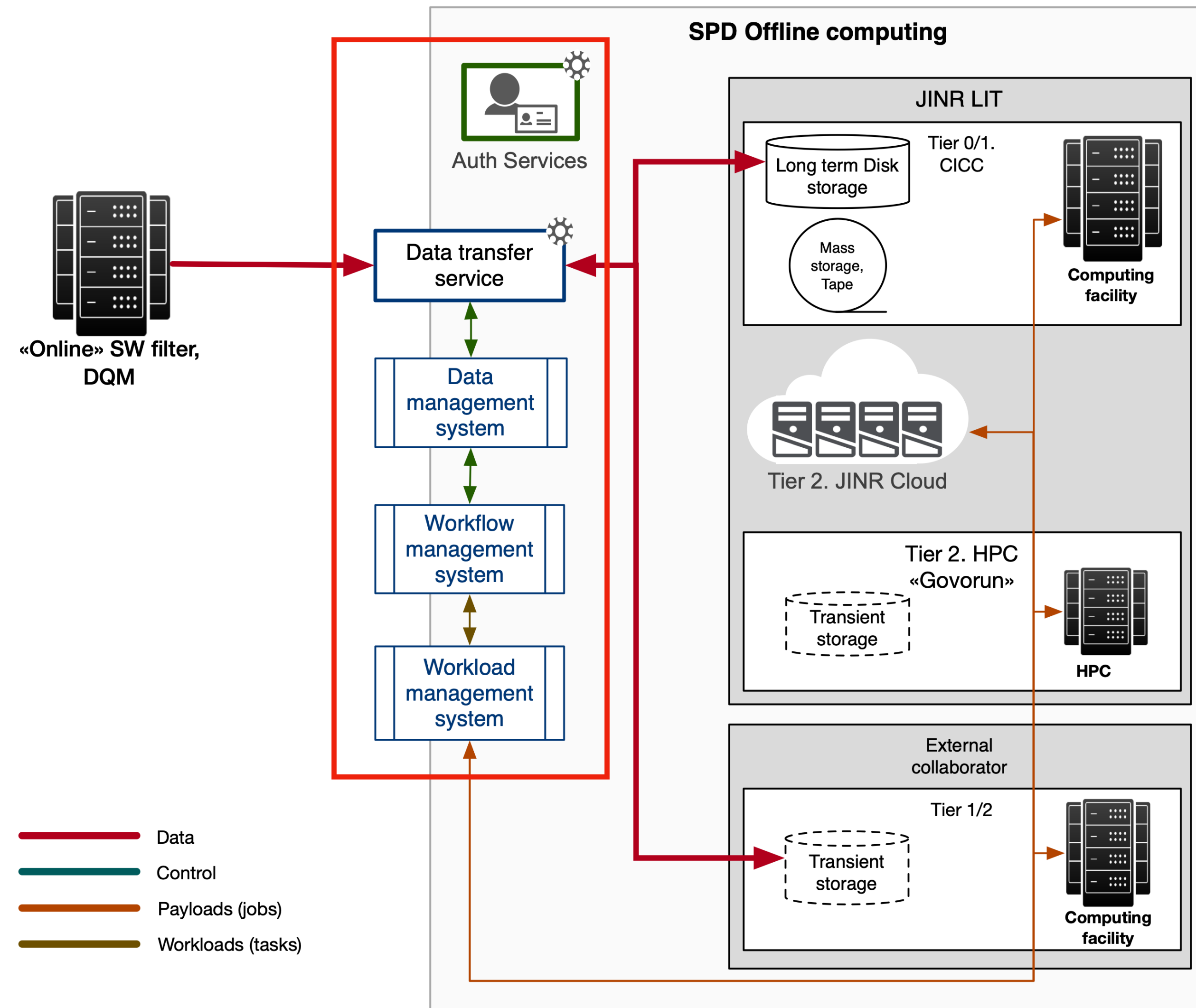


SPD Production System Status

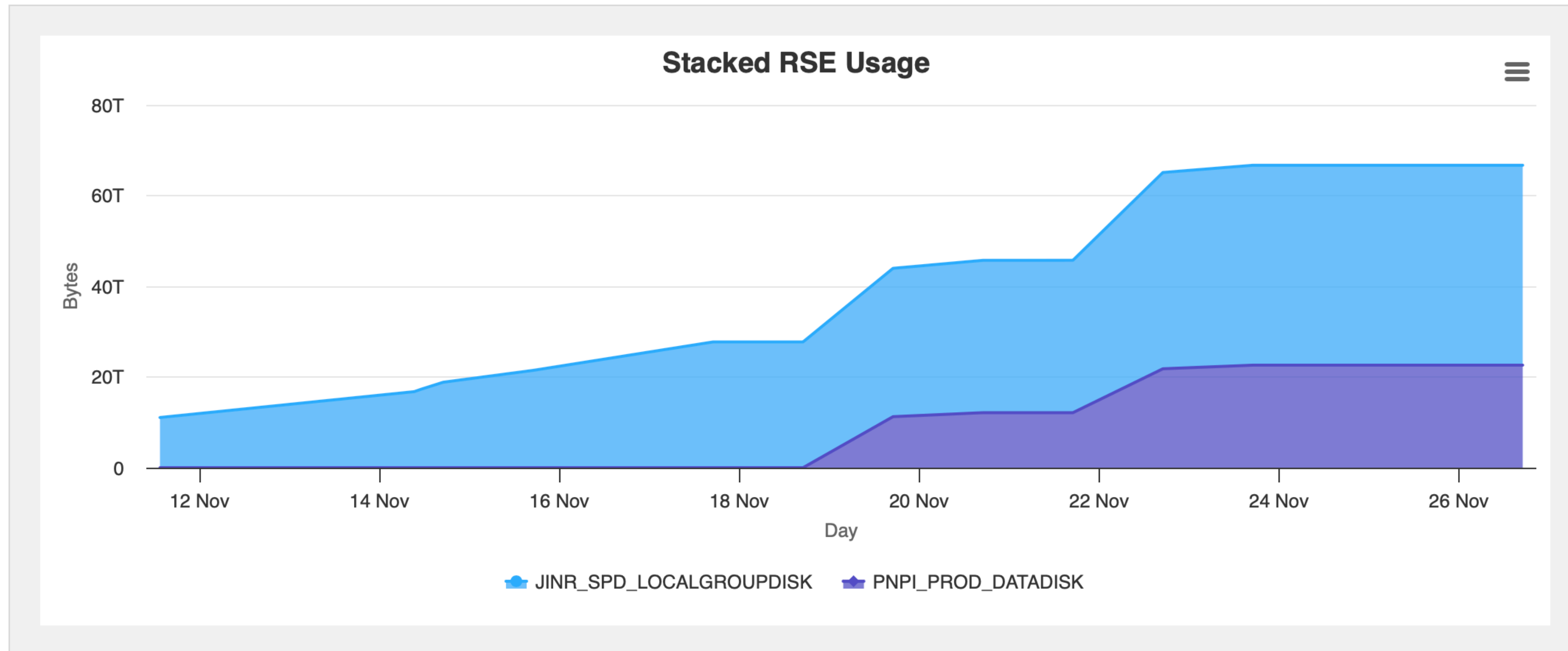
Artem Petrosyan, MLIT, JINR
SPD S&C Weekly
November 26, 2024

Components of the Production System

- IAM — an entry point to all members of the computing services of the collaboration: stores user profiles, their roles and rights to perform certain actions
- CRIC information system — the main integration component of the computing system: contains info about all computing and storage resources, access protocols, entry points, and many other things in one place and distributes this info via API to all other components mentioned below
- PanDA WFMS/WMS — manages data processing at the highest level of chains of tasks and datasets or periods and campaigns, finds the best computing resource for task to be executed on, manages individual jobs (usually 1 job means 1 input file) processing
- Rucio DMS — responsible for data management, including data catalog, data integrity and data lifetime management strategies
- FTS DTS — enables massive data transfers
- Control panel as top level orchestrator and user interface

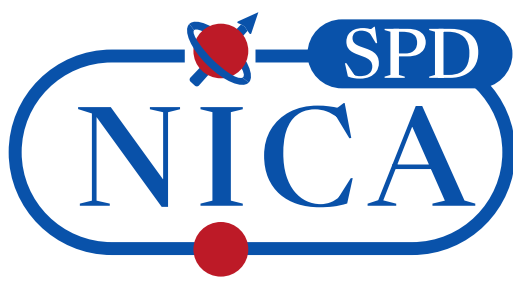


SPD Data Overview



- Recently we moved to the production Rucio server
- With data stored at the test Rucio server we managed to generate ~130TB of data
- We send to PNPI only final results

Production Manager Control Panel 1/3



SPD NICA

Welcome, Artem Petrosyan Create task Tasks Log out

Select field Filter

Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓
257	testProdsysExecution10	257	Monakov Nikita Glebovich	registered	0	1000/1000	0	11:27, 25 Nov 2024	None
256	testProdsysExecution10	256	Monakov Nikita Glebovich	running	0	1000/1000	0	11:27, 25 Nov 2024	11:27, 25 Nov 2024
255	testProdsysExecution9	255	Monakov Nikita Glebovich	registered	0	1000/1000	0	11:25, 25 Nov 2024	None
254	testProdsysExecution9	254	Monakov Nikita Glebovich	running	0	1000/1000	0	11:25, 25 Nov 2024	11:25, 25 Nov 2024
253	testProdsysExecution8	253	Monakov Nikita	registered	0	1000/1000	0	11:12, 25 Nov	None

- Registered as a client in the IAM, can be accessed only by users who has Production role in the IAM
- A control panel for pipelines submissions under development, MC simu+reco pipeline now in testing

Production Manager Control Panel 2/3



SPD NICA

Welcome, Artem Petrosyan [Create task](#) [Tasks](#) [Log out](#)

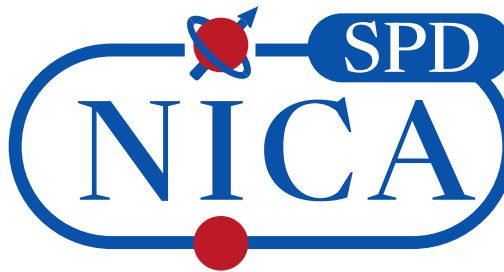
Select field

Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓
257	testProdsysExecution10	257	Monakov Nikita Glebovi	registered	0	1000/1000	0	11:27, 25 Nov 2024	None
256	testProdsysExecution10	256	Monakov Nikita Glebovich	ning		1000/1000	0	11:27, 25 Nov 2024	11:27, 25 Nov 2024
255	testProdsysExecution9	255	Monakov Nikita Glebovich	registered	0	1000/1000	0	11:25, 25 Nov 2024	None
254	testProdsysExecution9	254	Monakov Nikita Glebovich	running	0	1000/1000	0	11:25, 25 Nov 2024	11:25, 25 Nov 2024
253	testProdsysExecution8	253	Monakov Nikita	registered	0	1000/1000	0	11:12, 25 Nov	None

Processing type

MC DATA

Production Manager Control Panel 2/3



Welcome, Artem Petrosyan [Create task](#) [Tasks](#) [Log out](#)

Select field

Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓
257	testProdsysExecution10	257	Monakov Nikita	registered	0	1000/1000	0	11:27, 25 Nov 2024	None
256	testProdsysExecution9	256	Glebovich	running	0	1000/1000	0	11:27, 25 Nov 2024	11:27, 25 Nov 2024
255	testProdsysExecution9	255	Monakov Nikita Glebovich	registered	0	1000/1000	0	11:25, 25 Nov 2024	None
254	testProdsysExecution9	254	Monakov Nikita Glebovich	running	0	1000/1000	0	11:25, 25 Nov 2024	11:25, 25 Nov 2024
253	testProdsysExecution9	253	Monakov Nikita Glebovich	registered	0	1000/1000	0	11:12, 25 Nov 2024	None

MC task configuration

Simulation Reconstruction Simulation + Reconstruction

https://10-220-18-146.jinr.ru/pilot/create-task/naming/?processing_type=MC&data_type=RECO

Production Manager Control Panel 3/3



Task Creation

Data source:

Year:

Energy [GeV]:

Polarization:

Description:

Run number:

Data type:

Dataset name:

Version:

[Create task](#)

Grouping tier	Field	Description	Example
0	[YEAR]	Main Scope - the year of data production	2050
1	[MC DATA]	Real data or simulated data	DATA
2	[energy][polarization]		250LT
3	[desc]	Short name of physics aim	minbias
4	[RunNumber]	Run number for DATA, ID for MC	27189
5	[data type]	EVGEN, SIMUL, RECO....	RAW
6	[<u>DatasetUID</u>]	unique ID of the dataset	636763fd78df7d
7	[Version]	for reprocessing	0

- In order to ease metadata catalog navigation, data filtration, identification, etc., a datasets naming convention was proposed in November 2023
- Dataset name example: 2025.MC.250LT.minbias.27189.RAW.636763fd78df7d.0
- Control panel checks input parameters and ensures that dataset is created in accordance with the naming convention

Summary

- The system is in place, already works well as data generator
- Nearest development plan:
 - Production manager control panel enhancement, we want to allow production managers to define processings in one shot: not at the level of tasks, but at the level of full pipelines
 - Development of the tasks and jobs monitoring
- We are looking forward to start working with SAMPO-based payloads
- Services support and development:
 - Upgrade last remaining services (CRIC, PanDA) to AlmaLinux 9 and change their backend from MySQL to PostgreSQL
 - Continue transition from X.509 to JWT