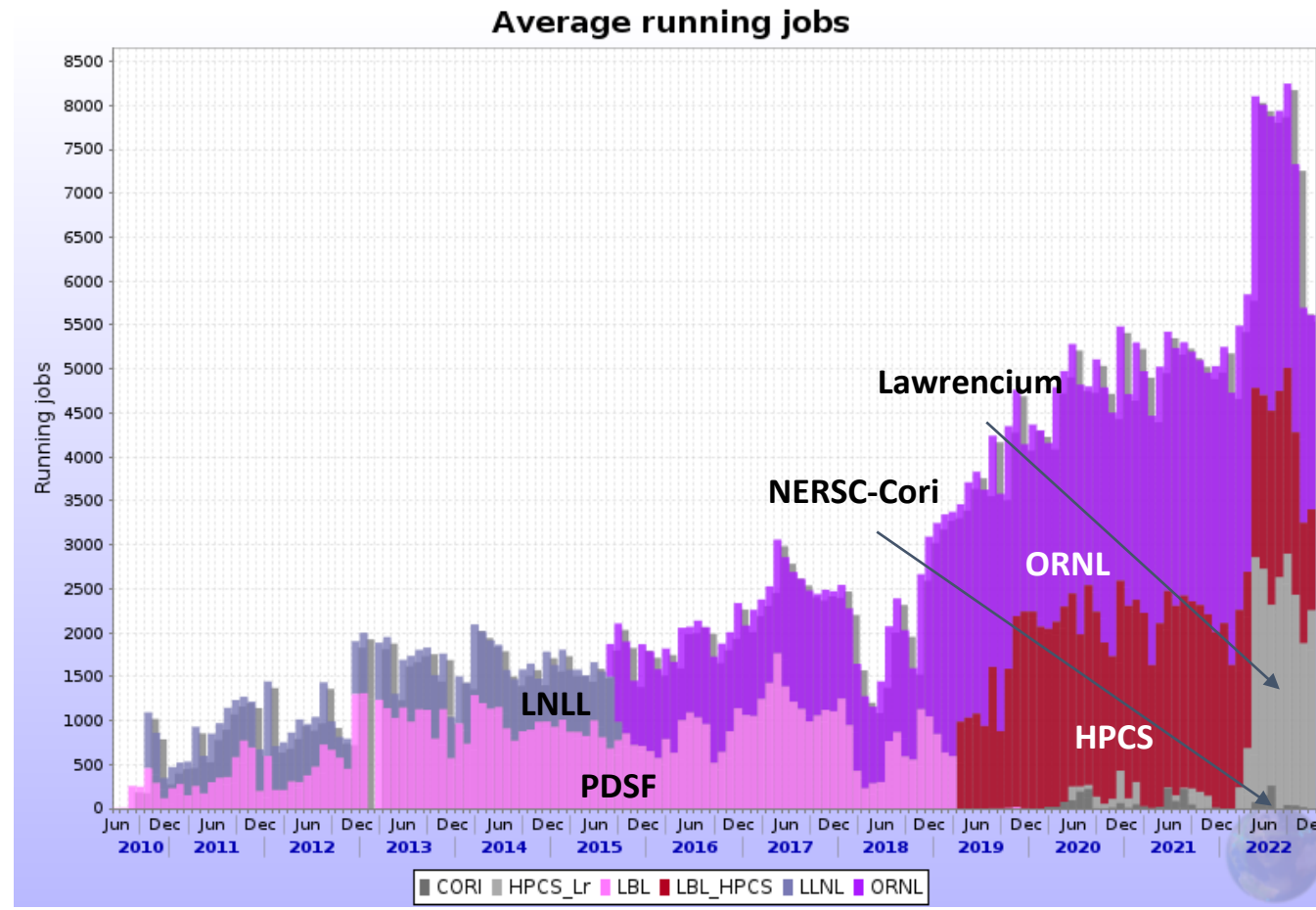


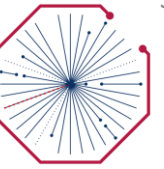
ALICE



Brief History

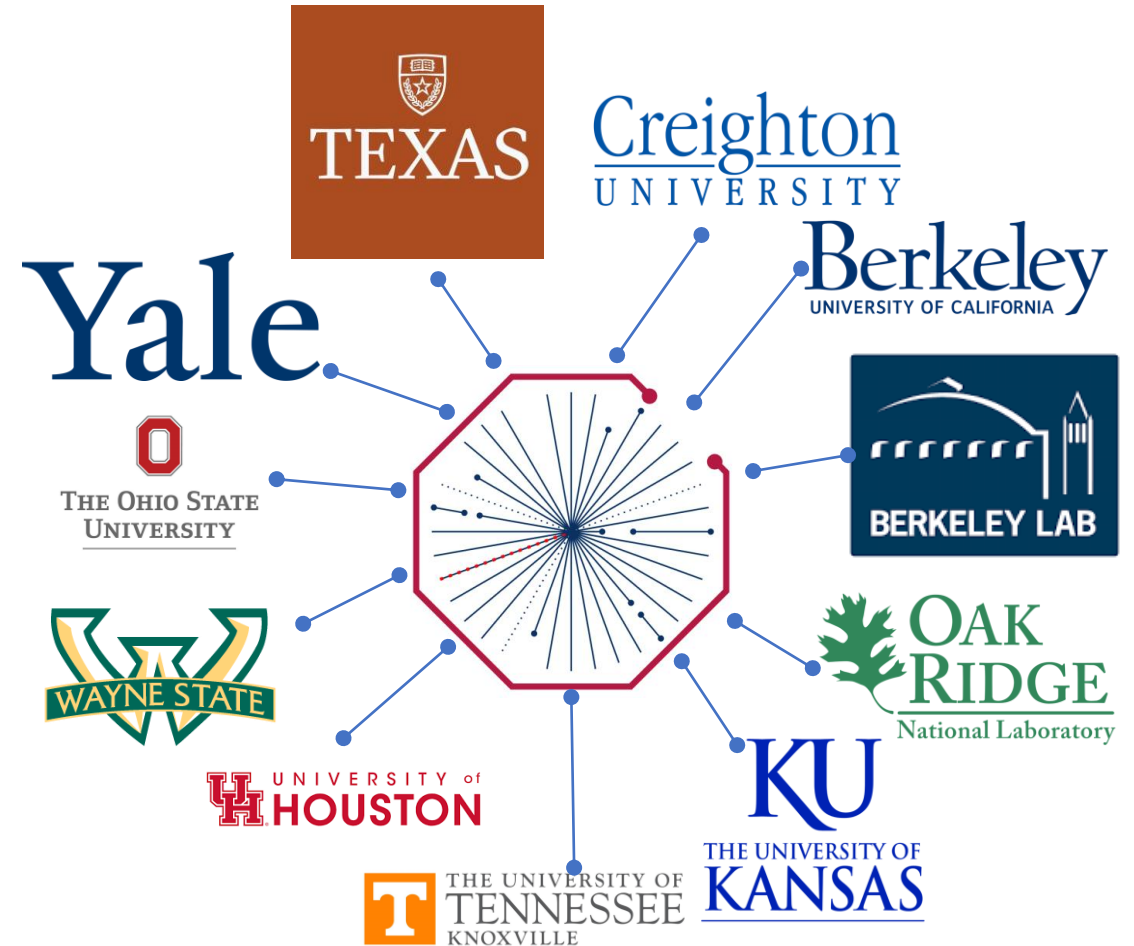
- Original Project Proposal - 2009
- In operational since 2010
- 3-Year Project Review Cycle
 - Initial plan requested in 2014
 - Latest review in July 2021
- Project Execution & Acquisition Plan (PEAP)
 - Multi-year plan containing resource delivery, milestones, budget estimates, operations guide
 - updated annually & submitted to DOE
 - Latest one submitted in October 2023
- Website: <https://sites.google.com/lbl.gov/alice-usa-computing/meetings?authuser=1>
 - I am adding useful info slowly but steadily





US Department of Energy Supported ALICE-USA Group

- The project's main goal is to fulfill DOE funded MoU-based ALICE USA obligations for
 - 11 Institutions
 - 46 M&OA
- Operates ALICE grid facilities at 2 DOE labs
- ALICE-USA Computing core group meets monthly, every third Tuesday of the Month
- We keep very close contact on our SLACK channel
- We try to hold two annual ALICE-USA Computing meetings
 - Spring meeting at LBL
 - Fall meeting at ORNL





US-ALICE Review Meetings

- External project review in July 2021
- Received favorable report

- US Operations review in May 2022

External DOE Review of LICE-USA Commuting Project

29-30 July 2021
Online
America/Los_Angeles timezone

- Overview
 - Timetable**
 - Scientific Programme
 - Contribution List
 - My Conference
 - My Contributions
 - Participant List
 - Organization
 - Contact
- iraklic@lbl.gov

Timetable

Thu 29/07 Fri 30/07 All days

Print PDF Full screen Detailed view Filter

07:00	Executive Session	
	Online	07:30 - 08:00
08:00	Introductions	Jefferson Porter
	Online	08:00 - 08:05
	ALICE-USA	Mateusz Ploskon

US-ALICE Grid Operations Review

4-6 May 2022
Lawrence Berkeley National Laboratory
America/Los_Angeles timezone

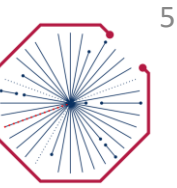
- Overview**
- Timetable
- Contribution List
- My Conference
- My Contributions
- Participant List
- Videoconference Rooms



After two years we get back together at the Lawrence Berkeley National Laboratory for the annual ALICE USA Grid Operations overview meeting.

Starts 4 May 2022, 07:00
Ends 6 May 2022, 19:00
America/Los_Angeles

Lawrence Berkeley National Laboratory
70-191
LBNL BLDG 70
[Go to map](#)



Status of operations | RRB 2022...



WLCG Summary

Target Availability for each site is 97.0%.

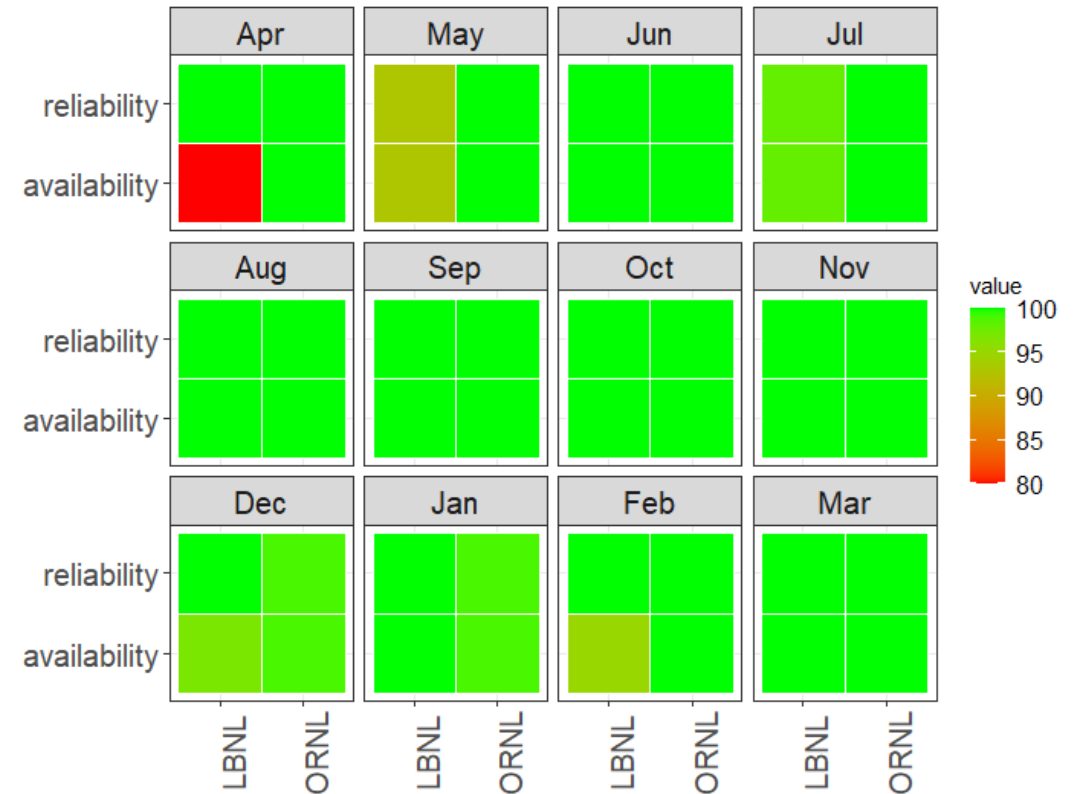
Availability Algorithm:

@ALICE_CE * @ALICE_VOBOX * all AliEn-SE

Since April 2022 we have reported only 24 hours downtime for the LBL_HPCS site [in October].

Both parameters **100%** for both sites in RRB 2022 so far.

RRB 2021 Numbers

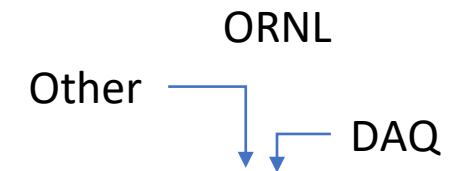
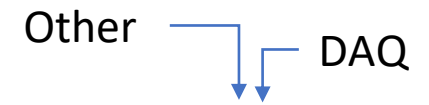
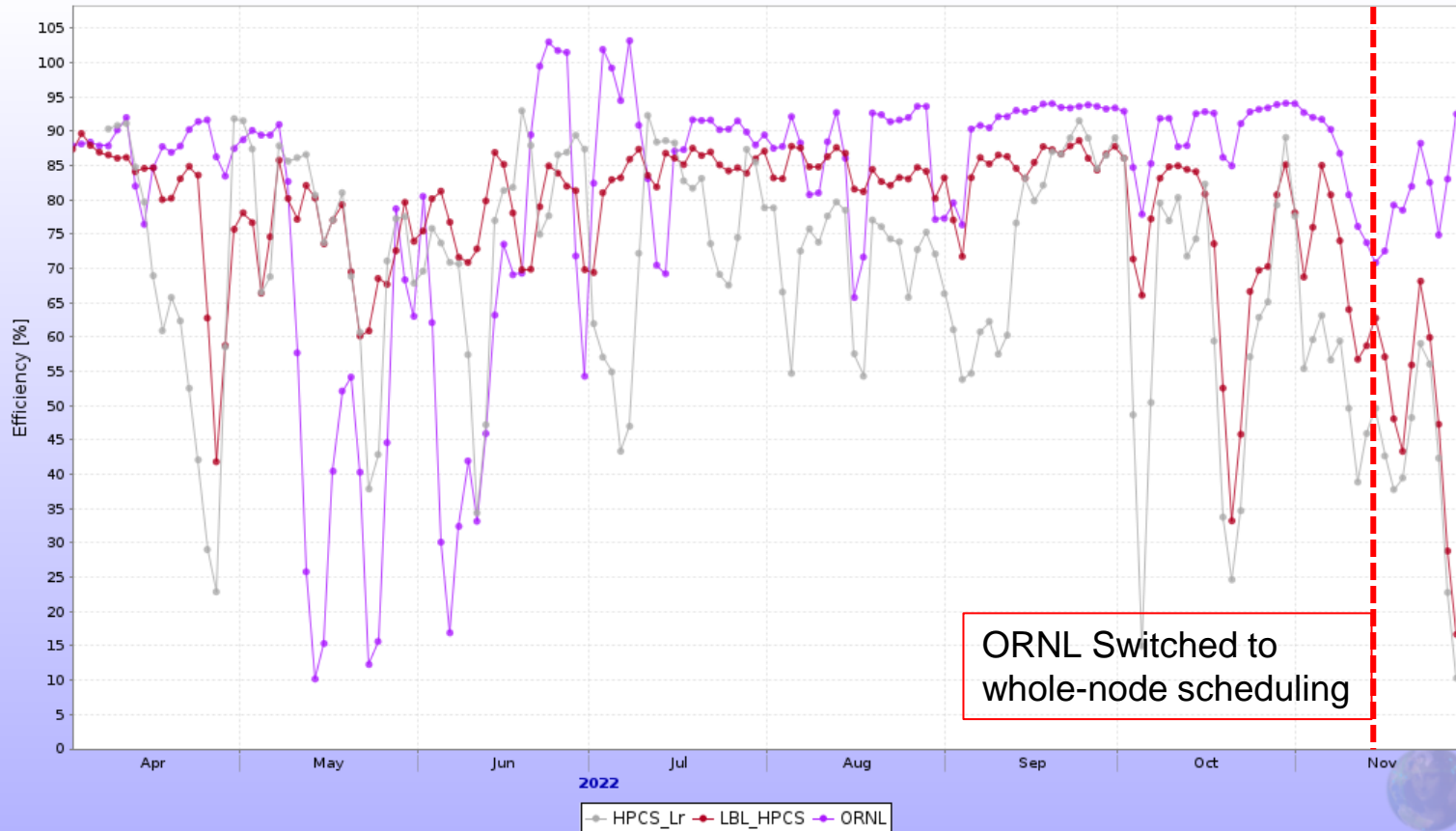




MonAlisa Summary | RRB 2022

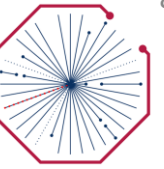
- Average job CPU efficiency over this period is about 70% at Lr, and about 80% on T2 sites
- JobMix defines the average CPU efficiency delivered

Jobs efficiency (cpu time / wall time)

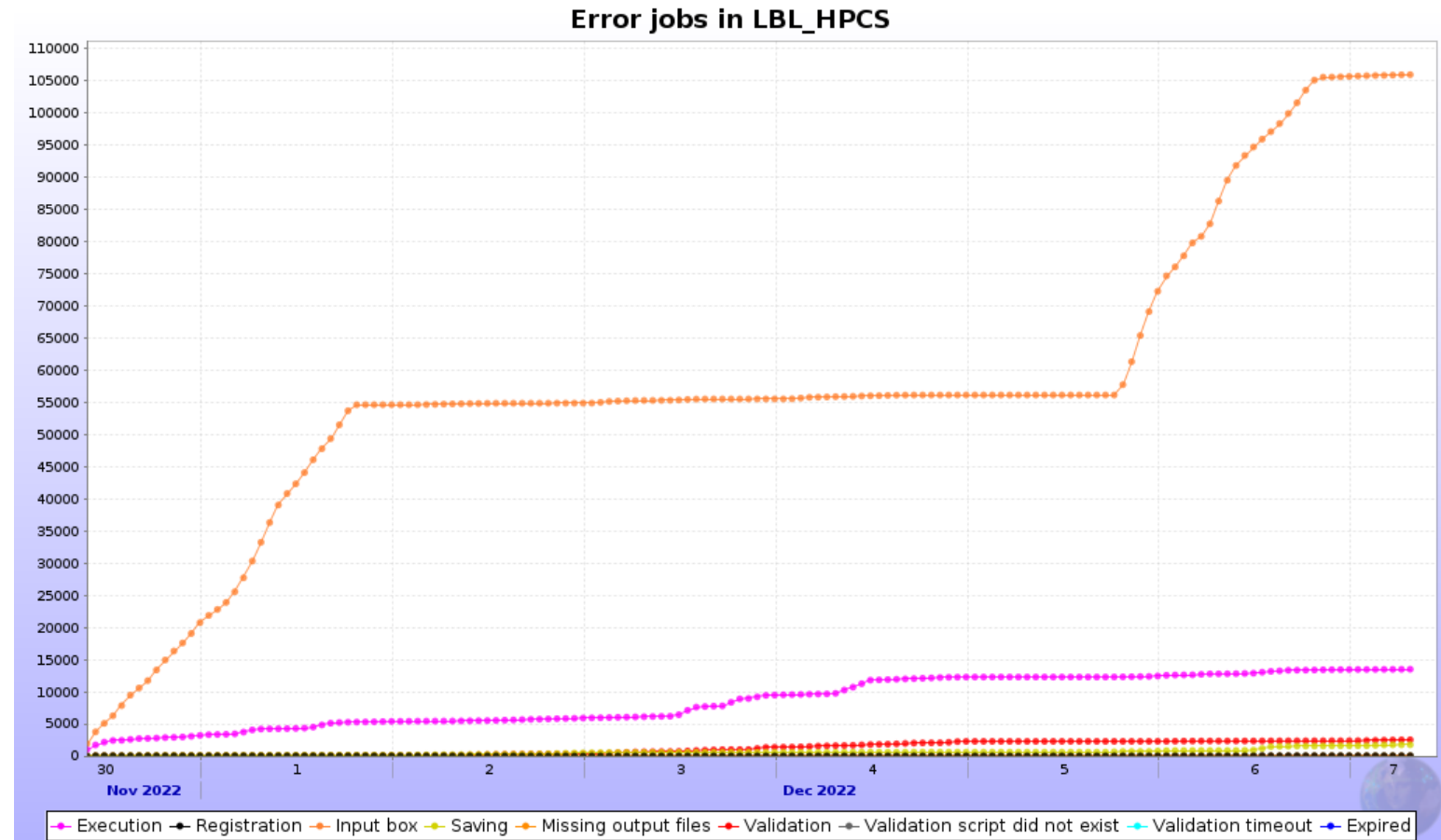


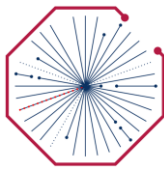
LBL_HPCS

Recent week



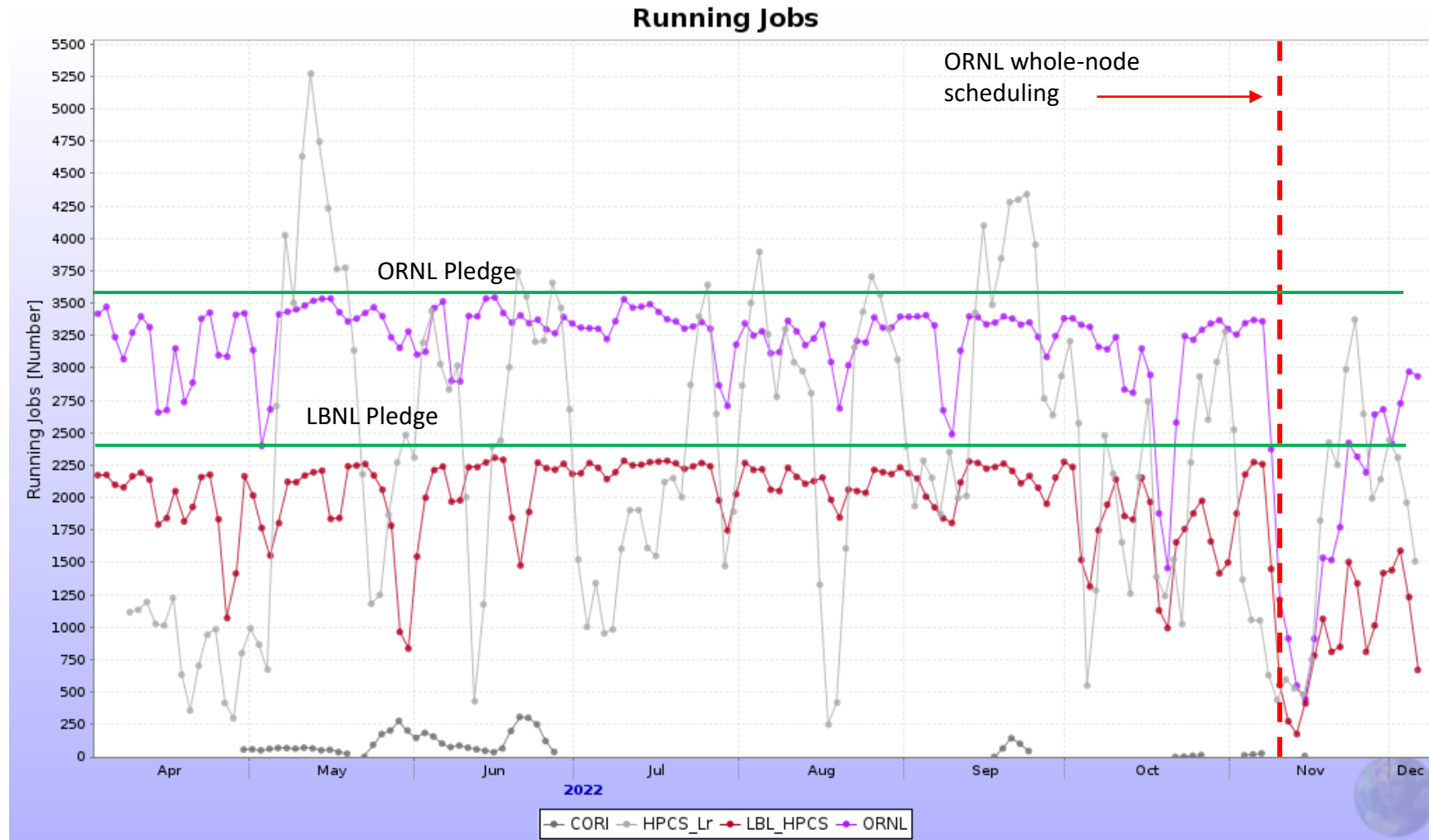
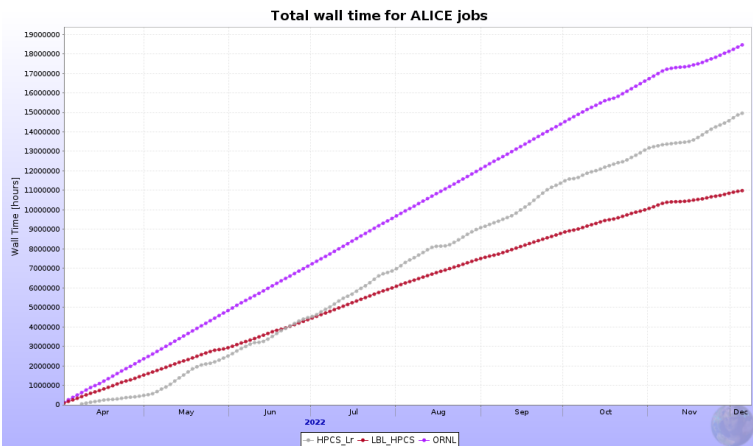
- As evidenced from the overall numbers, most of the issues are under control, but...
- One of the recurring issues in recent pas has been download issues of the OCDBSim.root [perhaps other large files too?!] file causing “input box errors”
- This issue is intermittent and usually goes away after some time
- Network issues? Too many accesses to the file?

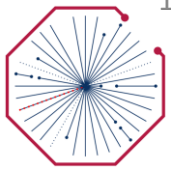




CPU Performance | RRB 2022

- Slightly under pledges
- Lawrenceium fills the gap
- Cori – R&D platform
- Almost no downtime





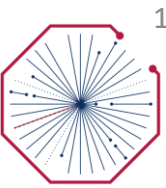
2022 RRB CPU Delivery vs. Obligations

- Due to some budget uncertainties the FY2021-2022 hardware acquisition was postponed and is in progress right now
 - LBNL requisitions are in
 - ORNL hardware is on site
- This causes slight under delivery for the ALICE-USA sites starting RRB2022
- The use of Lawrencium however picks up the slack
- Below are the number for the July-September 2022

CPU Obligations	
	kHS06
ALICE-USA	82
LBNL HPCS T2	41
ORNL T2	41

USA T2 Site	CPU/Core (HS06/Core)	CPU Delivered (MHS06*hr) ALICE-Monitored No Lawrencium	CPU Delivered (MHS06*hr) ALICE-Monitored	CPU Delivered (MHS06*hr) WLCG-Monitored	ALICE-USA Obligation (MHS06*hr)	Delivered per WLCG-Monitored Obligation (%)	Delivered per ALICE-Monitored Obligation (%) No Lawrencium
LBNL	16.7	73.9	189.6	227.1	90.5	250.8	81.6
ORNL	11.5	83.7	83.7	86.0	90.5	95.0	92.4
Total		157.5	273.3	313.0	181.1	172.9	87.0

Storage Performance | RRB 2022



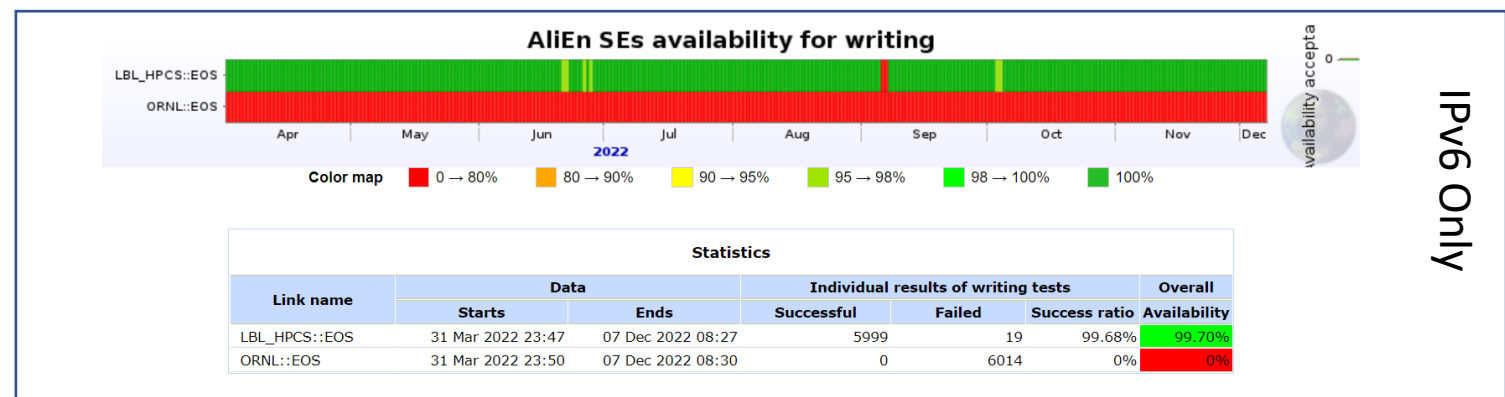
- SE availability – nearly perfect
- IPv6 has been up at LBNL since early RRB2021
- IPv6 at ORNL is looming on a horizon

READ WRITE

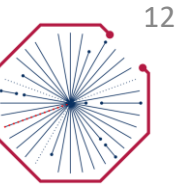
LBNL	99.7	99.8
ORNL	99.7	100



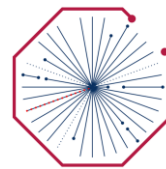
IPv4/IPv6



IPv6 Only



Project Execution and Acquisition Plan 2023

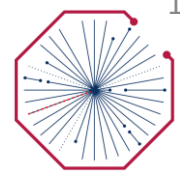


ALICE-USA Obligations and Acquisition Strategy

- ALICE-USA collaboration share of resources in respect to WLCG / C-RCG recommendations for the next few years and planned hardware acquisition strategy table

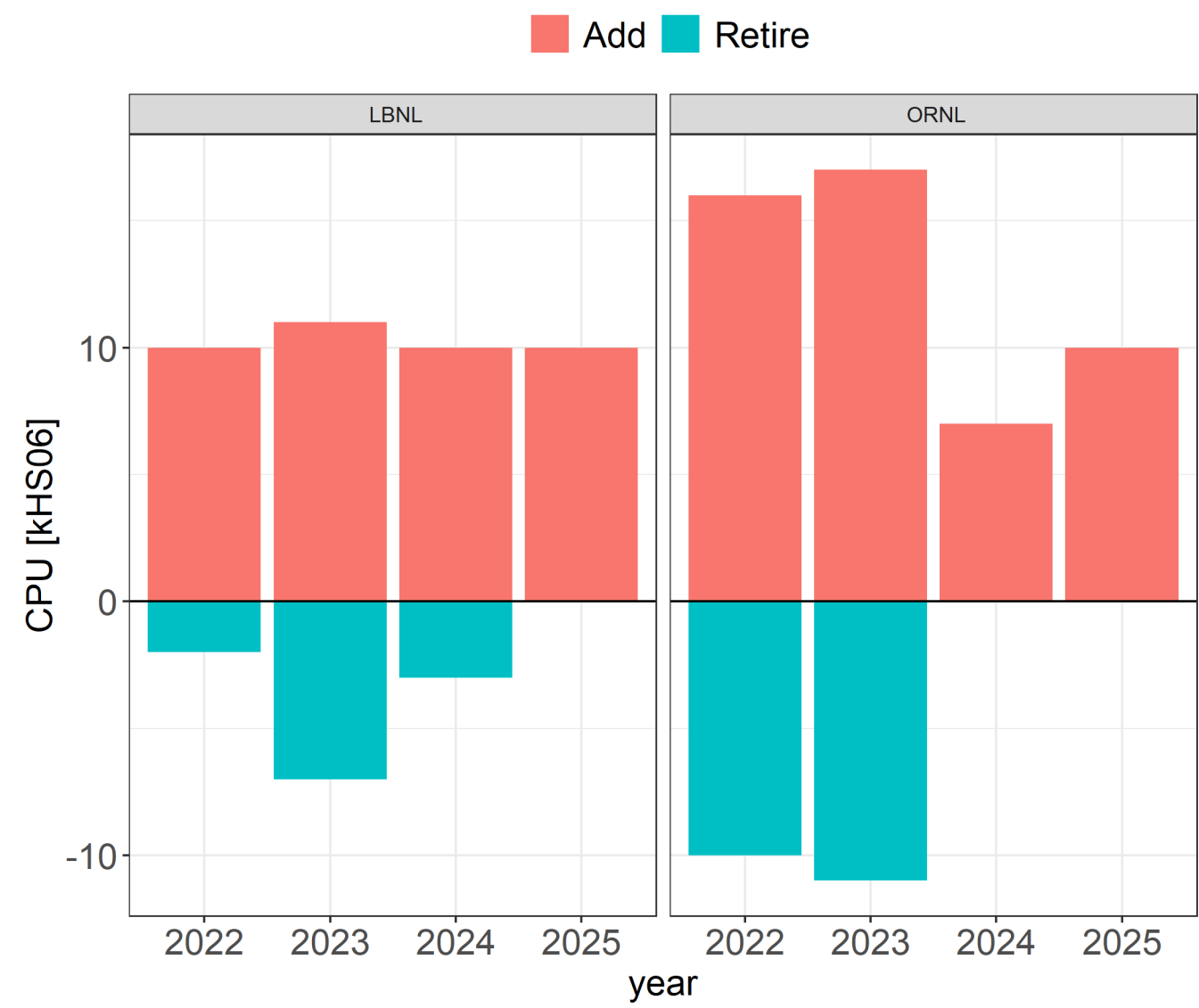
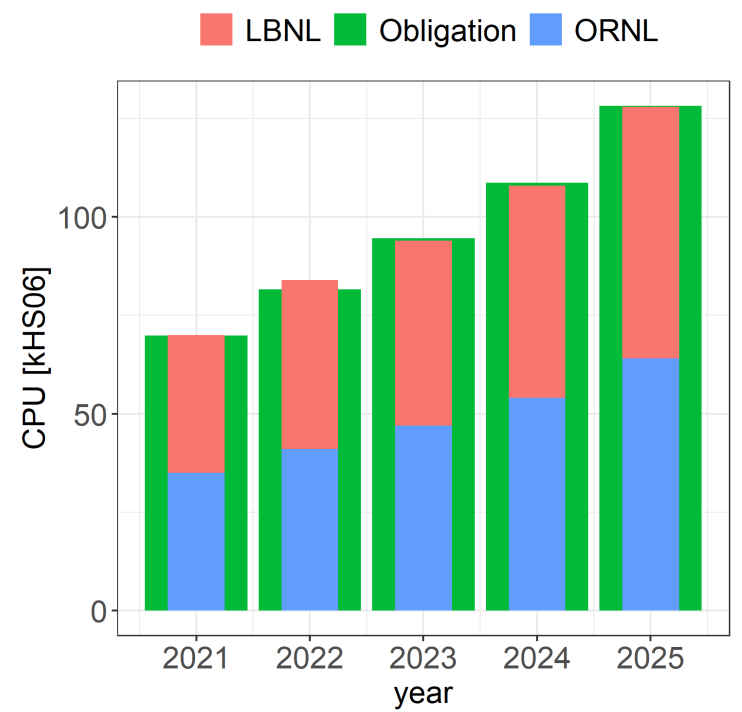
Year	FY2021	FY2022	FY2023	FY2024	FY2025
ALICE Requirements					
CPU (kHS06)	852	1013	1164	1338	1579
Disk (PB)	91	104	121	138	163
ALICE-USA Participation					
ALICE M&O-A	561	571	566	566	566
ALICE-USA M&O-A	46	46	46	46	46
ALICE-USA/ALICE (%)	8.20%	8.06%	8.13%	8.13%	8.13%
ALICE-USA Obligations					
CPU (kHS06)	69.9	81.6	94.6	108.7	128.3
Disk (PB)	7.5	8.4	9.8	11.2	13.2

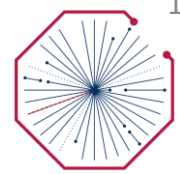
Resource	Installed	FY2022	FY2023	FY2024	FY2025
LBNL HW & Costs					
CPU change (+/- kHS06)		-2.0+10.0	-7.0+11.0	-3.0+10.0	0.0+10.0
CPU Installed (kHS06)	35	43	47	54	64
Disk change (+/- PB)		0.0+2.0	0.0+0.0	-1.6+2.1	0.0+0.9
Disk installed (PB)	2.75	4.75	4.75	5.25	6.15
ORNL HW & Costs					
CPU change (+/- kHS06)		-10.0+16.0	-11.0+17.0	0.0+7.0	0.0+10.0
CPU Installed (kHS06)	35	41	47	54	64
Disk change (+/- PB)		-2.75+3.0	0.0+2.0	0.0+1.2	0.0+0.9
Disk installed (PB)	2.75	3.00	5.00	6.20	7.10



CPU Acquisition

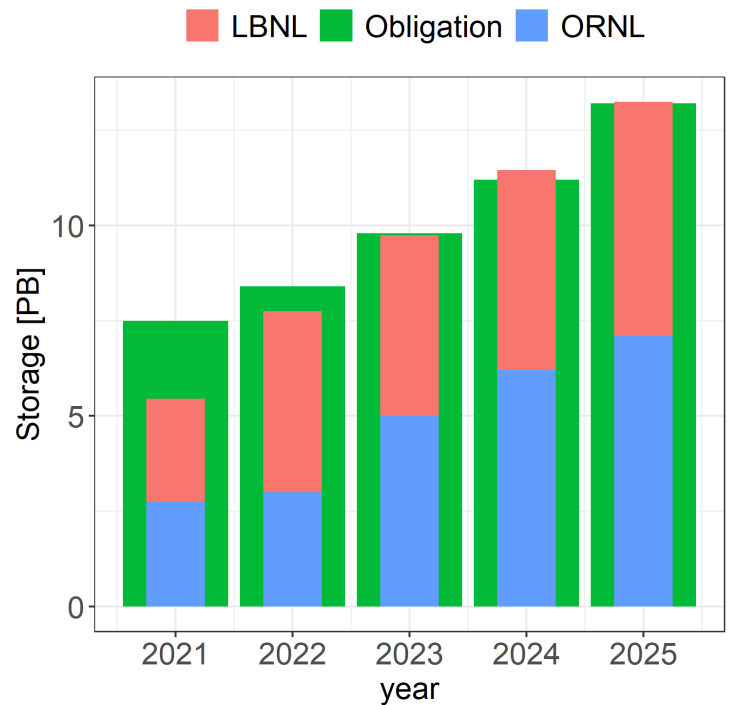
- We are gradually retiring out of warranty hardware
- HS06 for the planned CPU acquisition is roughly estimated based on the measurements we made on similar nodes





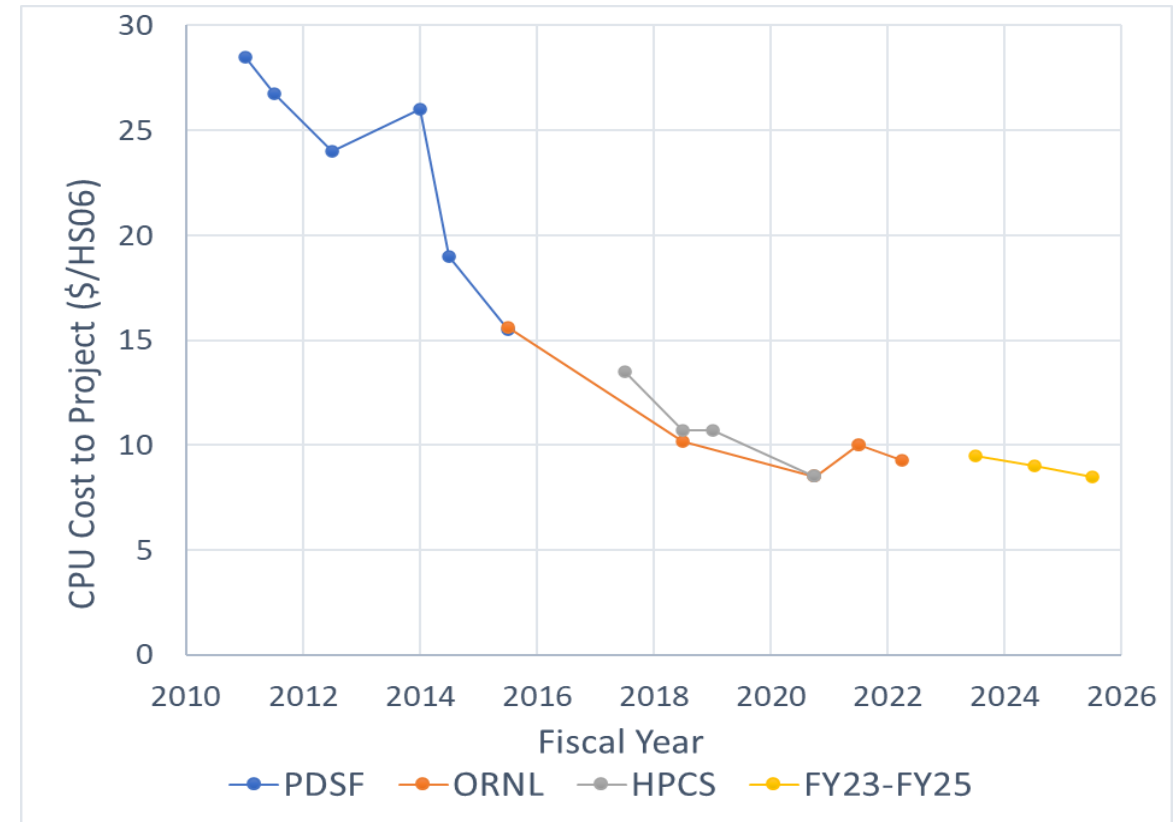
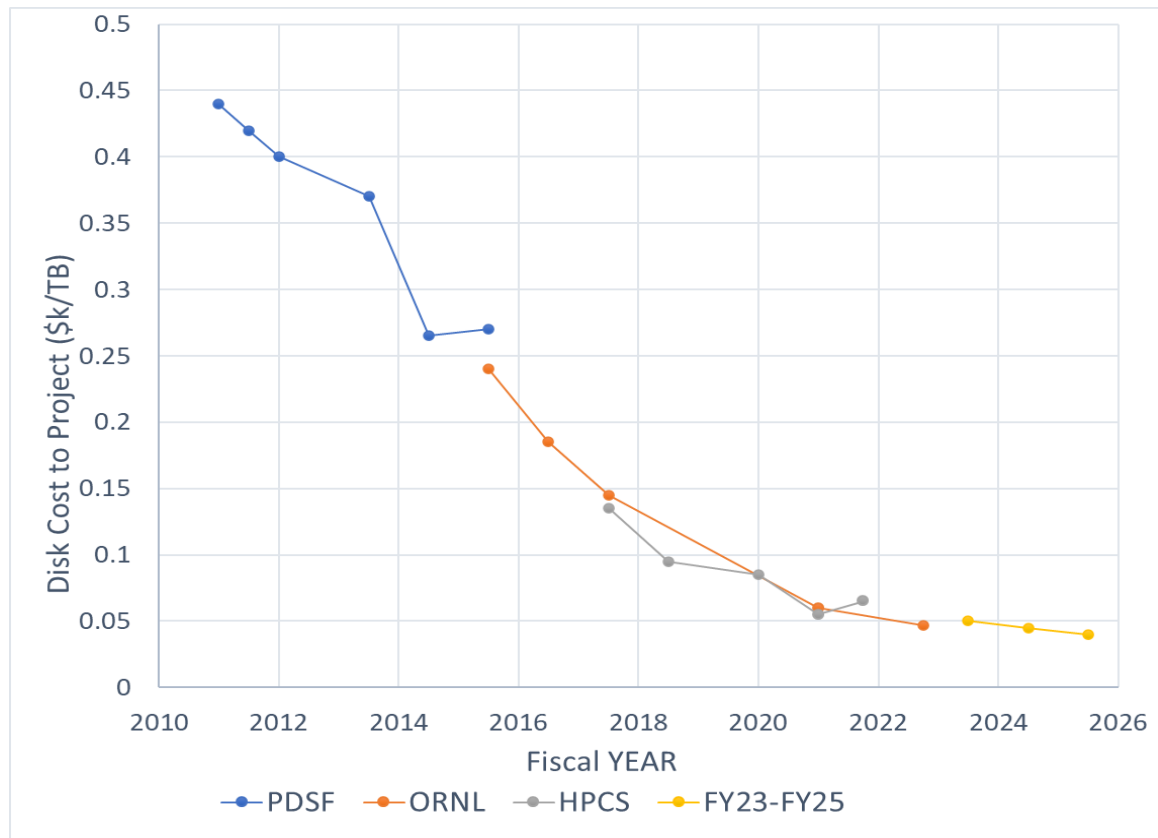
Storage Acquisition

- Our storage utilization has been on a lighter side
- We keep evaluating the utilization and purchase the new storage based on it and optimized for the available budget



Cost trends for the CPU and Storage

- CPU unit cost per HS06 is based on extrapolating the CPU clock frequency
- We'll measure the real HS06 values as soon as the new hardware is deployed



Delivered Resources by Year

- The result of the above-described plan is shown in this table

Resource	FY2021	FY2022	FY2023	FY2024	FY2025
ALICE-USA Obligations					
CPU (kHS06)	69.9	81.6	94.6	108.7	119.9
Disk (PB)	7.5	8.4	9.8	11.2	12.4
ALICE-USA Plan					
CPU (kHS06)	70.0	84.0	94.0	108.0	128.0
% CPU obligation	100.2%	102.9%	99.4%	99.3%	106.7%
Disk (PB)	5.5	7.8	9.8	11.5	13.3
% Disk obligation	73.7%	92.5%	99.1%	102.1%	107.1%

- At the moment we are still holding off the storage procurement
 - Due to budget uncertainties
 - Cancellation of 2022 HI run may not exert indented pressure on storage

PEAP 2023 Quarterly Milestones – Q1-Q2

Q1 FY2023

- [Submit 2023 WLCG pledges](#)
- Restore HTTPS support in EOS at ORNL
- [Present prototype AF configuration to ALICE grid team](#)
- Prepare storage for the prototype AF
- ORNL: IPv6 support, upgrade perfsonar system and register with OSG
- [Hold Fall project meeting at ORNL or virtually](#)
- Receive and review 2023 ERCAP allocations at NERSC
- [Migrate ALICE Grid at NERSC from Cori to Perlmutter](#)

Q2 FY2023

- [Test and move to whole-node scheduling at ORNL](#)
- Attend EOS community workshop
- Deploy LBNL/HPCS perfsonar, register with OSG, and upgrade dashboard
- Attend OSG All Hands meeting
- Set up analysis facility prototype at LBNL
- Initiate the larger usage of NERSC by the ALICE-USA physicists to run jobs with access to T2 storage
- Hold annual ALICE-USA/CERN resource review meeting at LBNL
- Purchase and deploy new CPU for the 2023 RRB year as needed

PEAP 2023 Quarterly Milestones – Q3-Q4

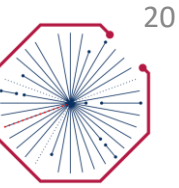
Q3 FY2023

- Update Analysis Facility Proposal to DOE
- Report status of ALICE-USA grid operations at annual ALICE T1/T2 Workshop
- Review updated ALICE requirements for 2024-2025
- Test and move to whole-node scheduling at LBNL

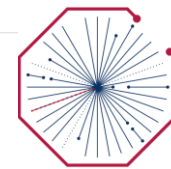
Q4 FY2023

- Update PEAP for FY2024
- Purchase and deploy remaining disk storage as needed depending on utilization
- Submit FY2024 ERCAP request for NERSC HPC CPU allocation

- Jeff will discuss some more details on Friday

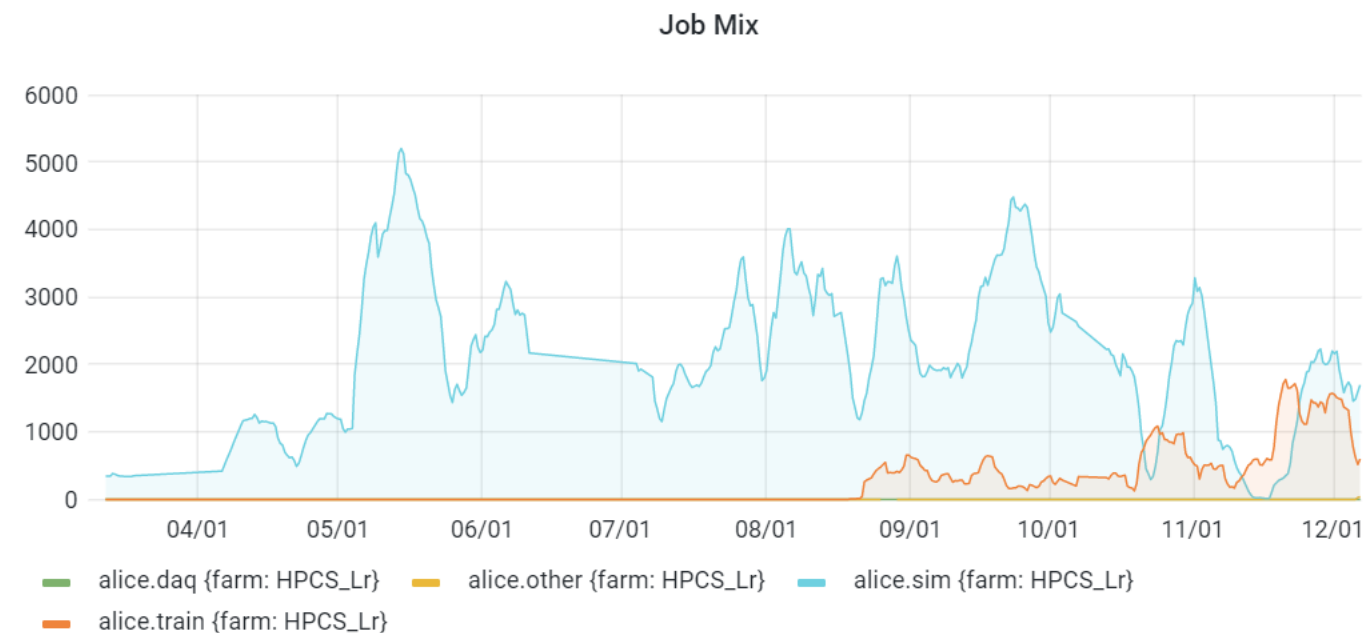
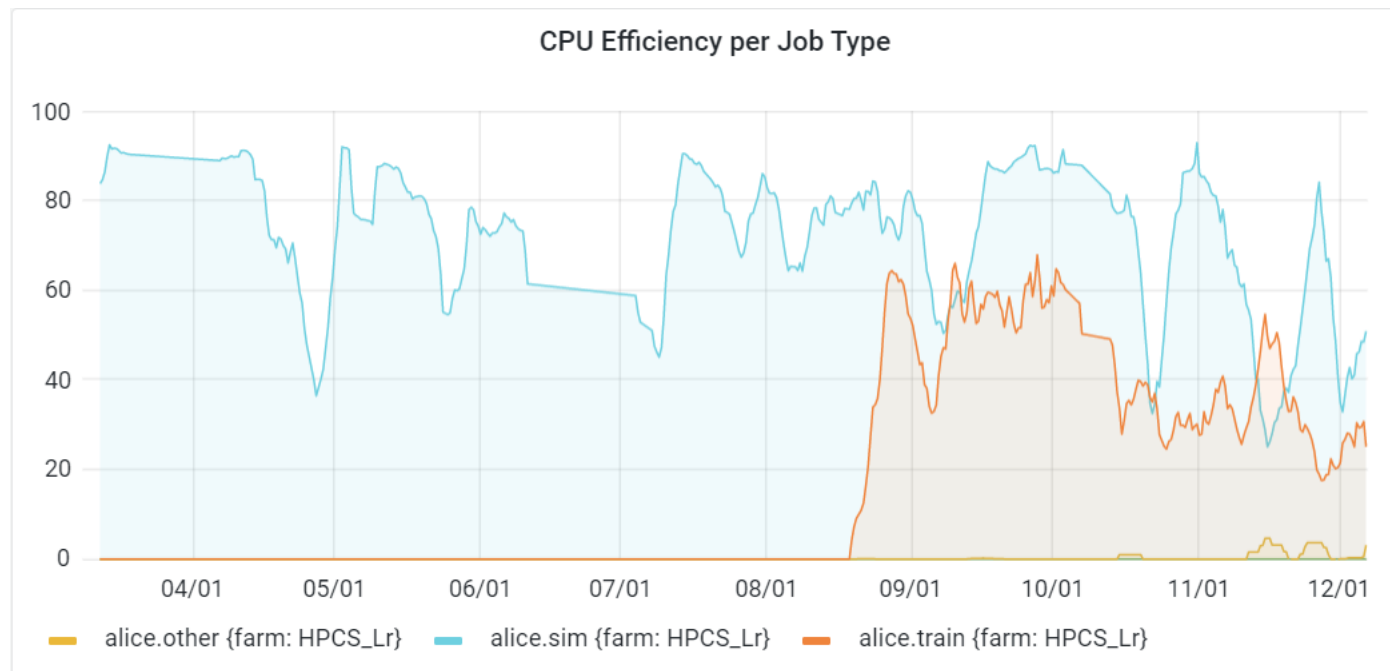
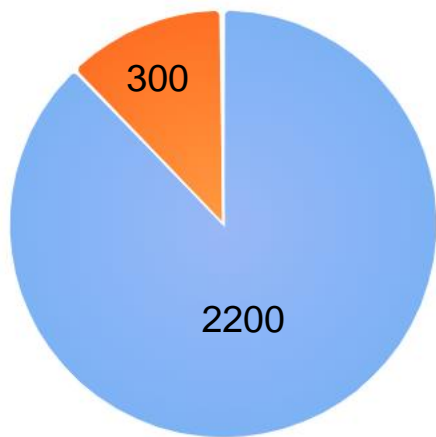


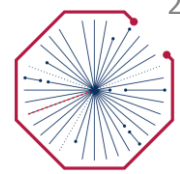
Extra Tier-estrial Activities



HPCS Lawrenceium

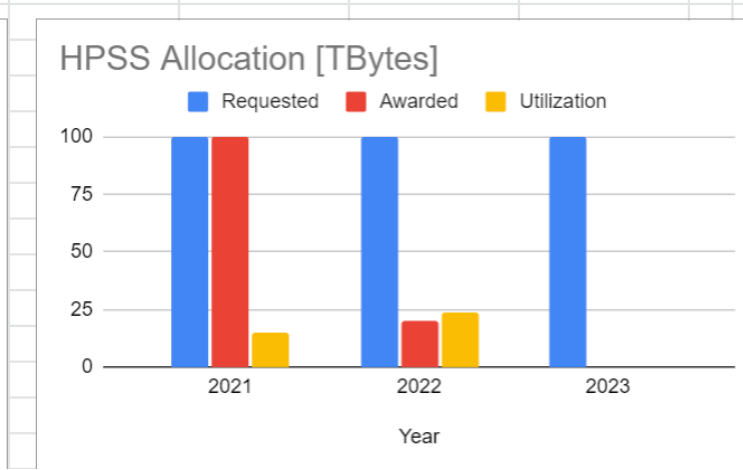
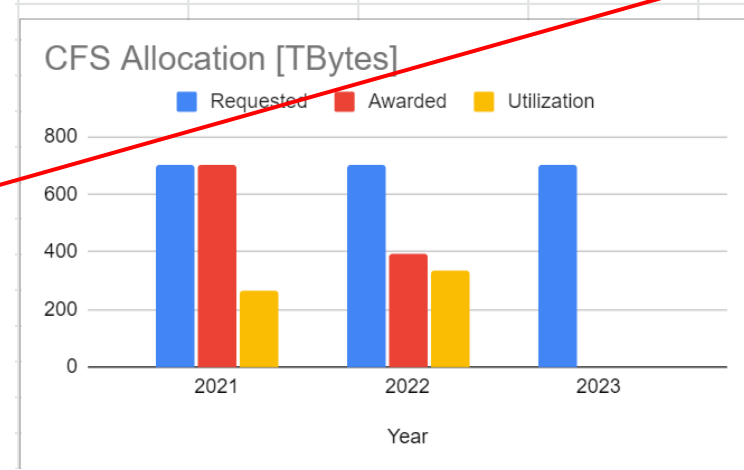
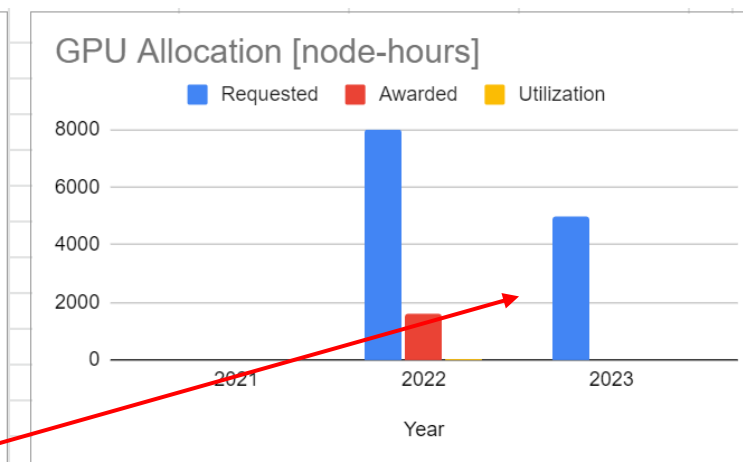
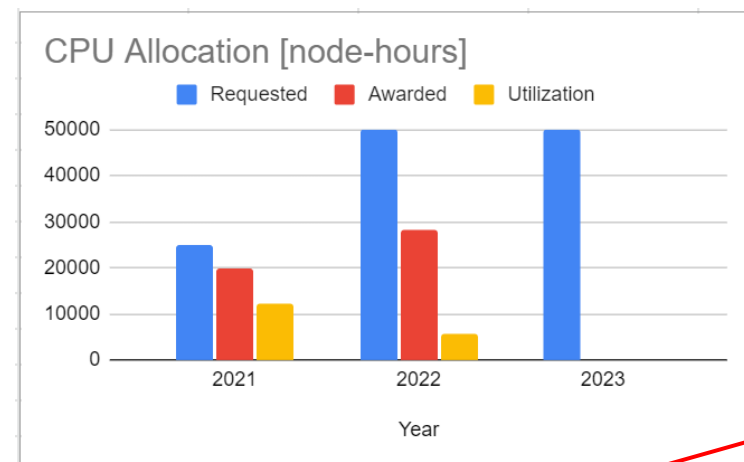
- R&D project for using opportunistic resources on HPC clusters is now fully operational
- Success with the Lawrenceium was recognized by the WLCG Computing Review
- There is a possibility to include additional partitions available at Lr cluster



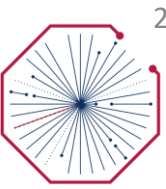


NERSC Allocation

- As part of the project, we request allocation on NERSC HPCs
- Cori is generally used as out R&D platform for HPCs
- Cori was announced to retire in February 2023
- We started working on moving the ALICE workflow to the new Perlmutter machine
- Explicate request of GPU in 2023



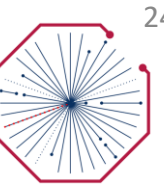
Cori → Perlmutter



Cori	Partition	# of nodes	Processor	Clock Rate	Cores Per Node	Memory Per Node
	Login	20	Intel Xeon Processor E5-2698 v3	2.3 GHz	32	515 GB
	Haswell	2,388	Intel Xeon Processor E5-2698 v3	2.3 GHz	32	128 GB
	KNL	9,688	Intel Xeon Phi Processor 7250	1.4 GHz	68	96 GB (DDR4), 16 GB (MCDRAM)
	Large Memory	20	AMD EPYC 7302	3.0 GHz	32	2 TB

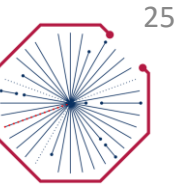
Perlmutter	Partition	# of nodes	CPU	Cores Per Node	GPU	Aggregated Memory [TB]
	GPU	1536	1x AMD EPYC 7763	64 for CPU	4x NVIDIA A100 (40GB)	240
	CPU	3072	2x AMD EPYC 7763	64	-	384
	Login	40	1x AMD EPYC 7713	64	1x NVIDIA A100 (40GB)	
	Large Memory	4	1x AMD EPYC 7713	64	1x NVIDIA A100 (40GB)	

Perlmutter



- By Switching from Cori to Perlmutter we lose access to the static IP node with long process running capability, e.g. cori21
- We took advantage of the NERSC offered Cori-to-Perlmutter workflow deployment officed hours
- Discussed, with the workflow experts, possible solutions and came up with the following:
 - Utilize NERSC Superfacility API to submit jobs to the scheduler on Perlmutter
 - Use NERSC SPIN system to run the VOBOX and access the API
- Superfacility API has been requested about a month ago and we were promised to get response sometime this week
- Meanwhile Sergiu has successfully tested CVMFS on SPIN and
 - More in hist talk

Perlmutter GPU



- Another interesting opportunity is the use of the GPU cluster of the new Perlmutter machine
- We requested 5,000 Node-hours with our 2023 ERCAP request
 - Decision pending
- Later today we will have a talk from David Rorh about ALICE GPU requirements
- The idea is to host ALICE GPU-intensive jobs at NERSC

Summary

- ALICE-USA Computing project grows along with ALICE to meet it's ALICE-USA obligations
- Both T2 sites, operate with very high availability, reliability, and efficiency
- PEAP 2023 update has been submitted and envisions meeting resources at 100%
- Opportunistic use of Lawrenceium is a highly dependable resource that supplemented the ALICE-USA pledged resources to meet the requirements
- CORI to Perlmutter switch has posed interesting challenges that are being tackled
- Perhaps we can report on Perlmutter GPU running conditions at the next meeting



Questions