ALICE Grid evolution

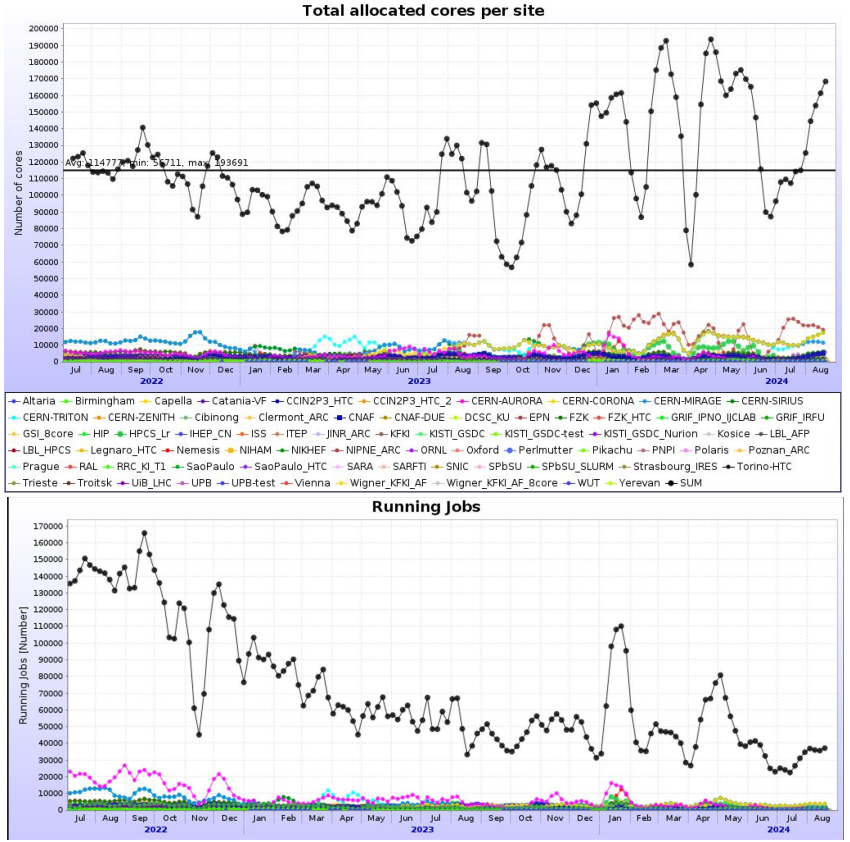
L. Betev

US operation and resources review @LBNL, September 2024

From jobs to cores

- New diverse landscape of 1- 2- 4- 8- 64- core jobs and accelerators in the mix
  - This was expected and planned to be addressed in the Grid middleware

=> One of the main tasks of the new middleware is to ***simplify the site-side ALICE workload management***

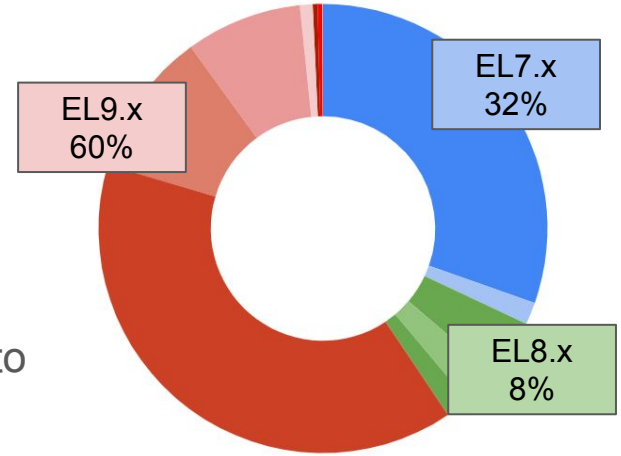


Payload scheduling

- Each batch slot can run different CPU requirements payload
  - Modulo priorities set in the central queue
  - WNs can provide from 8-core slots to full node, up to 128 cores
- Payloads are run concurrently and are mixed with respect to CPU, Memory and I/O requirements
  - MC + Data reconstruction + analysis
  - Orchestration is done in a feedback loop of the JobRunner and central services

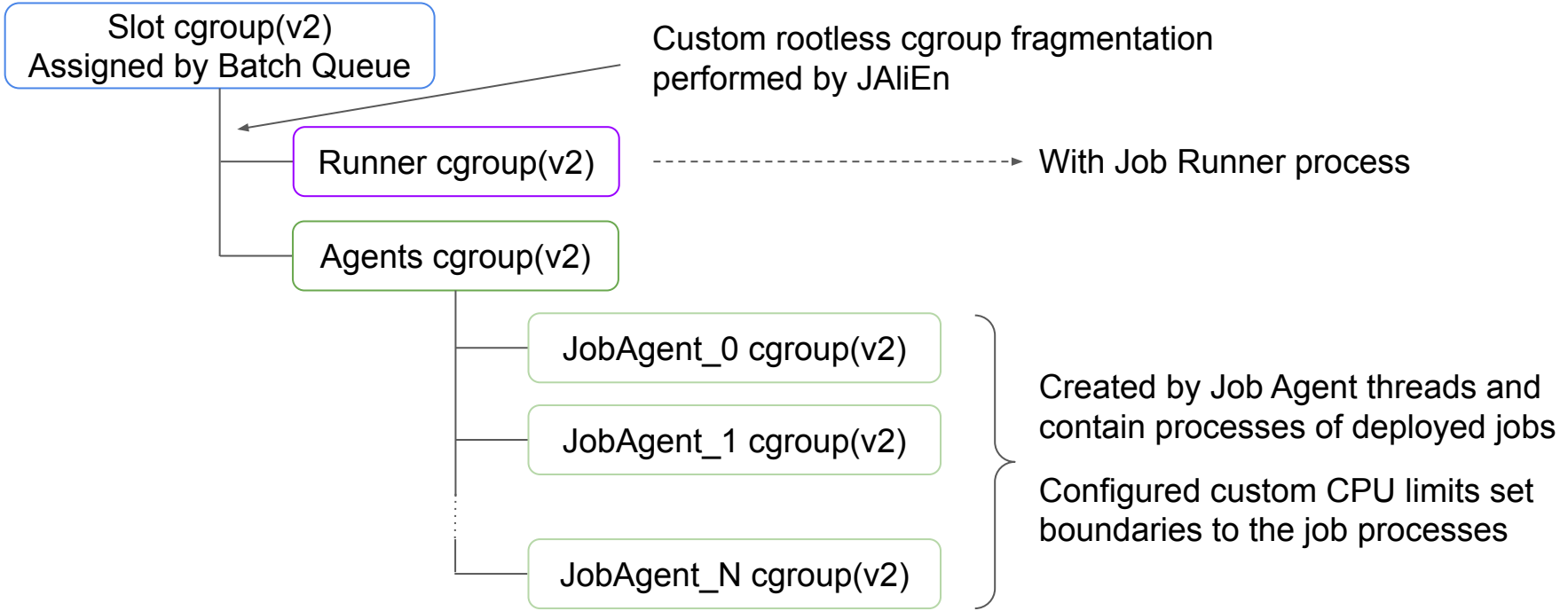
Resources use control

- Use of parameters control to avoid payload interference and unexpected resources overuse
- Implementation through ***cgroups v2*** to control workflow resource allocation
  - With different controllers to manage different resources (CPU, IO, memory…)
  - Lets unprivileged users divide the granted resources into new sub-cgroups
  - Partitioning the resources into the running jobs
- Most popular batch systems (HTCondor and Slurm) can already enable **rootless sub-division** into smaller sub-slots using *cgroups v2*



Distribution of OS versions among Grid hosts as of 19/08/2024

Cgroups v2 integration in JAliEn - cgroup tree

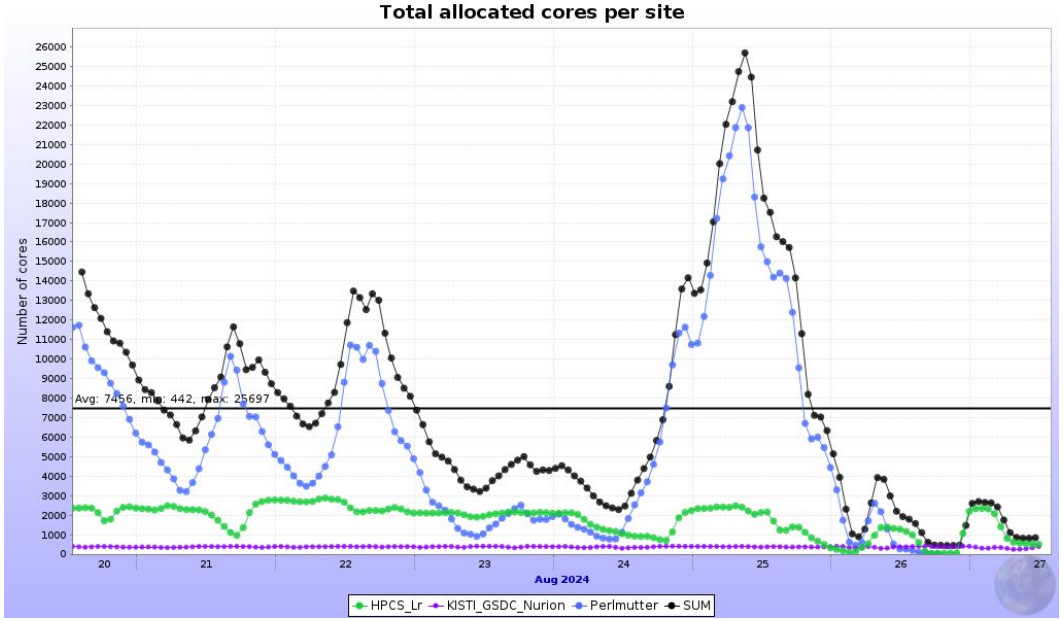


Alternative resources

- Increasing interest and deployment of ARM-based WNs
  - ARM clusters available in the UK, GridKA, CNAF, CERN
- Aarch64 architecture support incorporated in JAliEn
  - Automatic matching of binaries
  - Automatic matching of containers
  - Corresponding aarch64 versions of platforms requested by job
  - Monitoring adjusted to work across architectures
- Changes kept as generic as possible
  - Allows us to easily slot-in support for more architectures in future (e.g. RISC-V)
- Large-scale tests are already done on O2 software performance and compatibility
  - Several bugs in the code uncovered and fixed, more work needed

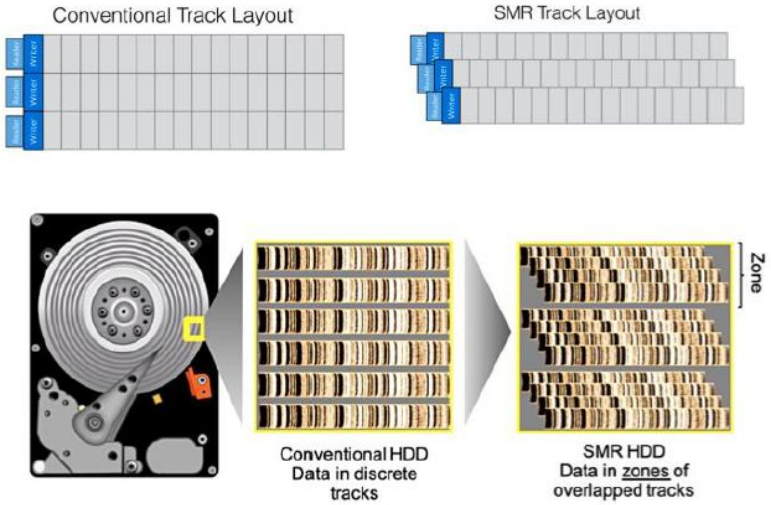
Supercomputers - briefly

- Three SC providing resources to ALICE:
  - LBNL - Lawrencium and Perlmutter
  - KISTI - Nurion
- 5% in average, 10% max contribution to CPU resources
- Incorporation of each supercomputer on the Grid is still an individual task

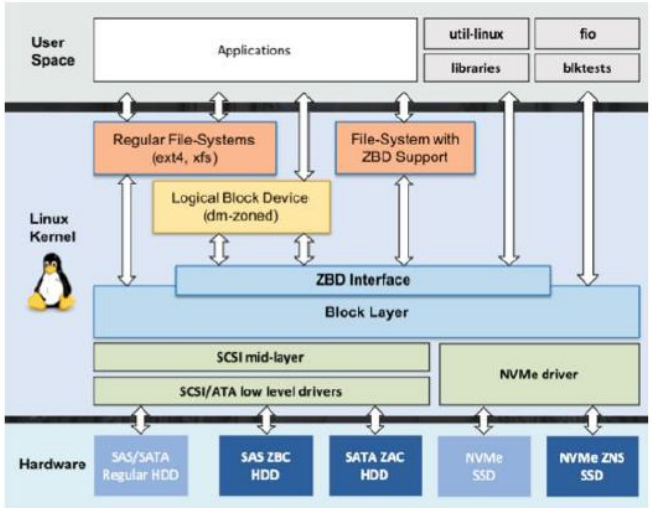


SMR disks study in EOS

- SMR allows increasing storage density by overlapping the data tracks on disks



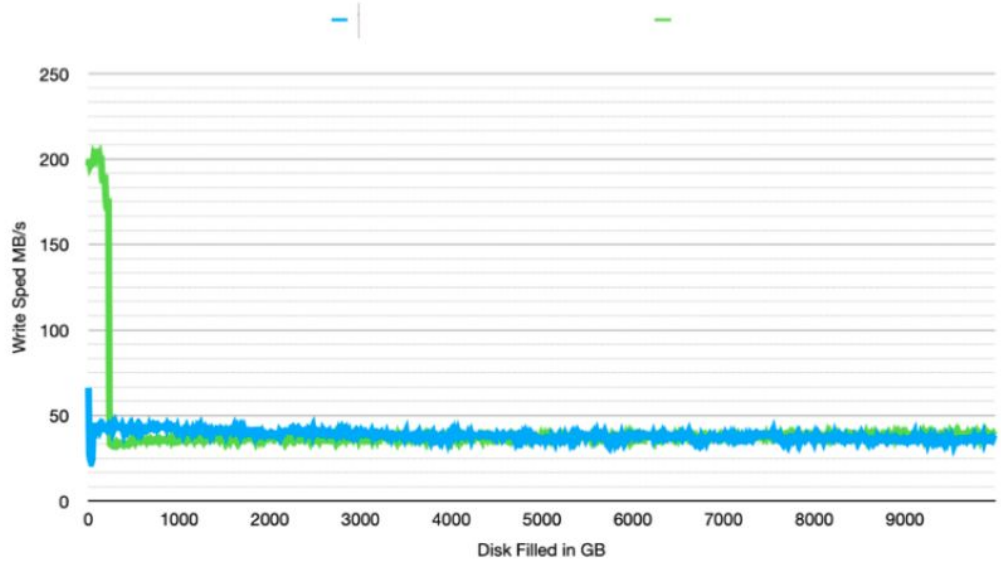
- Mainly two methods in Linux to use SMR disks
  - Conventional fs on top of *dm-zoned* logical device
  - Modern fs with native support for zoned block device



SMR disks study in EOS

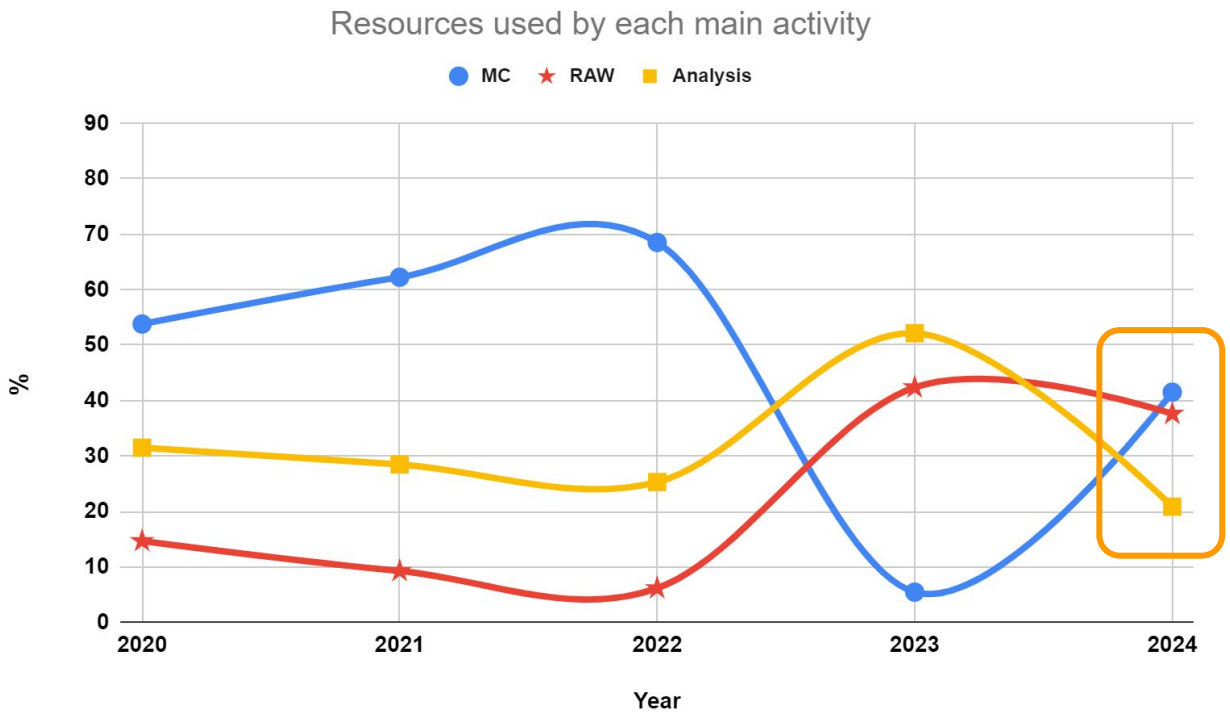
dm-zoned

- Poor sequential write performance with speed around 40 MB/s
- Good sequential read permormance around 200 MB/s



Part 2 - Activities and Operation

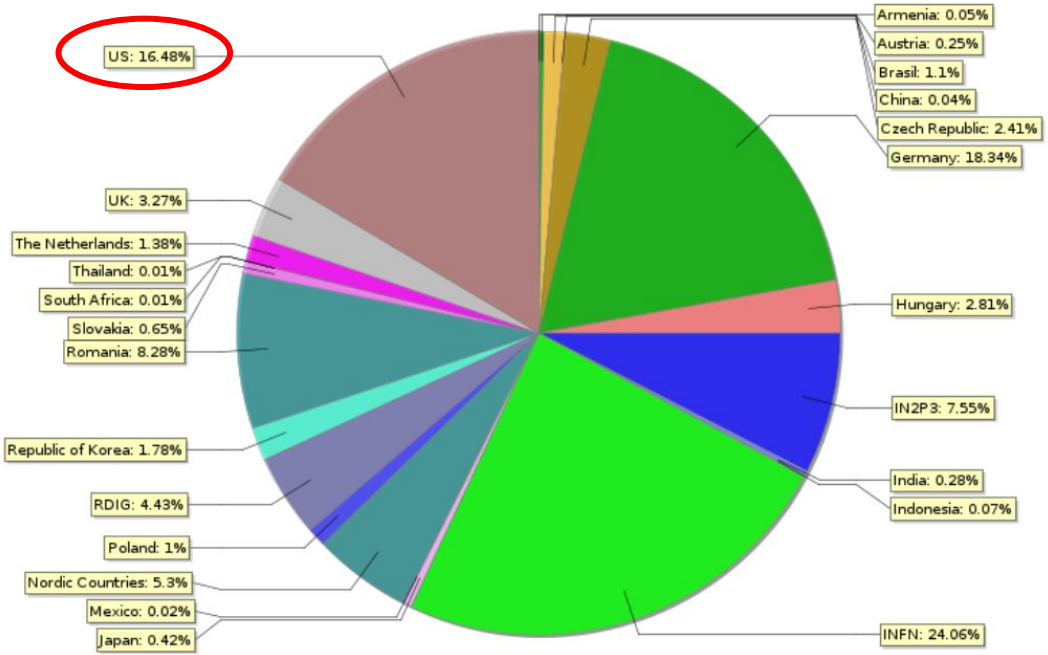
ALICE resources use



- Substantial decrease of MC as primary resource user
- RAW data reconstruction and analysis have taken lead
  => **more I/O intensive** tasks
- 2024 is representative for Run-3 type load on the Grid

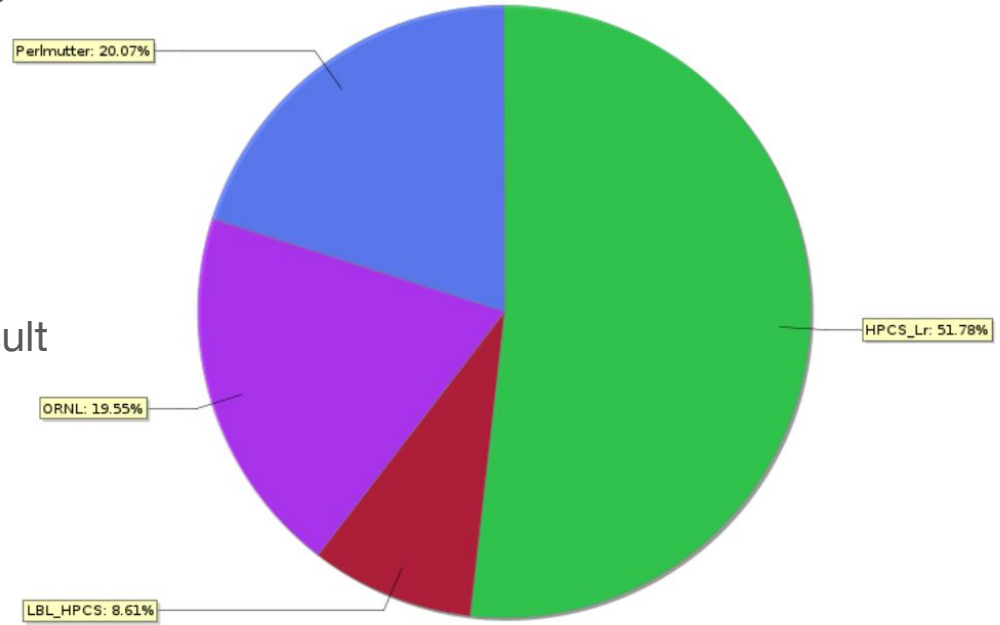
Regional contribution

- **16.5% US contribution** to CPU wall clock time
  - Comparable to the large T1s
- **8% US contribution** to disk
  - This can be increased substantially to balance the CPU/disk ratio

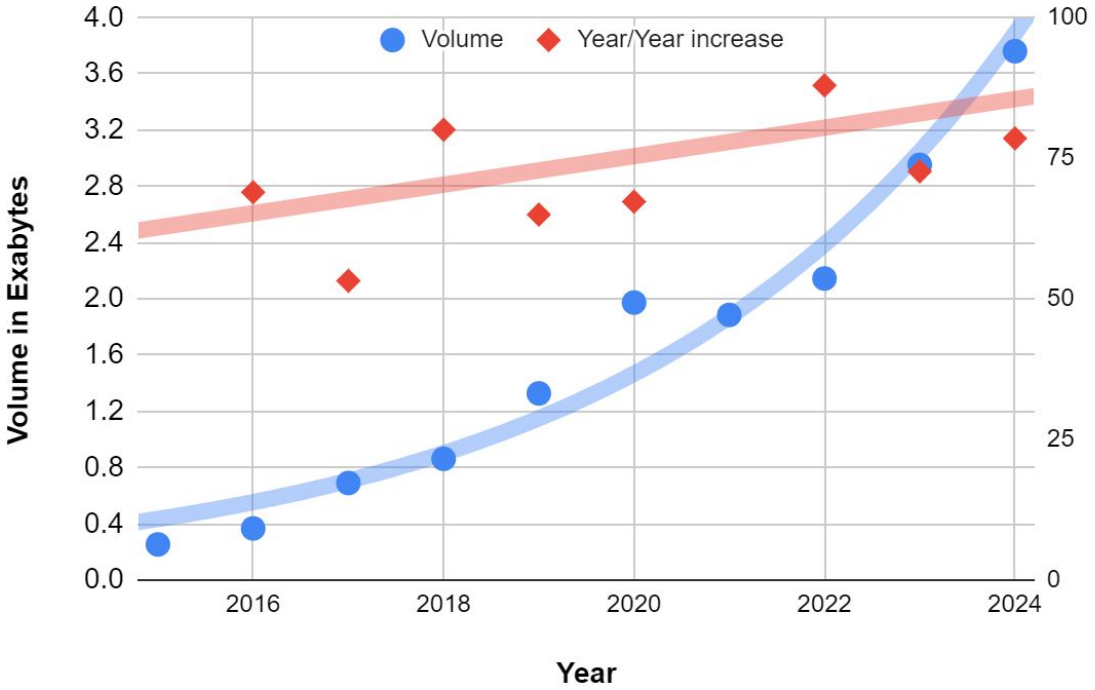


Contribution per centre

- Lawrencium contributes ~½ of the US resources (!)
- Perlmutter role is increasing
- The combination of the opportunistic resources is about ⅔ of the total US contribution
- This is excellent bottom line and a result of a lot of work from the ORNL and LBNL teams + UPB
-

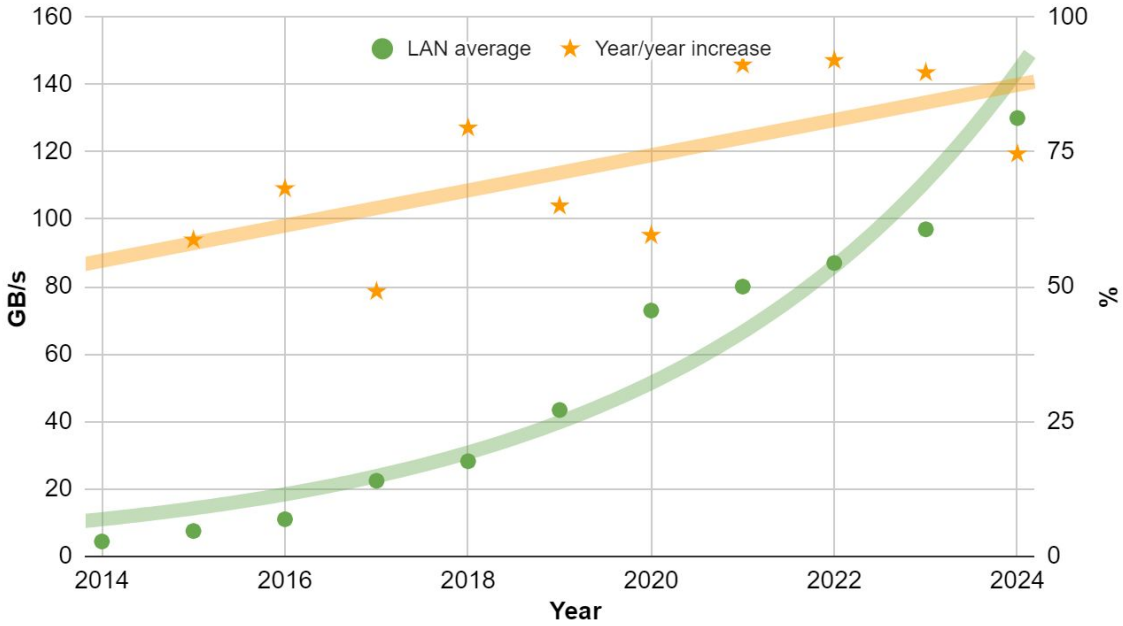


Data volume



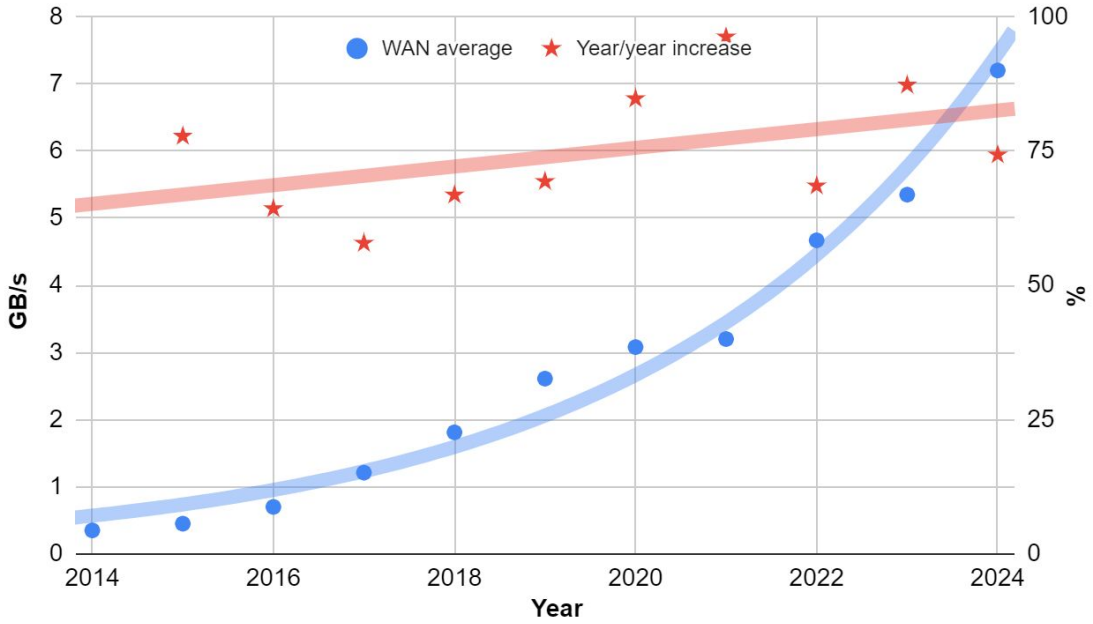
- Exponential growth of data access
- Surprising year/year increase
  - Volume growth ~15%,
  - Access growth **~75%**
  - Network growth and innovation fully supports the access
- The infrastructure moves further into the **HTC zone**

Data access - LAN



- LAN traffic has increased ~15x in the past 10 years
- Substantially above the storage capacity growth
  - Also seen from the previous slide
- SE resiliency and LAN infrastructure have largely followed the trend and have not become (yet) a blocking factor
  - This growth favours large storage capacity
  - Comes at substantial cost

Data access - WAN (LHCONE/LHCOPN)

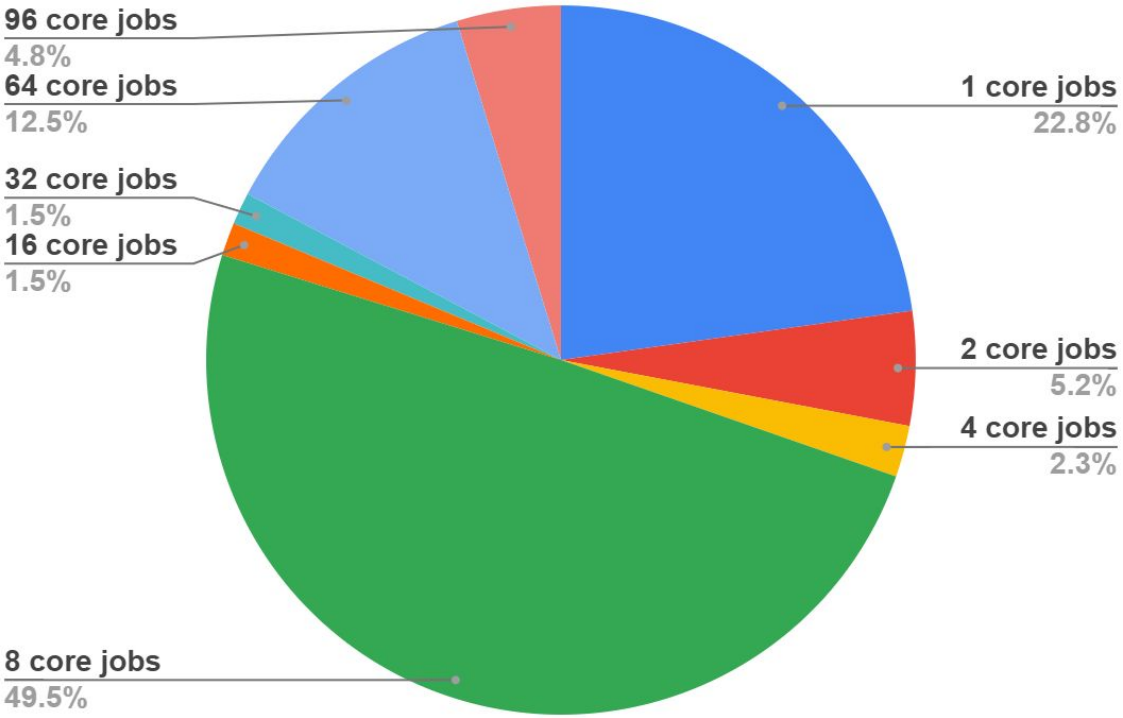


- WAN traffic is **~5% of the LAN**
- Comprises of data transfer between sites (about ½) and client access to remote storage
  - In case of local SE failure
- The increase is ~flat, corresponding to the storage capacity growth

Storage operation

- Alice’s presentation

Cores per job type in the past year



- 1-8 core jobs can be executed in a ‘standard’ WLCG queue
- Significant amount of jobs (**~20%**) require **>8 cores**
  - These are usually for payloads using GPUs and the associated higher memory needs
  - Only possible with whole node submission

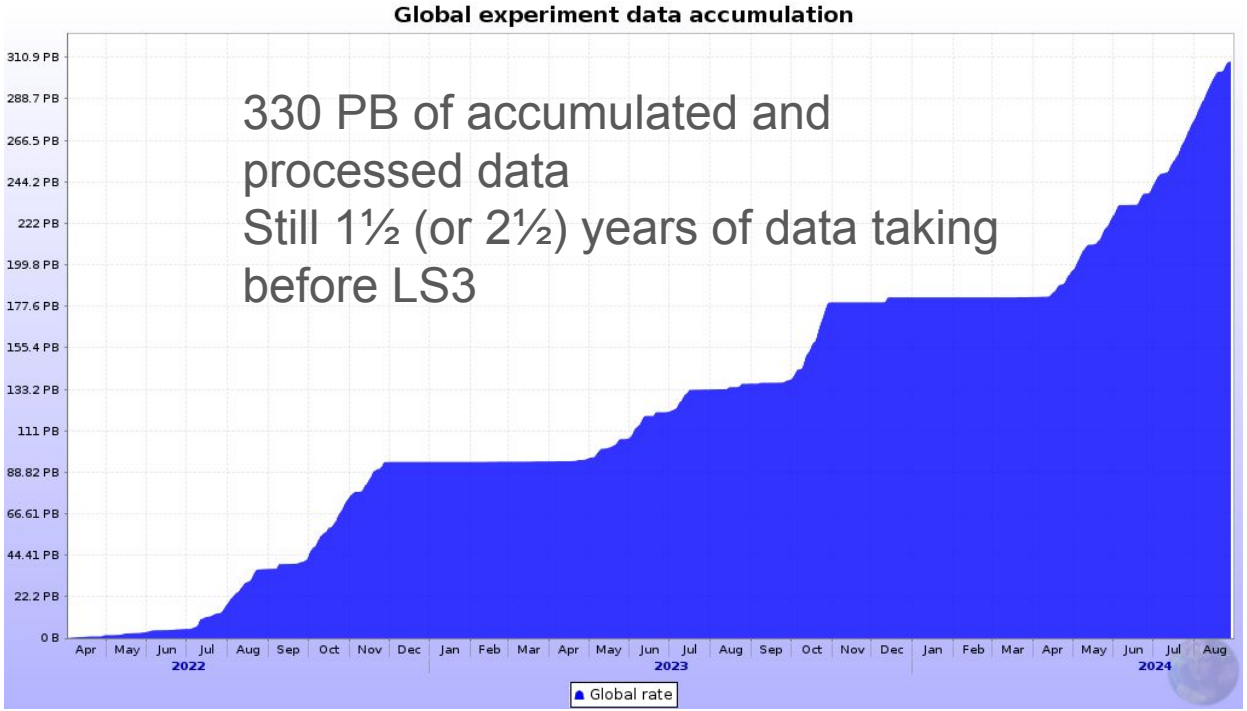
8-core queues versus whole node

- 8-core is the ‘magic’ number for multicore queue on the Grid
  - Advantageous for site uniformity and multi-VO compatibility
- However, 8-core is not a lot of real estate to work with
  - Cannot do ‘smart scheduling’ of jobs with different requirements (8-core slots are individually distributed on the same nodes)
  - <8 core jobs can block a 8-core slot, lowering efficiency
  - Cannot control fully the resources use on the WNs, only of the job slots we have - at the mercy of the OOM
  - Not suitable for all types of jobs, especially those using accelerators

  => **All ALICE-only sites (including US)** are already whole-node submission
- New WLCG workgroup on CPU management is being formed

2022-2024 data collection

- Record-breaking data volume in the past 2 ½ years of data taking
- LHC delivered luminosity in Run3 is already higher than Run1+2
- Processing of the data keeps pace with its accumulation
  - About 10PB of CTFs/week



Summary

- In Run3 ALICE collected record amount of p-p and Pb-Pb data with upgraded detector, new online, offline and Grid software
- The Grid sites are updated and continue to be the backbone of the ALICE data storage and processing
- Alternative resources are increasing in relevance and volume
  - JAliEn is able to integrate these with relative ease
  - Software validation requires a substantial effort
- The processing strategy continues to depend on good network connectivity for data exchange
  - The network progress is impressive and fully covers the needs
- ALICE computing requirements will increase
  - There is a high probability that LS3 will be delayed by one year => Full data taking in 2026