NERSC and the **ALICE** Computing Grid:

Challenges for integrating NERSC resources into an existing distributed and automated data processing model

R. Jeff Porter (LBNL)

NERSC SIG on Experimental Facilities

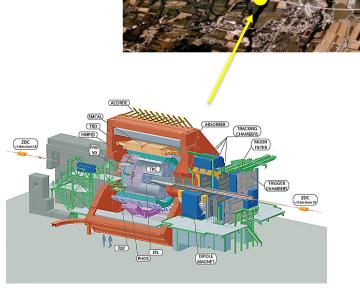
July 1, 2020



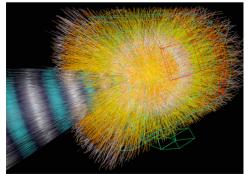
Outline



- ALICE @ NERSC
- Grid Computing Model
- NERSC & the ALICE Grid
 - Running jobs
 - Accessing storage
- Outlook



Pb+Pb collision in ALICE





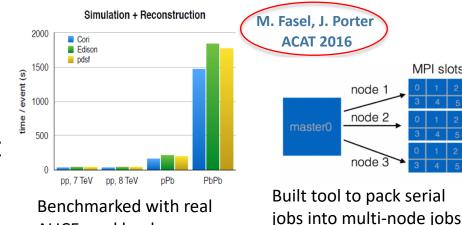
Early ALICE R&D activities on CORI



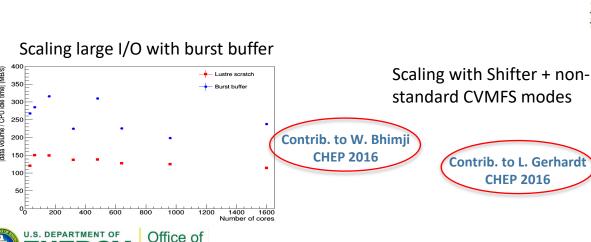
MPI slots

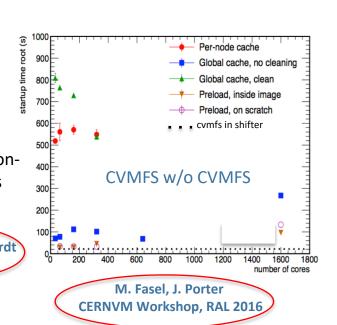
- LBNL ALICE group proactively R&D'd different models for using CORI
- 4 years later, CORI is lightly used by ALICE
 - mainly by local group for one-off tasks
 - Remains an outlier resource in ALICE

Science



What about the direct integration of **NERSC** into the ALICE computing model?





ALICE workloads

ALICE Computing Model



Grid Facility & Processing model

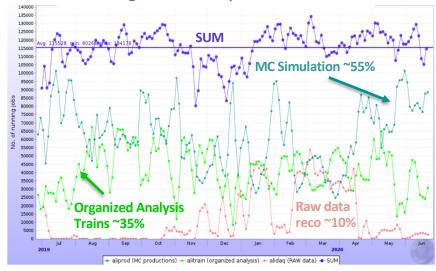
- ~80 active sites
 - CERN Tier-0, 8 Tier-1s, ~70 Tier-2 (T2)
- >120,000 serial jobs, 24 x 365
 - Fully automated, resubmit on job failures
- 110 PB distributed disk storage
 - Data distributed to multiple sites
 - Jobs run where the data is
- AliEn: software to connect distributed resources into single facility

Every T2 site supports ~90% of job types

- MC simulation jobs:
- Organized analysis trains



Running Jobs over past 12 months





Site requirements & NERSC CORI



Node level

- ✓ any modern Linux distribution we use a thin shifter image + CVMFS
- √ memory capacity of ~2.5 GB/useable-core
 - X Swap enabled is highly desirable ← this limits our full use of CPU cores
- ✓ outgoing network connectivity from worker node
- ✓ local disk (or performant scratch) for small block I/O Shifter's per-node-cache
- ✓ CVMFS for software distribution

Facility Level

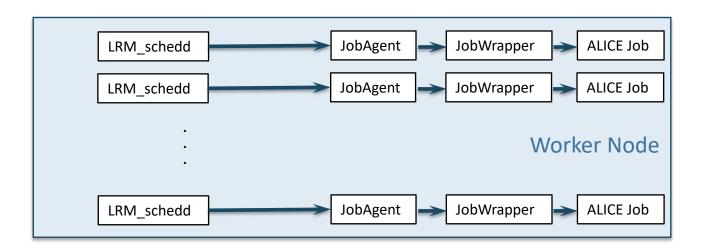
- ✓ workflow node (VOBox in WLCG-speak) as site point of contact
- ✓ most any LRM: LSF, PBS, SGE, SLURM, HTCondor, ..., ARC-CE, OSG-CE, ...
- optimally configured for serial jobs
- **X** large long term disk storage:
 - Grid enabled with EOS or XRootD
 - Incoming network with ALICE Token AuthN

We can try to address these









JobAgent:

- Gets full job definition from central services
- spawns JobWrapper thread with job definition & monitors resource usage
- Repeats when JobWrapper exits if enough time remains

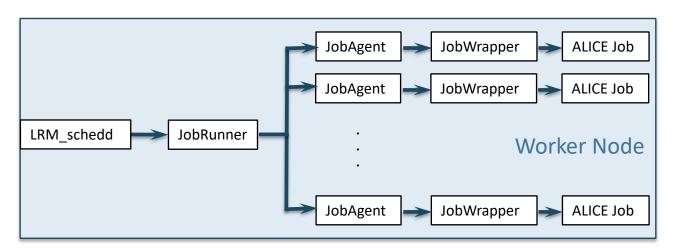
JobWrapper

- prepares sandbox and launches payload (ALICE Job)
- validates output and copies output to destination storages









JobRunner

Manages node resources and launches JobAgents as needed to optimize node usage

JobAgent:

- Gets full job definition from central services & reserves resources from JobRunner
- spawns JobWrapper thread with job definition & monitors resource usage
- Exits when JobWrapper exits

JobWrapper

- prepares sandbox and launches payload (ALICE Job)
- validates output and copies output to destination storages

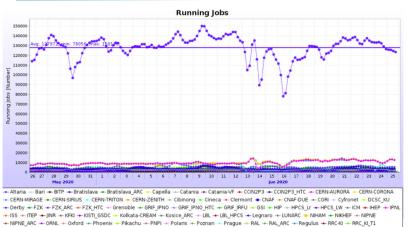
* Initial funding from LDRD with LBNL Physics (Z. Marshall PI)



Jeff Porter LBNL -7 -

Initial deployment in early 2020





Sources of low utilization rate:

Mem+no-swap limits use to ~45 jobs/node

+ SaoPaulo + SaoPaulo_HTC + SARA + SNIC + SPbSU + Strasbourg_IRES + Subatech + Subatech_CCIPL + SUT + Torino

◆ UiB ◆ UiB_LHC ◆ UNAM ◆ UNAM_T1 ◆ UPB ◆ Vienna ◆ WUT ◆ Yerevan ◆ ZA_CHPC ◆ SUM

- Inefficient packing into wall time of slurm job
- SLURM scheduling
 - jAliEn keeps 10 x 1-node jobs queued
 - Limited backfill with 48 hour jobs?
- High error rate (16%)

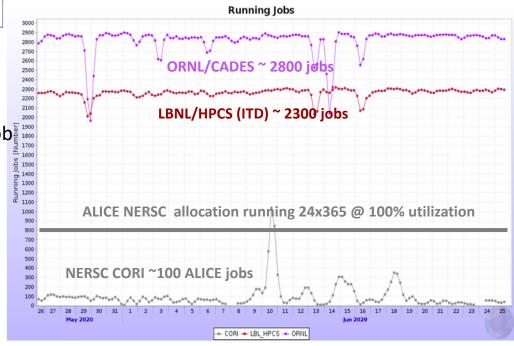
This gives us something to work with

ALICE Grid ~ 130,000 jobs

Retained fully automated workflow

- Late-binding of job to resource
- Auto cleanup & resubmit on job failure
- Useable in serial and whole node settings

ALICE USA ~5000 jobs



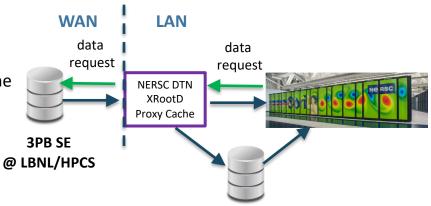


Optimize remote data access



XRootD Proxy Cache *

- Deployed in user space on DTN
- Fills request from (any) remote SE & adds data to cache
- returns local filename to client if data is in cache



Cache on \$SCRATCH

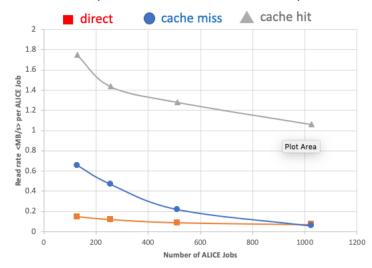
Initial tests use a single DTN node for XRootD Proxy Cache + (non-optimized) analysis code

- Cache hits show significant improvement
- Even cache misses are improved relative to direct access

XRootD is highly modular and open source

- "Return-local-file" developed at GSI by ALICE colleagues
- Extendable for our use case as warranted

Preliminary test results with XRootD Proxy Cache



^{*} Initial funding from LDRD with LBNL Physics (Z. Marshall PI)



Summary & Outlook



- Integrated NERSC resources into the ALICE Grid Facility
 - Leveraged NERSC supplied features:
 - · Outgoing network from worker nodes
 - Workflow nodes for ALICE VOBox
 - CVMFS on nodes via Shifter
 - Added whole-node scheduling to ALICE job management tools
- Workflow retains automation though not highly efficient for using NERSC
 - Extends without altering ALICE computing model
 - Overall throughput is low
 - Known steps may improve CPU utilization, throughput, & data access
- Effort was ALICE development, CORI was a use case
 - ALICE is experimenting with multi-core simulation jobs
 - Other accessible (HPC) sites have similar requirements



Features added for CORI are general: Lawrencium example





- Right - Ball - Bit - Balshara - Balshara - Balshara - Capella -

Lawrencium @ LBNL

- Requires whole(multi)-node jobs
- Preemptable scavenger queue

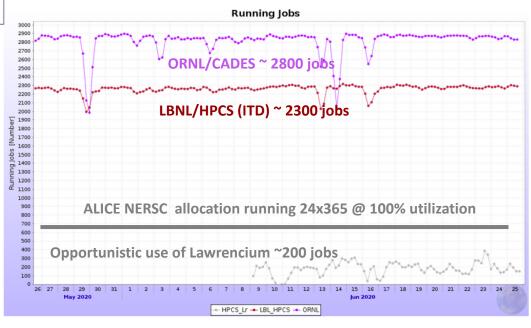
All other site requirements are met, so we just launched it

ALICE Grid ~ 130,000 jobs

Retained fully automated workflow:

- Late-binding of job to resource
- Auto cleanup & resubmit on job failure
- Useable in serial and whole node settings

ALICE USA ~5000 jobs





Jeff Porter LBNL - 11 -