ALICE feedback on Analysis Facilities

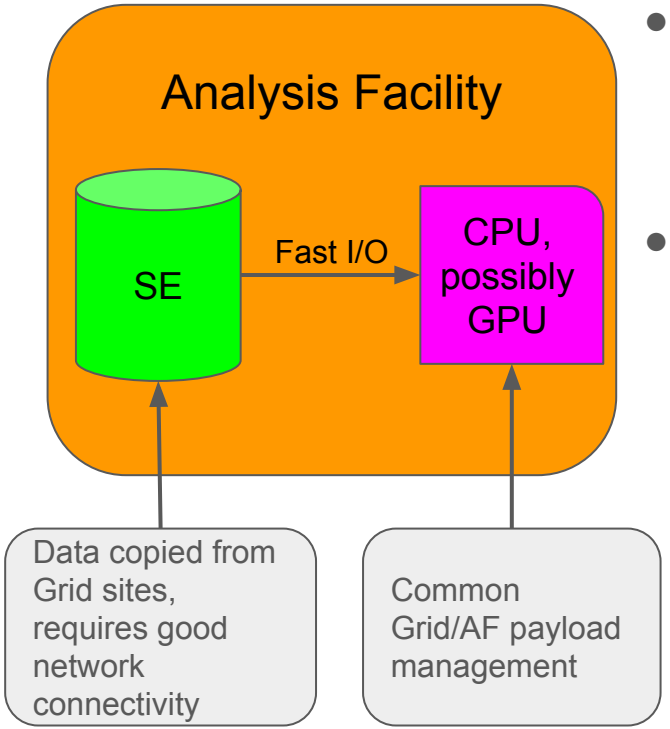
L. Betev

WLCG OTF #3, February 11, 2025

Scope of this presentation

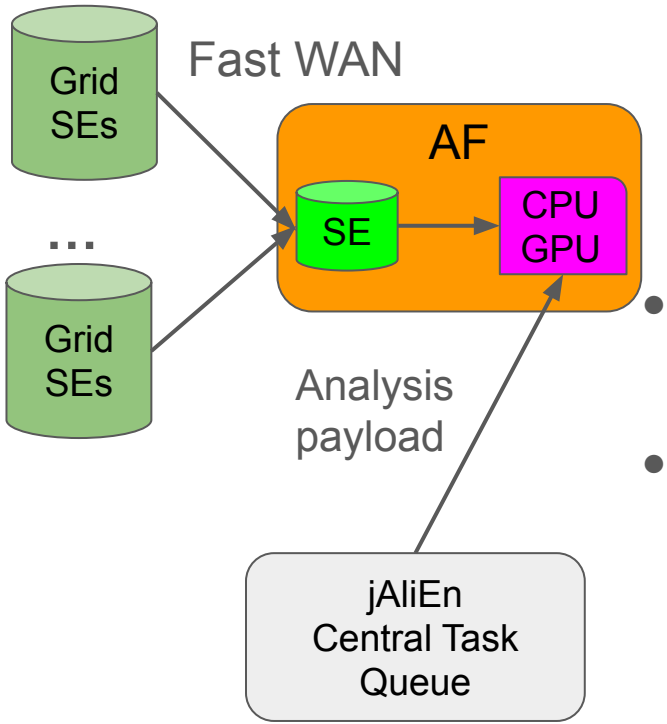
- Guided by the WLCG referees questions on analysis infrastructure
  - Not discussing **point 1** (*Description of the current Run-3 analysis model*) and **point 2** (*Future analysis model in Run-4 and Run-5*) - these are extensively covered in the ALICE upgrade TDR and elsewhere
- Focus on point 3 (*Managing the evolution of the Analysis Infrastructure)*
  - ALICE approach to Analysis Facilities
  - Operation and future plans
  - In the hope this may help and inspire…

Quick recap of ALICE Analysis Facilities vision



- Integral part of the ALICE computing model for Run3+
  - Bulk of data analysis through organized analysis trains on Grid
  - Final steps on physicist laptop/workstation
- Basic architectural considerations
  - Dedicated location with comprehensive data samples from asynchronous and MC data processing at **~10% statistics**
  - Fast I/O, ~15MB/s/core, total of ~10kcores, ~10PB of storage capacity
    - Predictability of analysis turnover
  - Leverage existing Grid sites and middleware
  - Easy functional expansion to include different computing elements, for example accelerators

AF implementation



- Identify AF suitable computing facilities
  - Enthusiastic local support - highly visible part of the analysis system, downtimes are undesirable
  - Adequate internal site structure - fast interconnect between SE and WNs
  - Good WAN connectivity for data transfers
  - Compatibility with Grid operation
- 3 T2s identified for full or partial conversion to AF
  - GSI Darmstadt (GreenCube), KFKI Budapest (Wigner Scientific Computing Lab), LBNL Berkeley (HPC systems)
- Choice of these T2s is not coincidental
  - Full compatibility with architectural requirements for storage, compute size and connectivity
  - Already existing centres with proven operational record
  - Added bonus - lower energy footprint

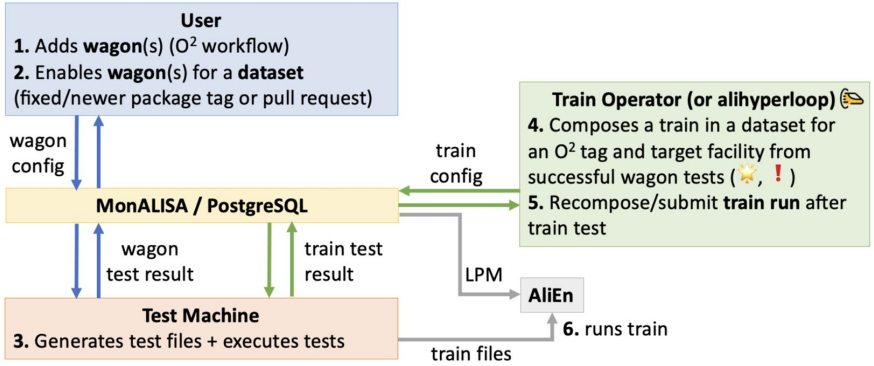
ALICE

AF use cases sample

- Develop and tune analysis algorithms rapidly using a data subset
  - Scale the analysis to the full dataset on the Grid using the same software
- Perform initial data exploration and low-statistics analysis
- Conduct detector calibration
  - Leveraging the suitability for specific datasets and the need for fast turnaround
- Ensure seamless transition between AF and Grid
  - Both run the same O2 code, guaranteed consistent execution across all types of resources

Hyperloop system

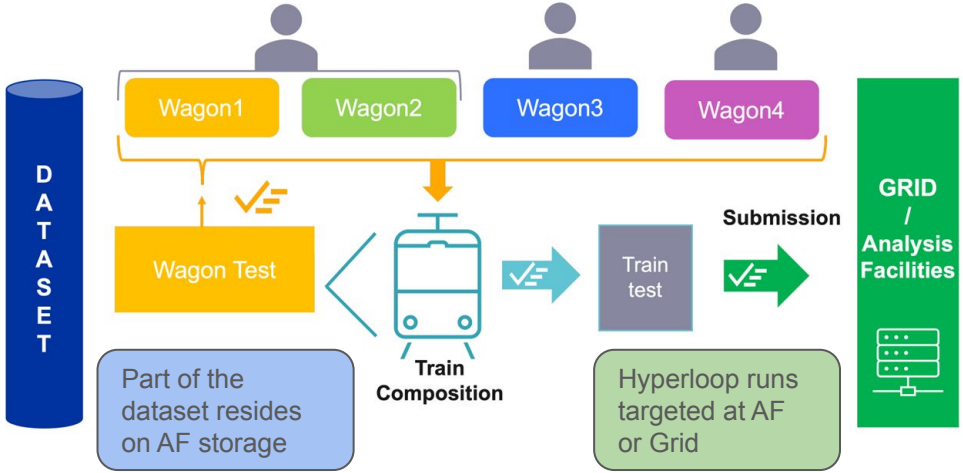
- Facilitates the organisation and execution of all large-scale analysis in ALICE
- Fully integrated with jAliEn workload management and MonALISA
- Individual physics tasks are organized into wagons
  - Wagons are assembled onto trains to run over the same dataset (minimize I/O, maximise efficiency)
  - Operates on daily builds of O2Physics code, distributed through CVMFS

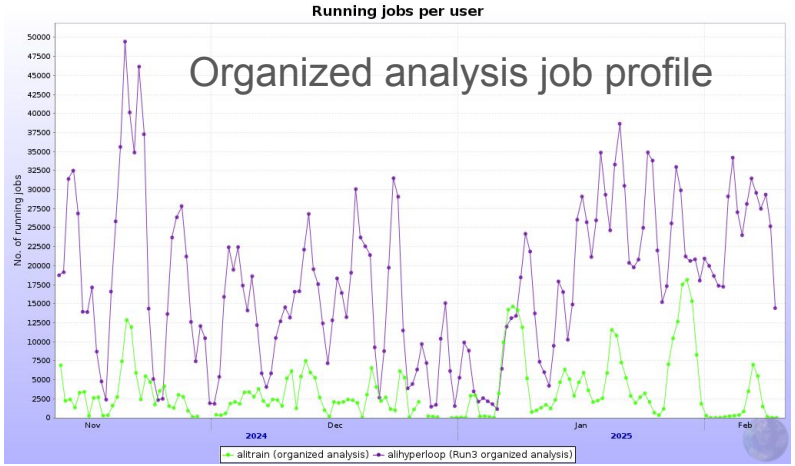


- Hyperloop offers
  - Advanced web interface with user and operator views
  - Immediate and automatic wagon test
  - Automatic train submission under certain defined conditions
  - Wagon and dataset bookkeeping

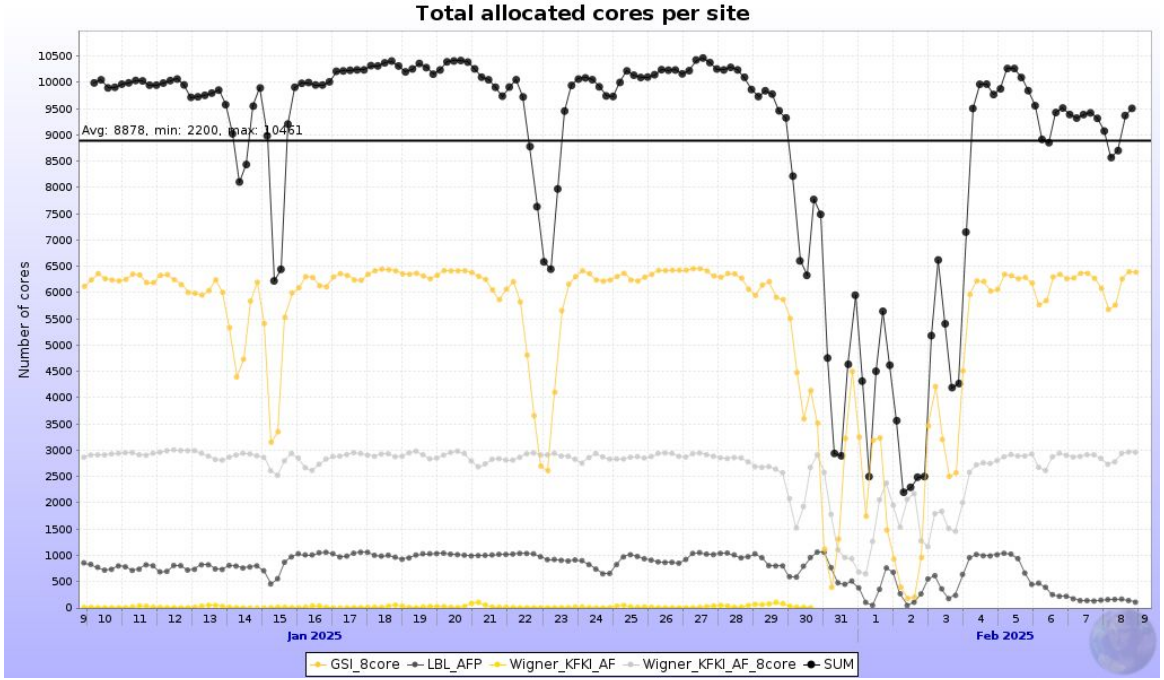
Role of AF

- Organized analysis runs on Grid or AF, depending on the dataset
  - Daily O2Phsyics builds, trains typically start later afternoon, results ready overnight
  - Analysis trains can run at any time, for example these which differ only by applied cuts





AF operation



- AFs are in operation since 3 years
- Execute analysis jobs with priority
  - If free resources are available - backfill with normal Grid payload
- Targeted distribution of datasets
  - Sites receive specific datasets, depending on availability of space

ALICE

Challenges and future work

- Data management per AF
  - Current dataset placement relies solely on space availability
    - This may result in excessive data movement over long distances
  - Implement a data proximity/throughput based system to optimize AF storage transfers
  - Develop a popularity service for automatic dataset removal
  - Explore and experiment with new storage technologies (if available)
- Consider the use of alternative resources
  - Deploy GPUs at one of the AFs in Europe
  - Leverage existing payload matching/container systems
- Hyperloop development
  - Organized analysis is tightly coupled with AF usage
  - Maintain continuous development to incorporate new use cases, improve the control system