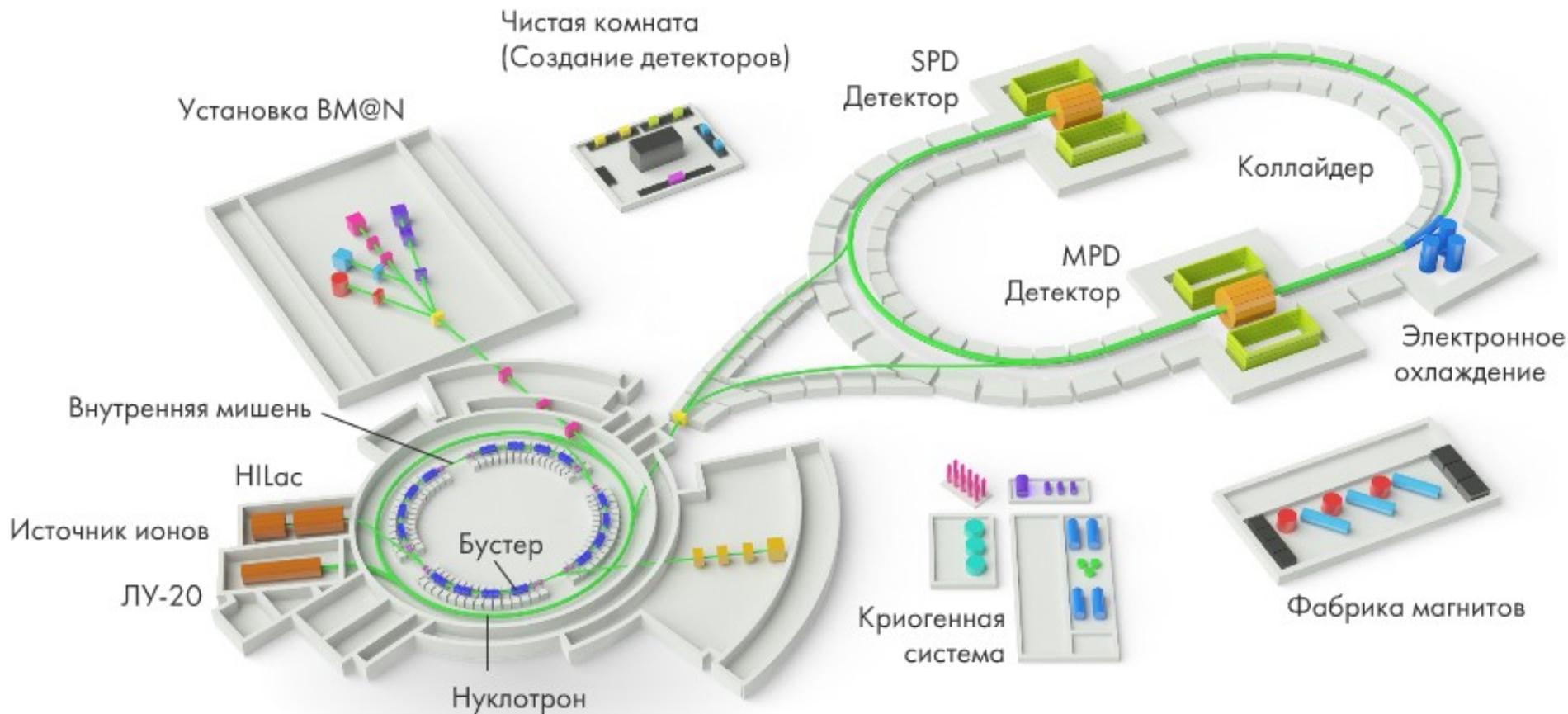
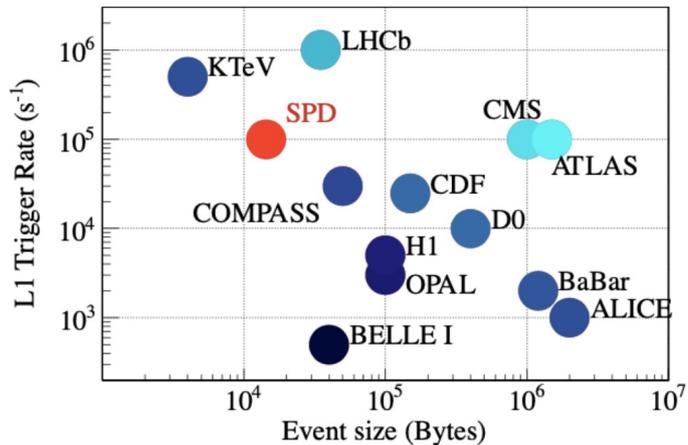




Разработка информационных систем для эксперимента SPD

Прокошин Федор Валерьевич





- Частота событий – 3 МГц
- При расчетной светимости 10^{32} см⁻²с⁻¹ для pp - столкновений при энергии 27 гэВ
- Поток данных 20 Гбит/с (200 ПБ /год при рабочем цикле детектора 0.3)

- Онлайн фильтр должен уменьшить поток в ~20 раз
- 10 ПБ или 10^{13} событий в год при размере события – 10-15 кБ
- + Сравнимые объемы моделированных данных и повторная обработка
- Всего в ходе эксперимента ожидаем получить до **200 ПБ** данных

- Эксперимент SPD в ходе работы произведет огромные объемы данных (порядка 200 пб), которые будут храниться в миллионах файлов разбросанных по узлам распределенных систем
- Для реконструкции и анализа требуется генерация сравнимых объемов данных
- Для обработки экспериментальных данных требуется широкий спектр вспомогательной информации
- Для успешной работы установки требуется контроль большого числа параметров установки и систем обработки данных
- Эти данные необходимо хранить, обрабатывать, и предоставлять пользователям — именно этой цели служат информационные системы эксперимента

- Базы данных не существуют сами по себе, f как правило являются ядром информационной системы
- Помимо БД и СУБД ИС могут содержать
 - Средства сбора информации
 - Средства транспортировки информации
 - Интерфейсы пользователей
 - Интерфейсы приложений (API)
 - Сервера приложений (например кэширующие прокси)
 - Оркестровщики (supervisor) и мониторинг
 - Интерфейсы внешних информационных систем
- ИС экспериментальной установки как правило тесно взаимодействуют между собой и с внешними системами
 - ИС ускорителя, технические службы, сетевые службы

- На установке SPD будут использоваться различные БД и ИС:

- 
 - Базы данных оборудования и подключений

- Базы данных условий и калибровок

- 
 - Системы мониторинга

- 
 - Системы ведения журналов и учета

- 
 - БД физических метаданных

- 
 - Каталог событий (Event Index)

- Системы управления распределенными вычислениями и хранением данных

- 
 - База данных персонала и публикаций

Установка

Данные

- Часть ИС создаются специально для эксперимента SPD
 - При разработке используется опыт полученный на других экспериментах
- Некоторые системы адаптированы для использования на SPD
 - Системы управления распределенным вычислениями и данными
- Для других используется общая структура при существенно отличной реализации компонент
 - Каталог событий, БД физических метаданных
- Некоторые ИС будут совместно разрабатываться совместно с другими экспериментами, или предоставлены ОИЯИ
- В качестве основной платформы - реляционная СУБД PostgreSQL

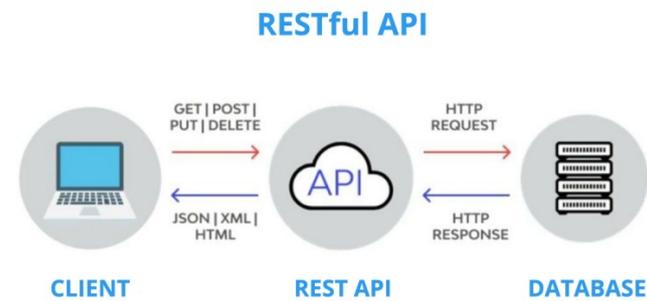
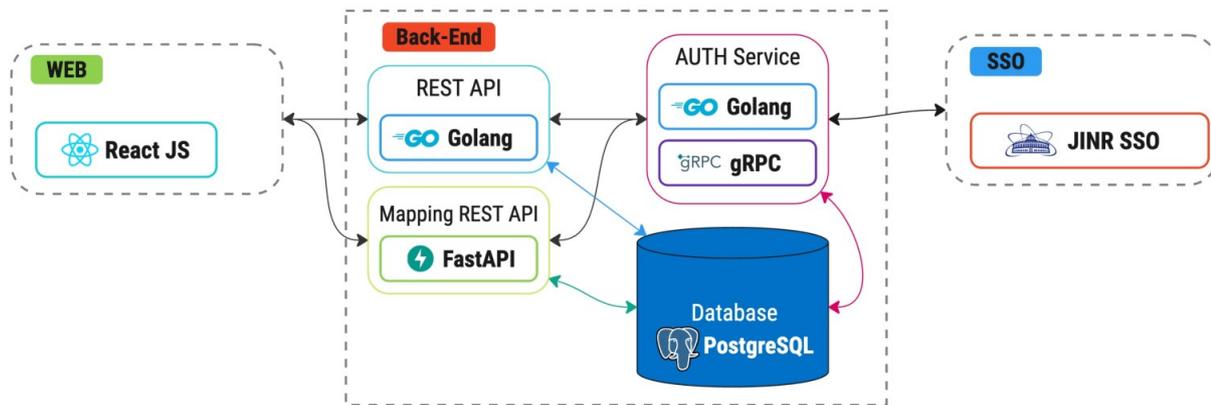
- Каждая детекторная система состоит из множества элементов, число которых может достигать сотен тысяч.
- Разрабатывается БД оборудования (HWDB) - каталог аппаратных компонентов, из которых состоит SPD.
- Будет содержать информацию о детекторных и электронных компонентах, а также их расположении на установке
 - Идентификаторы компонент, модели, различные параметры и другие (полу)постоянные характеристики
- Крайне полезна для разработки и успешной эксплуатации систем установки а также для обмена знаниями между членами команды.

- Количество каналов сбора данных порядка 200 — 300 тысяч
- Сигналы будут проходить через несколько устройств связи
- Необходима карта соединений (mapping) системы сбора данных, устанавливающая соответствие адресов каналов на выходах DAQ устройствам, с которых поступил этот сигнал
- Из-за большого количества элементов в системе практически невозможно построить карту соединений вручную
- Необходима процедура автоматического построения mapping
 - Устройство должен выдавать HWID по каналу передачи данных в ответ на специальный сигнал
 - Для устройств не поддерживающих эту функцию, должен быть предусмотрен интерфейс, позволяющий вводить данные по группам.



- Ожидается, что наполнение базы данных оборудования будет происходить постепенно, а обновления будут редкими
- Построение схемы подключения и ее изменения также будут выполняться редко (не чаще одного раза в неделю)
- Требования к скорости записи информации в БД невысоки
- Информация о сопоставлении может потребоваться при обработке каждого файла. Возможно, что десятки тысяч процессов попытаются одновременно получить доступ к системе.
- Необходимо обеспечить их обработку, избегая перегрузки базы данных из-за слишком высокой частоты запросов

- Разработан прототип системы
 - СУБД PostgreSQL используется для хранения данных
 - Серверная часть интерфейса (Back-End) представляет доступ через Rest API
 - Web интерфейс для ввода, поиска и отображения данных
 - Авторизация через JINR SSO, права доступа выдаются администратором
- Архитектура проекта построена на базе микросервисов
- Используется стек современных программных платформ





- Устройства организованы в группы, по моделям
- Для каждой группы устройства определен набор параметров, которые являются общими для всех устройств этого типа
 - Для каждого параметра задан формат (строка, число, список...)
 - Могут быть заданы значения по умолчанию и/или диапазон
- Каждому устройству присвоен уникальный идентификатор (HWID)
 - для устройства могут быть заданы конкретные значения параметров либо используются значения по умолчанию
- Устройства имеют особый параметр, *component ID*, определяющий “географическое” положение на установке
- Группы и устройства могут быть “клонированы”, позволяет существенно ускорить ввод информации

Hardware Database

prof



- Home
- All Data ▼
 - All groups
 - All devices
- Get Data ▼
 - Get group
 - Get device
- Admin

Create group

NAME	CREATED_AT	UPDATED_AT	ACTIONS				
PG2T70H	2024-11-13 18:48	2024-11-13 18:48	Show props				
PG2T160H							
PG2T390H							
10CX105YF780E6G							
AT5202							
DT5202							

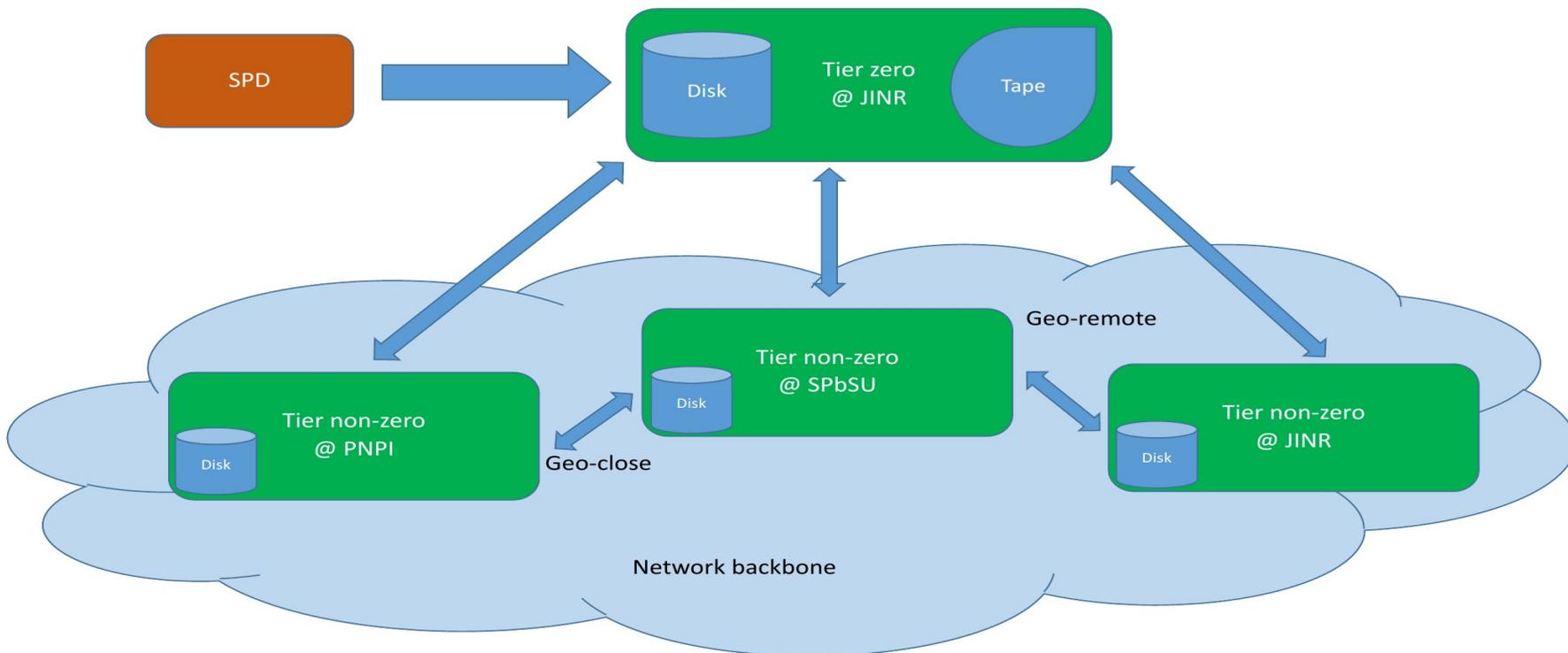
NAME	UNITS	PROPERTY_TYPE	MIN	MAX	DEFAULT_VALUE		
Inputs		int			64		
S/N		str					
Shaping Time Fast	ns	float			15		
production date		datetime					
status		enum					
Shaping Time Slow	ns	enum					
HG_Gain		int	1	63			
eth		ip					

Разработка ведется Б.Двизовым, С.Гуссаловым и Р.Магкаевым

- **Данные условий** - данные, представляющие состояние детектора (не относящиеся к отдельным событиям)
 - Состояние аппаратного обеспечения детектора
 - Калибровки детектора
 - Условия считывания показаний детектора
 - Выравнивания детектора
 - Калибровки по физике
 - Измерения светимости и поляризации
- **Использование данных условий**
 - Калибровка подсистем
 - Оперативная и повторная обработка данных, реконструкция
 - Физические анализы

- Различные фрагменты информации неоднородны как по типу данных, так и по степени детализации во времени
- Данные должны быть организованы по "интервалам достоверности" (**IOV**), которые представляют собой промежуток времени, в течение которого эти данные действительны
- Могут быть пересчитаны позже, если улучшится понимание поведения детектора или повысится качество входных данных.
 - за исключением конфигураций детектора и триггера
- Для воспроизводимости реализуется версионирование данных
- Данные обычно записываются один раз и часто считываются
 - для использования в распределенных вычислениях должна поддерживаться скорость считывания до нескольких кГц

- Для хранения и обработки данных используются распределенные хранилища и вычислительные сети

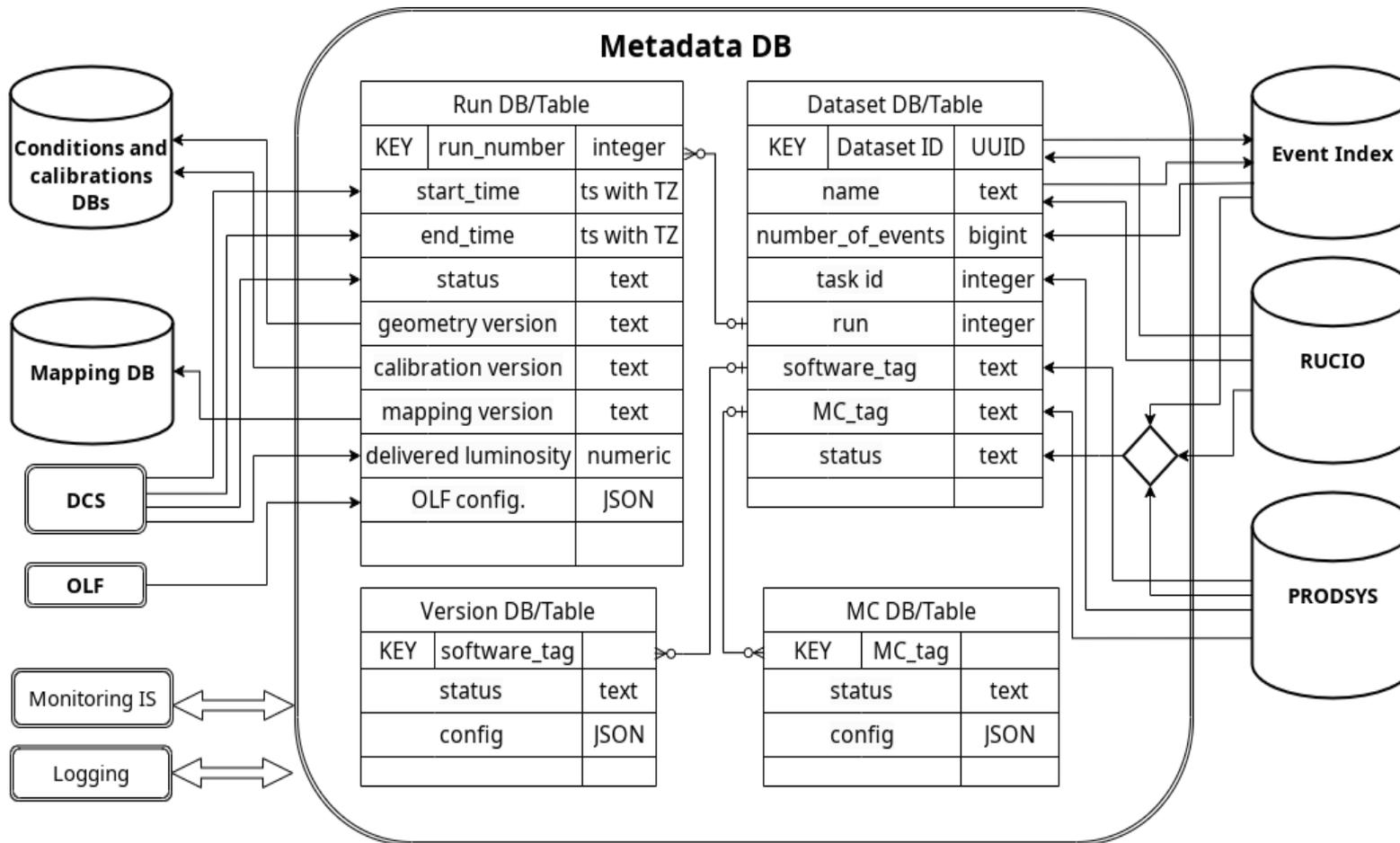


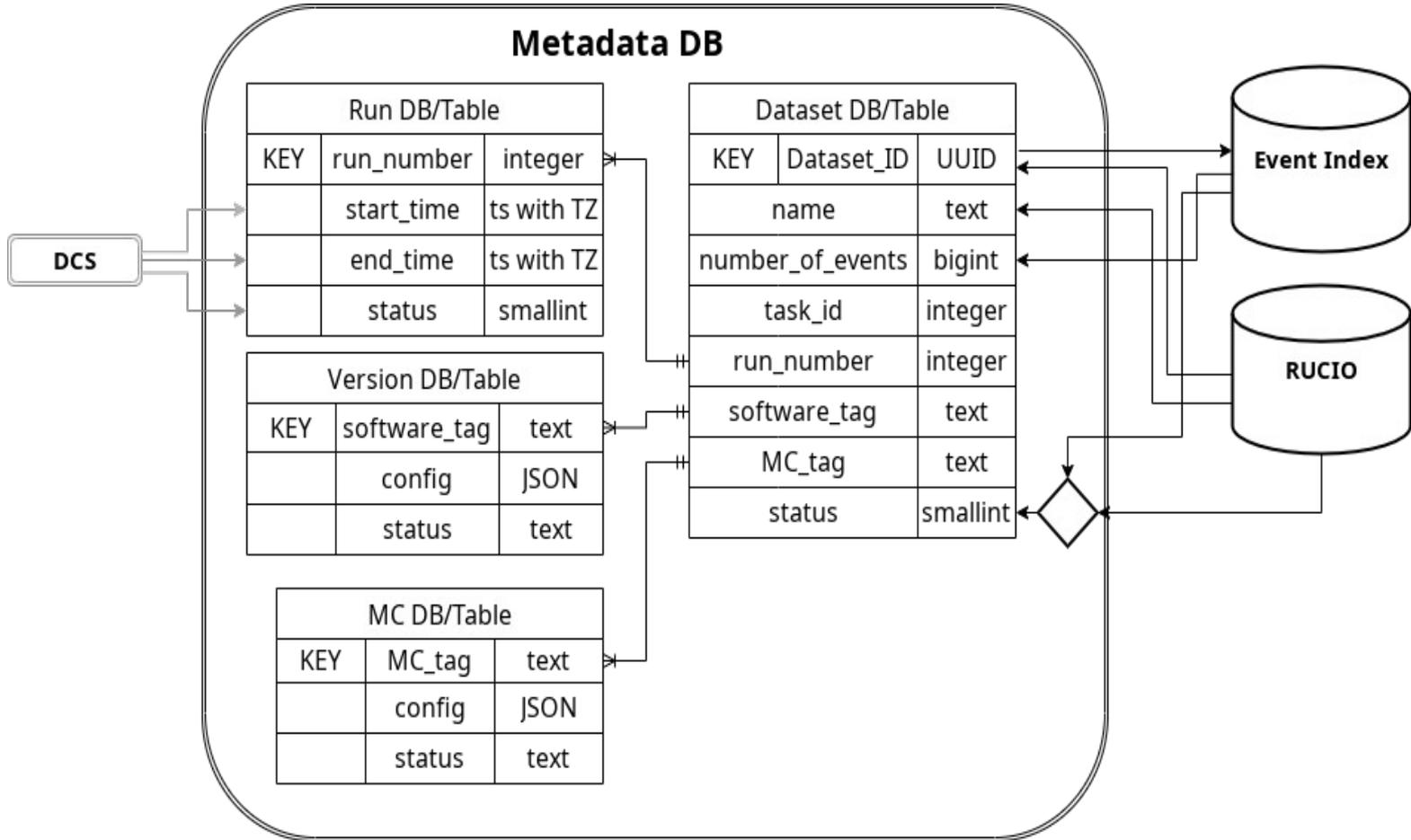
- **Computing Resource Information Catalog (CIRC)**
 - собирает информацию обо всех вычислительных ресурсах и хранилищах данных, протоколах доступа, точках входа, etc...
 - распространяет эту информацию через API
 - эволюционировал из ИС ATLAS Grid (AGIS)
- **ProdSys/JEDI/PanDA (Управление распределенными вычислениями)**
 - Список запрошенных задач (tasks) а также датасеты на входе и выходе этих задач, версии программного обеспечения, ...
 - Список заданий (jobs), их состояние и размещение, ...
 - Списки ресурсов обработки, состояние ...
 - Получение информации и управление задачами возможно через Web интерфейс, клиент в командной строке или API



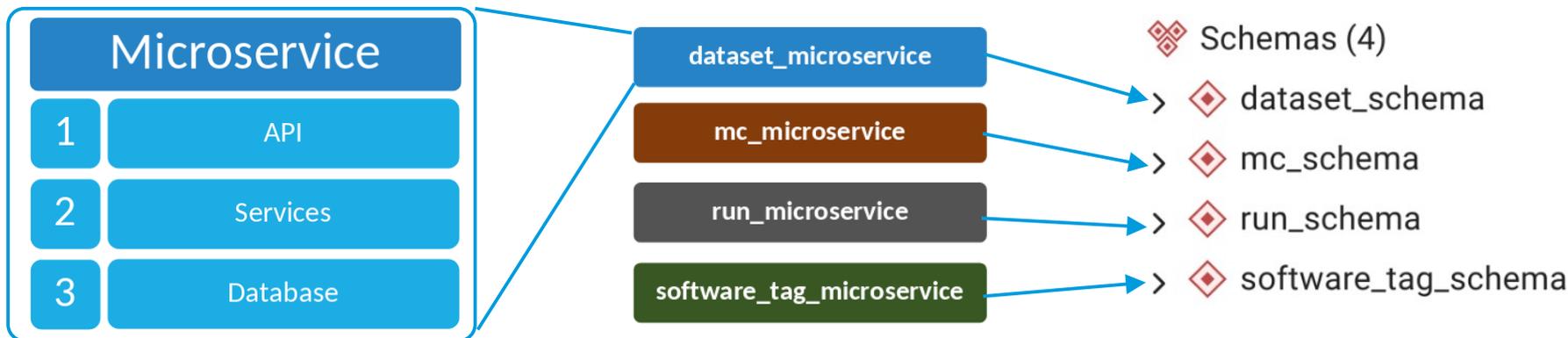
- Система управления данными Rucio
 - Каталог содержимого датасетов:
 - Список файлов, суммарный размер, принадлежность, происхождение, время жизни, состояние ...
 - Каталог файлов:
 - Размер, контрольная сумма, число событий ...
 - Каталог расположения датасетов:
 - Список копий датасетов на узлах сети Grid
 - Средства пересылки данных:
 - Очередь пересылаемых датасетов, состояние
 - Средства для удаления данных:
 - Список датасетов (или копий) для удалены, состояние
 - Списки и состояние ресурсов хранения (на узлах GRID)
- **FTS / DTS: Обеспечивает массовую передачу данных**

- Содержат информацию о
 - наборах данных (датасетах), включая наследование и ссылки на конфигурации задач, которые создали эти датасеты
 - версии ПО, используемые при обработке данных
 - сечения и конфигурации, используемые для моделирования
 - информация о параметрах пучка (светимость, поляризация)
 - настройки онлайн фильтра
- Большая часть информации информации будет поступать из других ИС эксперимента, анализа логов и конфигурационных файлов
- Номенклатуру датасетов, параметров Run, спецификации версий ПО и конфигураций MC еще предстоит разработать
- Создается концептуальный прототип ИС метаданных





- Архитектура проекта построена на основе микросервисов.
- Каждый микросервис представляет из себя структуру из трех слоев:
 - Слой API, слой сервисов и слой работы с базами данных
- Один микросервис — одна схема базы данных
- Каждый микросервис работает в Docker контейнере.
- Внутри Docker контейнера создаются экземпляры микросервисов, что позволяет обрабатывать запросы в параллельном режиме





- Слой API представляет для пользователя программный интерфейс, который работает через протокол HTTPs.
 - На этом уровне происходит валидация данных от клиента.
 - Запросы пользователей отправляется на следующий уровень – сервисов.
- Слой сервисов реализует логику проекта.
 - На этом уровне происходит обработка и преобразование данных, общение с внешними системами, реализуется внутренний функционал системы.
- Слой баз данных отвечает за настройку подключений, запись и чтение из баз данных.
 - Реализует ORM — преобразование объектов в таблицы баз данных

Разработка ведется Н.Шкериным и Р.Короткиным

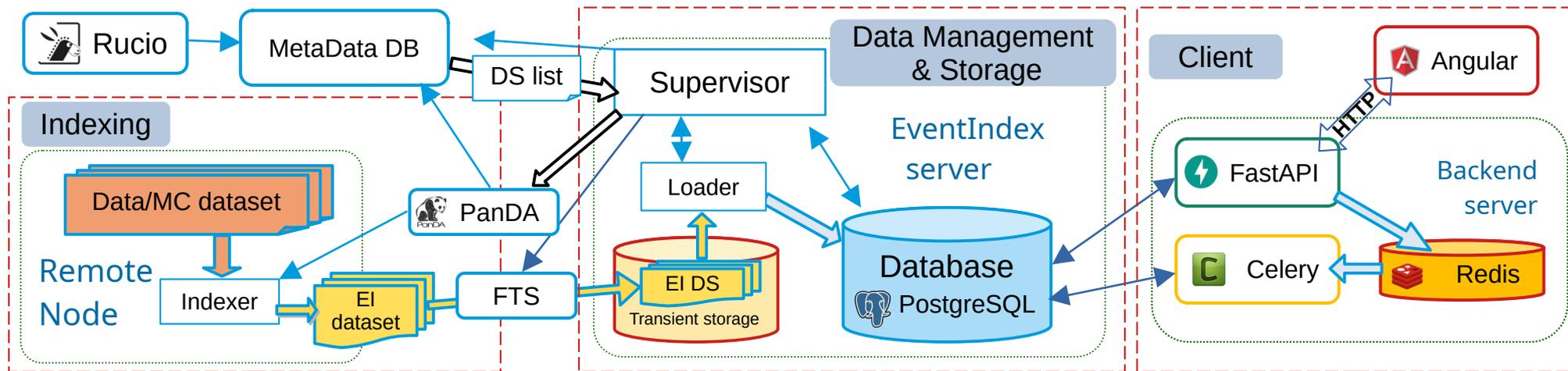


- Полный каталог событий SPD, реальных и смоделированных
- Область применения:
 - доступ по идентификаторам к наборам событий, полученным по результатам отбора в рамках физических анализов
 - подсчет и предварительный отбор событий для анализа
 - создание «виртуальных датасетов» из отобранных событий
 - проверка целостности и корректности наборов события в датасетах
- Запись в EventIndex для события:
 - Идентификаторы события (Run Number, EventNumber)
 - Идентификаторы файлов и датасетов с этим событием (все версии)
 - Результаты On-Line фильтра и другие важные параметры (TBD)



- **EventIndex** должен обеспечивать следующие задачи:
 - получение метаданных событий путем индексации файлов данных с этими событиями в распределенной системе хранения
 - передачу этой информации на сервер EventIndex
 - проверку полученных данных и загрузку в базу данных
 - доступ к информации через универсальный API
 - интерактивные и асинхронные интерфейсы пользователей
- **Данные индексируются на серверах где они расположены**
 - Программа индексирования запускается в PanDA по запросу EI
 - Либо она может быть запущена по окончании создания данных
- **Передача данных на сервер EventIndex осуществляется FTS**

- EventIndex использует микросервисную архитектуру
- Каждый компонент разрабатывается и устанавливается независимо
- Система состоит из трех основных блоков
 - Индексирование данных: извлечение метаданных событий
 - Управление и хранения данных: сбор данных, проверка и загрузка в БД
 - Клиентская часть: API, управление запросами и Web-интерфейс



- Для хранения и обработки данных была выбрана система управления базами данных (СУБД) PostgreSQL
 - Ожидаемый поток данных: до 150 тысяч событий в секунду
 - Исследованы различные методы оптимизации для ускорения процесса записи данных в БД
 - Использование асинхронной потоковой загрузки обеспечивает достаточную производительность
- Разработан программный интерфейс на основе RESTful API.
 - Графический Web клиент использует Angular framework
 - Для серверной части был выбран fastAPI: легкий асинхронный RESTful framework для Python
- Для асинхронной обработки задач использованы пакеты Redis и Celery
 - Celery управляет задачами выполняющимися в течение долгого времени
 - Redis служит брокером сообщений и временным хранилищем результатов

- Supervisor управляет сбором и загрузкой данных EventIndex
 - Отслеживает появление новых данных в системе хранения
 - Запускает в PanDA задачи индексирования метаданных событий
 - Управляет передачей данных на сервер EventIndex при помощи FTS
 - Запускает программы верификации и загрузки данных в БД
 - Контролирует процессы и собирает информацию для логов
- Проект SPD Event Index разрабатывается с использованием GitLab CI/CD (Непрерывная интеграция / развертывание).
 - CI/CD позволяет автоматизировать сборку, тестирование и развертывание, обеспечивая быструю установку обновлений.
- Собранные образы контейнеров хранятся в реестре Gitlab, что позволяет отслеживать версии и контролировать права доступа.

Разработка ведется З. Будтуевой, А. Газзаевым и Р. Гурциевым

- Сейчас в проекте SPD участвуют около 400 человек
 - число участников увеличится ближе к началу эксперимента
- Чтобы организовать эффективное сотрудничество с совместным использованием ресурсов, необходимо иметь ИС для
 - обработки данных персонала и организаций
 - поддержку рабочих групп: членство, права доступа
 - учет вклада (если будет реализовано)
 - формирование отчетов с разбивкой по различным параметрам
- Процедуры создания, утверждения и редактирования связанных документов
 - Регистрация и изменение членства в сотрудничестве
 - Создание и редактирование списков групп и привилегий
 - Включение в авторские списки

- SPD уже подготовил несколько публикаций и докладов
- Для подготовки и публикации результатов необходимо
 - Отслеживание прохождения публикаций
 - Обмен сообщениями между авторами, рецензентами и кураторами
 - Поиск по документам
 - Отслеживание публикаций SPD во внешних каталогах
 - Отчеты о количестве публикаций с разбивкой по авторам, темам...
- Организация презентаций и докладов на конференциях и совещаниях:
 - Составление списка конференций и доступных докладов
 - Организация вызова докладчиков и отбор
 - Прием, рецензирование и утверждение названий, тезисов и слайдов
 - Отслеживание публикации материалов

- Для авторизации сотрудников следует использовать службы аутентификации ОИЯИ
 - Предоставление возможности доступа внешним сотрудникам, у которых нет учетной записи ОИЯИ
 - Внедрение групповых учетных записей и учетных записей роботов для использования в автоматизированных задачах
 - Следует внедрить систему управления ролями и привилегиями, а также политики группового доступа
 - Процедуры включения пользователей в группы и отзыва членства
 - Реализуется ИТ-службами ОИЯИ.
 - Информация для должна быть предоставлена коллаборацией
- Ресурсы рабочих групп должны использовать единую систему аутентификации и авторизации



Welcome to **SPD**

Sign in with your SPD credentials

Sign in

[Forgot your password?](#)

Or sign in with

Your X.509 certificate

JINR SSO

Your institutional account

Not a member?

Apply for an account

Users Users

Search.. Show all

Pic	Name ^	Active	E-mail	Created	Groups	Actions
	Admin User	●	admin@iam.test	5 days ago		
	Aleksandr Vladimirovich	●	baranov@jinr.ru	2 days ago	spd.nica.jinr/VO-Admin spd.nica.jinr	
	Alexey Konak	●	konak@jinr.ru	2 days ago	spd.nica.jinr/production spd.nica.jinr	
	Alexey Zhemchugov	●	zhemchugov@jinr.ru	2 days ago	spd.nica.jinr spd.nica.jinr/VO-Admin	
	Andrey Kiryanov	●	Kiryanov_AK@pnpi.nrcki.ru	2 days ago	spd.nica.jinr spd.nica.jinr/production	
	Andrey Zarochentsev	●	andrey.zar@gmail.com	2 days ago	spd.nica.jinr	
	Artem Ivanov	●	arivanov@jinr.ru	2 days ago	spd.nica.jinr	
	Artem Petrosyan	●	artem.petrosyan@jinr.ru	2 days ago	spd.nica.jinr/production spd.nica.jinr spd.nica.jinr/pilot spd.nica.jinr/VO-Admin	
	Danila Oleynik	●	danila@jinr.ru	2 days ago	spd.nica.jinr spd.nica.jinr/production spd.nica.jinr/VO-Admin	
	Dzmitry Yermak	●	dmierk@hep.by	2 days ago	spd.nica.jinr	

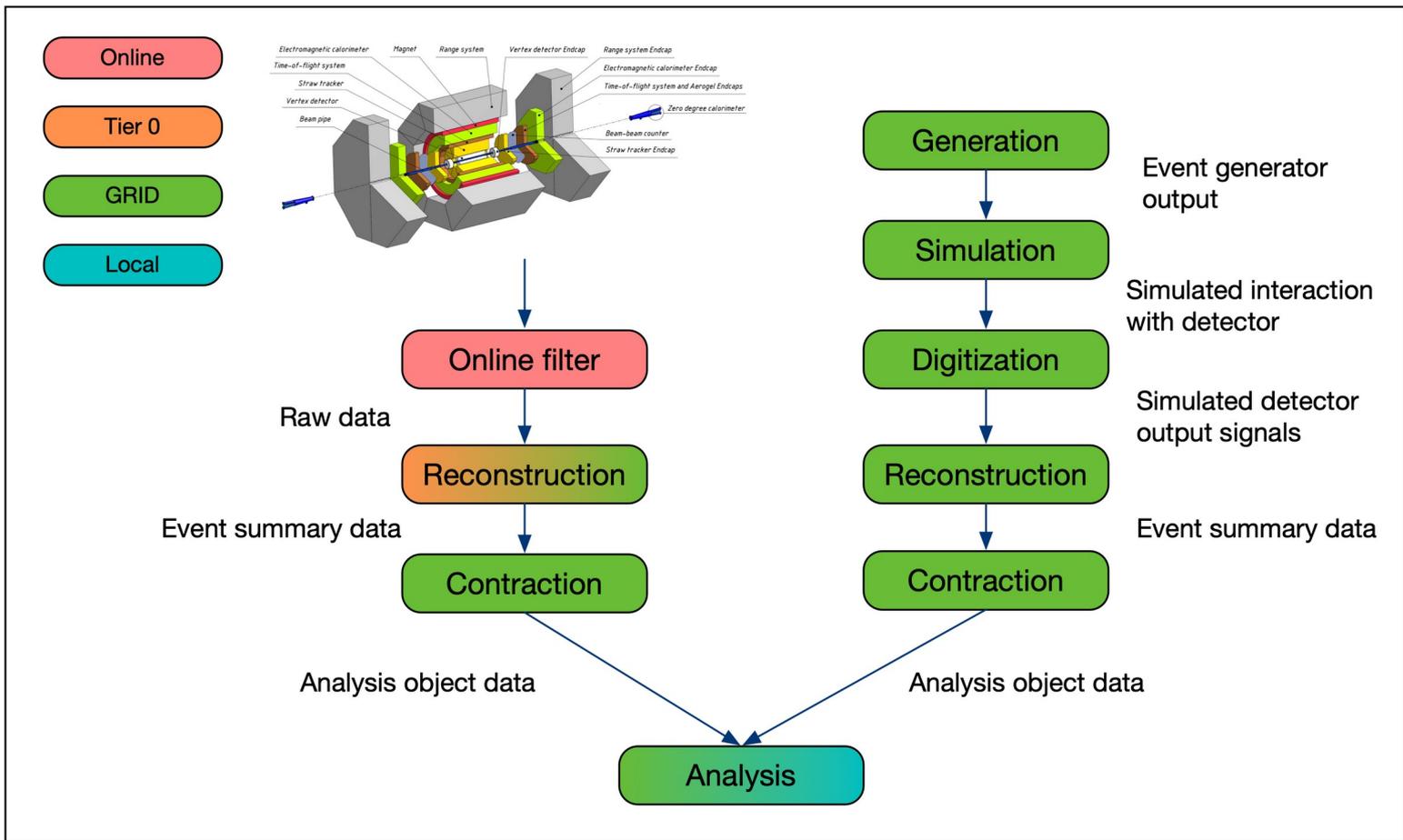
1 2

[+ Add User](#)

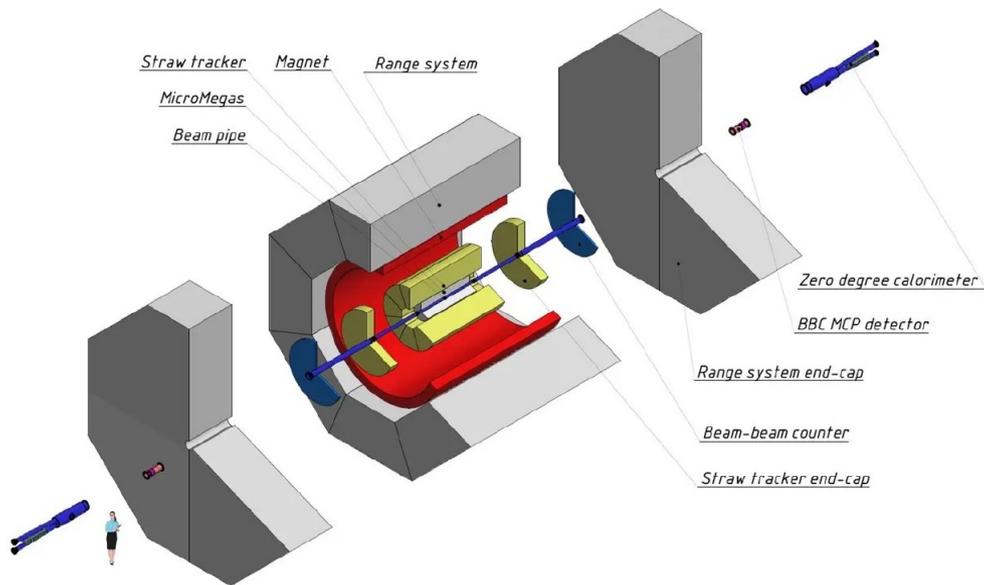
1g OAuth 2.0

- Разработка информационных систем установки SPD продолжается
- В ней участвуют специалисты лабораторий ОИЯИ а также сотрудники организаций — участников эксперимента
- Информационные системы разрабатываются под требования эксперимента, исходя из объема и содержания данных, с учетом опыта накопленного в предыдущих проектах
- Разработка в создании информационных систем тесно связан с прогрессом в создании компонент установки и систем обработки и анализа данных.

BACKUP



- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Nisl tincidunt eget nullam non.
- Quis hendrerit dolor magna eget est lorem ipsum dolor sit:
 - Volutpat odio facilisis mauris sit amet massa.
 - Commodo odio aenean sed adipiscing diam donec adipiscing tristique
 - Mi eget mauris pharetra et. Non tellus orci ac auctor augue. Elit at imperdiet dui accumsan sit
- Ornare arcu dui vivamus arcu felis. Egestas integer eget aliquet nibh praesent. In hac habitasse platea dictumst quisque sagittis purus. Pulvinar elementum integer enim neque volutpat a

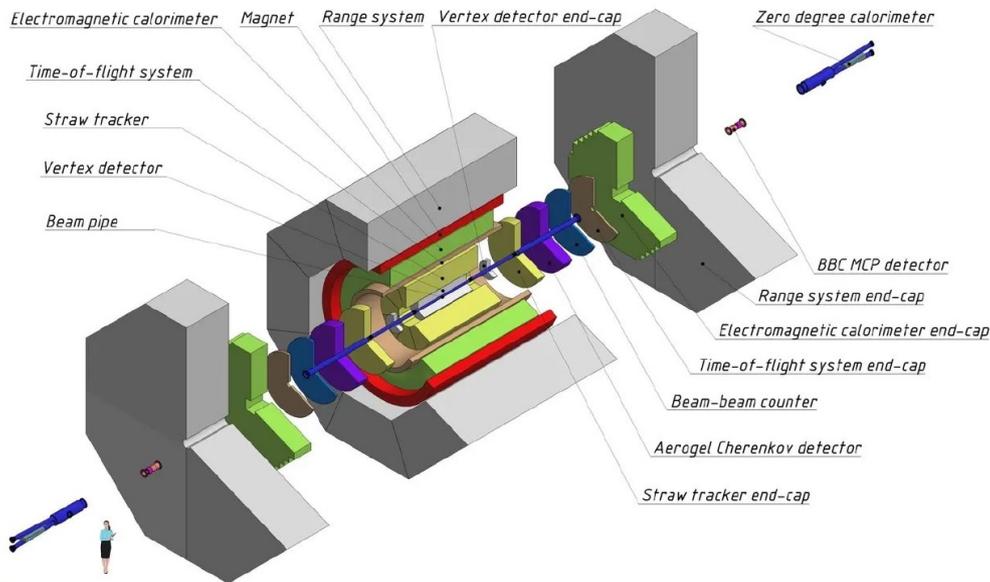


- Подготовка эксперимента
- Монте-Карло моделирование: 2 ПБ / год. Всего 10 ПБ
- 1500-3000 вычислительных задач одновременно 24/7

- **Этап I: работа на низкой светимости NICA (3 года)**

- Монте-Карло моделирование + набор и обработка данных: 4 ПБ / год
- Повторная обработка: 2 ПБ/год
- Всего за этап — 18 ПБ, порядка 1 триллиона событий, 5000-6000 задач

Установка SPD как источник данных



- Модернизация для работы при высокой светимости
- Монте-Карло : 2 ПБ / год.
Повторная обработка: 2 ПБ/год.
Всего 8 ПБ
- 1500-3000 задач

- **Этап II: работа на высокой светимости коллайдера (3 года)**
 - Монте-Карло моделирование + набор и обработка данных: 20 ПБ / год
 - Повторная обработка: 10 ПБ/год
 - Всего за этап — 120 ПБ, порядка 5 триллионов событий, до 60000 задач