

# ALICE-USA Streamlined-T1 *Just an Idea*

T1/T2 Workshop, Seoul, Korea

17 April 2024

Irakli Chakaberia

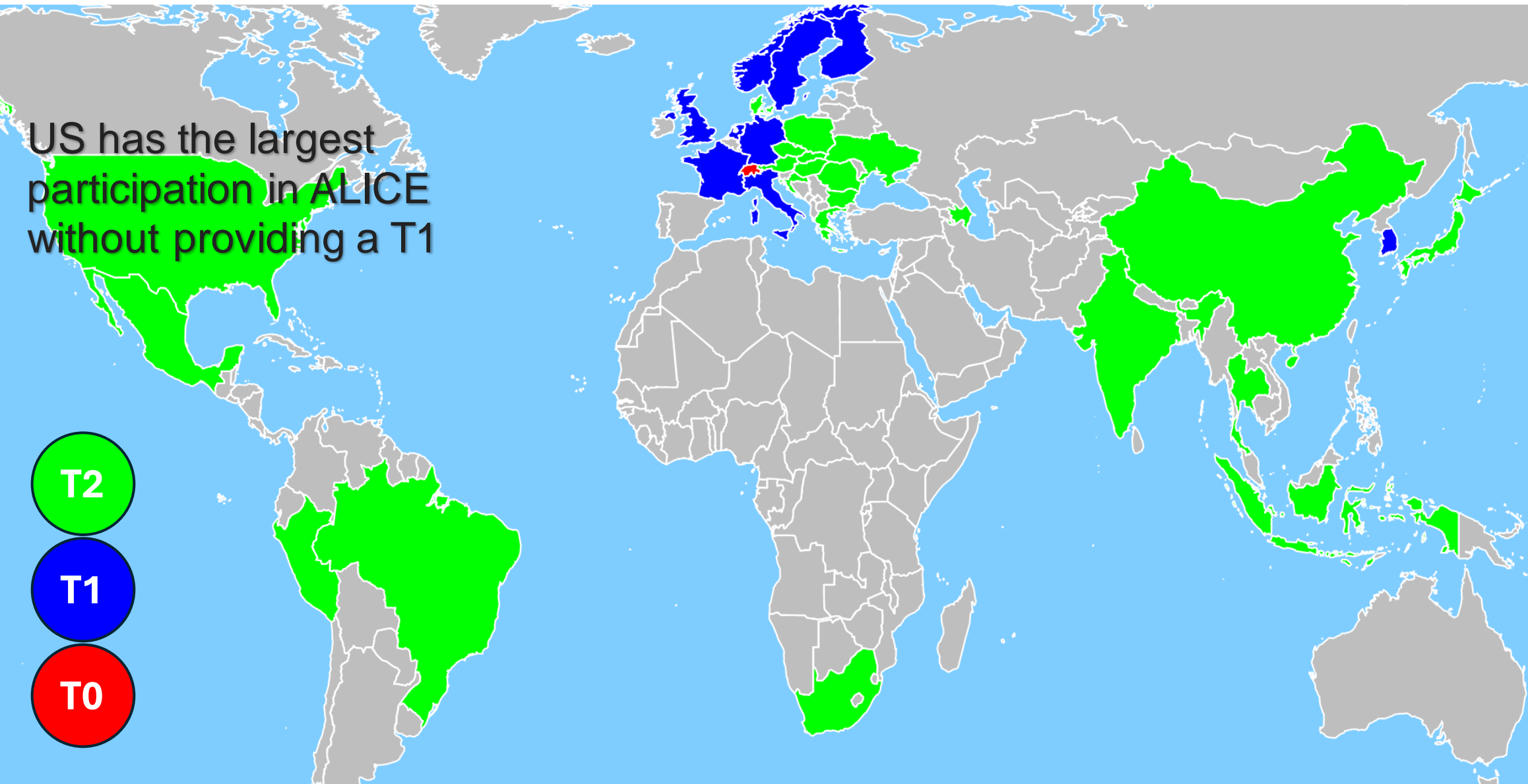


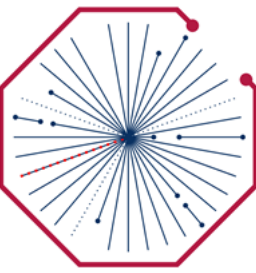
# ALICE Computing Grid Map



US has the largest participation in ALICE without providing a T1

- T2
- T1
- T0

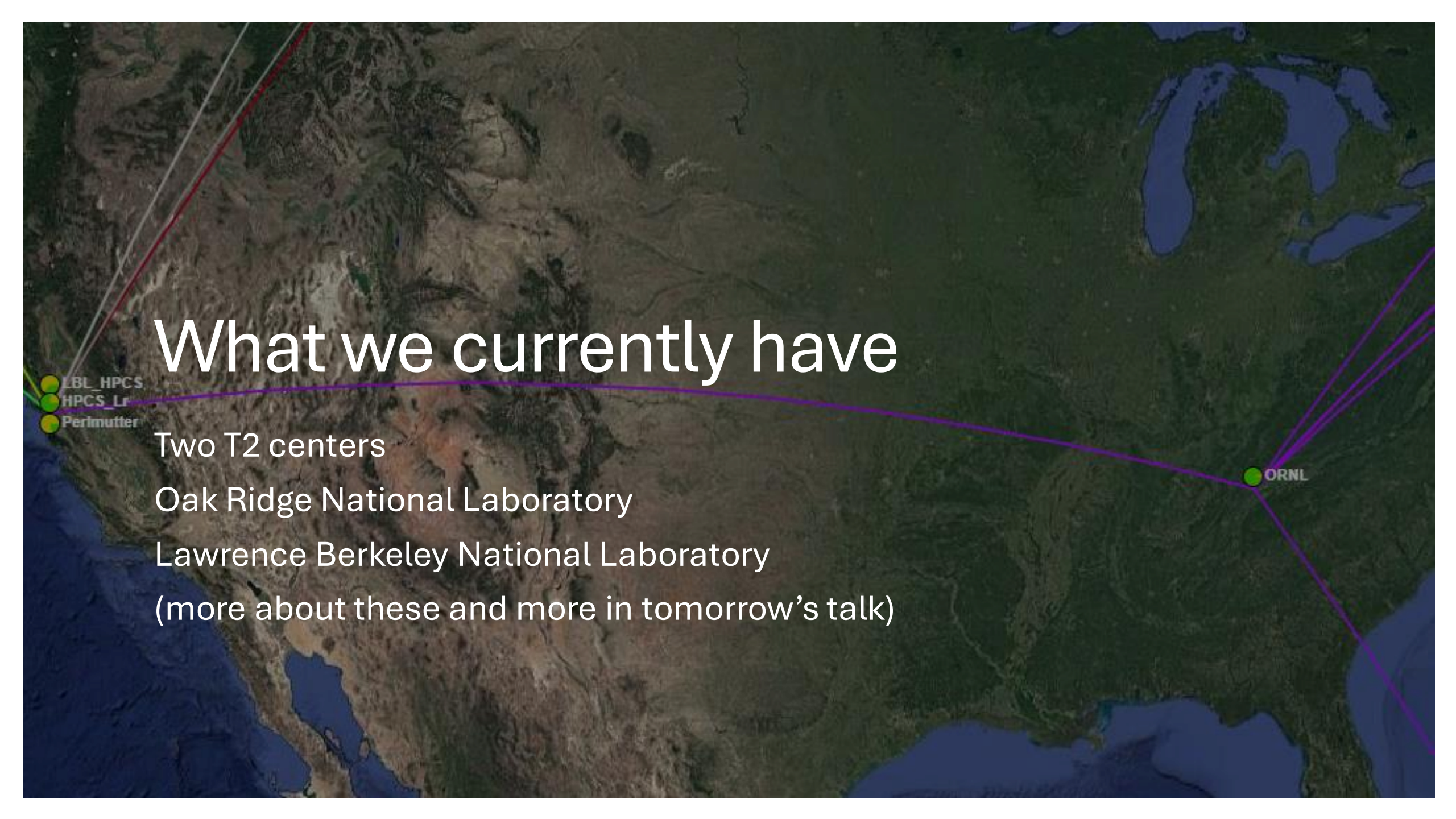




# Pressing Needs for ALICE

- T1 resources become scarce
- The 2025 request is above the flat-budget scenario

ALICE		2024					2025			
		Req.	C-RSG	Pledge	RU + JINR pledge	Pledge - (RU+JINR)	Prev. Est.	Req.	Req. 2025 / C-RSG 2024	Req. 2025 / (Pledges - RU) 2024
CPU [kHS23]	Tier-0	600	600	600		600	690	680	113%	113%
	Tier-1	630	630	540	0	540	725	690	110%	128%
	Tier-2	650	650	641	33	608	750	730	112%	120%
	<b>Total</b>	<b>1880</b>	<b>1880</b>	<b>1782</b>	<b>33</b>	<b>1748</b>	<b>2165</b>	<b>2100</b>	<b>112%</b>	<b>120%</b>
Disk [PB]	Tier-0	67.5	67.5	67.5		67.5	78.5	78.0	116%	116%
	Tier-1	71.5	71.5	61.9	0.0	61.9	82.5	79.0	110%	128%
	Tier-2	66.5	66.5	69.8	3.2	66.5	77.5	77.0	116%	116%
	<b>Total</b>	<b>205.5</b>	<b>205.5</b>	<b>199.2</b>	<b>3.2</b>	<b>195.9</b>	<b>238.5</b>	<b>234.0</b>	<b>114%</b>	<b>119%</b>
Tape [PB]	Tier-0	181.0	181.0	181.0		181.0	226.0	220.0	122%	122%
	Tier-1	107.0	107.0	102.4	0.0	102.4	135.0	123.0	115%	120%
	<b>Total</b>	<b>288.0</b>	<b>288.0</b>	<b>283</b>	<b>0</b>	<b>283.4</b>	<b>361.0</b>	<b>343.0</b>	<b>119%</b>	<b>121%</b>

A satellite-style map of the United States with a dark overlay. Three green circular markers on the West Coast represent Berkeley-based computing centers: 'LBL\_HPCS', 'HPCS\_Lf', and 'Perlmutter'. A single green marker on the East Coast represents 'ORNL' at Oak Ridge National Laboratory. A purple line connects the Berkeley centers to ORNL. A red line and a grey line also originate from the Berkeley area and extend towards the top-left of the frame.

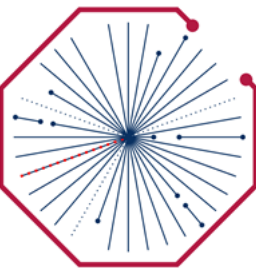
# What we currently have

Two T2 centers

Oak Ridge National Laboratory

Lawrence Berkeley National Laboratory

(more about these and more in tomorrow's talk)

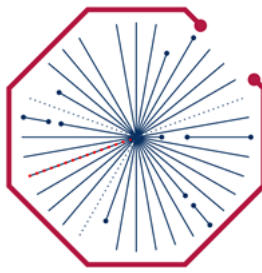


# What else we have (available)

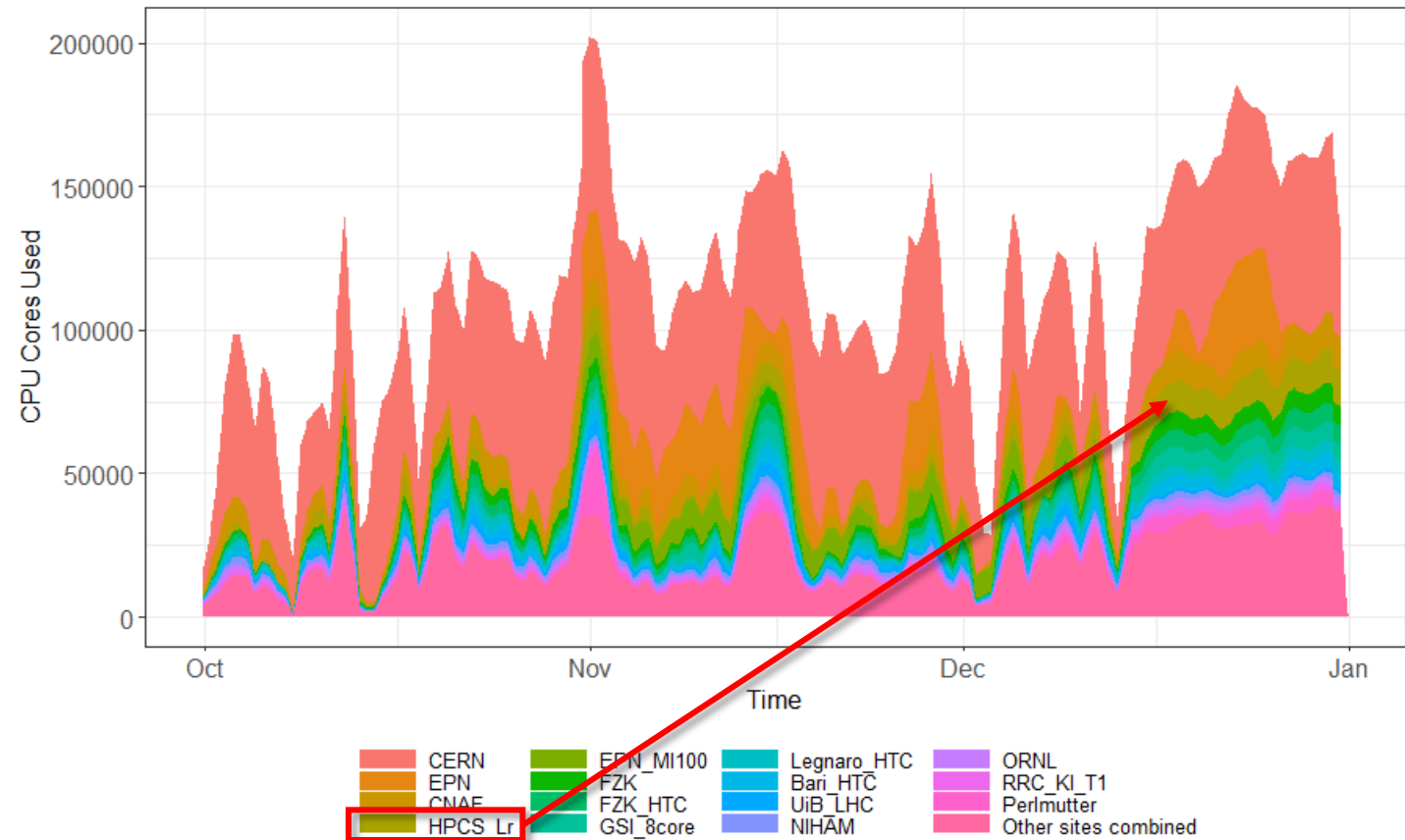
- ITD Lawrencium
- NERSCS (Cori to Perlmutter to NERSC 10 (an exaSystem))



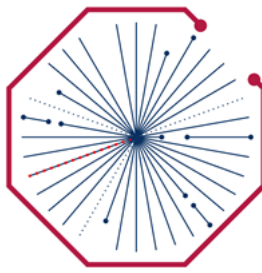
# Lawrencium



- We are using LBL ITDs Lawrencium HPC's scavenging queue
- Technically the assigned job slots are preemptable, but this resource has showed great reliability since very deployment

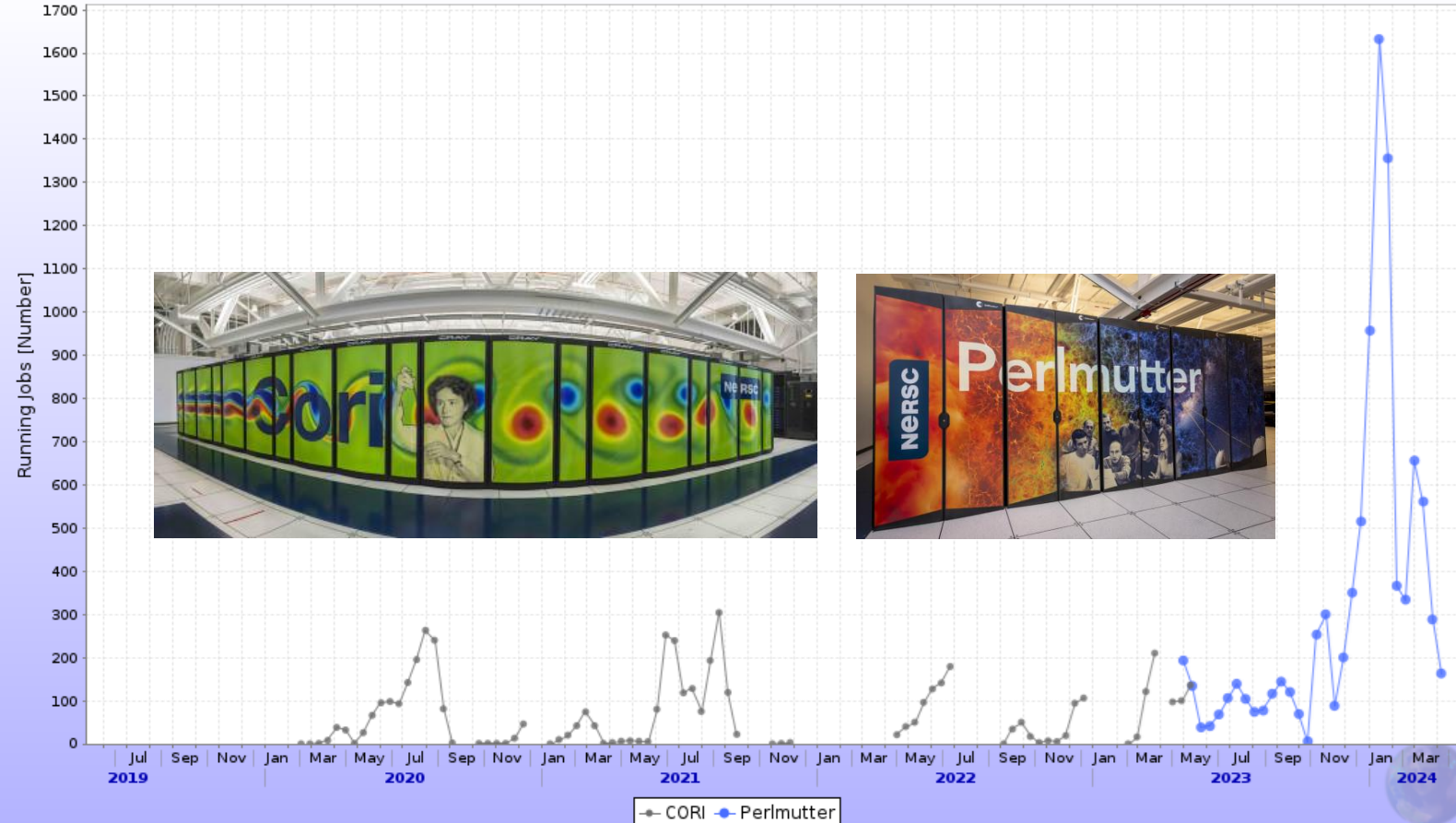


# Perlmutter



- We apply for and maintain resource allocation on Perlmutter distributed by the DOE annually
- So, this is not quite “opportunistic” in a sense
- More about this in Sergiu’s talk

Running Jobs



# GPU Cluster



- One of the major parts of this idea is enabling GPU-based reconstruction at NERSC
- GPU resource allocation is also requested annually

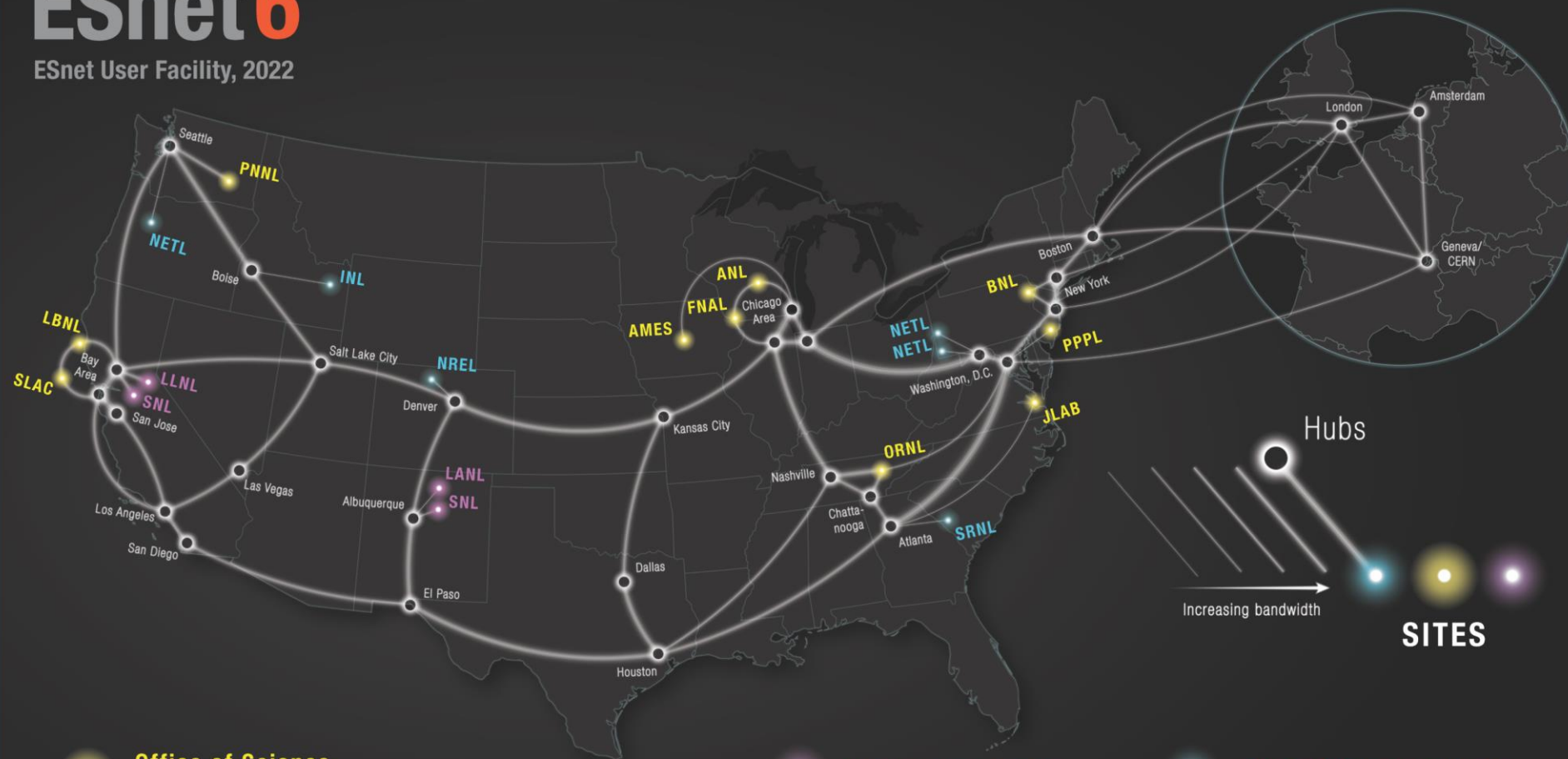
Partition	# of nodes	CPU	GPU	NIC
GPU	1536	1x <a href="#">AMD EPYC 7763</a>	4x <a href="#">NVIDIA A100</a> (40GB)	4x <a href="#">HPE Slingshot 11</a>
	256	1x <a href="#">AMD EPYC 7763</a>	4x <a href="#">NVIDIA A100</a> (80GB)	4x <a href="#">HPE Slingshot 11</a>
CPU	3072	2x <a href="#">AMD EPYC 7763</a>	-	1x <a href="#">HPE Slingshot 11</a>
Login	40	1x <a href="#">AMD EPYC 7713</a>	1x <a href="#">NVIDIA A100</a> (40GB)	-



Something else we have

# ESnet6

ESnet User Facility, 2022



**Office of Science National Laboratories**

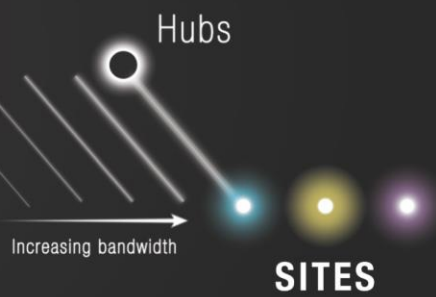
- AMES** Ames Laboratory (Ames, IA)
- ANL** Argonne National Laboratory (Argonne, IL)
- BNL** Brookhaven National Laboratory (Upton, NY)
- FNAL** Fermi National Accelerator Laboratory (Batavia, IL)
- JLAB** Thomas Jefferson National Accelerator Facility (Newport News, VA)
- LLNL** Lawrence Livermore National Laboratory (Livermore, CA)
- SNL** Sandia National Laboratory (Albuquerque, NM; Livermore, CA)
- LANL** Los Alamos National Laboratory (Los Alamos, NM)
- ORNL** Oak Ridge National Laboratory (Oak Ridge, TN)
- PNNL** Pacific Northwest National Laboratory (Richland, WA)
- PPPL** Princeton Plasma Physics Laboratory (Princeton, NJ)
- SLAC** SLAC National Accelerator Laboratory (Menlo Park, CA)
- LBL** Lawrence Berkeley National Laboratory (Berkeley, CA)
- ORNL** Oak Ridge National Laboratory (Oak Ridge, TN)
- PNNL** Pacific Northwest National Laboratory (Richland, WA)
- PPPL** Princeton Plasma Physics Laboratory (Princeton, NJ)
- SLAC** SLAC National Accelerator Laboratory (Menlo Park, CA)

**NSNS Laboratories**

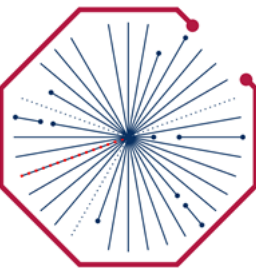
- LANL** Los Alamos National Laboratory (Los Alamos, NM)
- LLNL** Lawrence Livermore National Laboratory (Livermore, CA)
- SNL** Sandia National Laboratory (Albuquerque, NM; Livermore, CA)

**Other DOE Laboratories**

- INL** Idaho National Laboratory (Idaho Falls, ID)
- NETL** National Energy Technology Laboratory (Morgantown, WV; Pittsburgh, PA; Albany, OR)
- NREL** National Renewable Energy Laboratory (Golden, CO)
- SRNL** Savannah River National Laboratory (Aiken, SC)

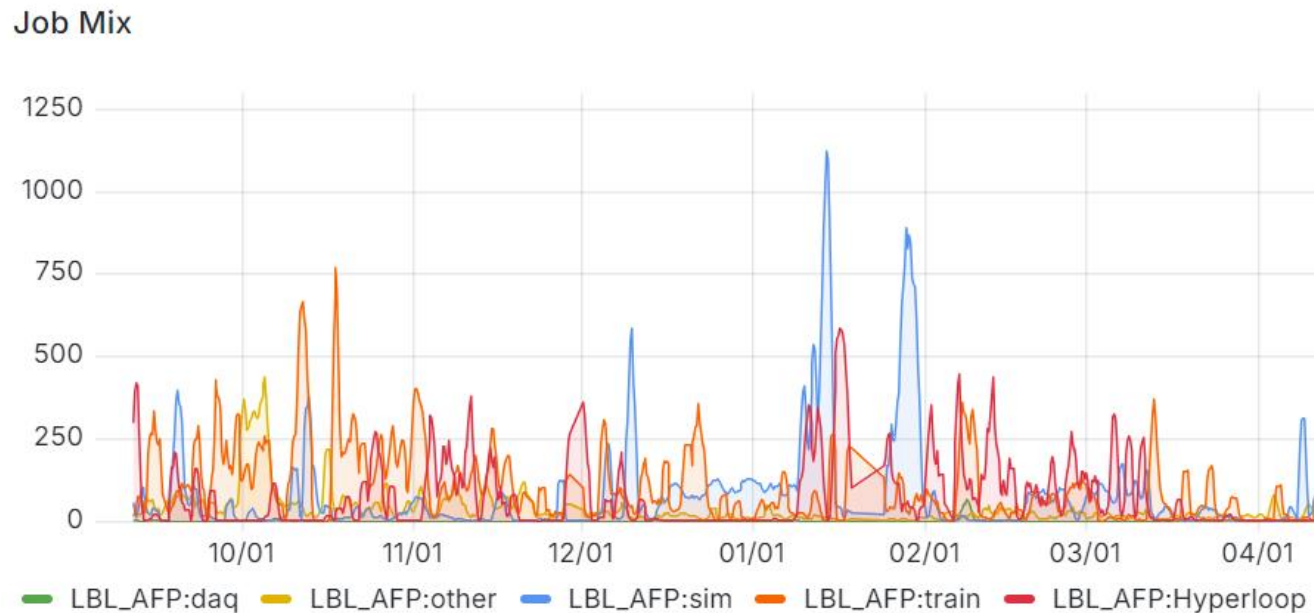


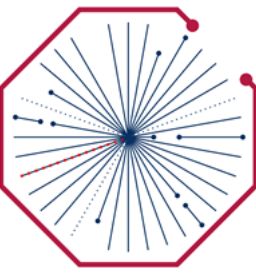
Accelerating Scientific Discovery



# Analysis Facility (Prototype)

- Last September we deployed a prototype AF as a proof-of-concept (640 cores; 1.1 PB)
- It's been operational virtually without downtime since
- There are interesting things we learned about the differences between AF and T2 type of loads
- We seek funding to expand it to a larger AF



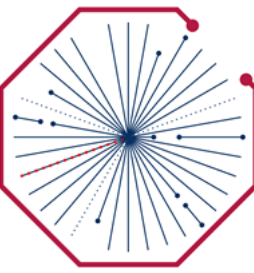


# An Idea

- Deploy T1-like capabilities (without commitment to the custodial storage)
- Expand O2 GPU capabilities to enable it at NERSC (perhaps further)
- Deploy full scale AF

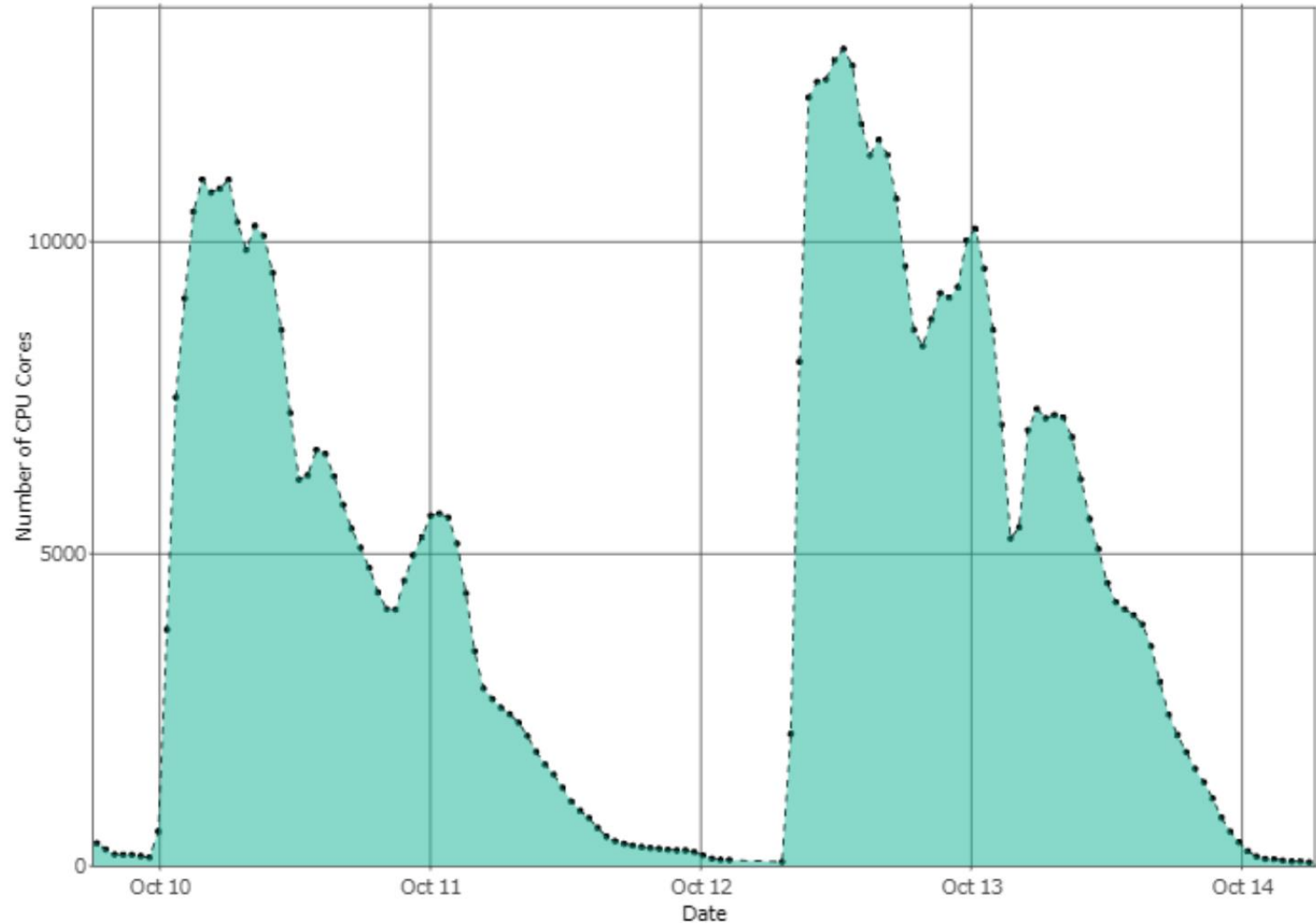
**Single partitioned EOS storage with RAIN (?) configuration.**

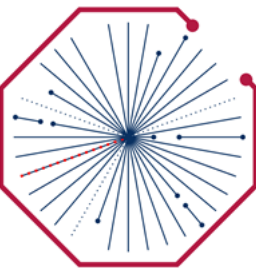




# Something we already tried

- We ran a data reconstruction tests on Lawrencium
- Data as streamed directly from CERN
- There is network, there is CPU capability
- Adding a dedicated T1 type of storage will take further advantage of this





# Summary

- A very compelling idea that can augment current ALICE computing capabilities
  - T1 disk storage
  - “Opportunistic” and “dedicated” HPC resources
  - Bringing more GPUs into the game
  - Expanding the Analysis Facility

