GPU DAY 2023

Hyperloop Train System for ALICE Run 3

Raluca Cruceru - CERN

CERN

Detector Upgrades for Run 3

- **Main detector parts have been upgraded**

  - New Inner Tracking System (ITS2)

  - New Fast-interaction Trigger system

  - Upgraded Time Projection Chamber (TPC)

  - O^2 - Data taking and Processing Infrastructure

- **Implications**

  - **100 times more recorded collisions** compared to Runs 1 and 2

  - An increased data-taking capability **by two orders of magnitude**

  - Resulting data throughput is estimated to be greater than **1 TB/s for Pb-Pb collisions**

  - 1 month of Pb-Pb data would create ~ 4 PB of AODs (reconstruction output)
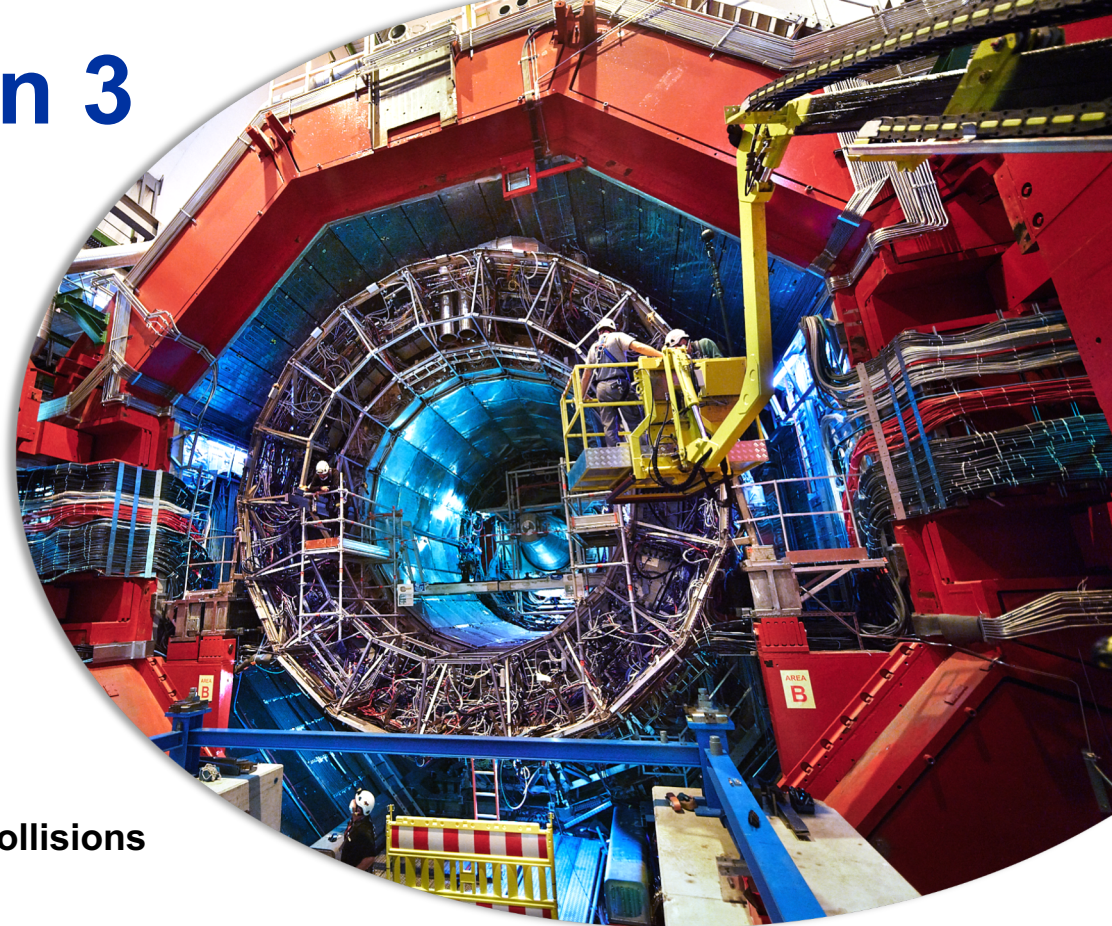
➤ **Run 3 analysis requirements**

  ○ Processing this unprecedented amount of data

  ○ Analysis infrastructure needs to cope with 100 times more data with more efficient algorithms and techniques

➤ **How do we achieve this?**

Online-Offline Software Framework (O^2) – allows for distributed and efficient processing of the new amount of data

Hyperloop Train System – allows fast and demanding analysis workflows on Grid and Analysis Facilities (specialized Grid sites with CPU and disk resources adjusted for analysis needs)
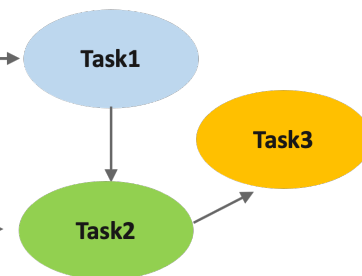
O2 and Hyperloop

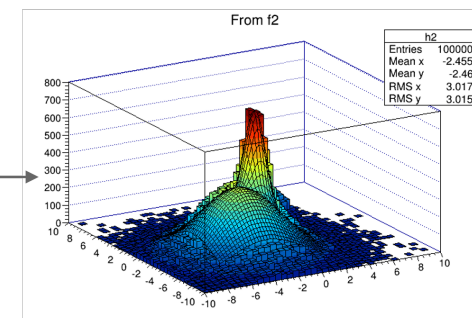
Online-Offline Software Framework (O^2)

- Data taking and processing infrastructure
- Collisions and tracks are represented in trees (flat tables)
- Trees are connected through indices passed through shared memory
- Users define tasks with callbacks and I/O declarations (table subscriptions)
- These are then combined into workflows representing a particular analysis / group of analyses
- O^2 Data Processing Layer translates the defined workflows to an actual FairMQ topology of devices

On Disk
In Memory
Monte Carlo
HF & Jets
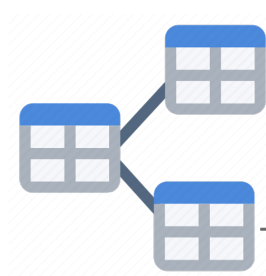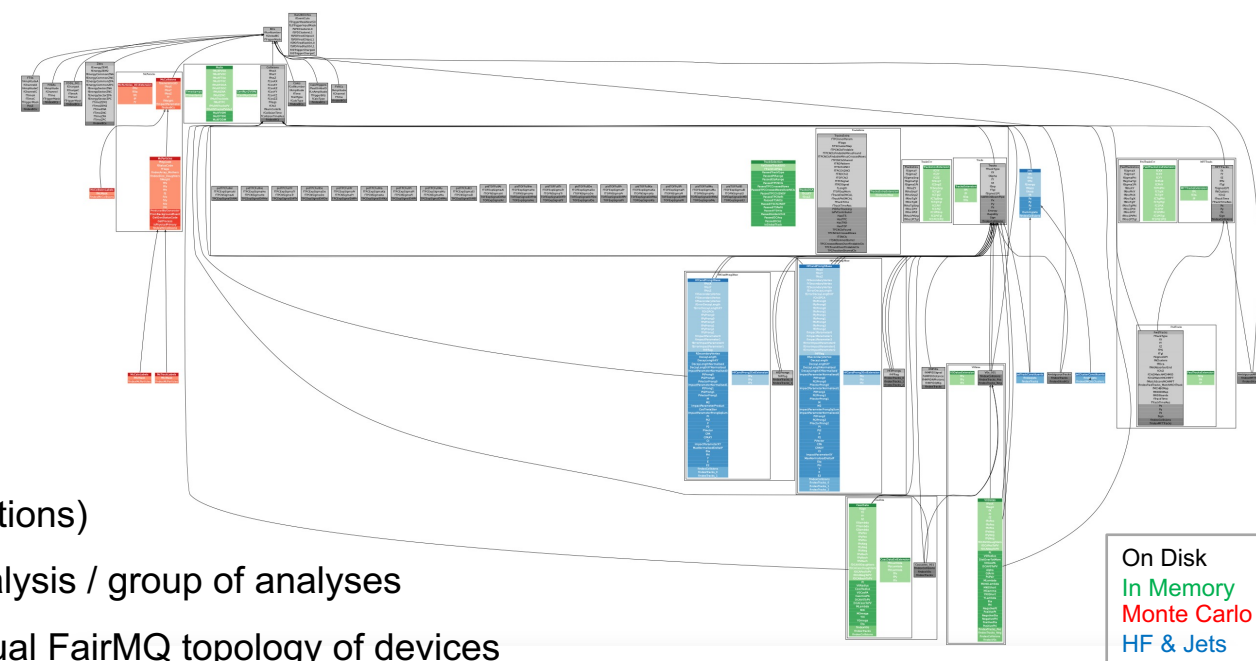
Data Model   **Analysis Tasks**   **ROOT serialized output**

O2 and Hyperloop

Data Model   **Analysis Tasks**   **ROOT serialized output**

Online-Offline Software Framework (O^2)

- Data taking and processing infrastructure
- Collisions and tracks are represented in trees (flat tables)
- Trees are connected through indices passed through shared memory
- Users define tasks with callbacks and I/O declarations (table subscriptions)
- These are then combined into workflows representing a particular analysis / group of analyses
- O^2 Data Processing Layer translates the defined workflows to an actual FairMQ topology of devices
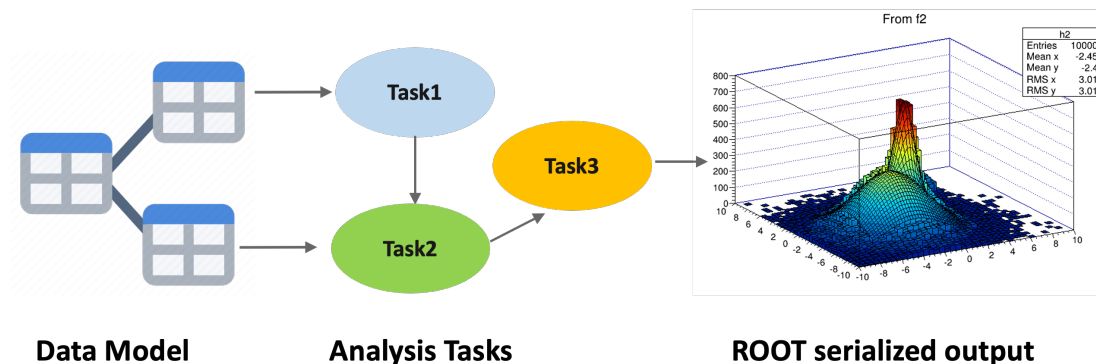
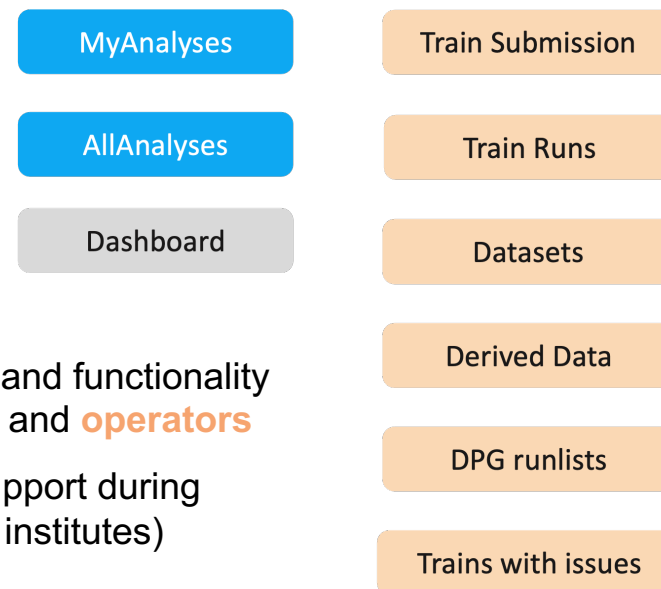## Hyperloop Train System - Concept of analysis trains to optimize the usage of computing resources

- **Built upon tools used in Run 1 and Run 2**
- Re-written with a modern reactive front-end technology: **React**
- Allows organized analysis on the **Grid** and **Analysis Facilities**
- Fully integrated with O^2, allowing **task configuration**
- Individual workflows - known as wagons - are combined into trains
- Skimmed / **Derived data** stored for further processing in subsequent trains
- Data available: converted Run 2 data, Run 3 data and MC, Derived data
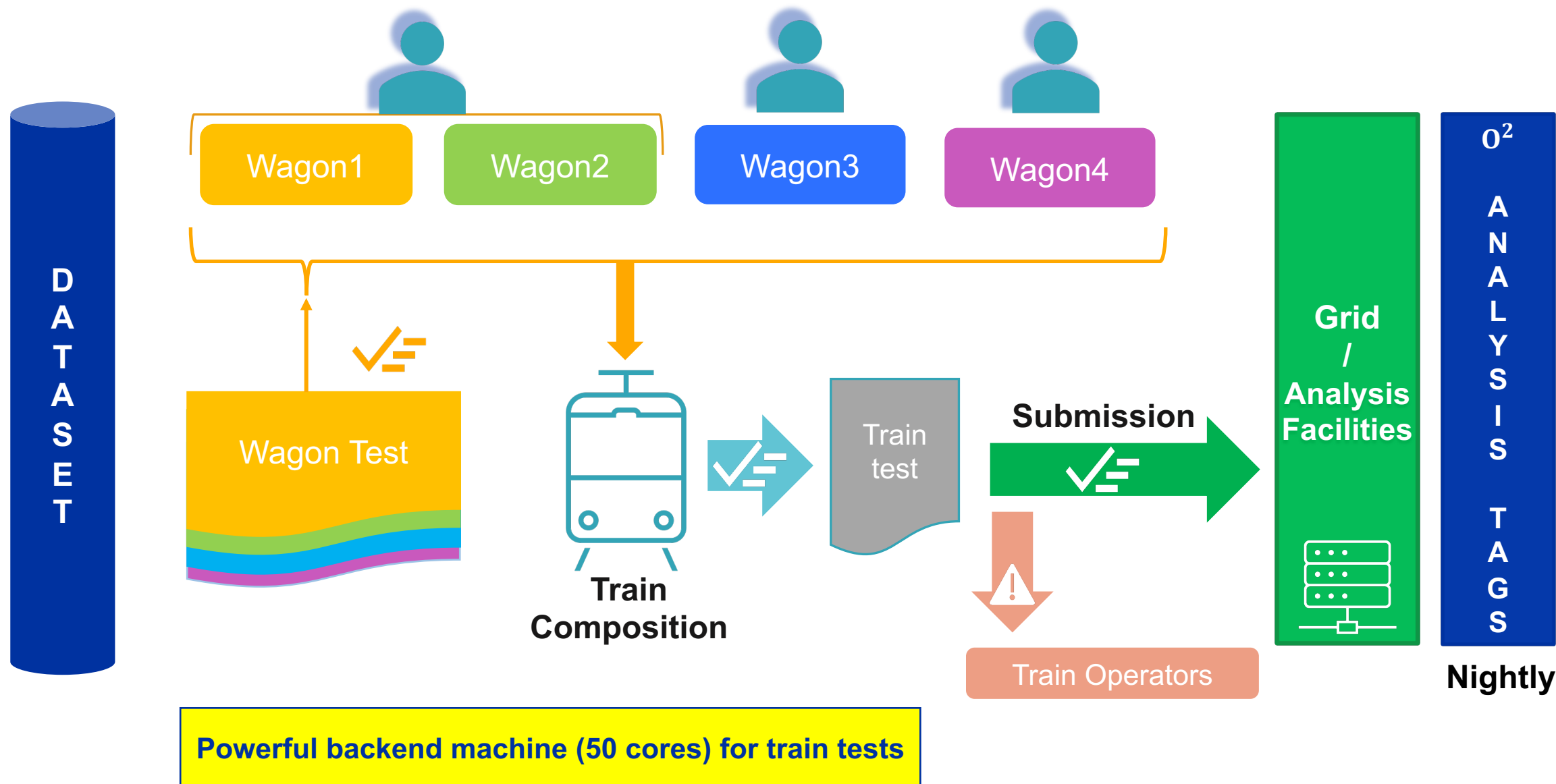- Full bookkeeping, changelog and several comparison tools

Workload for Run 2 analysis:
- ◆ **40 000 cores** utilized on average at any time
- ◆ Used by more than **90% of the analysis**

USER

| MyAnalyses |
| AllAnalyses |
| Dashboard |

OPERATOR

| Train Submission |
| Train Runs |
| Datasets |
| Derived Data |
| DPG runlists |
| Trains with issues |

- Dedicated views and functionality for regular **users** and **operators**
- 24/5 shift-type support during working hours (4 institutes)

Hyperloop Train System

Powerful backend machine (50 cores) for train tests

Hyperloop Train System

- **Analyses defined in JIRA, shared by users**
- **Datasets are enabled per analysis**
- **Analyzers can add, configure, compare, clone or remove wagons**
- **Wagons created from available O^2 workflows or pull request**
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- Analyzers can enable / disable wagon tests
- Immediate testing and overview
- Per device (reader, workflows, writer) and total performance metrics
- Expected resources and Interactive graphs

Hyperloop Train System

- Analyses defined in JIRA, shared by users
- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- Per device (reader, workflows, writer) and total performance metrics
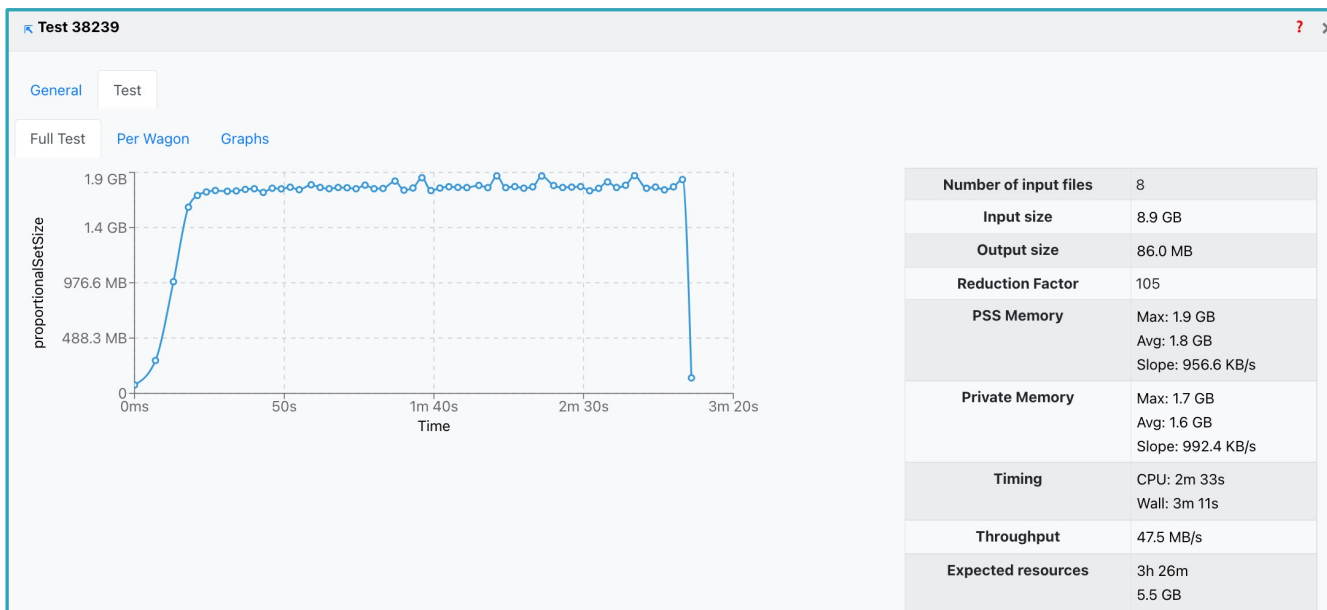- Expected resources and Interactive graphs

Hyperloop Train System

- Analyses defined in JIRA, shared by users
- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- Per device (reader, workflows, writer) and total performance metrics
- Expected resources and Interactive graphs
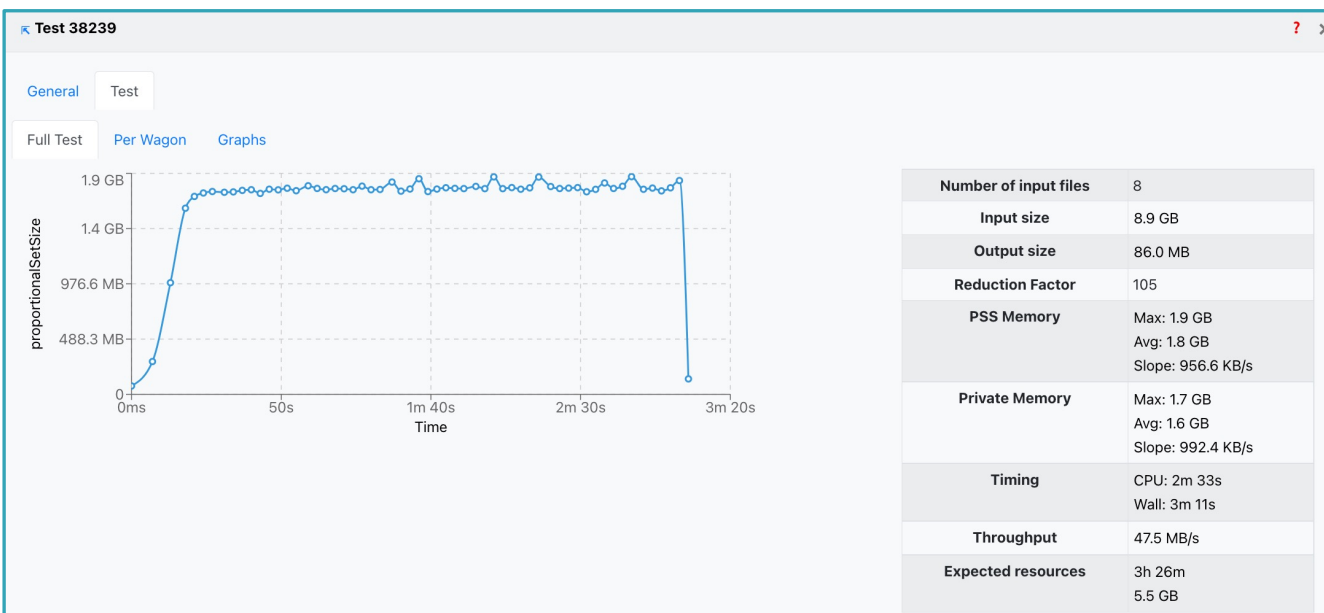
Hyperloop Train System

- Analyses defined in JIRA, shared by users
- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- Per device (reader, workflows, writer) and total performance metrics
- Expected resources and Interactive graphs
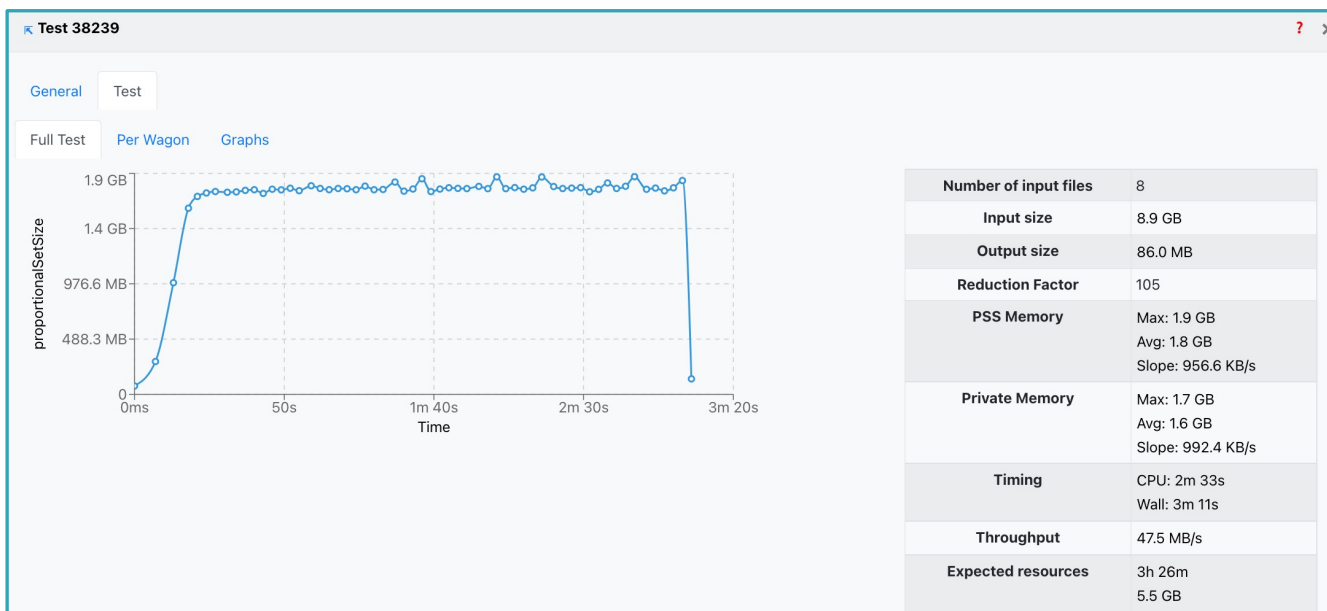
Hyperloop Train System

- Analyses defined in JIRA, shared by users
- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- Per device (reader, workflows, writer) and total performance metrics
- Expected resources and Interactive graphs
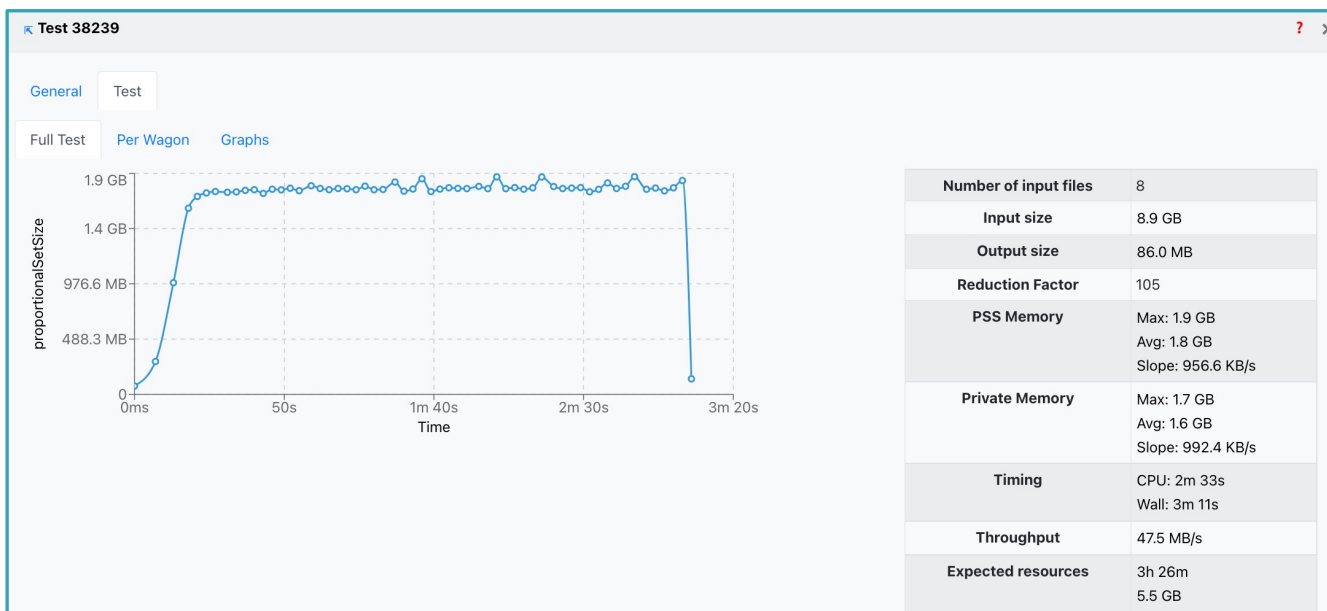
Hyperloop Train System

- Analyses defined in JIRA, shared by users
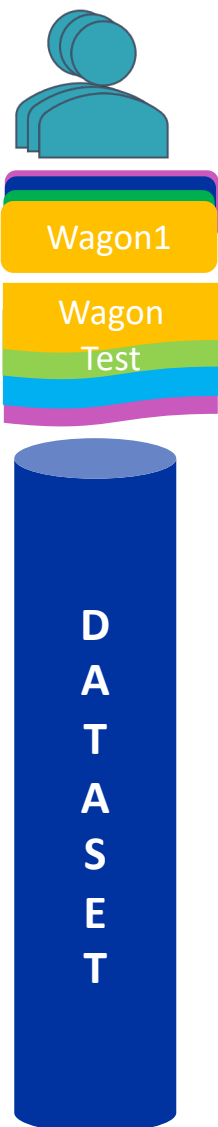- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
  - Allows **subwagons** creation – these will run the same task with different parameter values
  - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- **Per device (reader, workflows, writer) and total performance metrics**
- **Expected resources and Interactive graphs**
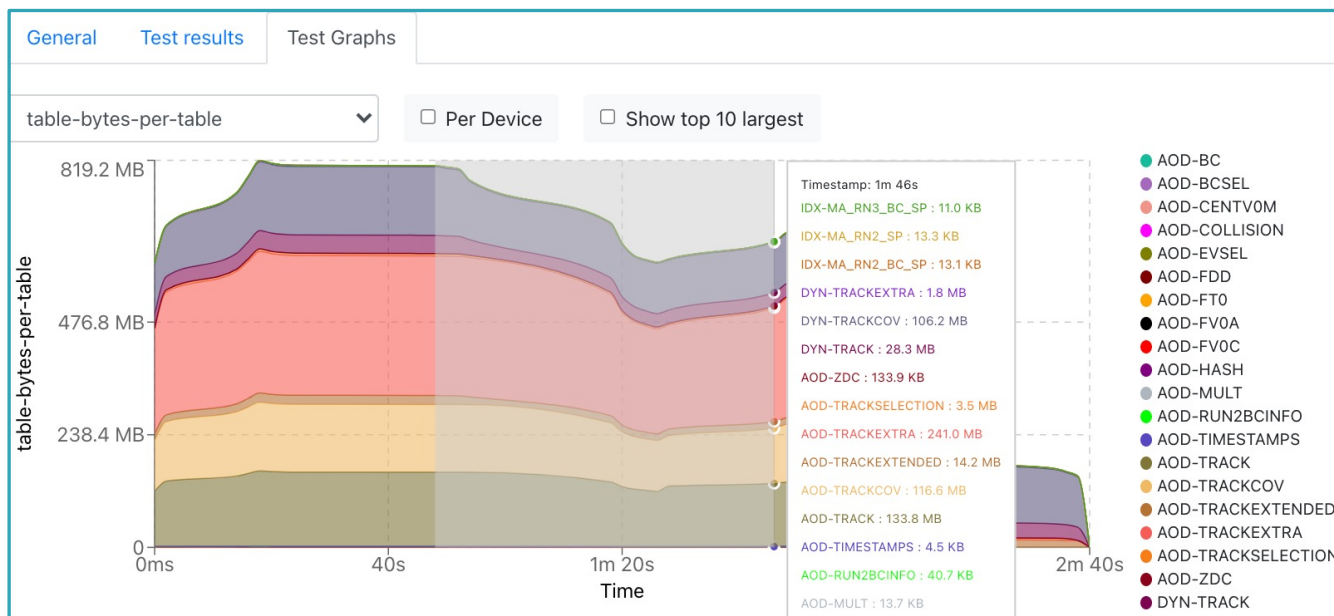
Hyperloop Train System

- Analyses defined in JIRA, shared by users

- Datasets are enabled per analysis

- Analyzers can add, configure, compare, clone or remove wagons

- Wagons created from available O^2 workflows or pull request

  - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning

  - Allows **subwagons** creation – these will run the same task with different parameter values

  - Output can be stored as derived data

- **Analyzers can enable / disable wagon tests**

- **Immediate testing and overview**

- **Per device (reader, workflows, writer) and total performance metrics**

- **Expected resources and Interactive graphs**
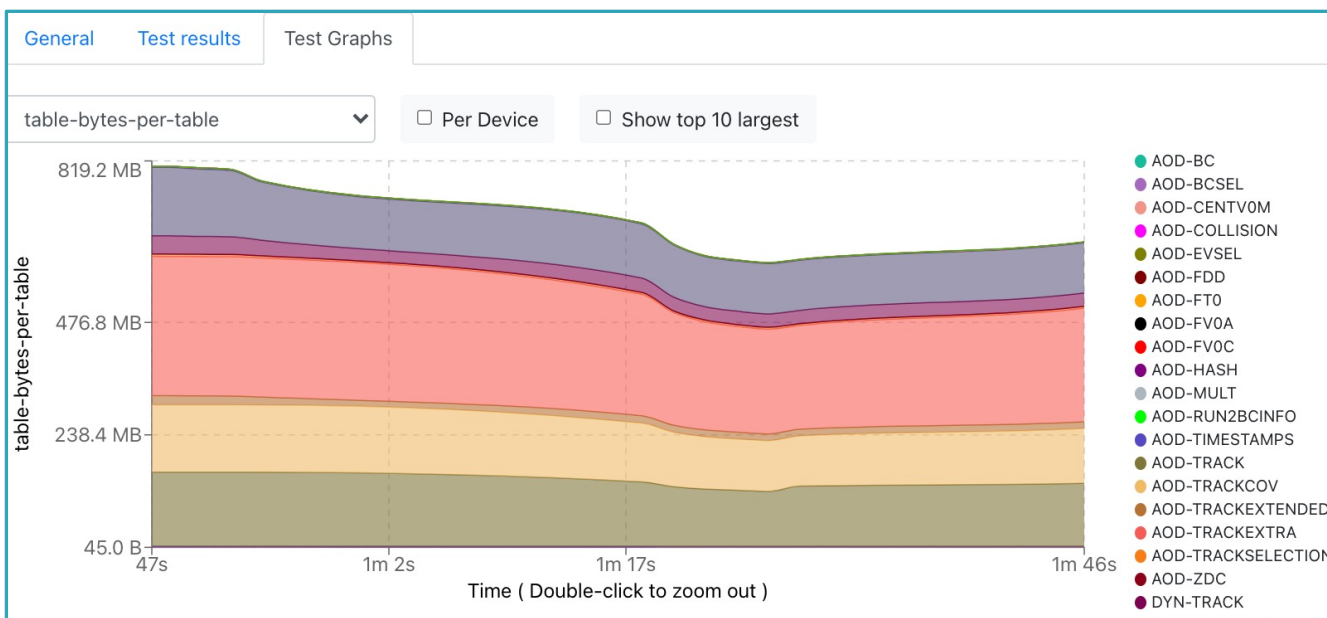
Hyperloop Train System

- Analyses defined in JIRA, shared by users
- Datasets are enabled per analysis
- Analyzers can add, configure, compare, clone or remove wagons
- Wagons created from available O^2 workflows or pull request
    - Supports a variety of parameter types such as primitive types, arrays, matrices, labelled matrices, histogram binning
    - Allows **subwagons** creation – these will run the same task with different parameter values
    - Output can be stored as derived data
- **Analyzers can enable / disable wagon tests**
- **Immediate testing and overview**
- **Per device (reader, workflows, writer) and total performance metrics**
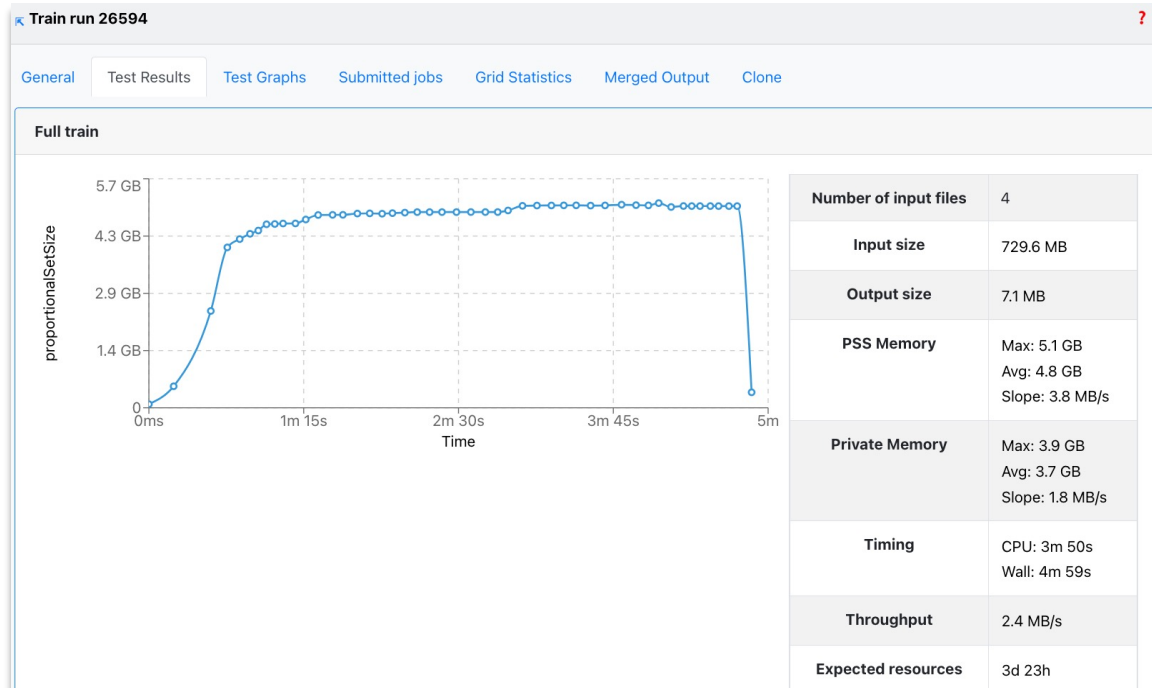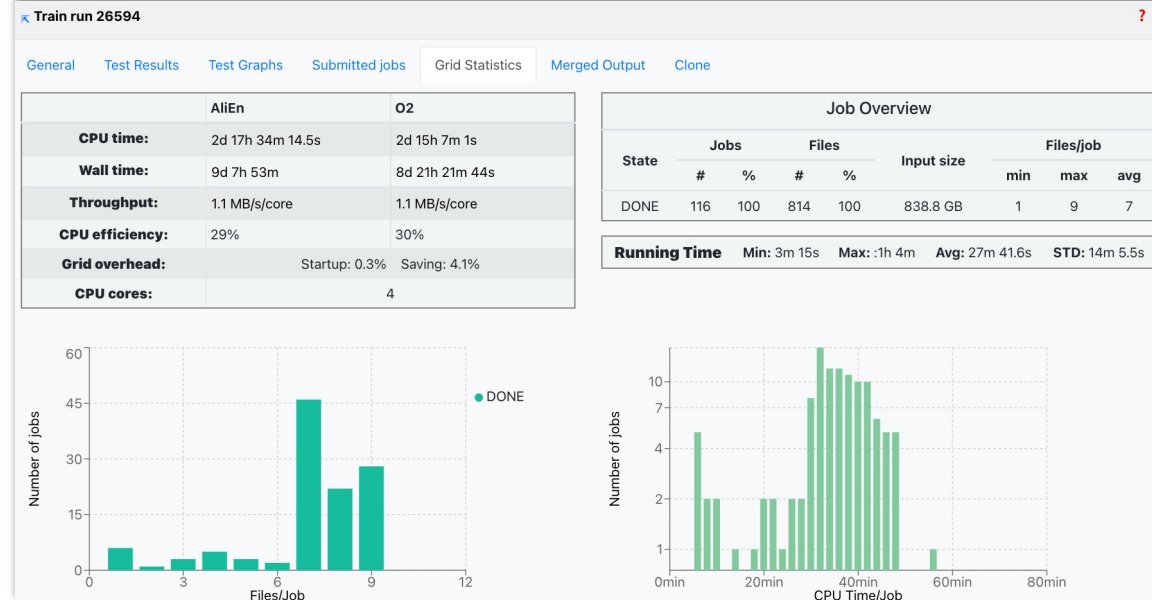- **Expected resources and Interactive graphs**

Hyperloop Train Runs

- Automatic **train composition** is scheduled per dataset
  - Composition based on target memory, wagon configuration and dependencies (only wagons with successful tests considered)
  - Dataset specific restrictions: CPU usage and maximum number of analysis trains allowed per week
- Automatic **train submission** to **Grid** or **Analysis Facilities**
- Train test results, submitted jobs and Grid statistics
- **Clone train** – runs with the same wagon timestamp but updated dataset configuration
- **Staged submission** for large data samples – run first on a smaller dataset before approval to run on bigger dataset **(approved by PWGs)**

Train Support

- 24/5 Operation (different timezones)
- Institutes: 1 in Americas, 2 in Europe, 1 in Asia
- Dedicated channel for users' request and issues
- Shift-type support during working hours
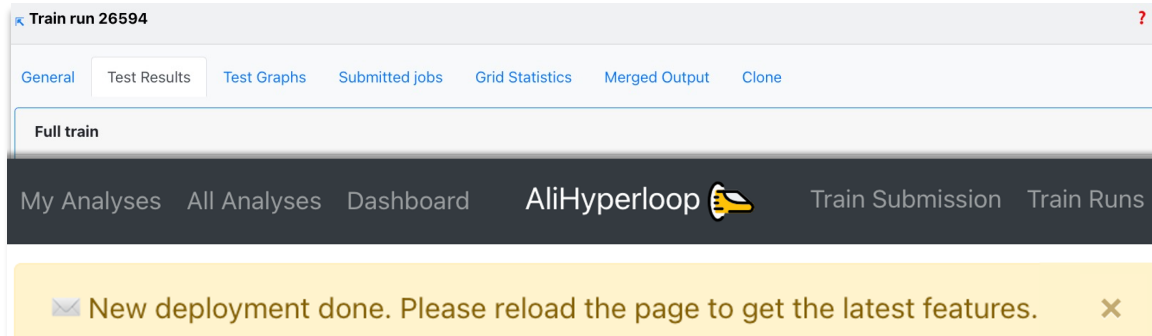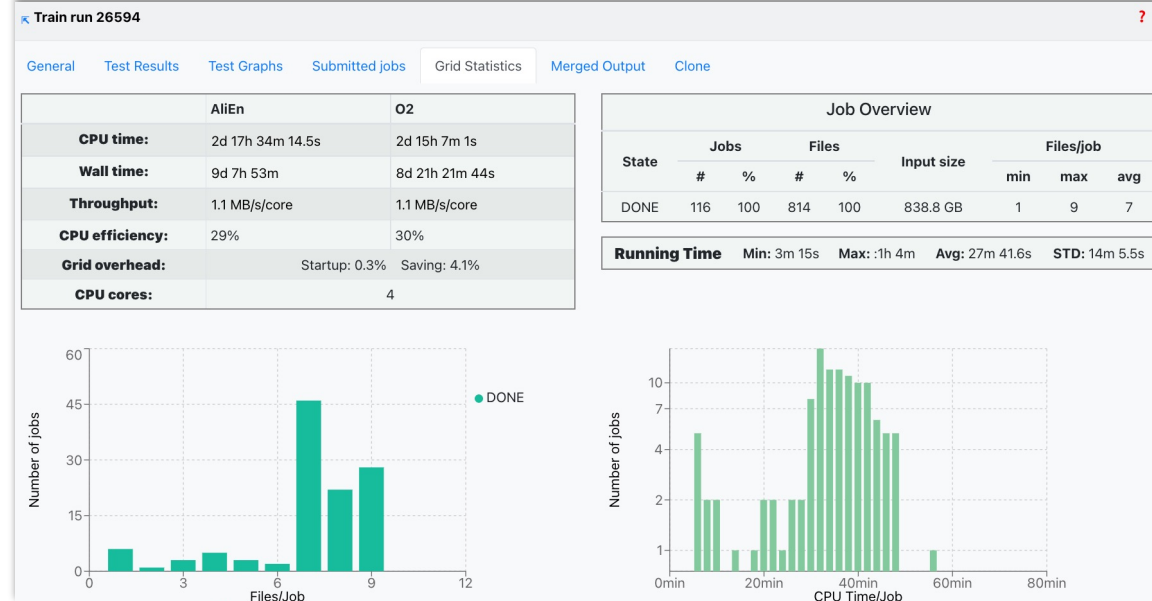- Organized feedback sessions

Hyperloop Support

- Automatic **train composition** is scheduled per dataset
  - Composition based on target memory, wagon configuration and dependencies (only wagons with successful tests considered)
  - Dataset specific restrictions: CPU usage and maximum number of analysis trains allowed per week
- Automatic **train submission** to **Grid** or **Analysis Facilities**
- Train test results, submitted jobs and Grid statistics
- **Clone train** – runs with the same wagon timestamp but updated dataset configuration
- **Staged submission** for large data samples – run first on. a smaller dataset before approval to run on bigger dataset
  (approved by PWGs)

Train Support

- 24/5 Operation (different timezones)
- Institutes: 1 in Americas, 2 in Europe,1 in Asia
- Dedicated channel for users' request and issues
- Shift-type support during working hours
- Organized feedback sessions

Monitoring & Bookkeeping

Dedicated Pages

- Displays relations between the datasets

- Group by **Dataset, PWG** or **Analysis**

- **Direct links** to derived data train, analysis and HY dataset

- Can be marked for deletion

  - If derived data train is not included in activated datasets

Derived Data

By Dataset   By Analysis   By PWG

Clear all filters

Dataset	Train	Analysis	PWG	Produced on	Size	In datasets ✅/❌	Last used	Last month	Delete
Search 7 records...	Search	Search 7 records...	All	05/10/20, 15:29 CEST Off		Search 7 recor Search	Not used Off	Search	
📥 Run 2 LHC15o	1					0	1 / 1		
	883	Hyperloop Framework Test Analysis	CF	05 October 2020, 15:29:36 CEST	0	0	1 / 1	0	
📥 LHC15o_benchmark	5					0	0 / 1		
	1038	Hyperloop Framework Test Analysis	CF	30 October 2020, 14:48:12 CET	0	0	0 / 1	0	🗑
	3602	Hyperloop Framework Test Analysis	CF	19 January 2022, 12:21:47 CET	0	0	0 / 0	0	🗑
	3713	Hyperloop Framework Test Analysis	CF	03 February 2022, 21:35:30 CET	0	Included in activated dataset(s): LHC15o_derived_correlations.		0	
	4059	Hyperloop Framework Test Analysis	CF	18 March 2022, 15:29:49 CET	0	1 / 0	23 March 2022, 11:29:14 CET	0	
	4063	Hyperloop Framework Test Analysis	CF	18 March 2022, 15:47:56 CET	0	0 / 1		0	🗑
📥 LHC15o_rcruztor	1					0	0 / 0		
	3727	hyperloop_test	JE	03 February 2022, 22:12:17 CET	0	0 / 0		0	🗑

Datasets

↖ Derived data   ☑ Show staging status   ☐ Show removed datasets   ➕ Add Dataset

Clear all filters

Name ▲	Description	Type	Size	Staging status
pass3	Search 6 records...	DATA	Search 6 records...	Completed in warning state(99.63% success)
LHC22m_pass3_relval_cpu2	https://alice.its.cern.ch/jira/browse/O2-3540	DATA	13.0 TB	ALICE::GSI::SE2 99.63%
LHC22p_pass3_subset	https://alice.its.cern.ch/jira/browse/O2-3562	DATA	20.8 TB	ALICE::KFKI::EOS 100%
LHC22q_pass3	https://alice.its.cern.ch/jira/browse/O2-3562	DATA	46.1 TB	ALICE::GSI::SE2 100%
LHC22r_pass3_subset	https://alice.its.cern.ch/jira/browse/O2-3562	DATA	29.7 TB	ALICE::GSI::SE2 98.33%
LHC22t_pass3_small	https://alice.its.cern.ch/jira/browse/O2-3562	DATA	25.4 TB	ALICE::KFKI::EOS 100%
PilotBeam_pass3	async pass 3	DATA	80.1 GB	ALICE::GSI::SE2 100% / ALICE::KFKI::EOS 99.99%

- Summarizes the **staging status** of datasets to Analysis Facilities

- Direct link to **transfer requests** and details

- Once staging is done for one site, trains can be submitted to that Analysis Facility

Bookkeeping & comparison tools

- Changelogs for wagons, datasets and runlists

- Trains and wagons comparison tools

Dashboard & Statistics

- 200 analyses, 240 datasets
- Total input size 160 PB (5600 train runs)
- **Converted Run 1 and 2 data to O^2 format**
  The new framework supports analysis on old data
- Ongoing QA of the MC data

- Real-time updates through WebSocket
- **Personalized notifications** (also by email)
- Documentation accessible from the UI

Dashboard

Number of wagons waiting to be tested: **0**

Tests ongoing: **1**

Number of wagons waiting to be included in a train: **81**

Number of running trains: **13**

Number of trains ready to be submitted: **1**

Test Daemon: **OK**   Uploader Daemon: **OK**

Number of trains waiting for Grid job submission: **0**

AliEn metrics:

Number of running jobs: **5379**   Jobs ran in the last week **248388**   Number of waiting jobs: **2330**   Jobs in INSERTING state: **0**

Last Week ✎

Avg completion time: **7h 56m**   Avg submission time: **3m 2.01s**   Avg running time: **7h 53m**   Number of finished trains: **155**   Number of wagons tests: **1177**

Grid jobs	Total	Done	Active	Waiting	Error
Grid - Single core	168612	147414	4692	0	16506
Grid - 2 core	64025	57226	28	2330	4441
Grid - 4 core	2483	2440	0	0	43
Grid - 8 core	0	0	0	0	0
GSI - 2 core	0	0	0	0	0
GSI - 4 core	0	0	0	0	0
GSI - 8 core	0	0	0	0	0
Wigner - Single core	19672	17404	763	0	1505
Wigner - 2 core	1353	1321	0	0	32
Wigner - 4 core	0	0	0	0	0
Wigner - 8 core	0	0	0	0	0
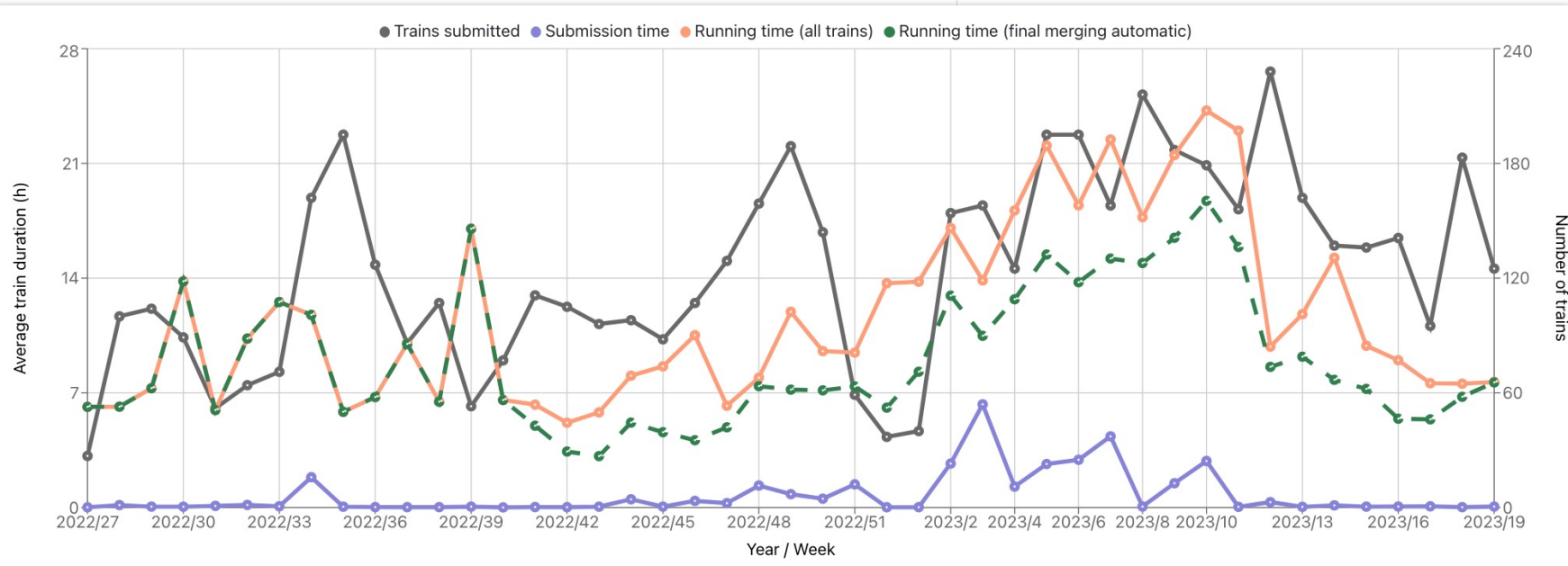
Summary

- ALICE detector received **major upgrades**, allowing for 100 times **more data**

- New tools were developed: O^2 Framework and **Hyperloop Train System**

- O^2 **Framework** allows distributed and efficient processing of the unprecedented data

- **Hyperloop** enables analysis workflows to be run on the Grid and Analysis Facilities (specialized Grid sites with CPU and disk resources adjusted for analysis needs)

- **Automatic activity** (e.g. train composition and submission) → less actions to be taken by the operators

- **Real-time updates** and personalized notifications → users can focus on analysis

- Modern **interactive** User Interface tested with new advanced frameworks (Jest & Puppeteer)

- Easy access to **statistics**, status **overview** and **documentation**

- **User support 24/5**

Thank you!

Ongoing ALICE Run 3 Analysis