



СИСТЕМА УПРАВЛЕНИЯ ПРОЦЕССАМИ ОБРАБОТКИ ДЛЯ СПЕЦИАЛИЗИРОВАННОГО ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА SPD ONLINE FILTER

Студент

Научный консультант

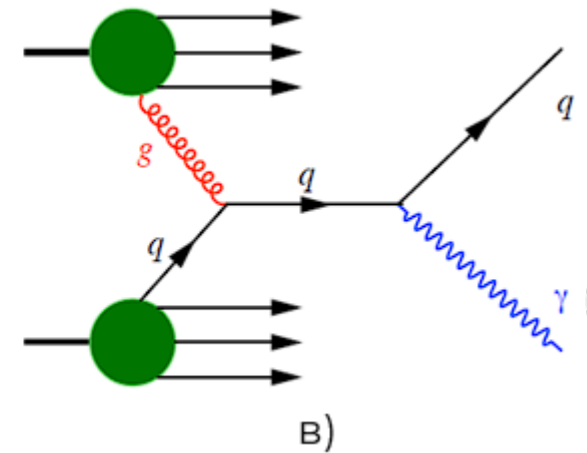
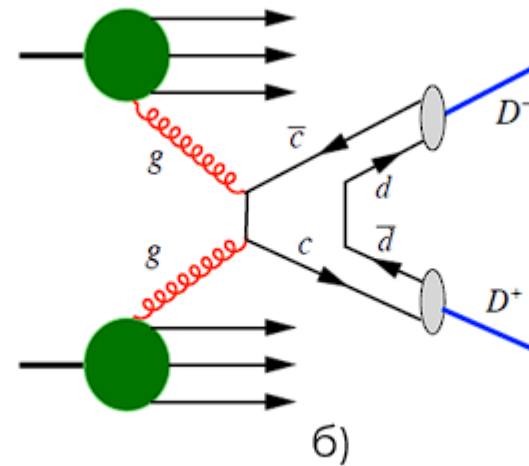
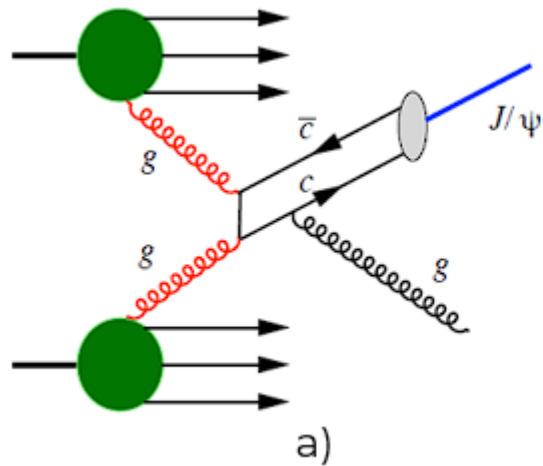
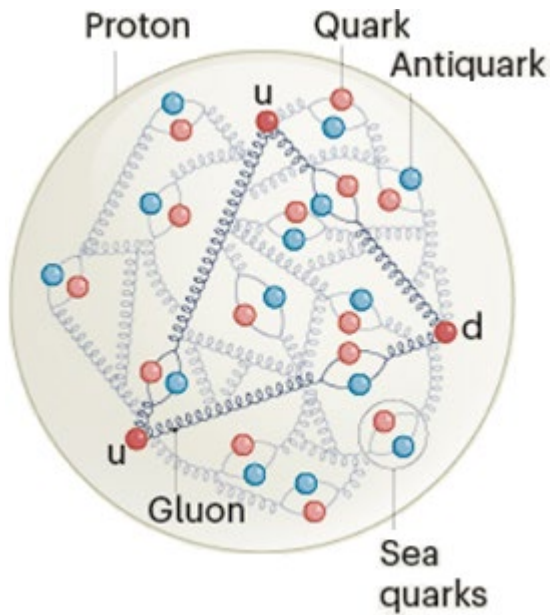
А. В. Плотников

Д. А. Олейник

Москва 2023



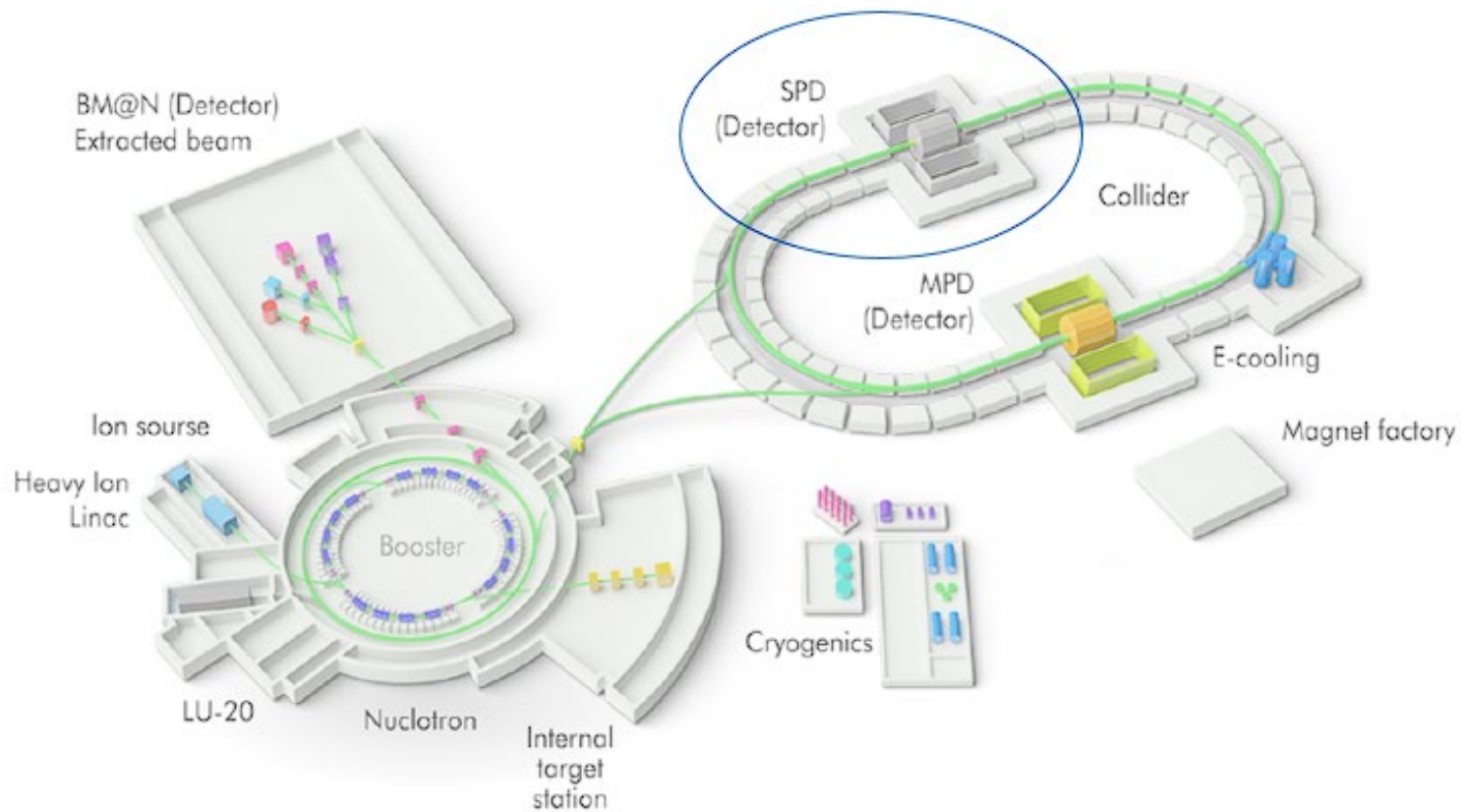
Изучение спиновой структуры нуклонов



Комплексная структура
протона

Изучение поляризованной глюонной компоненты в реакциях инклюзивного рождения:
а) чармониев; б) открытого чарма; в) прямых фотонов.

Spin Physics Detector



Комплекс NICA, Дубна

Цель: всестороннее изучение спиновой структуры протона и дейтрона в поляризованных p-p, d-d и p-d-столкновениях

Объём данных, получаемых с детектора

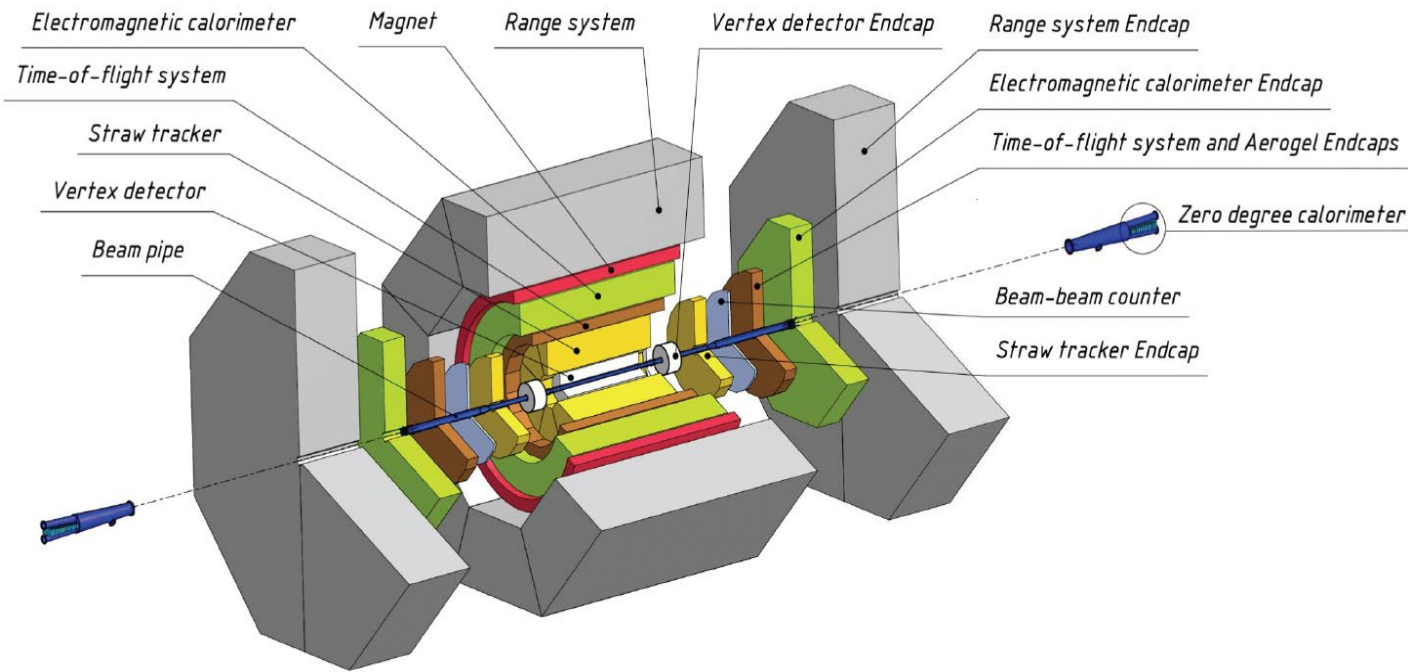


Схема 4 п детектора

- Пересечение пучков с частотой 13 МГц.
- Предполагаемая частота регистрируемых событий ~ 3 МГц (при проектной светимости $10^{32} \text{ см}^{-2} \text{ с}^{-1}$)
- Из приблизительной оценки получаем 20 ГБ/с (или 200 ПБ/год), что соответствует $\sim 3 \cdot 10^{13}$ событий/год.

При этом ввиду сложности и широты изучаемых процессов (выбор физического сигнала требует реконструкции импульса и вершины) невозможно построить критерии отбора данных на аппаратном уровне



Бестриггерная система сбора данных

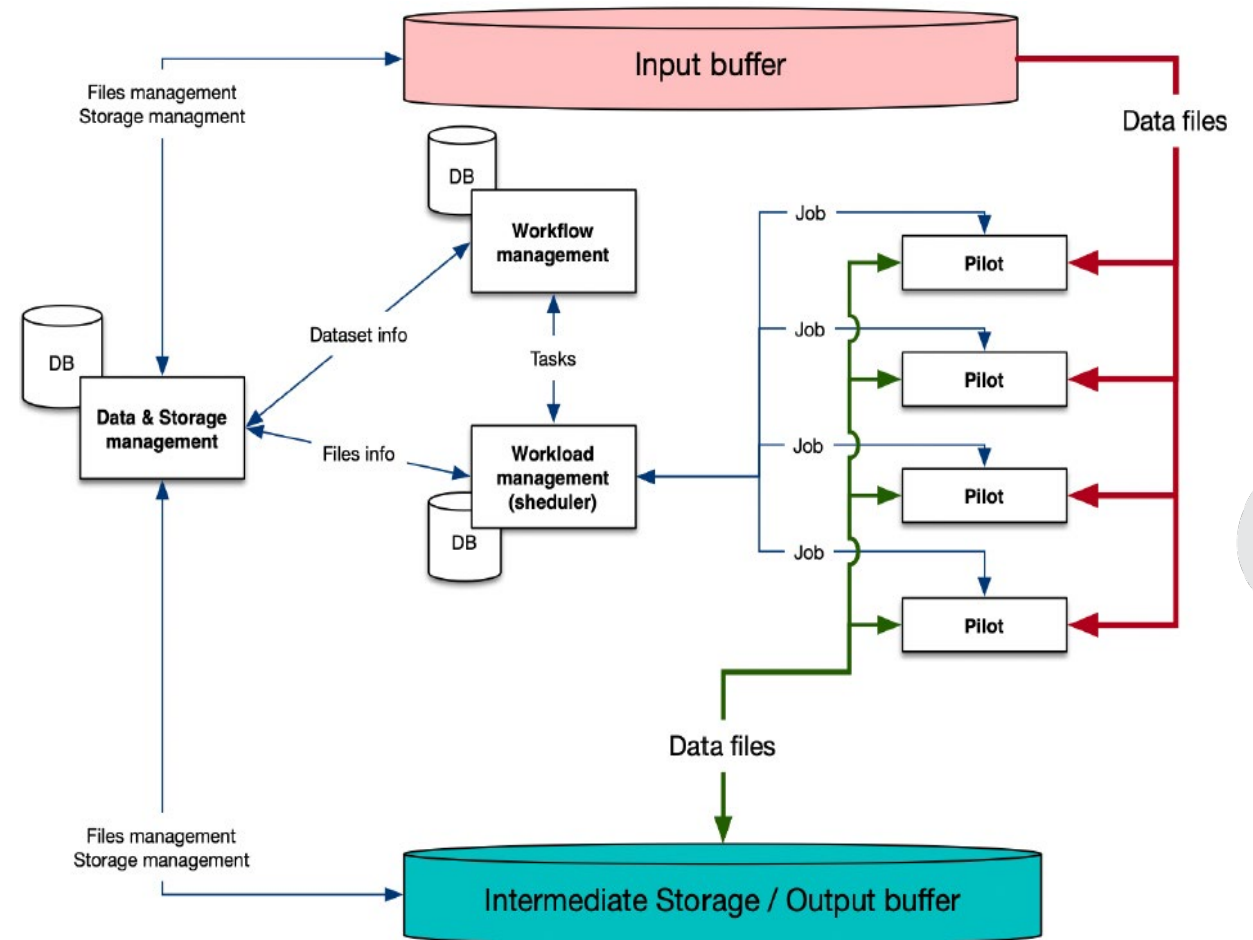
SPD Online Filter

Цель:

Отобрать и сохранить только нужные данные.

Задачи:

- 1) подготовка входных данных;
- 2) частичная реконструкция;
- 3) сравнение данных после частичной реконструкции с данными, полученными после прогонки через обученную модель;
- 4) фильтрация;
- 5) подготовка выходных данных.



Архитектура Online Filter

Управление процессами обработки данных

Концепция базовой обработки:

- Реконструкция треков и связывание их с вершинами;
- Связывание срабатываний электромагнитного калориметра и пробежной системы с каждой вершиной по времени;
- Определение несвязанных срабатываний детектора;
- Объединение с необработанными данными от других субдетекторов;
- Формирование блоков данных и сохранение частично реконструированных событий.

Система должна одновременно и быстро обрабатывать большой объём данных



Для каждого из шагов обработки нужен свой процесс



При этом каждый из шагов образует промежуточные данные, которые нужно перенаправлять на дальнейшую обработку



Нужна система управления процессами обработки

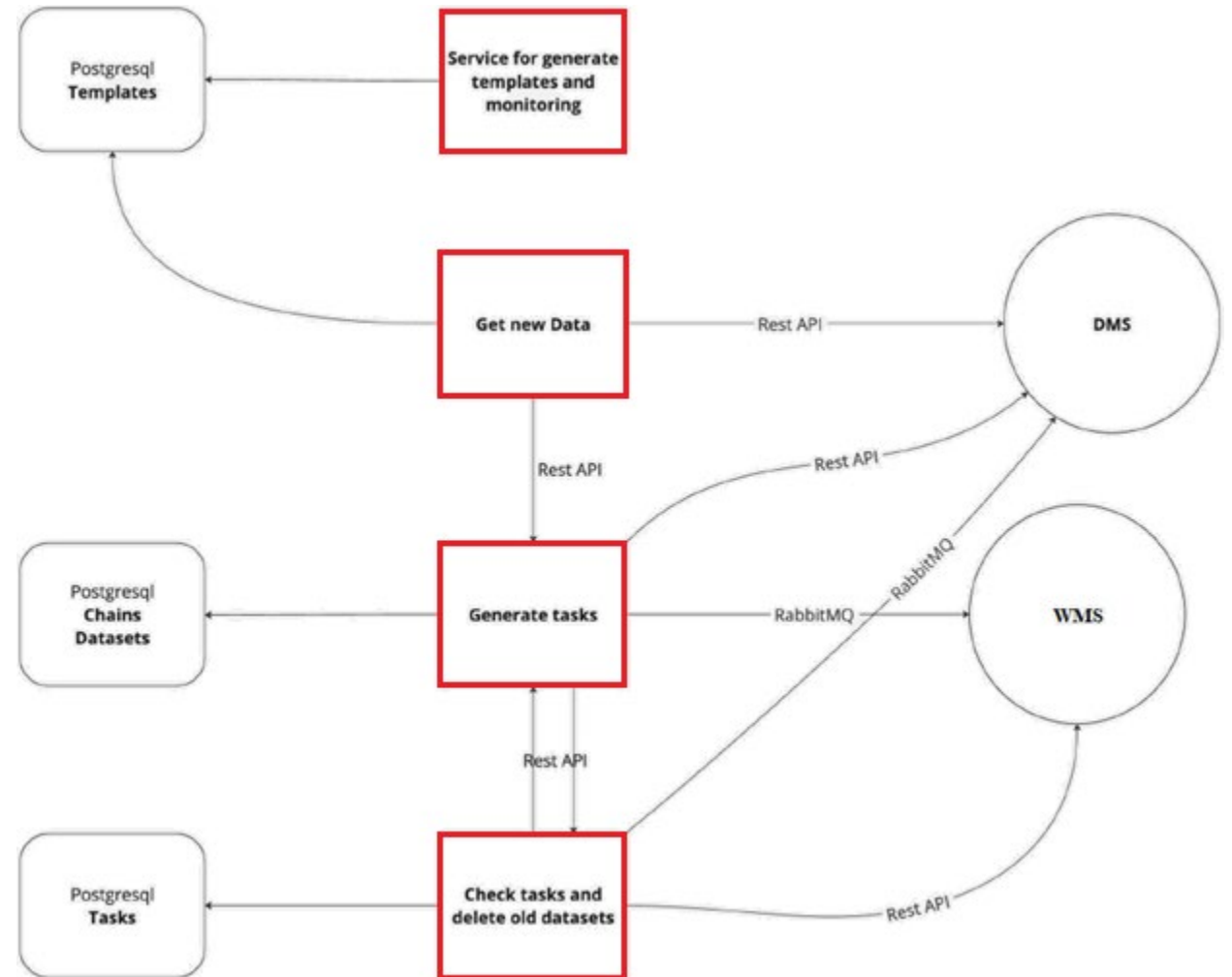
Система управления процессами обработки (WfMS)

Основная цель системы:

параллельное и наиболее результативное управление большим количеством процессов обработки данных и контроль статусов выполнения цепочек обработки для физического эксперимента SPD.

Основные задачи системы:

- Создание формального описания процесса;
- Контроль за выполнением обработки.



Сервисы WfMS

- Сервис для взаимодействия с оператором обработки данных:
 - 1) вывод информации о заданиях;
 - 2) изменение приоритетов заданий или их отмена.
- Сервис для опроса DMS:
 - 1) получение данных из DMS (информация о датасетах);
 - 2) передача промежуточных и выходных данных, а также логов в DMS.
- Сервис генерации заданий:
 - 1) определение последовательностей обработки данных;
 - 2) запрос шаблонов;
 - 3) создание заданий (tasks) по шаблонам для последовательностей.
- Сервис для работы с WMS:
 - 1) отправление заданий на обработку в WMS;
 - 2) периодический опрос WMS для отслеживания статуса.