The ALICE Grid upgrade, Grid methods and tools for LHC Run 3 and beyond

Antonio Uras, Latchezar Betev

Journées LCG-France, CC-IN2P3, 4-6 Dec 2024

Part 1 - Grid middleware development

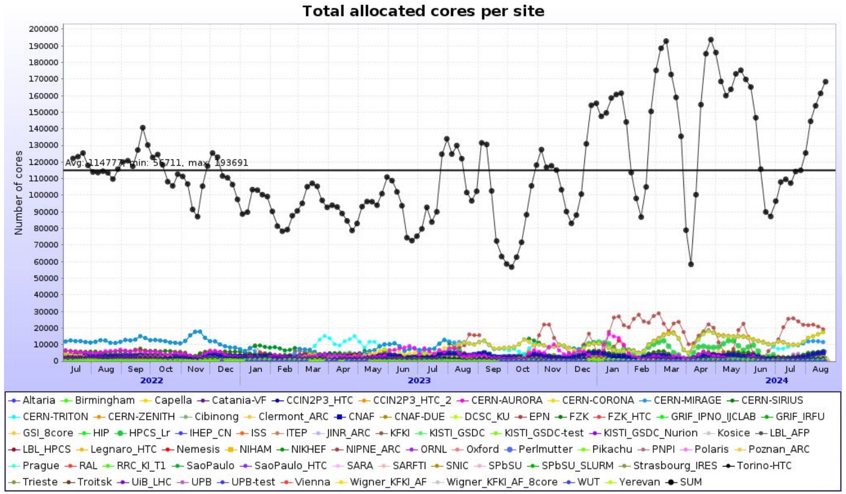
The ALICE Grid in Run3

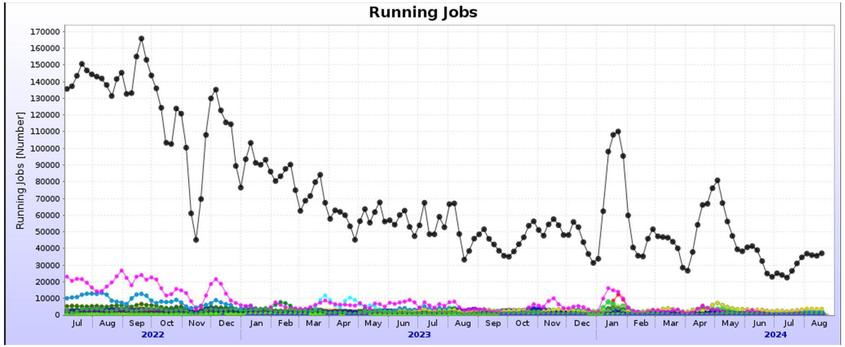
- ALICE upgrade in brief
  - Entirely upgraded detector with 10x faster readout
  - New DAQ capable of 3.5TB/s data throughput
  - New multiprocess/multicore GPU-enabled online and offline (O2) software
  - New Grid middleware in Java capable of managing the above
- The Grid did not stand still
  - The usual 15% CPU/storage growth
  - More sophisticated and capable network
  - New type of resources, including new supercomputer sites
- In the next slides
  - Outline of some of the challenges we are facing
  - Projects under development and possibilities for collaboration

From jobs to cores

- Gone are the simple days of 1 job = 1 CPU core
- New diverse landscape of 1- 2- 4- 8- 64- core jobs and accelerators in the mix
  - Complicated to keep all x-core combination queues at all sites
  - SuperComputers do not provide resources this way
  - This was expected and planned to be addressed in the Grid middleware

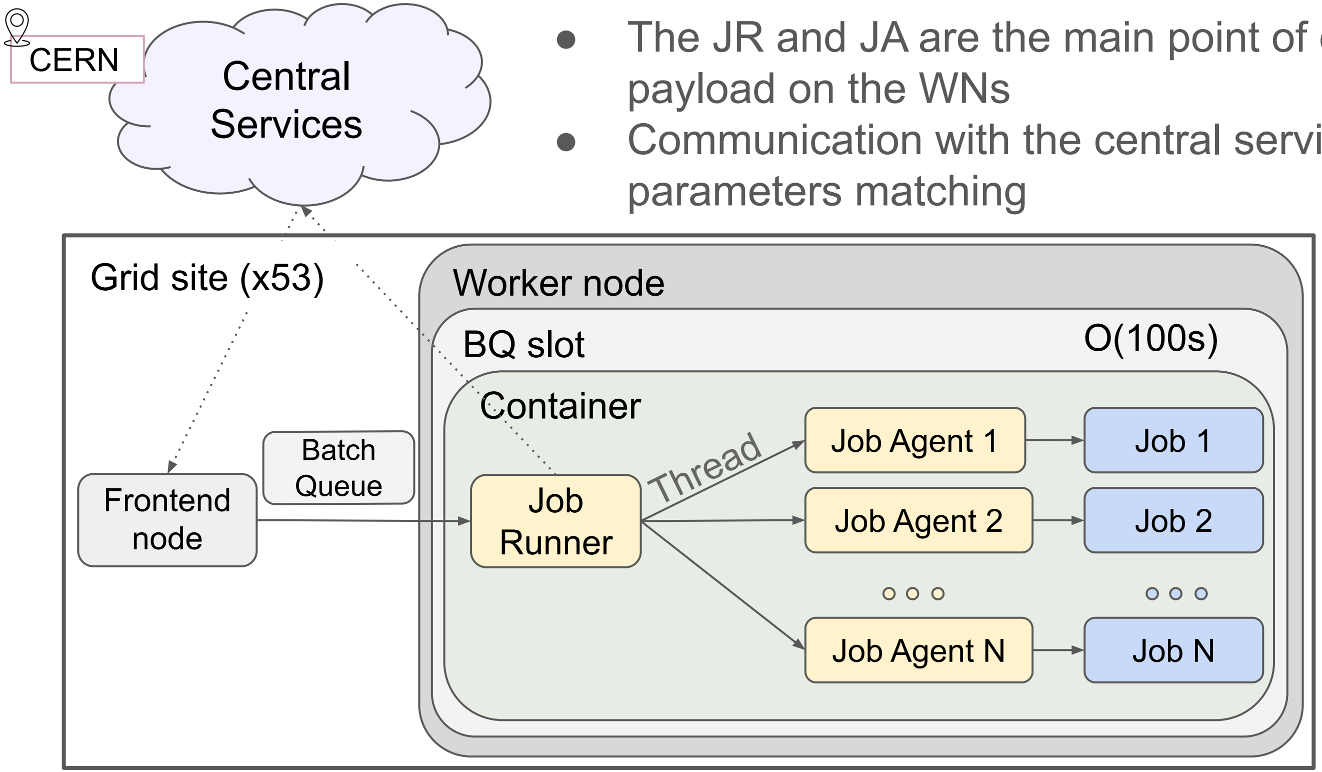
=> One of the main tasks of the new middleware is to ***simplify the site-side ALICE workload management***





JAliEn - The JobRunner and JobAgent

- The JR and JA are the main point of control of the payload on the WNs
- Communication with the central services is critical for parameters matching



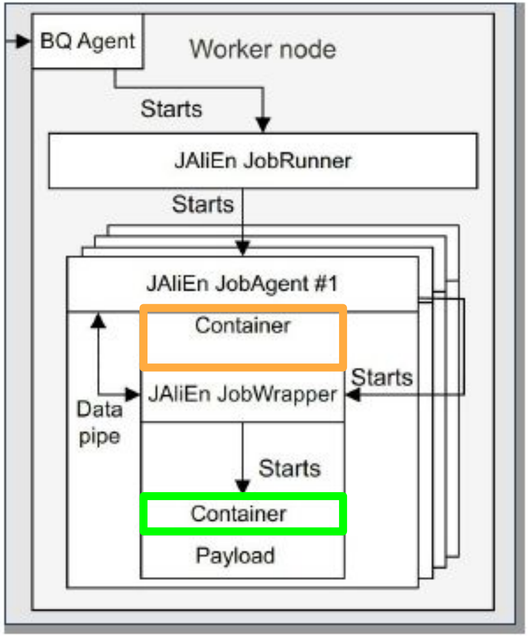
Worker nodes are typically shared with other workloads

Batch Queue slots of custom resource allocations per site

Different workloads have different resource requirements

Payload isolation

- Another main function of the JAliEn middleware is to provide payload isolation and control
- All Grid jobs are wrapped by a top-level container - provides a tried-and-tested environment across sites/nodes
  - Additional isolation from WN host
  - Images built and tested centrally and put in CVMFS
  - Specific image selected by JAliEn based on required packages for job
  - Currently supporting CentOS 7, Alma 8 or Alma 9
  - Multiple containers can be used in the same batch job slot
  - Full GPU support



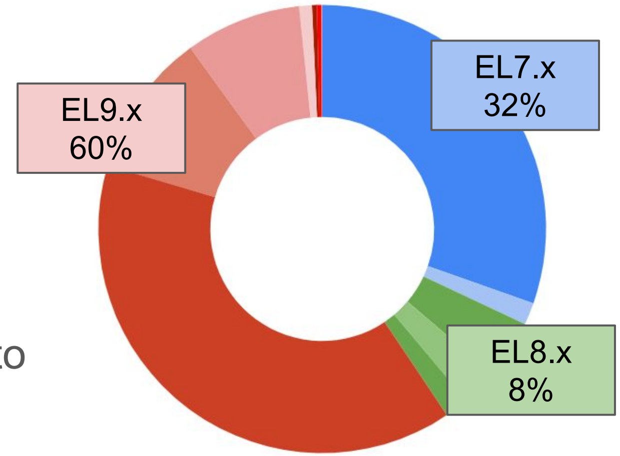
Payload scheduling

- 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

ALICE

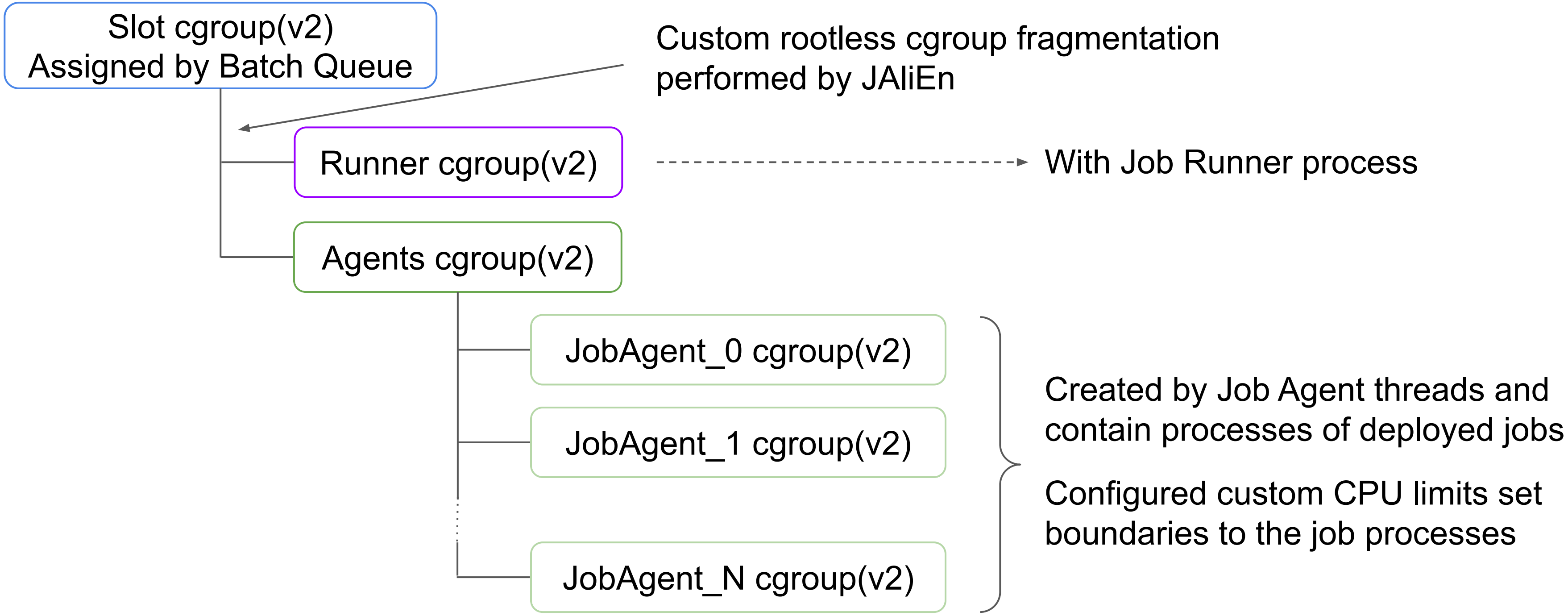
Resources use control

- 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



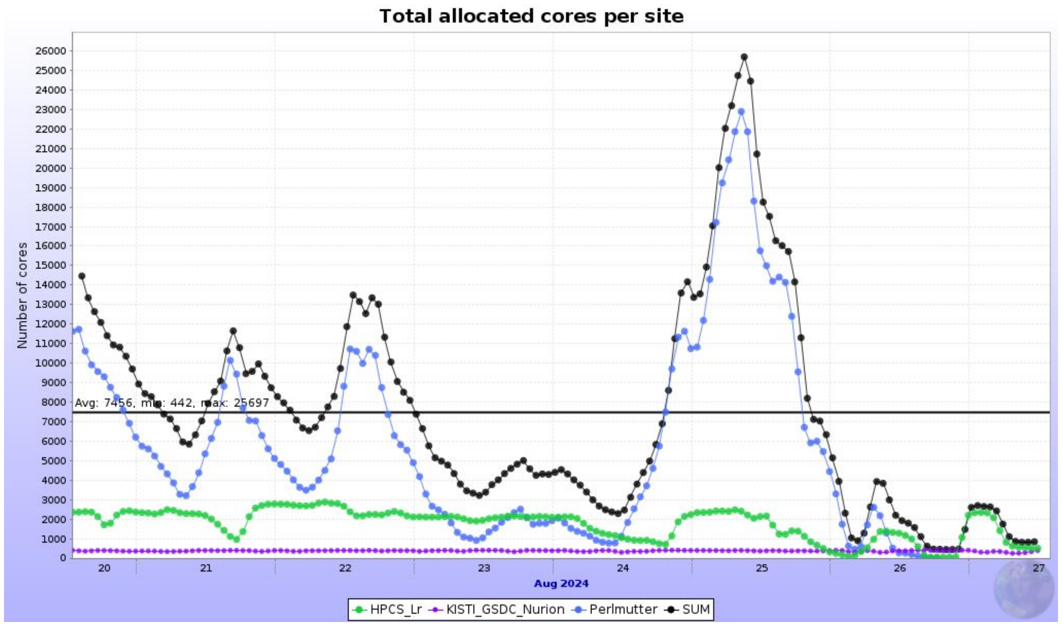
Alternative resources

- 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

ALICE

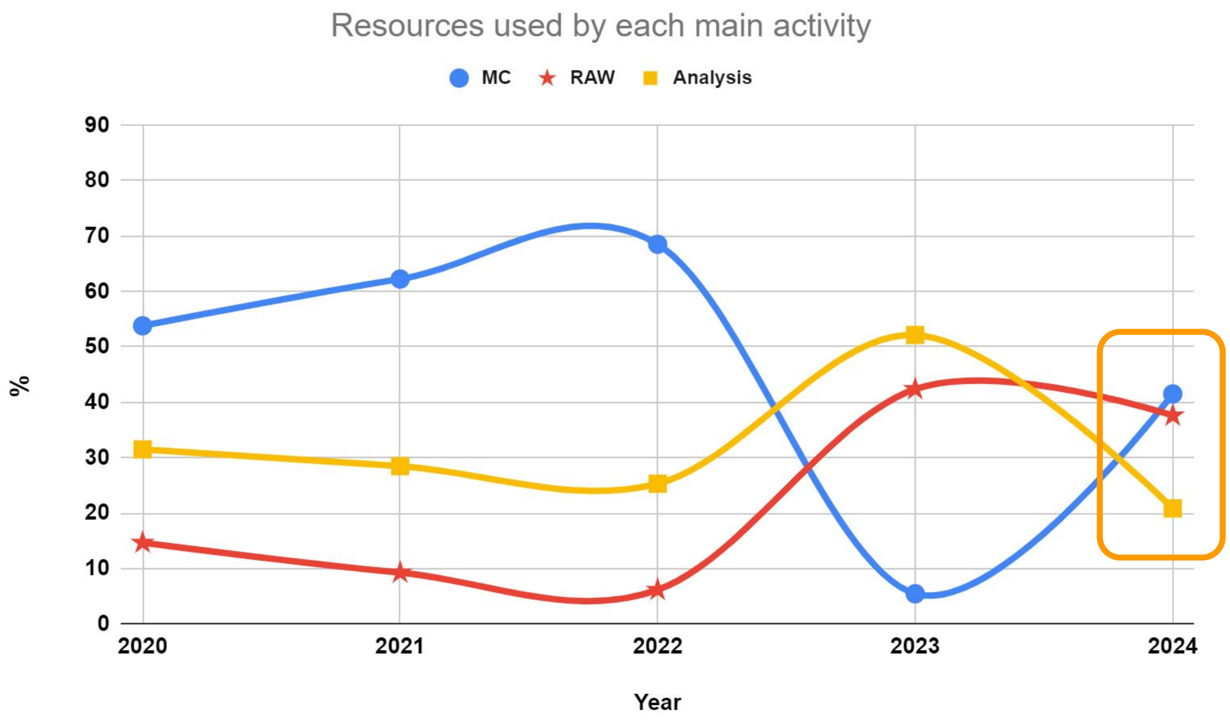
Supercomputers - briefly

- Three HPCs 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



Part 2 - Activities and Operation

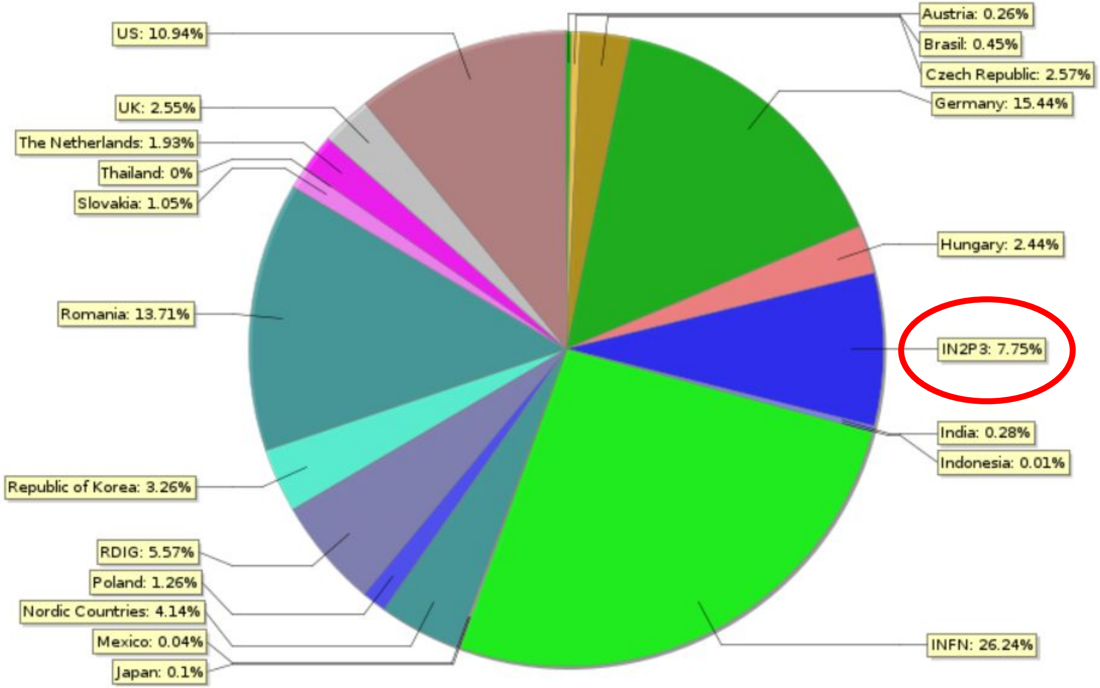
ALICE resources use



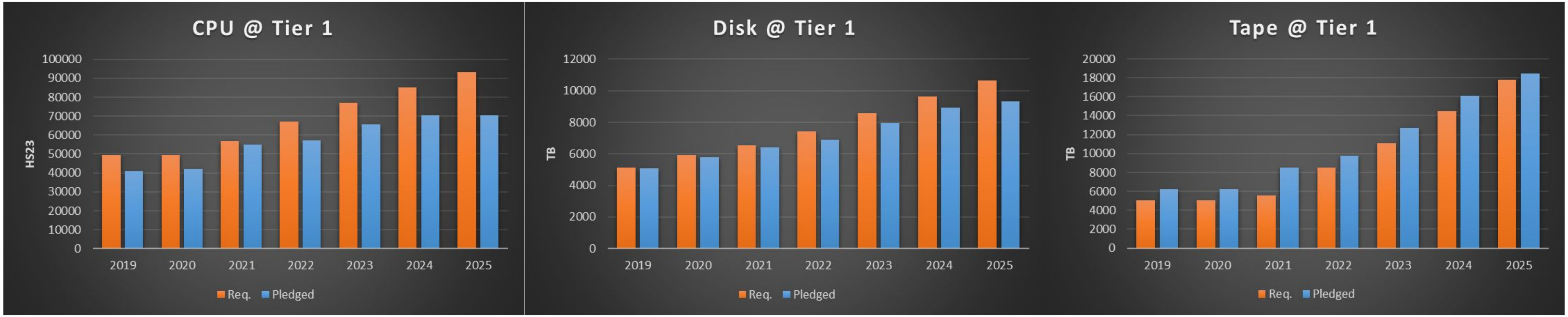
- Substantial decrease of MC as primary resource user
- RAW data reconstruction and analysis have taken lead
  => **more I/O intensive** tasks
- 2024 is representative for Run-3 type load on the Grid

Regional contribution

- ~8% FR contribution
- Diminishing role of smaller T2 centres - this is an unfortunate global trend
  - ALICE lost Subatech

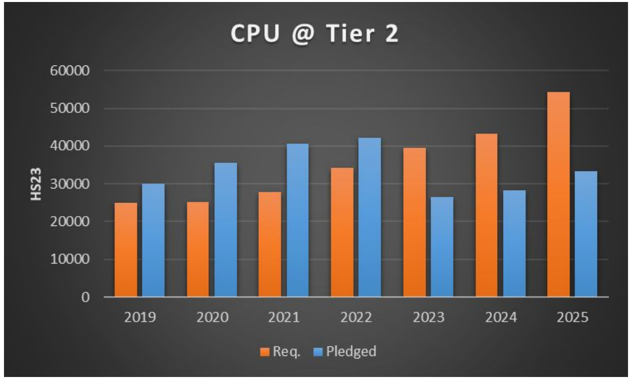


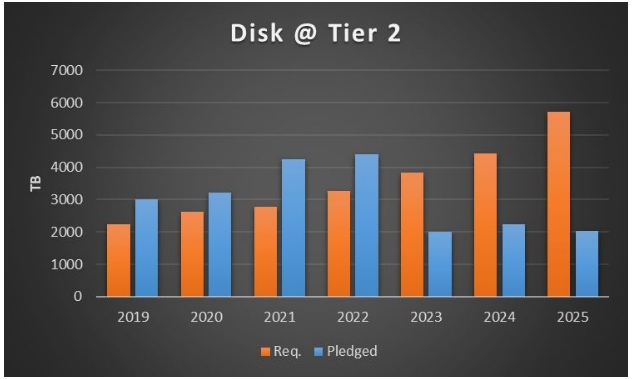
Computing resources @ CCIN2P3



- CCIN2P3 provided reliable resources to ALICE
- Resource growth aligned with ALICE’s priorities during data-taking years:
    - Tape -> disk -> CPU
    - Expected lower growth during LS3
    - This will help offset the small deficit in disk and CPU

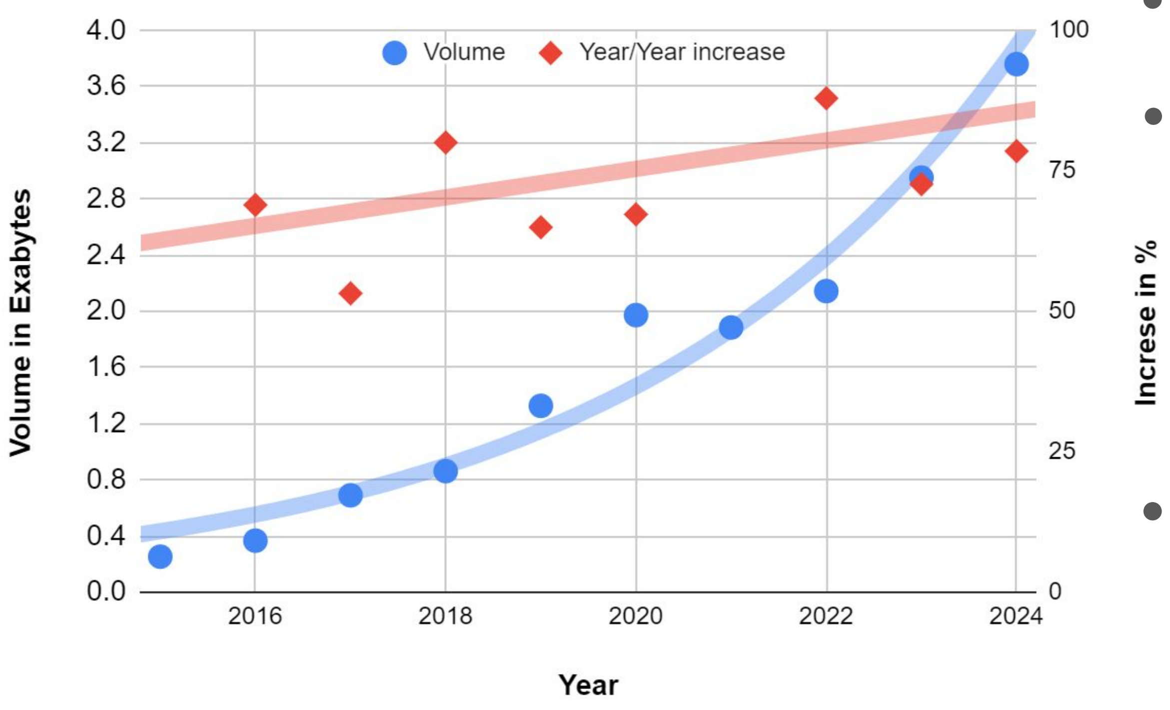
Tier 2 computing resources





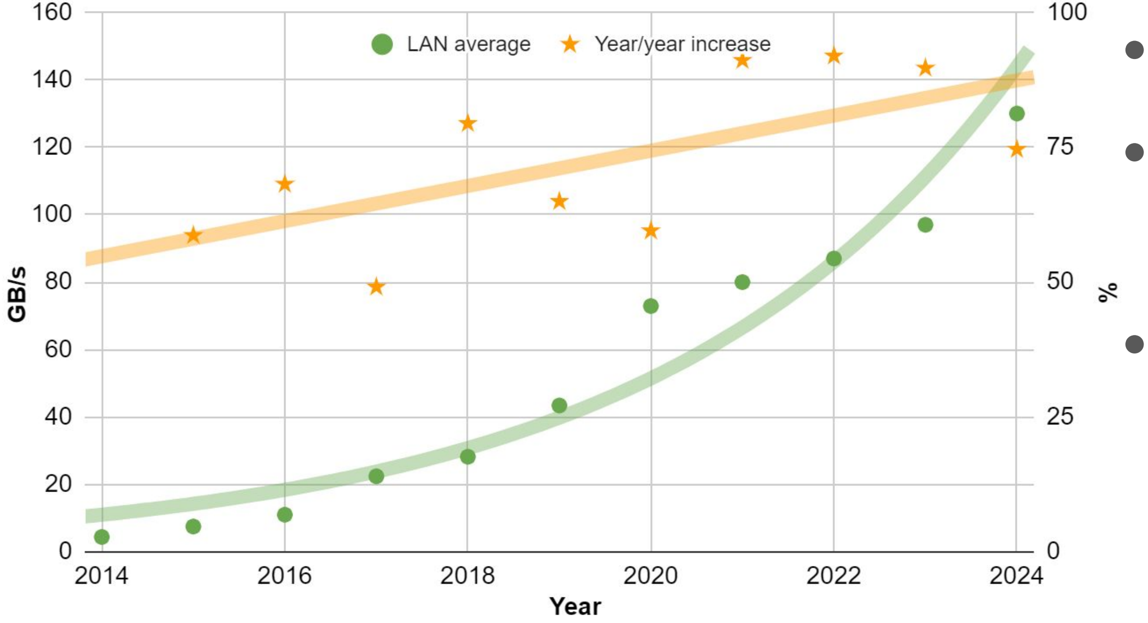
- Considered only IN2P3 excluding CEA
- Until 2022 more resources were provided than the expected share
- After decommissioning of the Subatech site:
  - No new resources were added to compensate for Subatech's resources
  - Particularly for disk resources, where small and fragmented disks are harder to manage efficiently
  - Diskless site should be well connected to other sites with storage elements that have sufficient capacity

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**

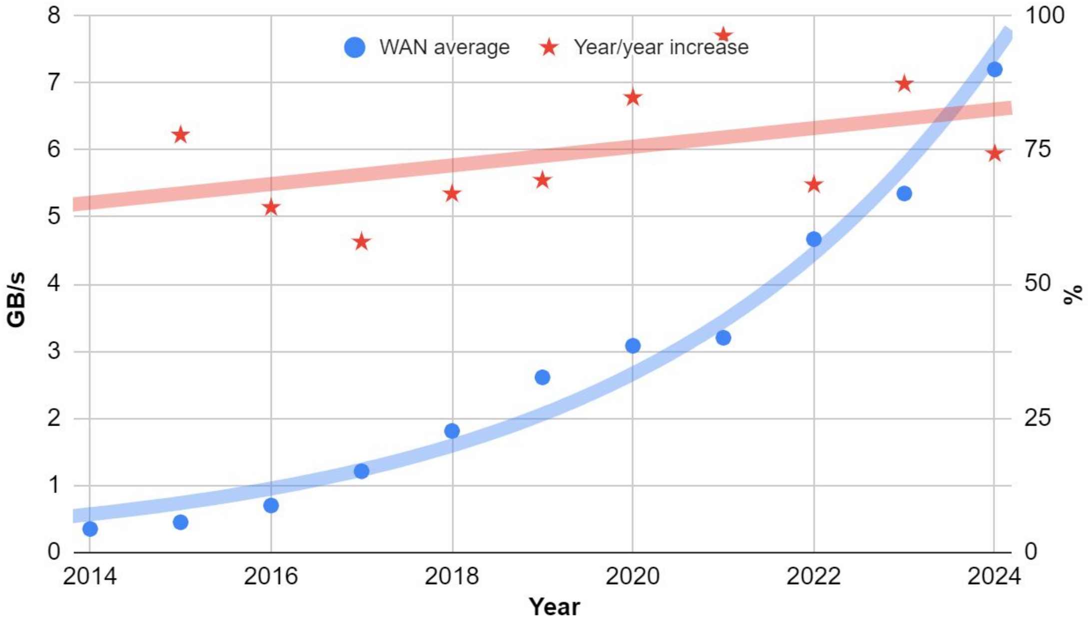
Data access - LAN



- 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

ALICE

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 monitoring and testing

- Testing the integrity of the SEs
  - File crawler - Mimicking normal jobs, sorts the failure rates; Reporting on file health, throughput and accessibility
  - Dark data detector - compares file catalogue list with storage content, uses EOS built-in tools; keeps the dark data on <.% level

Status codes extracted from the crawler

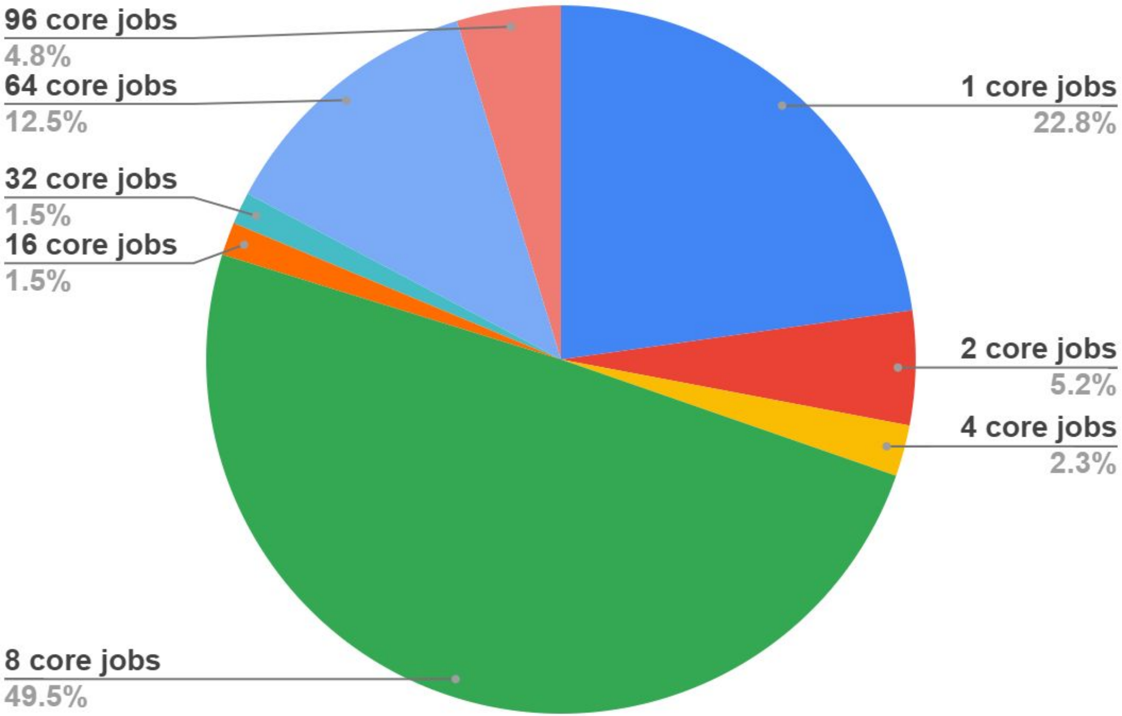
SE Name: ALICE::HIROSHIMA::EOS **Interval:** Last week

Status Type	Status Code	Status Count	Status Code Ratio	Download throughput
FILE_OK	S_FILE_CHECKSUM_MATCH	26972	99.79 %	21.97 Mb/s
	E_CATALOGUE_MD5_IS_BLANK	2	0.01 %	19.04 Mb/s
INTERNAL_ERROR	XRDFS_CANNOT_CONFIRM_UPLOAD	21	0.08 %	
FILE_INACCESSIBLE	XROOTD_EXITED_WITH_CODE	35	0.13 %	
TOTAL		27030	100 %	

Averaged metrics for the selected interval

SE Name	Start	End	Success ratio	Corrupt ratio	Inaccessible ratio	Internal error ratio
SARA::DCACHE	18 Oct 2022 06:08	17 Nov 2022 10:32	99.87 %	0.09 %	0.05 %	0.00 %
Hiroshima::EOS	18 Oct 2022 06:08	17 Nov 2022 10:33	99.73 %	0.00 %	0.18 %	0.09 %
SNIC::DCACHE	18 Oct 2022 06:12	17 Nov 2022 10:28	99.68 %	0.02 %	0.27 %	0.03 %
Vienna::EOS	18 Oct 2022 06:07	17 Nov 2022 10:38	99.60 %	0.24 %	0.16 %	0.00 %
NIPNE::EOS	18 Oct 2022 06:09	17 Nov 2022 13:03	99.58 %	0.03 %	0.37 %	0.03 %
Trieste::SE	18 Oct 2022 06:11	17 Nov 2022 12:11	99.54 %	0.11 %	0.35 %	0.00 %
Bari::SE	18 Oct 2022 06:04	17 Nov 2022 12:22	99.50 %	0.08 %	0.42 %	0.00 %
IHEP::SE	18 Oct 2022 06:07	17 Nov 2022 10:20	99.35 %	0.11 %	0.53 %	0.01 %
Torino::SE2	18 Oct 2022 06:09	17 Nov 2022 11:07	99.34 %	0.13 %	0.53 %	0.00 %
Troitsk::SE	18 Oct 2022 06:04	17 Nov 2022 10:43	99.26 %	0.54 %	0.19 %	0.01 %
CERN::EOS	18 Oct 2022 06:12	17 Nov 2022 10:47	99.19 %	0.08 %	0.65 %	0.07 %
CNAF::SE	18 Oct 2022 06:10	17 Nov 2022 10:35	99.06 %	0.02 %	0.92 %	0.00 %
FZK::SE	18 Oct 2022 06:11	17 Nov 2022 10:33	98.86 %	0.06 %	1.07 %	0.01 %
Legnaro::SE	18 Oct 2022 06:04	17 Nov 2022 10:26	98.54 %	0.03 %	1.34 %	0.09 %
UPB::EOS	18 Oct 2022 06:08	17 Nov 2022 10:32	98.49 %	0.07 %	1.44 %	0.00 %
ORNL::EOS	18 Oct 2022 06:06	17 Nov 2022 10:31	98.18 %	0.46 %	1.36 %	0.00 %
NDGF::DCACHE	18 Oct 2022 06:04	17 Nov 2022 10:30	97.89 %	0.23 %	1.87 %	0.00 %
NIHAM::EOS	18 Oct 2022 06:08	17 Nov 2022 10:49	97.75 %	0.12 %	2.13 %	0.00 %
GRIF::EOS	18 Oct 2022 06:05	17 Nov 2022 10:31	97.75 %	0.05 %	2.20 %	0.00 %
Subatech::EOS	17 Oct 2022 17:38	16 Nov 2022 16:28	97.46 %	0.06 %	0.91 %	1.57 %
JINR::EOS	18 Oct 2022 06:11	17 Nov 2022 12:13	95.93 %	0.13 %	3.92 %	0.03 %
RRC_KI_T1::EOS	18 Oct 2022 06:06	17 Nov 2022 10:28	95.86 %	0.09 %	1.47 %	2.57 %
KISTI_GSDC::EOS	18 Oct 2022 06:07	17 Nov 2022 10:57	95.04 %	3.49 %	1.47 %	0.01 %
CCIN2P3::SE	18 Oct 2022 06:11	17 Nov 2022 10:37	94.27 %	0.02 %	5.69 %	0.02 %
Kosice::EOS	18 Oct 2022 06:07	17 Nov 2022 11:40	93.05 %	0.11 %	6.84 %	0.00 %
Prague::SE	18 Oct 2022 06:06	17 Nov 2022 10:44	90.18 %	0.02 %	9.79 %	0.01 %
Birmingham::EOS	18 Oct 2022 06:05	17 Nov 2022 10:26	87.70 %	0.06 %	12.23 %	0.01 %
Strasbourg_IRES::SE2	18 Oct 2022 06:04	17 Nov 2022 12:46	87.68 %	0.03 %	12.26 %	0.03 %
Catania::SE	18 Oct 2022 06:07	17 Nov 2022 10:23	86.12 %	0.03 %	13.84 %	0.00 %
KISTI_GSDC::SE2	18 Oct 2022 06:07	17 Nov 2022 10:41	86.03 %	0.17 %	13.80 %	0.00 %
LBL_HPCS::EOS	18 Oct 2022 06:04	17 Nov 2022 10:23	85.88 %	1.21 %	12.90 %	0.00 %
Poznan::SE	17 Oct 2022 23:50	17 Nov 2022 10:42	79.63 %	0.33 %	20.04 %	0.00 %
ISS::FILE	18 Oct 2022 06:07	17 Nov 2022 05:19	78.76 %	0.07 %	21.12 %	0.04 %
Kolkata::EOS2	18 Oct 2022 06:09	17 Nov 2022 12:09	68.71 %	0.61 %	30.57 %	0.12 %

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

ALICE

8-core queues versus whole node

- 8-core has emerged as 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

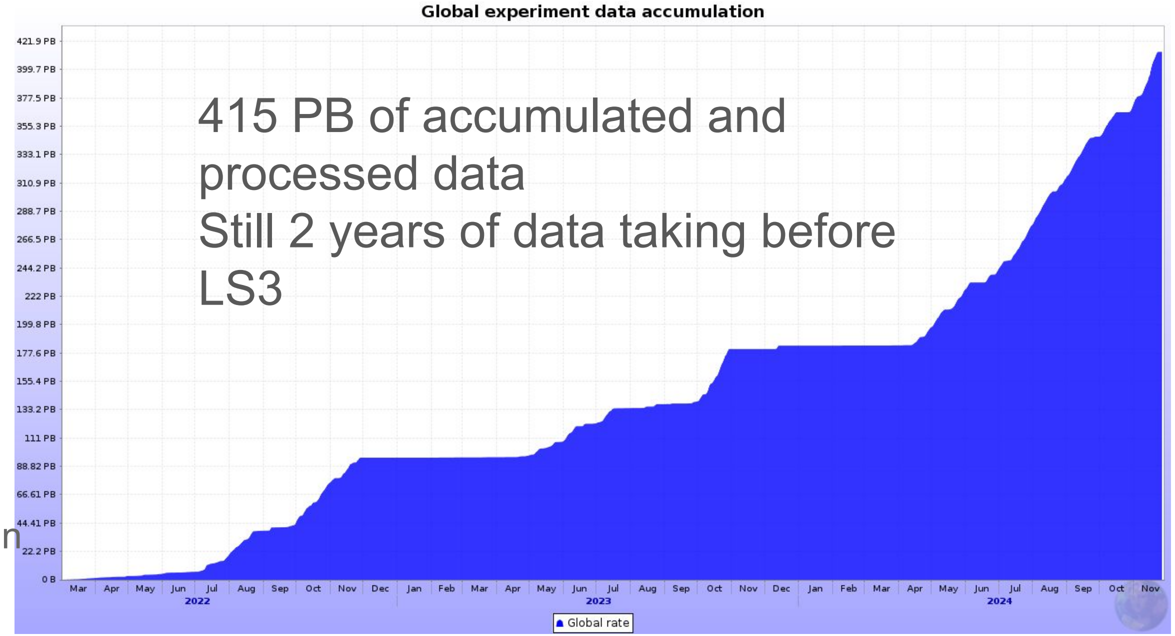
=> **All ALICE-only sites** are already on or being moved to whole-node submission, about 30% of the total capacity

- New WLCG workgroup on CPU management is being formed

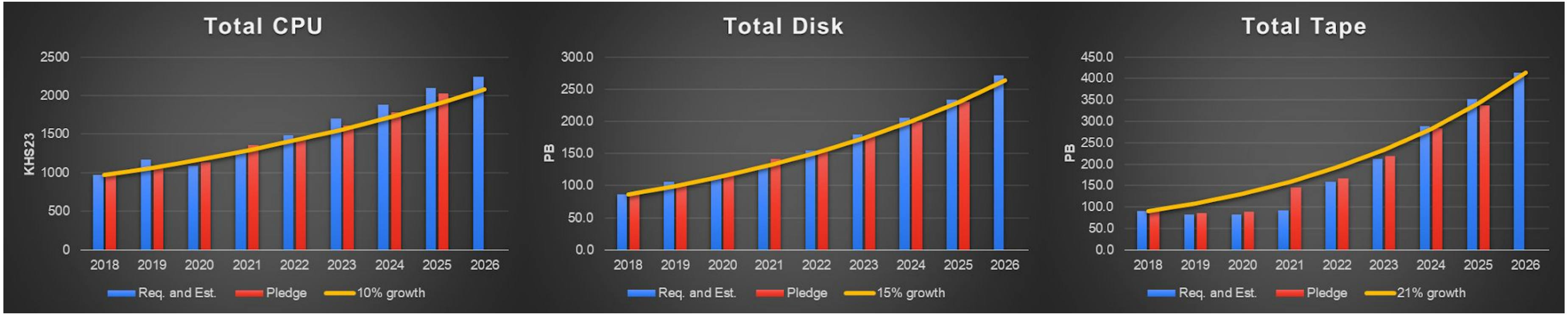
ALICE

2022-2024 data collection

- Record-breaking data volume in the past 3 years of data taking
- Processing of the data keeps pace with its accumulation
  - Use of GPUs is indispensable in this process
  - Alternative resources gain further traction



Outlook for computing resources and IN2P3's share



- During LS2 and Run 3, ALICE's growth was moderate for CPUs, within flat budget for disk, and 20% for tape. A similar growth is expected for LS3+Run 4, with growth flattening during LS3 and then increasing during the data-taking years

- Considering the current projections for the M&O-A, the IN2P3 share excluding CERN will gradually decrease from 7.4% in 2025 to 6.8% in 2028

ALICE

Summary

- ALICE upgrade for Run3 and beyond
  - ALICE collected record amount of p-p and Pb-Pb data from 2022 onward 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
  - Number of ongoing projects to increase its efficient use and include new resources, including HPCs, GPUs and alternative resources
- 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 and we count on the French sites to continue their historically solid support