

Весенняя школа по информационным технологиям ОИЯИ

Система управления заданиями на обработку данных эксперимента SPD

Докладчик: Монаков Никита Глебович
магистрант НИЯУ МИФИ
лаборант ЛИТ ОИЯИ

Руководитель: Петросян А.Ш.
кандидат технических наук

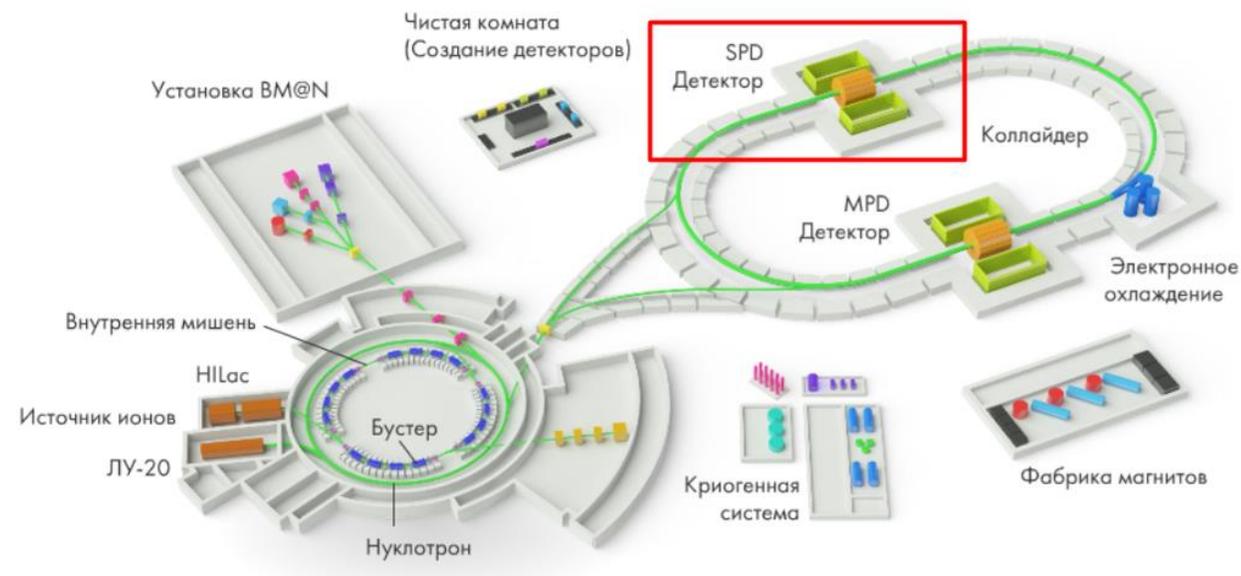
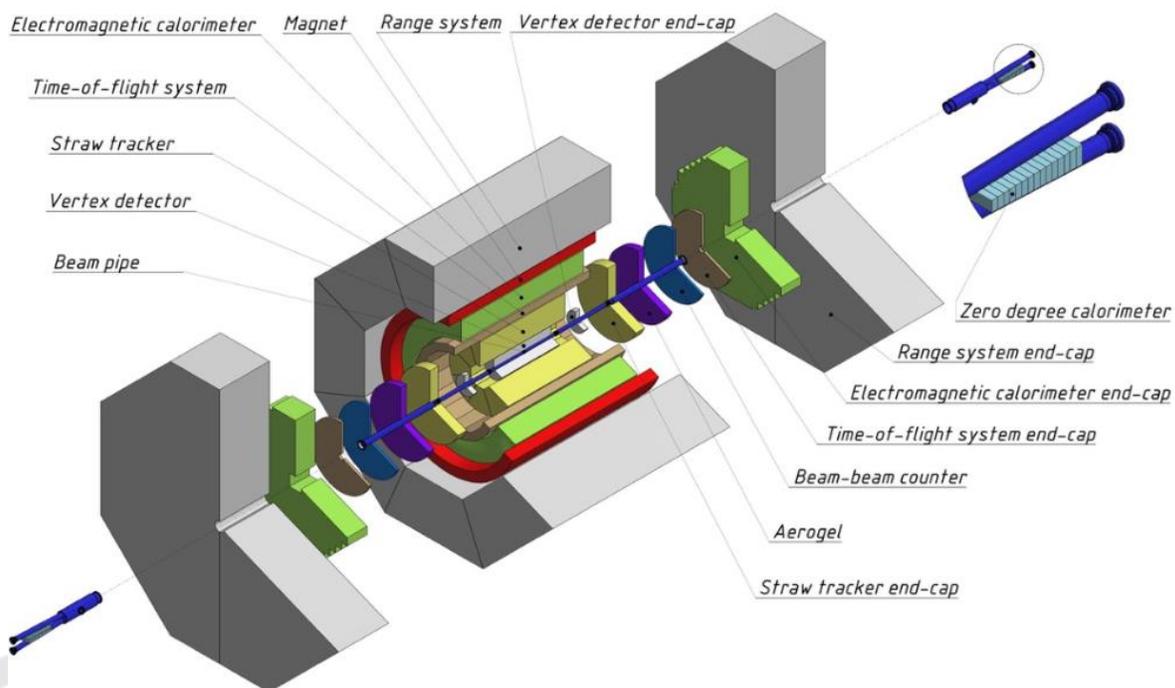


Эксперимент SPD



Установка SPD, размещаемая в одной из двух точек пересечения пучков коллайдера NICA, предназначена для всестороннего изучения **спиновой структуры протона и дейтрона**.

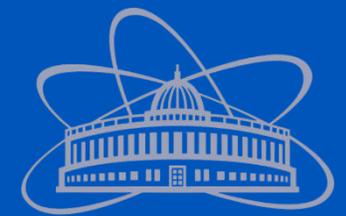
- \sqrt{s} до 27 ГэВ
 - L до $10^{32} \text{ см}^{-2} \text{ с}^{-1}$
- \Rightarrow 10^{13} событий \times 10-15 КБ \approx 10^2 ПБ/год



Основное внимание будет уделено изучению

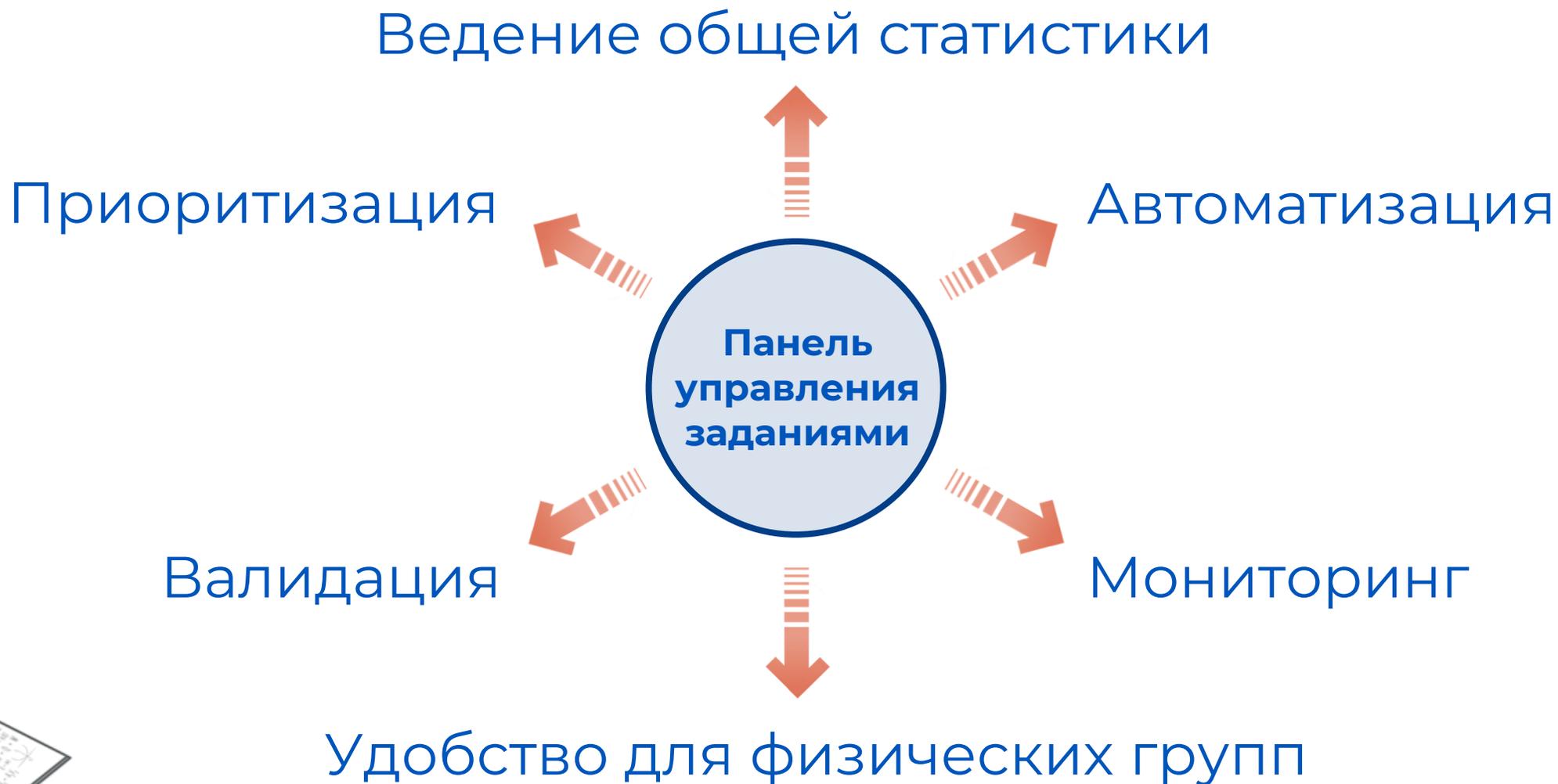
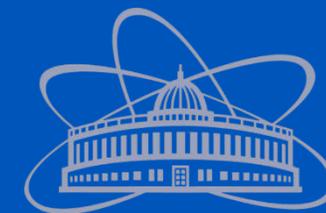
- поляризованной глюонной компоненты в реакциях инклюзивного рождения чармониев
- открытого чарма
- прямых фотонов
- прочих спин-зависимых явлений в p - p / d - d столкновениях

Мотивация создания системы

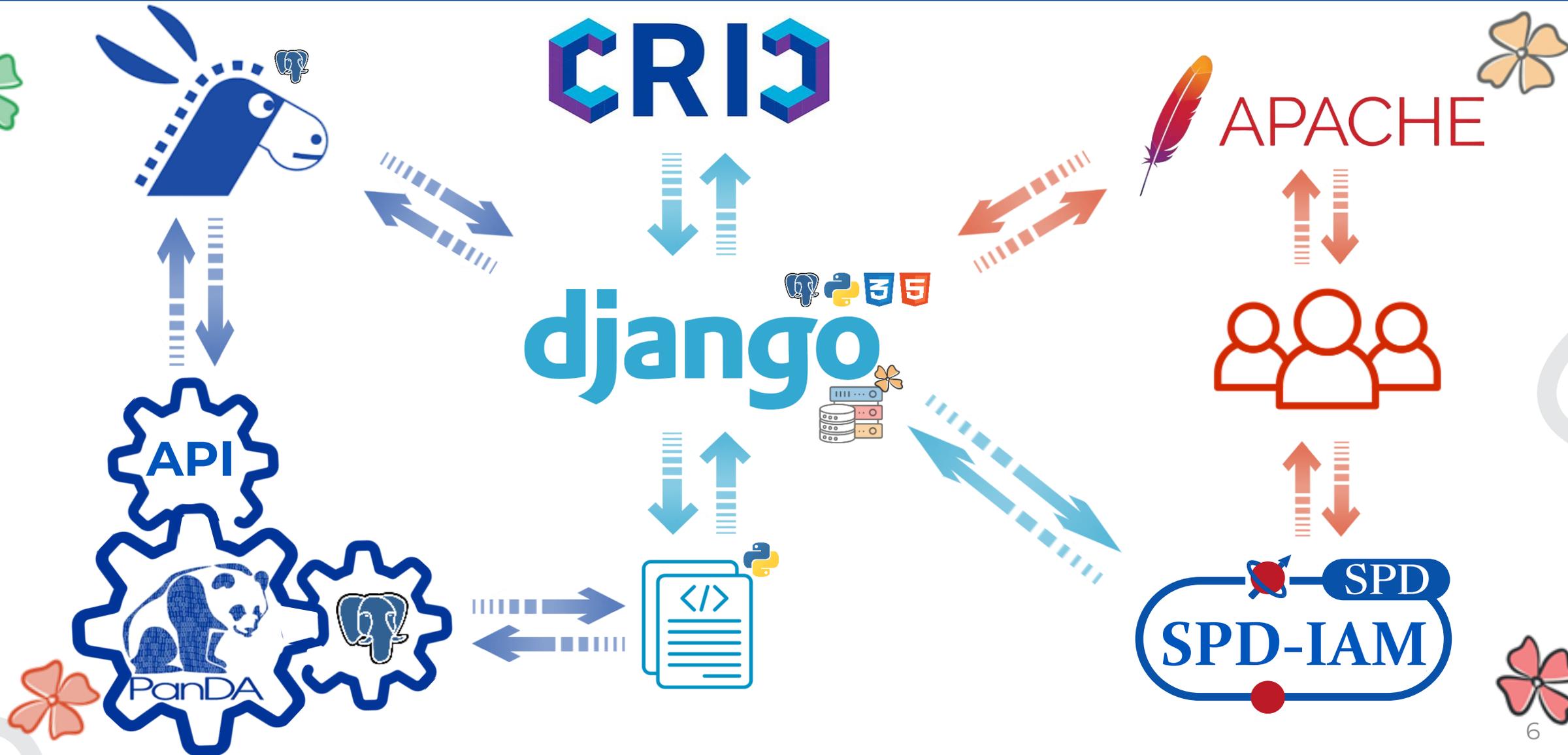


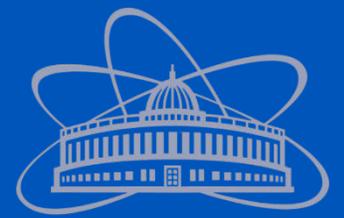
Многokратное повторение однотипных задач

- Необходимость постоянного сбора параметров для запуска задач
- Необходимость генерировать прокси сертификаты или получать токены
- Постоянно существующая вероятность ошибки

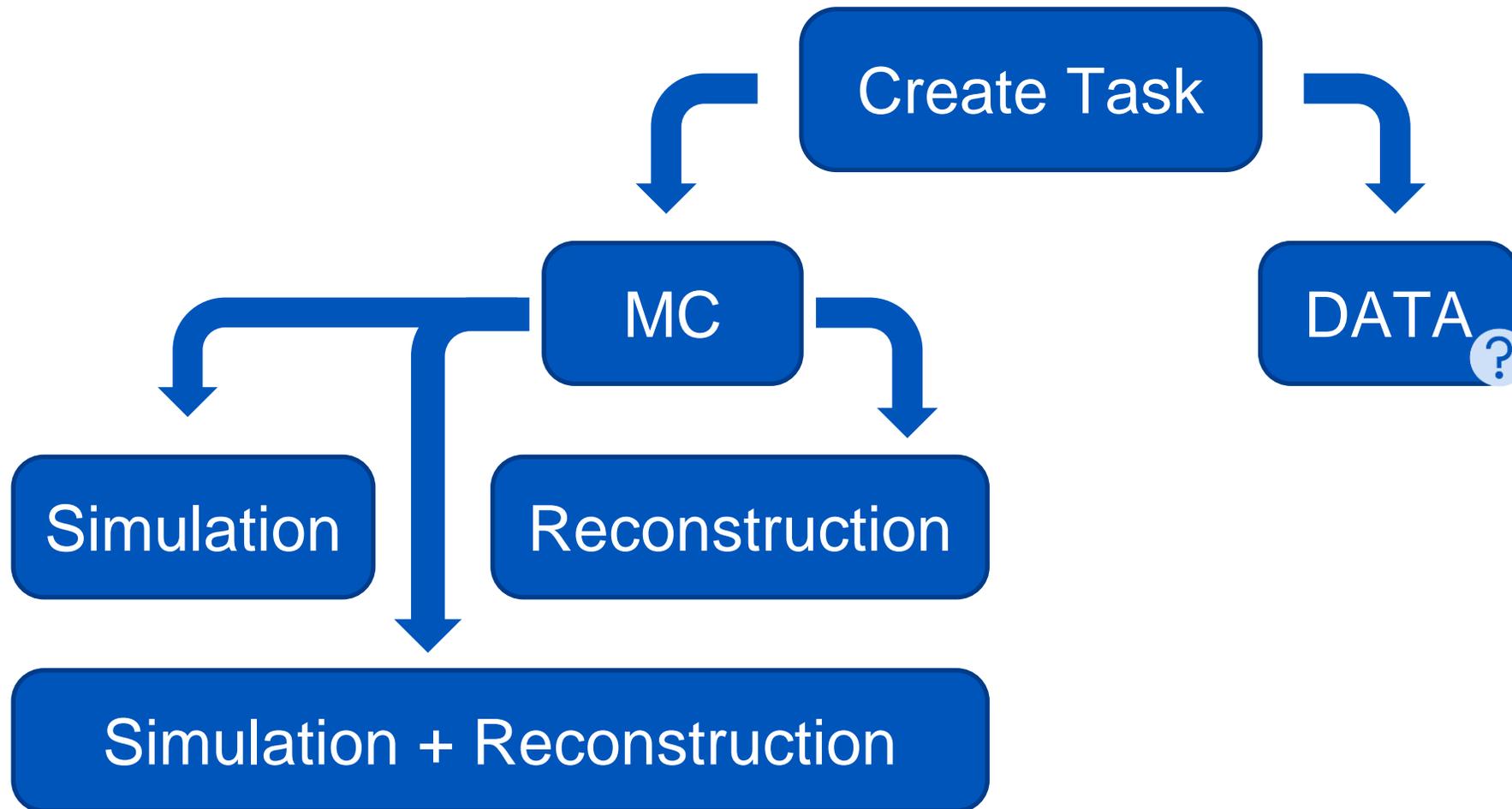
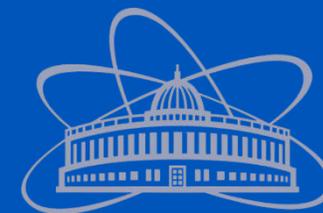


Взаимосвязи компонентов и систем

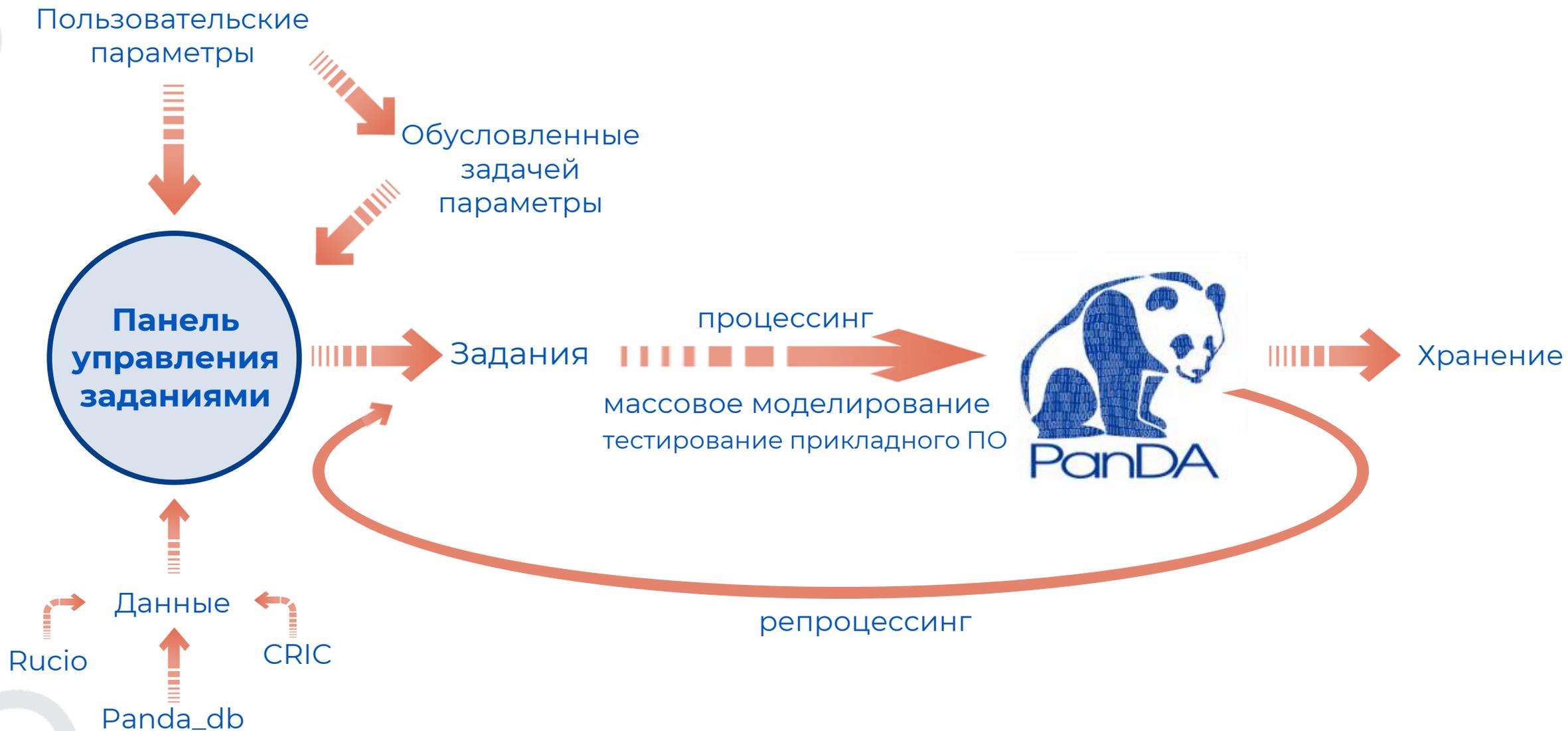


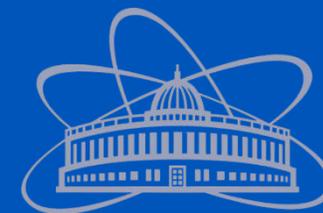


- Создан прототип веб-приложения
 - аутентификации пользователя с использованием `spd-iam`
 - создание заданий в соответствии с принятыми конвенциями
 - отслеживание состояния заданий с возможностью фильтрации и сортировки
 - настроены декораторы и `middleware`
- Начата реализация автоматизирующих скриптов
- Налажено взаимодействие с PanDA WMS с использованием JWT
- Логирование взаимодействий с PanDA WMS



Создание заданий 2/3





Simulation

Task Creation

Task name:

Output dataset name:
[Naming convention here](#)

Total events:

Events per job:

Cloud:

Data disk:

Skip scout:

Offset:

Path to execution files:
`smth like -> /cvmfs/spd.jinr.ru/production/MC/minbias-P8-spdroot417-dev.10GeV.V01`

Path to container:
`smth like -> /cvmfs/spd.jinr.ru/images/spdroot-dev-4.1.7.sif`

Create task

Reconstruction

Task Creation

Task name:

Input dataset name:
[Naming convention here](#), note that no extension expected

Output dataset name:
[Naming convention here](#)

Files per job:

Cloud:

Data disk:

Skip scout:

Offset:

Path to execution files:
`smth like -> /cvmfs/spd.jinr.ru/production/MC/minbias-P8-spdroot417-dev.10GeV.V01`

Path to container:
`smth like -> /cvmfs/spd.jinr.ru/images/spdroot-dev-4.1.7.sif`

Create task

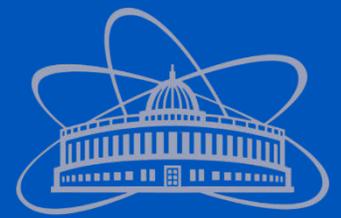


Текущие задания

Select field

Task ID	Task name ↑ ↓	Parent ID	Creator	Status	Done jobs	Default/Current priority	Total events	Submit time ↑ ↓	Start time ↑ ↓	End time ↑ ↓
343	PROD2025-008.RECO	343	Elena Zemlyanichkina	finished	9997	900/900	0	12:40, 06 Apr 2025	12:48, 06 Apr 2025	03:56, 07 Apr 2025
342	PROD2025-008.SIM	342	Elena Zemlyanichkina	done	10000	900/900	0	03:06, 05 Apr 2025	06:08, 05 Apr 2025	11:38, 06 Apr 2025
341	PROD2025-007.RECO	341	Elena Zemlyanichkina	finished	9993	900/900	0	03:40, 04 Apr 2025	06:46, 04 Apr 2025	01:41, 05 Apr 2025
340	PROD2025-006.RECO	340	Elena Zemlyanichkina	aborted	0	900/900	0	03:21, 04 Apr 2025	None	03:35, 04 Apr 2025
339	PROD2025-007.SIM	339	Elena Zemlyanichkina	done	10000	900/900	0	12:42, 03 Apr 2025	12:50, 03 Apr 2025	22:57, 03 Apr 2025
338	PROD2025-006-RECO	338	Elena Zemlyanichkina	done	1535	900/900	0	04:08, 03 Apr 2025	04:11, 03 Apr 2025	12:04, 03 Apr 2025
337	PROD2025-006-SIM	337	Elena Zemlyanichkina	running	1231	900/900	0	12:22, 01 Apr 2025	12:24, 01 Apr 2025	None
336	PROD2025-006	336	Elena Zemlyanichkina	aborted	0	900/900	0	11:27, 01 Apr 2025	None	12:10, 01 Apr 2025
335	PROD2025-005-RECO.1	335	Artem Petrosyan	finished	1221	900/900	0	12:15, 25 Mar 2025	12:18, 25 Mar 2025	16:35, 25 Mar 2025
334	PROD2025-005.1	334	Artem Petrosyan	done	1250	900/900	0	10:16, 25 Mar 2025	10:18, 25 Mar 2025	12:25, 25 Mar 2025

Page: 1 2 3 ... 25 26 27



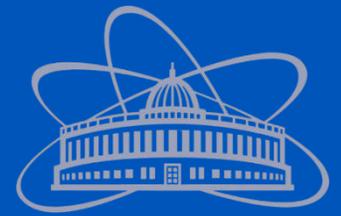
Успешные задания

Select field

Task ID	Task name ↑ ↓	Status	Start date	End date	Walltime	Total events	Events per job	Total jobs	Out DS size, GB	Out Log size, GB
343	PROD2025-008.RECO	finished	06 Apr 2025	07 Apr 2025	30	None	None	9997	3390.70	3.94
342	PROD2025-008.SIM	done	05 Apr 2025	06 Apr 2025	0	40000000	4000	10000	3214.03	1.06
341	PROD2025-007.RECO	finished	04 Apr 2025	05 Apr 2025	38	None	None	9993	3389.51	3.97
339	PROD2025-007.SIM	done	03 Apr 2025	03 Apr 2025	0	40000000	4000	10000	3214.65	1.07
338	PROD2025-006-RECO	done	03 Apr 2025	03 Apr 2025	15	None	None	1535	1138.59	0.62
335	PROD2025-005-RECO.1	finished	25 Mar 2025	25 Mar 2025	26	None	None	1221	413.70	0.48
334	PROD2025-005.1	done	25 Mar 2025	25 Mar 2025	0	5000000	4000	1250	401.83	0.13
333	PROD2025-004-RECO.1	finished	24 Mar 2025	25 Mar 2025	17	None	None	9993	7869.24	4.04
331	PROD2025-004-1	done	20 Mar 2025	22 Mar 2025	10524	40000000	4000	10000	7674.93	1.59
329	TEST_S1-001-SIM.1	done	18 Mar 2025	18 Mar 2025	0	100	10	10	0.03	0.00

Page: 1 2 3 4 5

Содержание заданий



Моделирование и реконструкция выполняются для p-p столкновений с использованием SpdRoot.

Фреймворк включает:

- геометрическое описание SPD-детектора
- распространение частиц
- упрощенное моделирование отклика детектора
- алгоритмы реконструкции

Во время моделирования и реконструкции были включен:

- вершинный детектор
- straw – трекер
- система определения времени полета
- электромагнитный калориметр
- система определения дальности полета

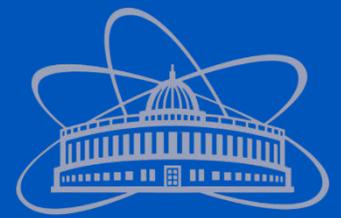
На текущей стадии разработки сохраняется избыточная информация о взаимодействии частиц внутри детектора, которая является предметом будущей оптимизации размера события, а также на этапах моделирования и реконструкции

	Всего событий	Событий/файл	Initial seed	Краткое описание		
PROD2025-001 <input type="button" value="Done"/> →	5M	4000	1-250	minbias-P8-spdroot417-dev test		
PROD2025-002 <input type="button" value="Done"/> →	20M	4000	1-5000	minbias-P8-spdroot417-dev	Этап	S1
PROD2025-003 <input type="button" value="Done"/> →	20M	4000	5001-10000	minbias-P8-spdroot417-dev	Энергия	10 ГэВ
PROD2025-004 <input type="button" value="Done"/> →	40M	4000	1-10000	minbias-P8-spdroot417-dev	Поляризация	UU
PROD2025-005 <input type="button" value="Done"/> →	5M	4000	1-250	minbias-FTF-spdroot4171-dev		

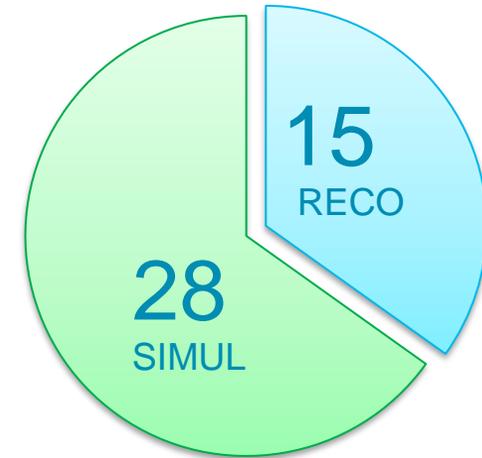
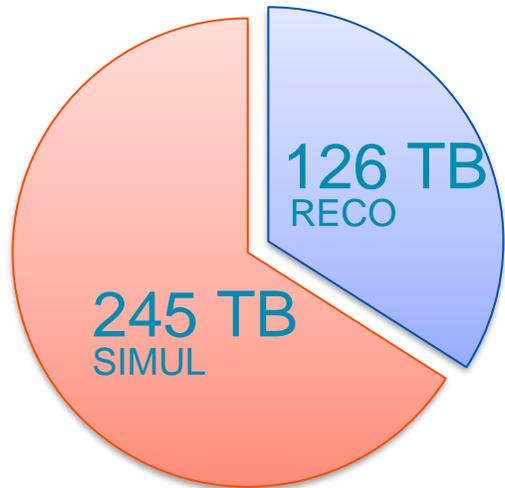
Шаблон наименования датасетов

(DATA SOURCE)(YEAR)_(STAGE):(SHORT DESCRIPTION).(ENERGY)-(POLARIZATION).(PRODUCTION NAME).(TASK TYPE).(VERSION).(EXTENSION)

https://docs.google.com/spreadsheets/d/10uG1p6wPQ_GWe7wkZ_t4RPFLmsHuHxnth5MqM49uOYQ/

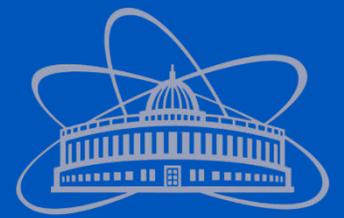


- Успешно обработано около 240 тысяч задач в рамках 43 заданий



- Суммарный объем выходных датасетов – около 370 ТБ





- Реализация механизма клонирования заданий*
- Дальнейшая автоматизация создания и валидации заданий
- Расширение мониторинга заданий



**Благодарю за
внимание**



