# Operation of the ALICE Hyperloop analysis train system


Vladimir Kovalenko (Saint Petersburg State University)
for the ALICE Collaboration

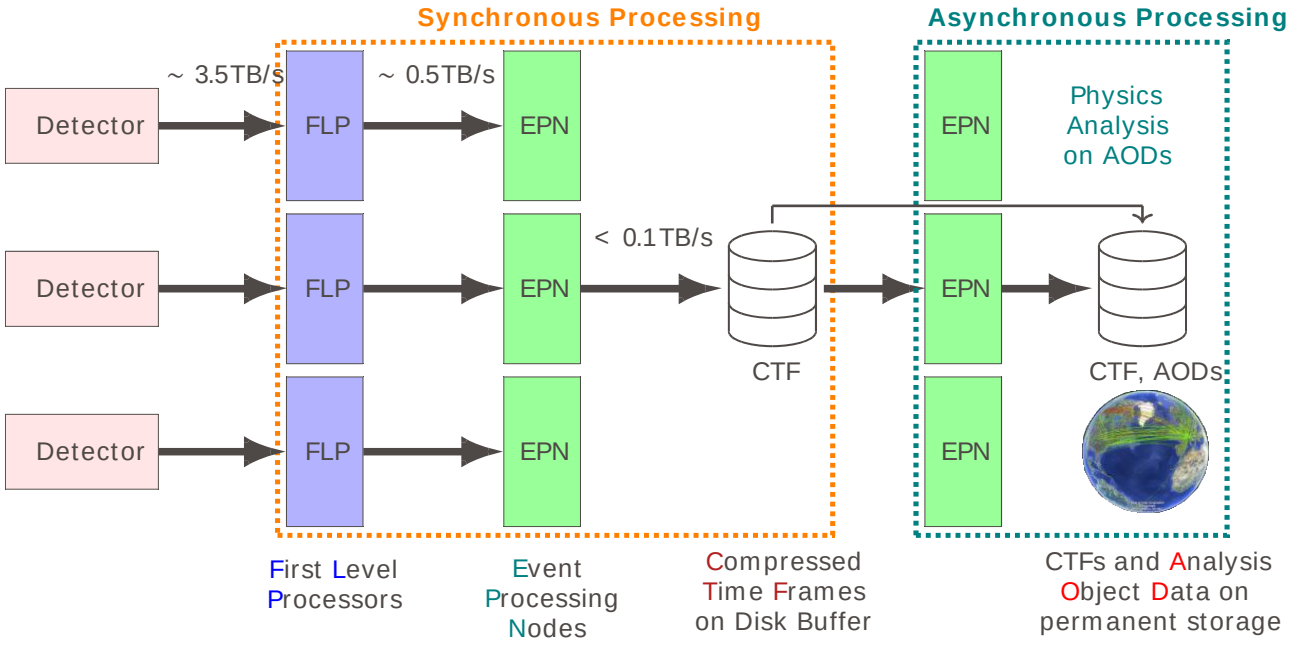
Seminar: The use of new methods for processing data of a physical experiment. Application of machine learning methods on the NICA complex. 28–29 Aug 2023, St. Petersburg, Nevsky 1

Supported by Saint Petersburg State University, project ID: 94031112.

# Outline

- WLCG infrastructure and AliEn framework in Run 3
- O2 Analysis Framework
- Hyperloop wagons and trains
- Comparison with LEGO trains
- Type of wagons: Service wagons and user analysis wagons
- Hyperloop web application
- Automatized wagon test with and train test / submission
- Bookkeeping and preservation
- Operation experience of the Hyperloop system
- Current status of the analysis in Run 3

# LHC Run 3 challenges of Data Processing




1 month of Pb–Pb data would produce several PB of final AO2Ds

# WLCG infrastructure and AliEn framework

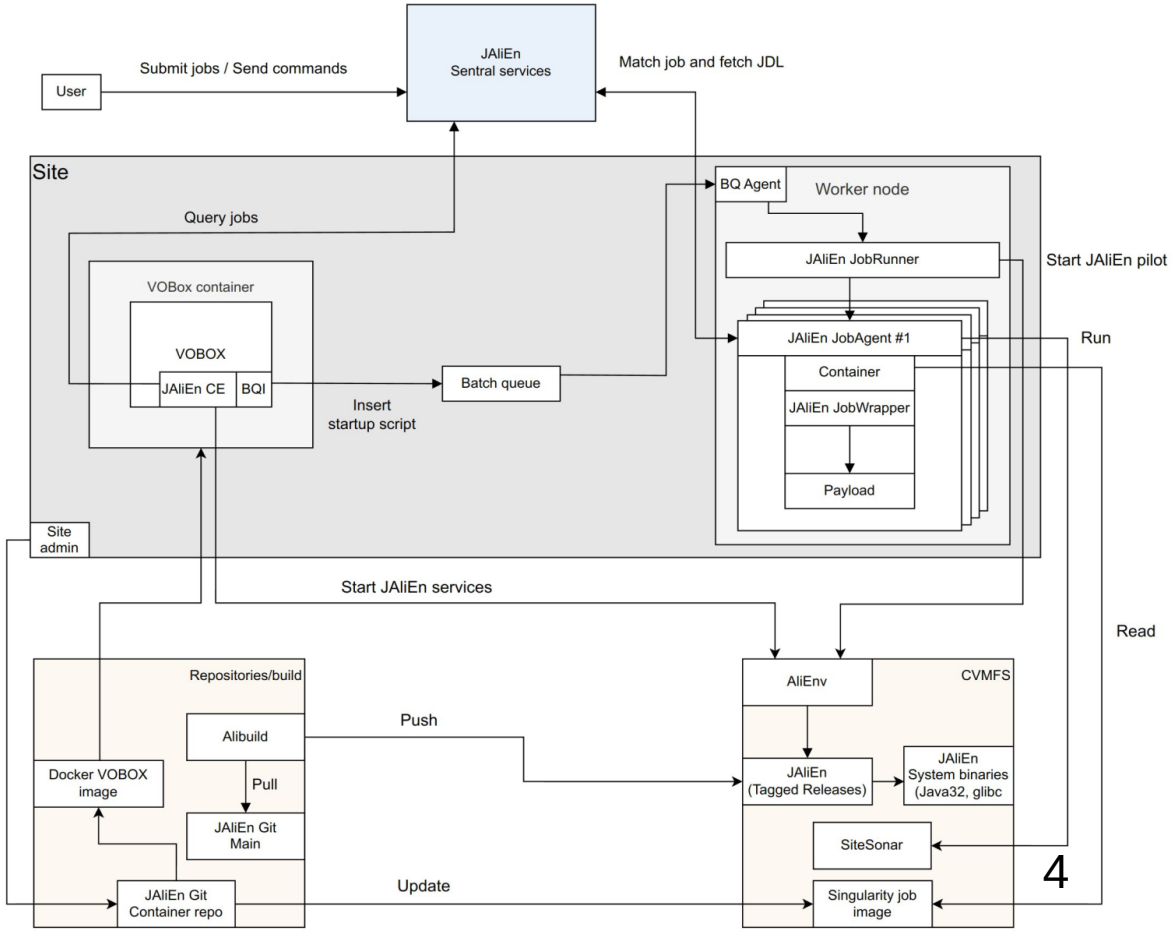
- JAliEn Middleware

  Grid Framework - combination of a Web Service and Distributed Agent Model. Data storage and job management.

- CVMFS central repository

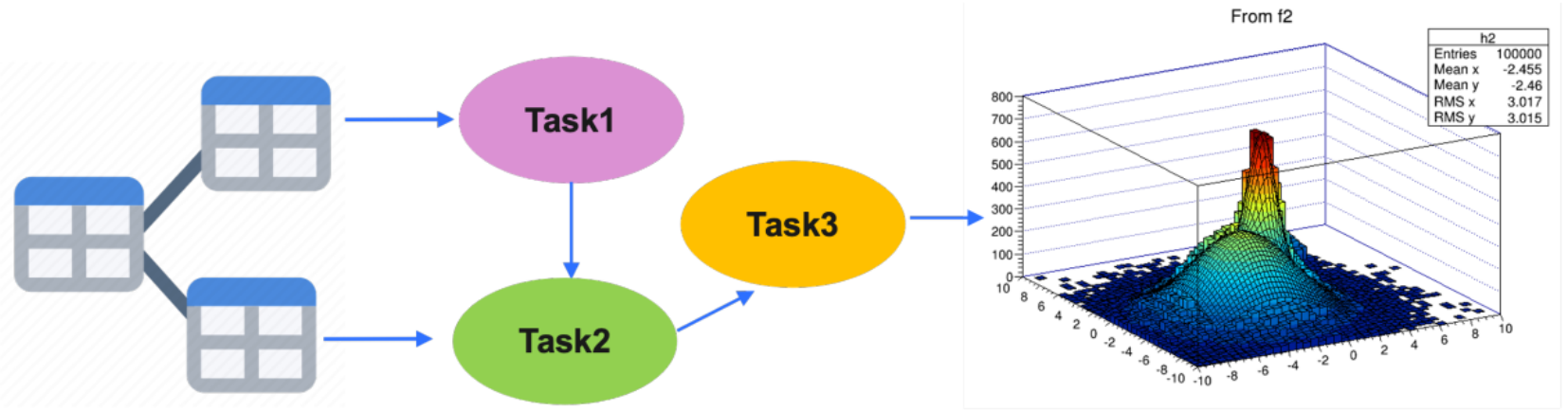
  ALICE analysis and required supplementary software delivery to run in Grid sites in containers. Integrated with build system for continuous deployment.

- Analysis software:
  - User
  - Centralized (Trains)
  - Data and MC production



# O2 Analysis Framework

- General structure of Data processing




**Data Model**

**Interconnected tables**
Based of Apache Arrows

**Analysis Tasks**

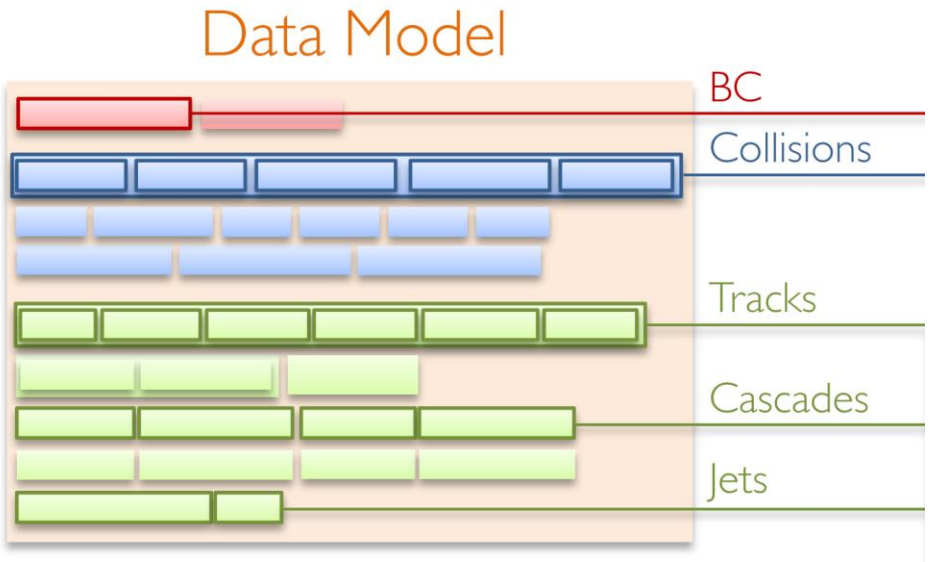
**User Tasks**
workflows~wagons

**ROOT serialized output**

**AnalysisResults.root**
+AO2Ds (derived data)

# O2 Analysis Data Model

Apache Arrow Tables




General tables

Indexed to bunch crossings

Indexed to collisions

Additional objects

Each analysis task is an executable →All the required are run in command line with pipe “|”

Plenty of Helper tasks → Produce required data tables on the fly

# O2 Analysis Model: Types of Wagons

**User wagons:**
*Spectra*
*Correlations*
*etc.*

**Core Service wagons**: helpers, dependencies of user wagons:
*Centrality*
*Event Selection*
*Multiplicity*
*Timestamp Creator*
*Track Propagation*
*etc.*

**Wagon – analyzer**
Creates and stores user defined analysis histograms

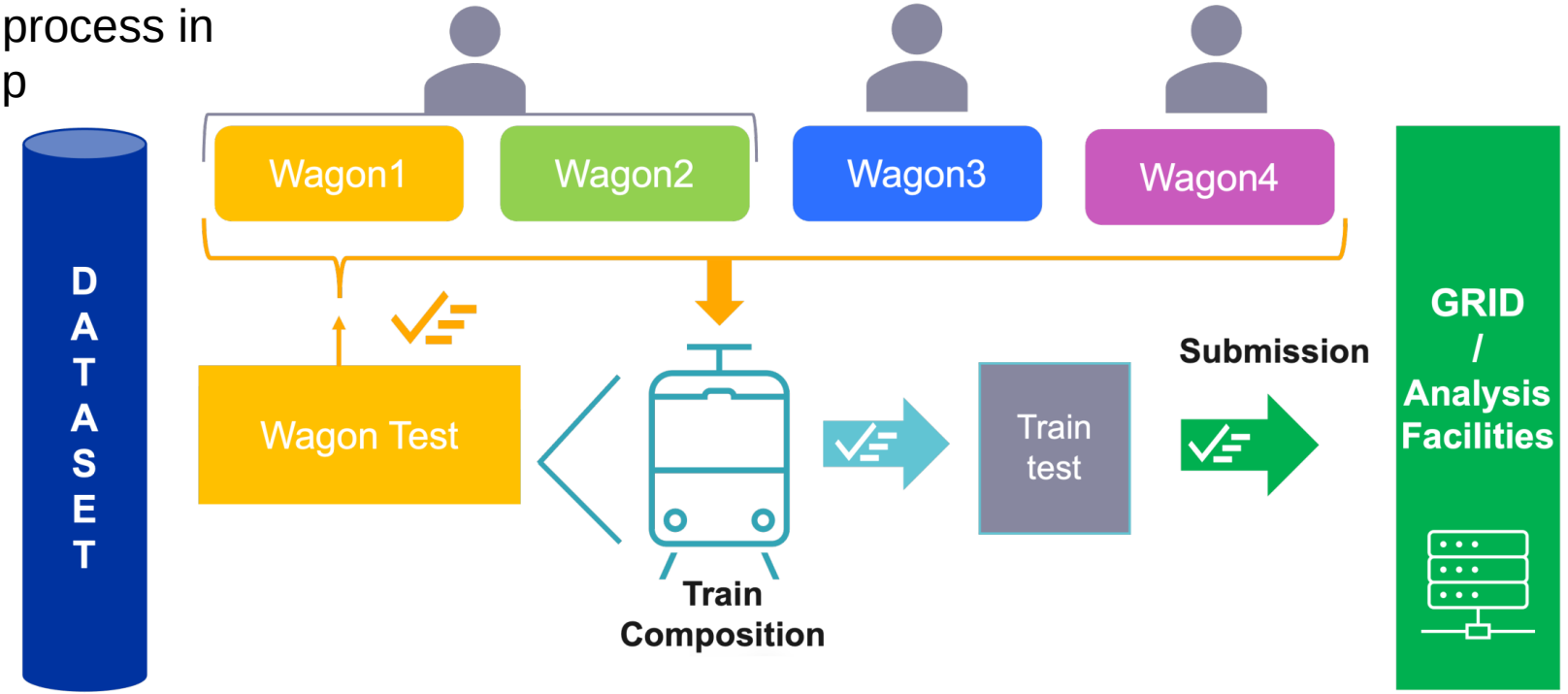
**Wagon – producer**
Mostly intended for generation of the derived data, which is used by other wagon or saved into AO2D files

**Wagon – reader**
Analysis of the derived data

**Wagon with parent level access**

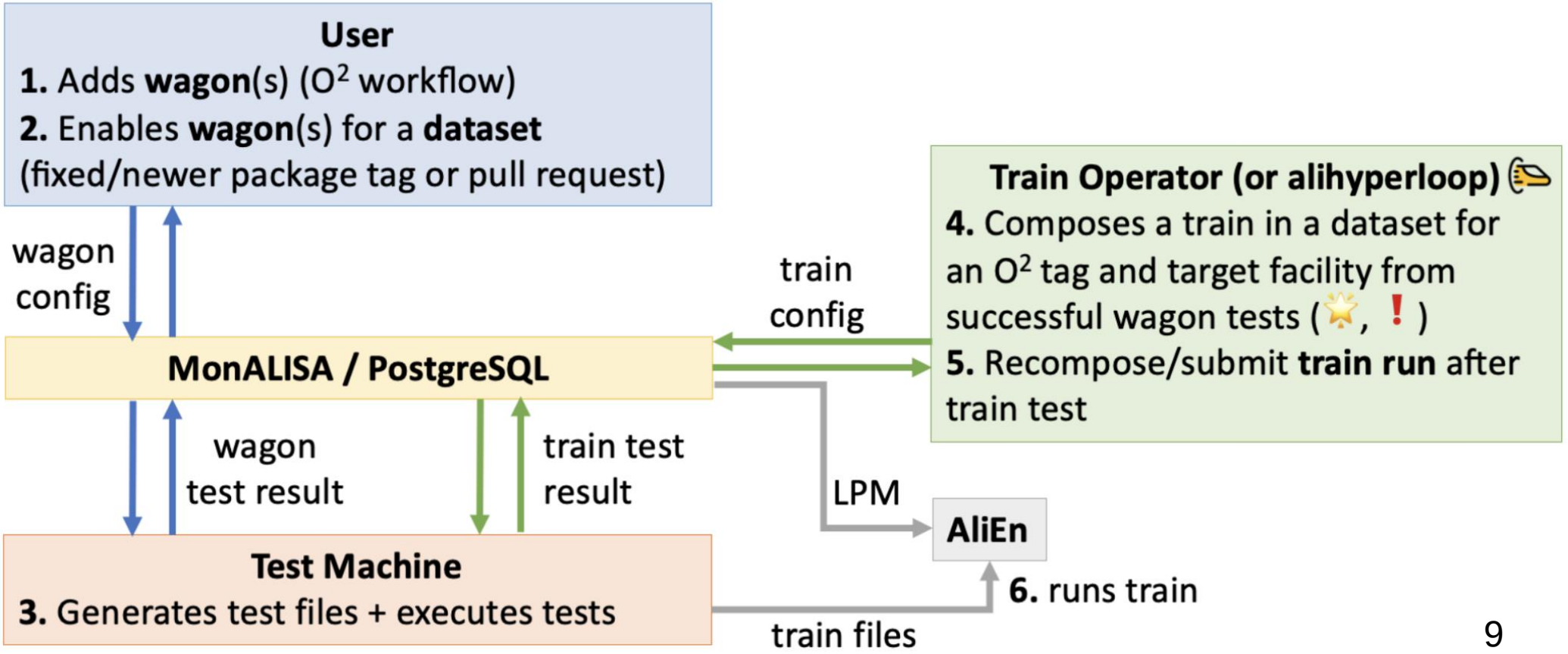
# Hyperloop trains

Analysis process in Hyperloop




Config of each wagon is saved and stored in JSON file
Configs of all wagons are merged into general train's config

# Process to submit hyperloop train



# Comparison of Hyperloop trains and LEGO trains (Run 2)

## ALICE LEGO train

- Analysis train framework for Run 2
- Analysis code is contained in an AliEn package, AliPhysics, and delivered via CVMFS
- Trains are defined per Physics Working Group (PWG), data type and collision system (~100)
- Analysis tasks (wagons) using the same dataset are run together
- Requires train operators (per PWG) to test, compose and submit train runs
- Main workhorse for Run 2 analysis:
  - 2020: 16 000 trains, 172 million Grid jobs

## ALICE Hyperloop

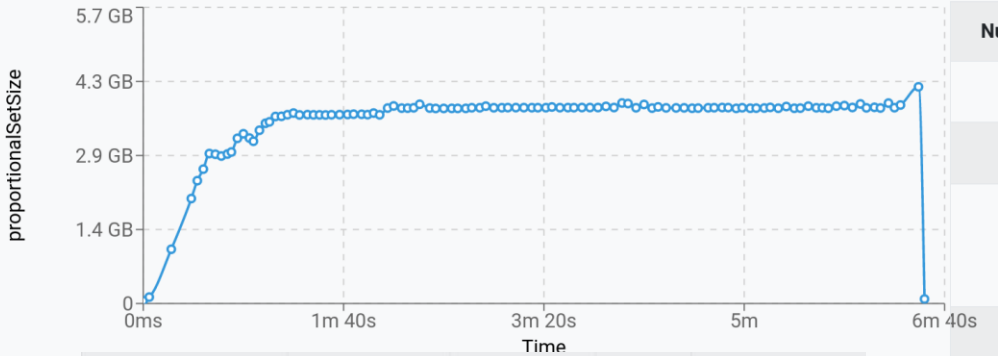
- Analysis train framework for Run 3
- Analysis code is built into O2Physics available on CVMFS
- Advanced web-interface (frontend: React.js)
- Unified trains throughout PWGs
- Personalized user and operator interfaces
- Immediate and automatic wagon test
- Automatic train submission under certain defined conditions
- Wagon and dataset bookkeeping
- Usage Last year 2023: ~ 9000 trains, 24 million Grid jobs

# Wagon and train run test performance results

**Train run 117133**

General | Test | Submitted jobs | Grid Statistics | Wagon resources | Merged Output | Clone

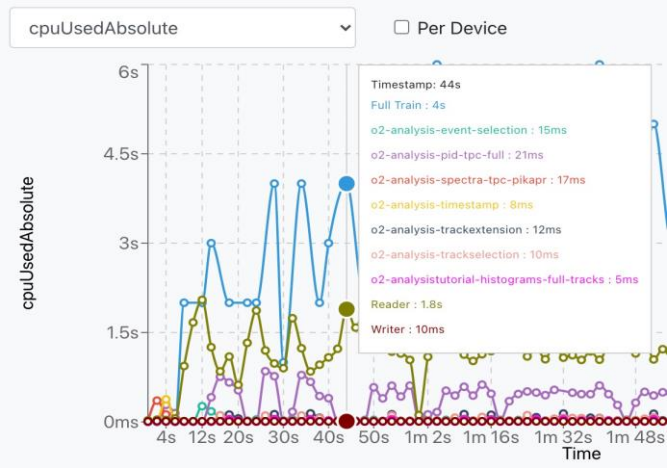
Full Test | Per Wagon | Graphs



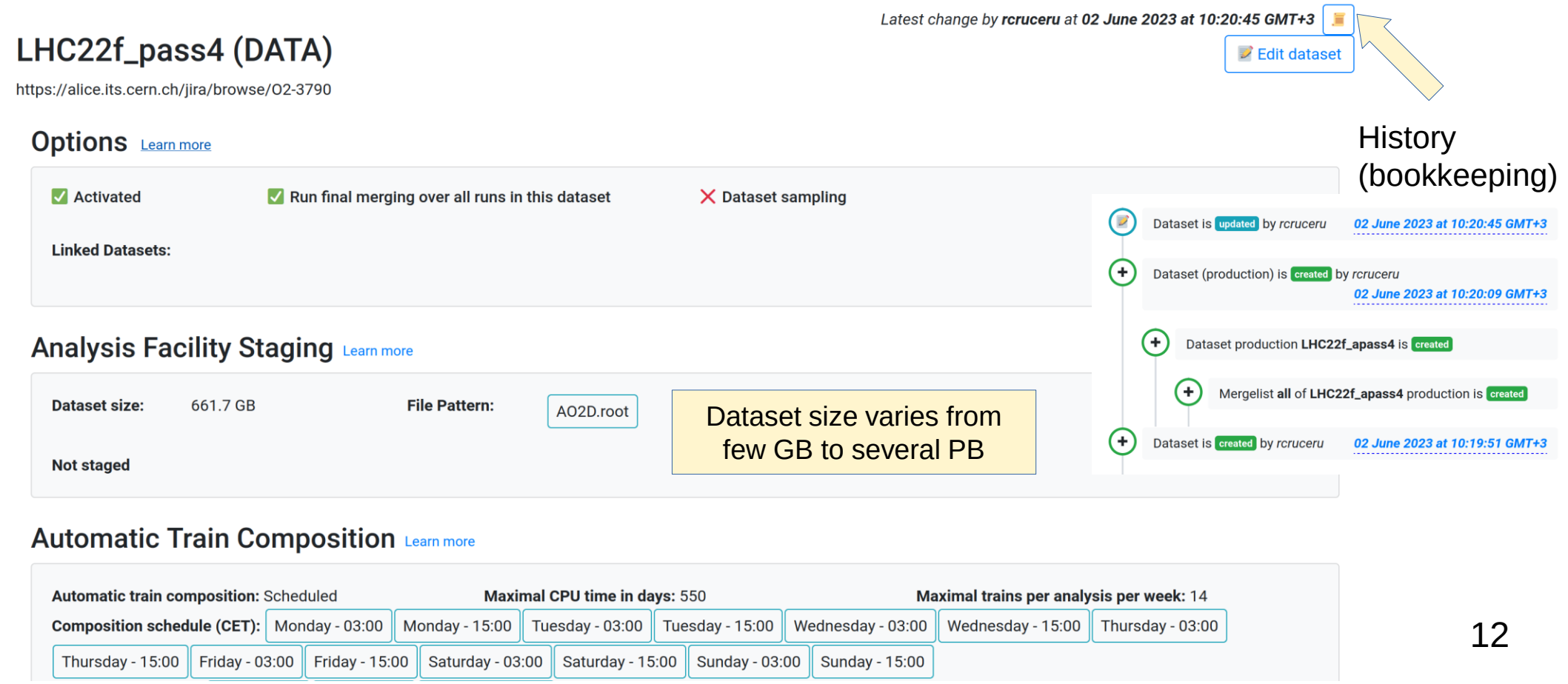

| Wagon | | PSS Memory | Private Memory | CPU Time | CCDB Calls | CCDB Transfer size |
|---|---|---|---|---|---|---|
| Reader | Max | 1.6 GB | 1.1 GB | 1m 51s | | |
| | Avg | 1.2 GB | 794.8 MB | | | |
| | Slope | 3.9 MB/s | 1.8 MB/s | | | |
| EventSelection | Max | 308.7 MB | 120.5 MB | 3s | | |
| | Avg | 150.4 MB | 114.7 MB | | | |
| | Slope | -360.1 KB/s | 179.5 KB/s | | | |
| HistogramsFull | Max | 25.7 MB | 9.6 MB | 2s | | |
| | Avg | 19.6 MB | 9.5 MB | | | |
| | Slope | 87.7 KB/s | 157.4 B/s | | | |
| PIDTPCFull | Max | 226.3 MB | 158.3 MB | 38s | | |
| | Avg | 185.5 MB | 120.2 MB | | | |
| | Slope | 368.5 KB/s | 11.7 KB/s | | | |

| | |
|---|---|
| **Number of input files** | 4 |
| **Input size** | 7.0 GB |
| **Output size** | 4.8 MB |
| **PSS Memory** | Max: 3.9 GB<br>Avg: 3.6 GB<br>Slope: 2.4 MB/s |
| **Private Memory** | Max: 3.4 GB<br>Avg: 3.2 GB<br>Slope: 1.2 MB/s |
| **Timing** | CPU: 9m 57s<br>Wall: 6m 29s |
| **Throughput** | 9.2 MB/s/core |
| **Expected resources** | 19h 49m |

Per wagon:



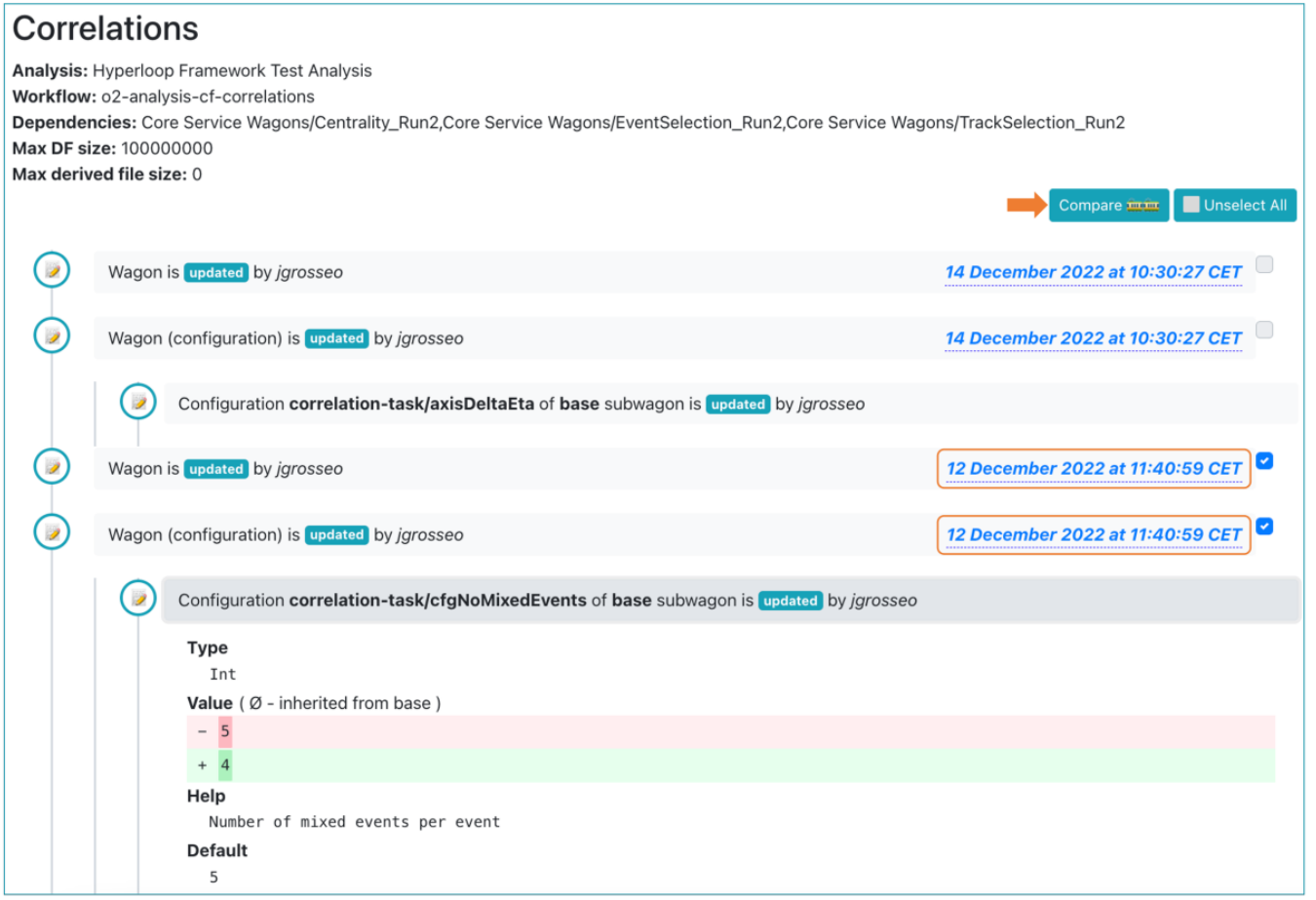
# Datasets




History (bookkeeping)

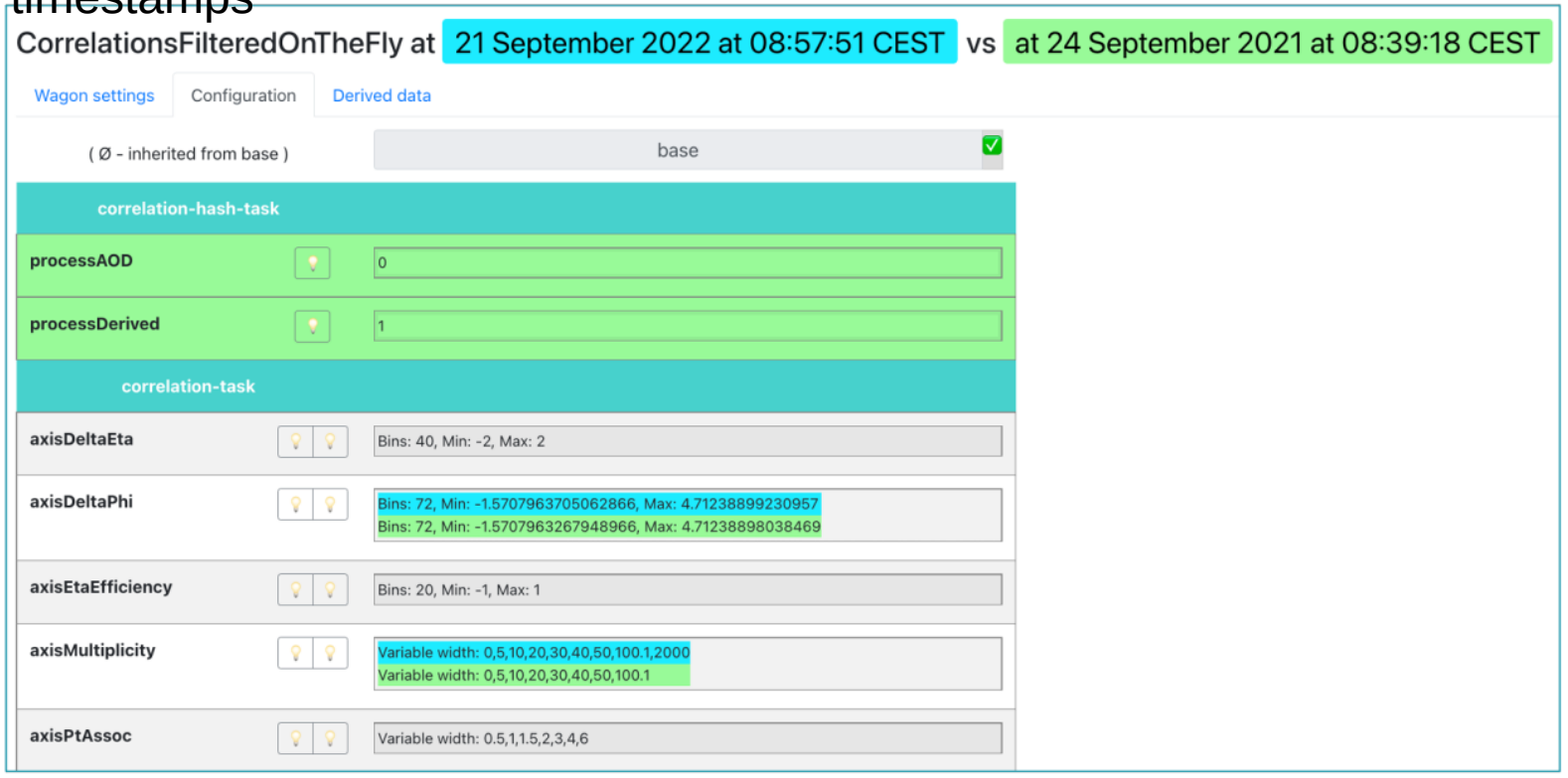
# Bookkeeping

Wagon changelog



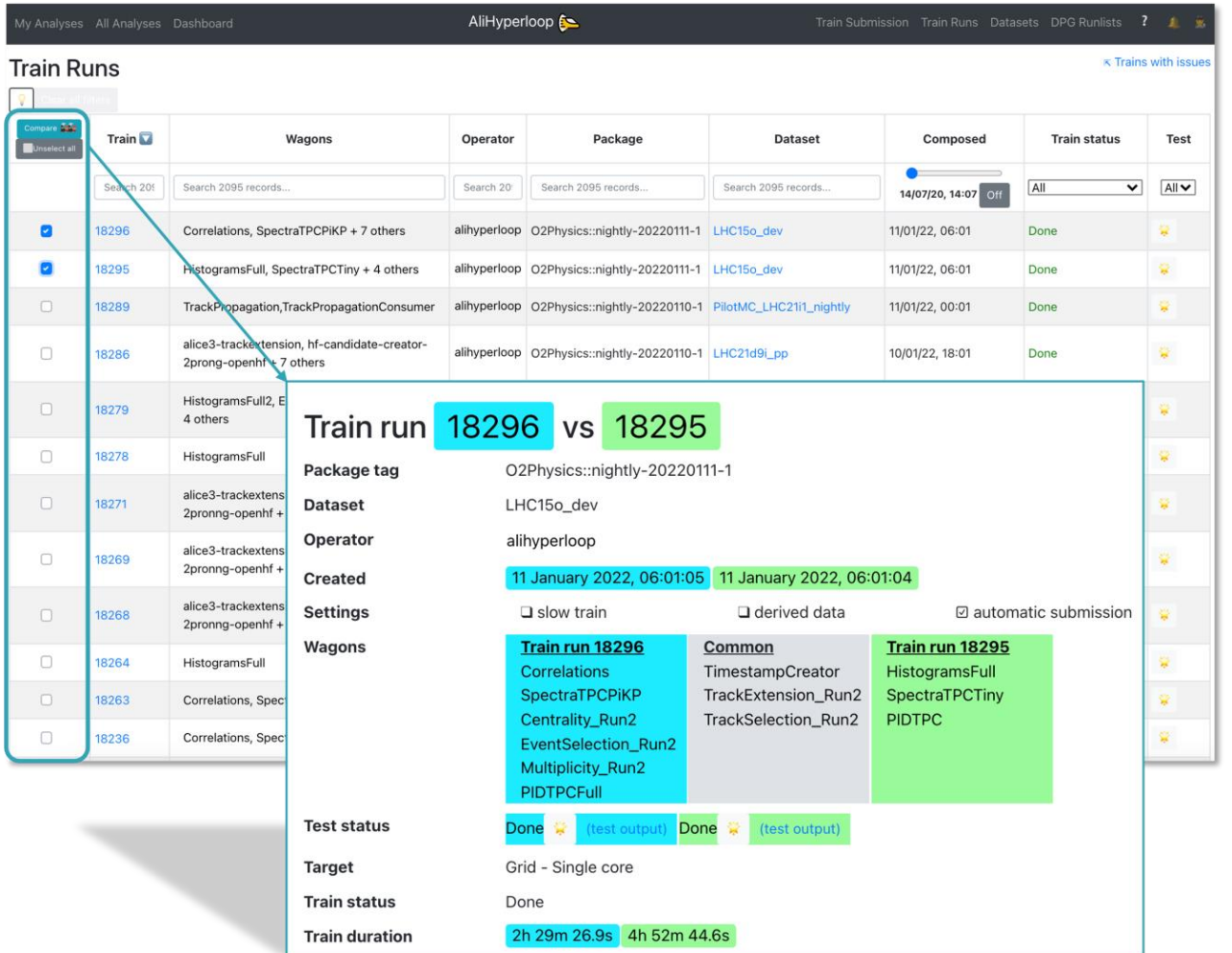
# Bookkeeping

Wagon comparison at different timestamps



# Bookkeeping

Train comparison



# Roles of Analyzer and Operator

**ANALYZER**

My Analyses

All Analyses

Dashboard

- Creates and configure wagons
- Runs wagon tests
- Studies test results
- Makes use of automatic train composition or ask to compose train
- Studies the resource consumption
- Stores derived data to be used in subsequent trains
- Makes use of history and statistics views

**OPERATOR**

Train Submission

Train Runs

Datasets

Derived Data

DPG runlists

Trains with issues

- Runs the system on a daily 24/5 basis
- Ensures efficient usage of the resources
- Follows up on overall system status
- Investigates issues and delegate to experts
- Responds to user requests
- Submits trains to the Grid or AFs
- Manages datasets
- Creates datasets of the derived data

- PWG Convener: approves long trains (> 200 Tb)

# Current status

- Hyperloop is in production since early 2022.
- Run 3 data and MC are available for the analysis.
- All Run 2 data and considerable amount of MC data converted to AO2D format and available on Hyperloop
- Operator support on a daily 24/5 basis by four institutional clusters in different timezones

| | |
|---|---|
| NISER | 2:00-9:00 UTC |
| St. Petersburg | 9:00-13:00 UTC |
| Brescia/Pavia | 13:00-17:00 UTC |
| US cluster | 17:00-1:00 UTC |

- ~300 datasets available (including derived data)
- ~300 users of hyperloop system
- More than 100k wagon tests done, ~12000 trains run on Grid or AF
- The average completion time of the trains is between 7 and 14 hours

# Conclusion

• Hyperloop is a modern innovative analysis train system for ALICE LHC Run 3 and 4, replacing the LEGO Train system used in Run 1 and 2.

• It has been fully integrated with the O2 analysis framework to enable analysis on the exceptionally large data samples

• It enables fast deployment and run of user analysis allowing efficient share of the resources to run multiple workflows of many users on the Grid.

• It provides full archival of past configurations, changelogs and comparison tools.

# Thank you