



ALICE

ALICE-USA

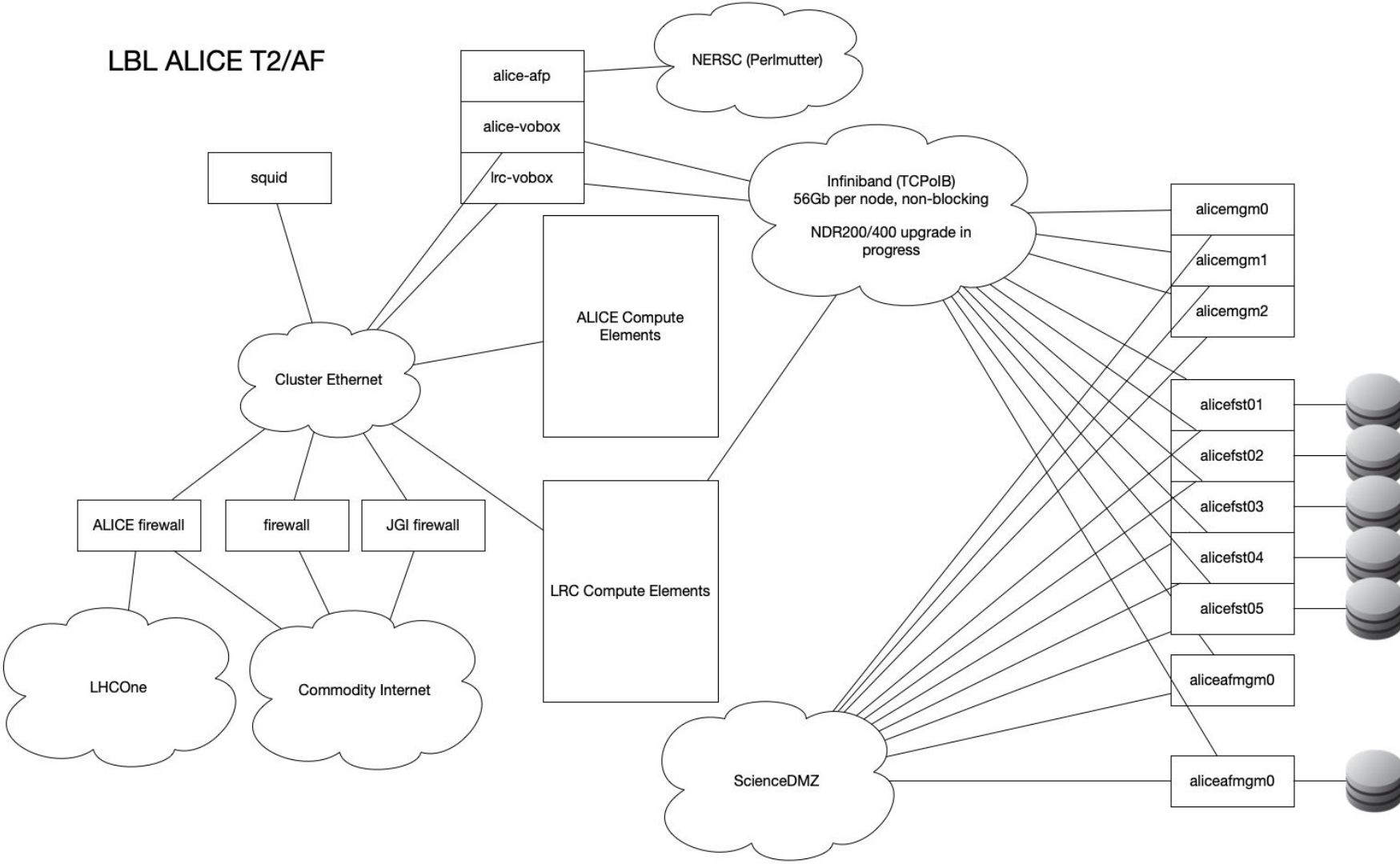
T2 and AF Computing @ LBNL

ALICE-USA Annual Meeting
Sept 16-18 2025

Lawrencium Overview

- **Lawrencium:** Institutionally-supported HPC Linux Cluster
 - ~2,680 nodes/ ~80,000 cores
 - 7.4PB Lustre parallel file system for Scratch supporting condo-style buy-in
 - Rocky 8, SLURM job scheduler, Singularity containers, Open OnDemand 4
 - Nvidia/AMD GPUs
 - MI50/100, A100, H100 & H200 HGX (and plenty of older stuff)
 - New NDR200/400 Infiniband Fabric
- Four models of access
 - PAYG \$0.01/core-hr
 - Condo Cluster
 - Dedicated cluster support
 - Opportunistic
- PCA - PI Computing Allowance
 - Free - 300K SUs per year
 - Can be shared with your staff or pooled together with other PIs

LBL ALICE T2/AF



Storage

- 2x40Gb routed link from LBLNet to HPC network
 - Currently providing a 8.4PiB CEPH platform
 - S3/RADOS
 - Cephfs
 - NFS
 - VAST offering
 - Home, group, software
 - Infiniband, New NDR fabric
 - S3, Application Hosting
- Scratch File system
 - DDN SFA 18K
 - 180GB/s theoretical peak R/W
 - First 64MB of every file on NVME Flash (Progressive File Layout)
 - Rest on Spindle
 - 440TB Buy-in available (non-purged space)

Cluster Futures

- Upgrades in the near future
 - Warewulf 4
- LDAP migration in progress
- More Sane MFA
 - <https://commons.lbl.gov/spaces/SCIEN/pages/315689584/SSH+Certificates+More+Sane+MFA+MSM>
- IPV6!!!!
 - waiting on timeline
- New LRC Generation

LR8 Nodes

- 51 Node Institutional Purchase
- Moved from 4nodes in 2u to 1n1u (air cooled)
- AMD dual-socket 9534 (64 core, 128 threads x 2)
- 768GB (24 x 32GB, 4800MT/s)
- 3.84TB local storage (Micron 7450, U.3)
- ConnectX-6 (Connected at 100Gb)

AF EOS

- 1 MGM node
 - 10Gb external, 56Gb internal
 - HTTP re-enabled
- 1 FST node
 - 1 JBOD, 125 drives
 - Again, 10Gb external, 56Gb internal
- All being folded back into T2

HPCS EOS

- 3 MGM nodes
 - 10Gb external, 56Gb internal
 - Need to re-re-enable HTTPS interface
- 5 FST nodes
 - Again, 10Gb external, 56Gb internal
 - 485 total drives, 1 warm spare per node
- EOS 5.2.31
 - "Stable Enough"
- Currently at ~77% Use, 5.47PB out of 7.09PB

EOS Merge + Server Count Increase

- new servers on-site - will "simply" split off existing 2nd jbods to new servers
- Will be added to existing T2
- Create new dir for AF on T2, set quota
- ask you to copy AF to T2
 - a. We don't really have the spare space but would like to stripe the data...
- fold existing AF back into T2 (was previously T2)
- Start Experiments with Erasure Coding

HPCS EOS Current Issues

- Rocky8 Bug still exists, papered over
 - XrdZMQ client errors until MGM reset
 - Post-Rocky 8 Upgrade
 - XrdZMQ::client looping since 24480.00 seconds ...
 - EOS team previously claimed out of scope, just restart mgm on monitor+timer

Projects

- Owncloud + Cephfs
 - Kubernetes OCIS
 - Stand-alone REVA instance
 - Custom Public OIDC forthcoming
 - Ongoing Collaboration with Giuseppe and team
- Monitoring Project
 - Vast S3 backend
 - LGTM or ELK
 - Alloy migration