

Вычислительная инфраструктура «Супер чарм-тау фабрики»

Виталий Воробьев

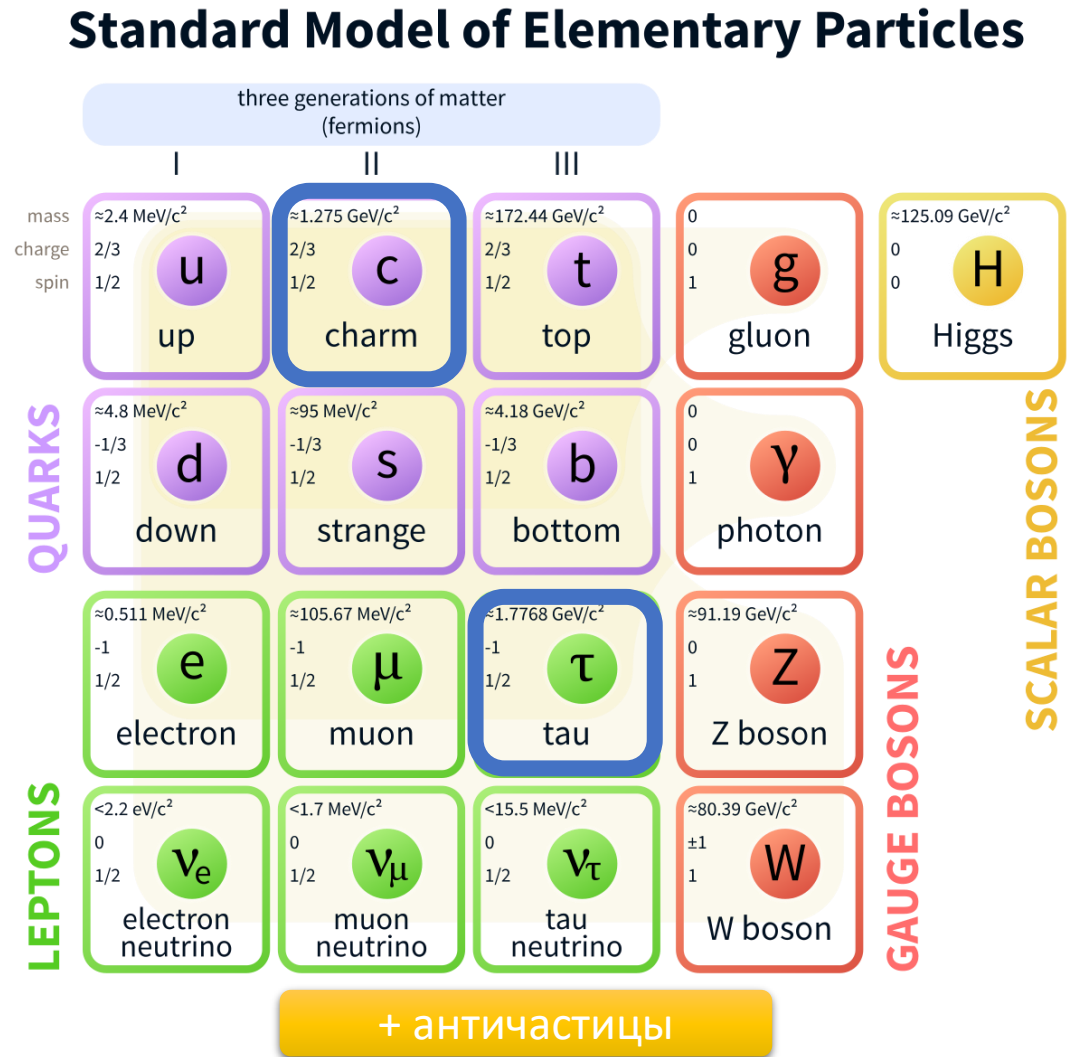
ИЯФ СО РАН



Семинар в ЦКП ССКЦ ИВМиМГ СО РАН

29 октября, 2020

Стандартная модель элементарных частиц



- Компоненты стандартной модели (СМ)
 - Теория сильного и электрослабого взаимодействий
 - Набор фундаментальных полей (частиц)
- Проявляется на характерных масштабах
 - $l < 10^{-15} \text{ м}$
 - $\tau \sim 10^{-24} \div 10^{-8} \text{ с}$
 - $E > 1 \text{ ГэВ}$
- Самая успешная физическая теория в истории
 - Согласуется с огромным количеством наблюдений

Открытые вопросы

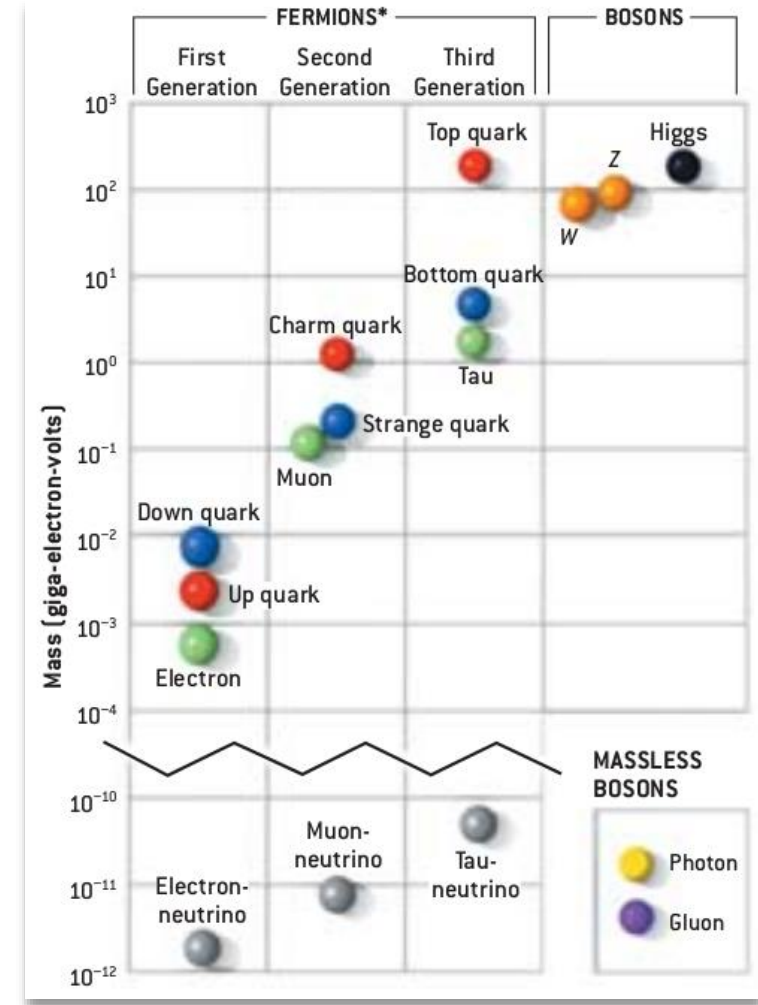
- Проблема иерархии масс
- «Тонкая настройка Вселенной»
- Почему электрический заряд квантован?
- Почему кварки имеют электрический заряд $+2/3$ и $-1/3$?
- Почему есть три поколения лептонов и кварков?
- Почему вокруг нас почти нет антиматерии?
- Почему только слабое взаимодействие различает материю и антиматерию?
- Существует ли теория, которая единообразно описывает все взаимодействия, включая гравитацию?
- Что такое темная материя?
- ...

$$V = -\mu^2 H^\dagger H + \lambda (H^\dagger H)^2$$

$$M_H^2 = 3.273, 459, 429, 634, 290, 543, 867, 496, 473, 159, 645 - 3.273, 459, 429, 634, 290, 543, 867, 496, 473, 159, 643$$

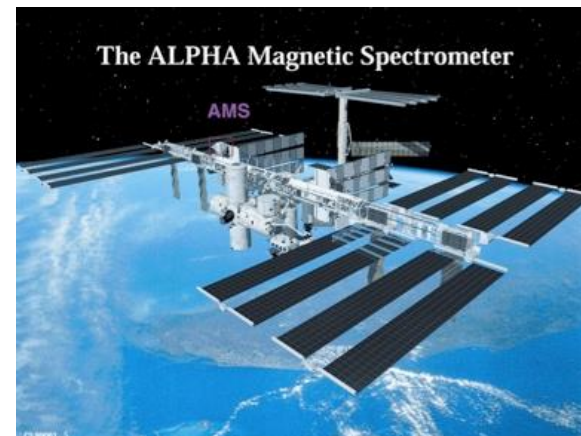
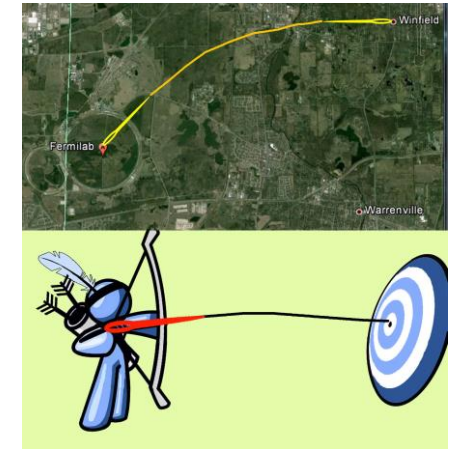
$$M_H^2 = M_{\text{tree}}^2 + \left(\text{loop with } H \right) + \left(\text{loop with } t \right) + \left(\text{loop with } W, Z \right)$$

СМ – низкоэнергетический предел более общей теории?



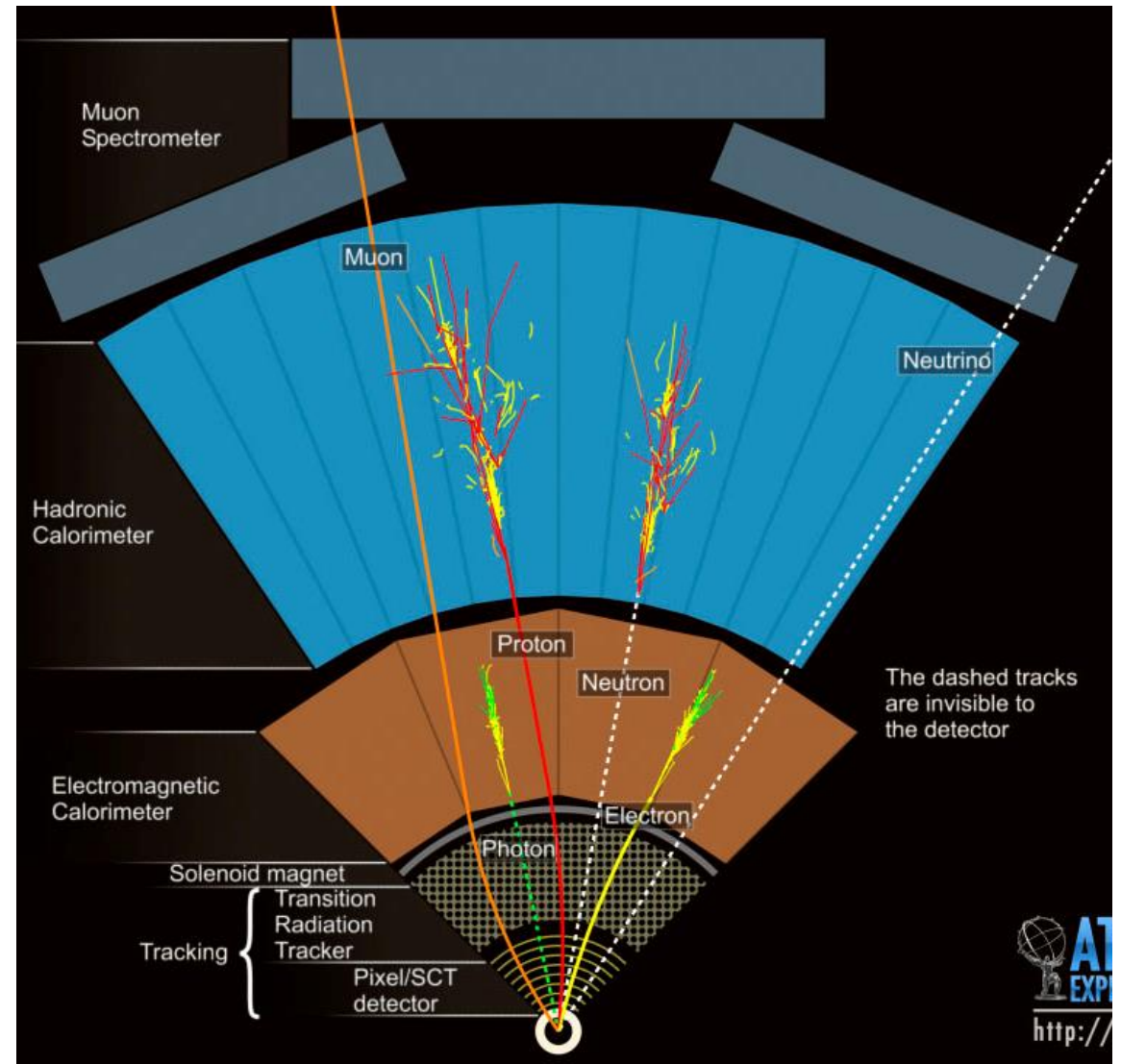
Эксперименты в физике частиц

- Общий принцип: изучаем рождение и распад тяжелых частиц в столкновениях легких
- Чем больше энергия сталкивающихся частиц, тем больше типов частиц доступно для прямого рождения
- Изучение редких процессов и прецизионные измерения требуют большой статистики
- Energy frontier vs. precision frontier
- Основные типы экспериментов
 - Эксперименты с фиксированной мишенью
 - Коллайдеры
 - Космические детекторы
 - Изучение атмосферных ливней



Регистрация частиц

- Взаимодействие частиц с веществом
 - Ионизация
 - Черенковское излучение
 - Электромагнитные и адронные ливни
 - ...
- Подсистемы детекторов
 - Соленоид, создающий магнитное поле
 - Трековая система (координата и импульс)
 - Калориметр (энергия)
 - Система идентификации заряженных частиц
 - Мюонная система

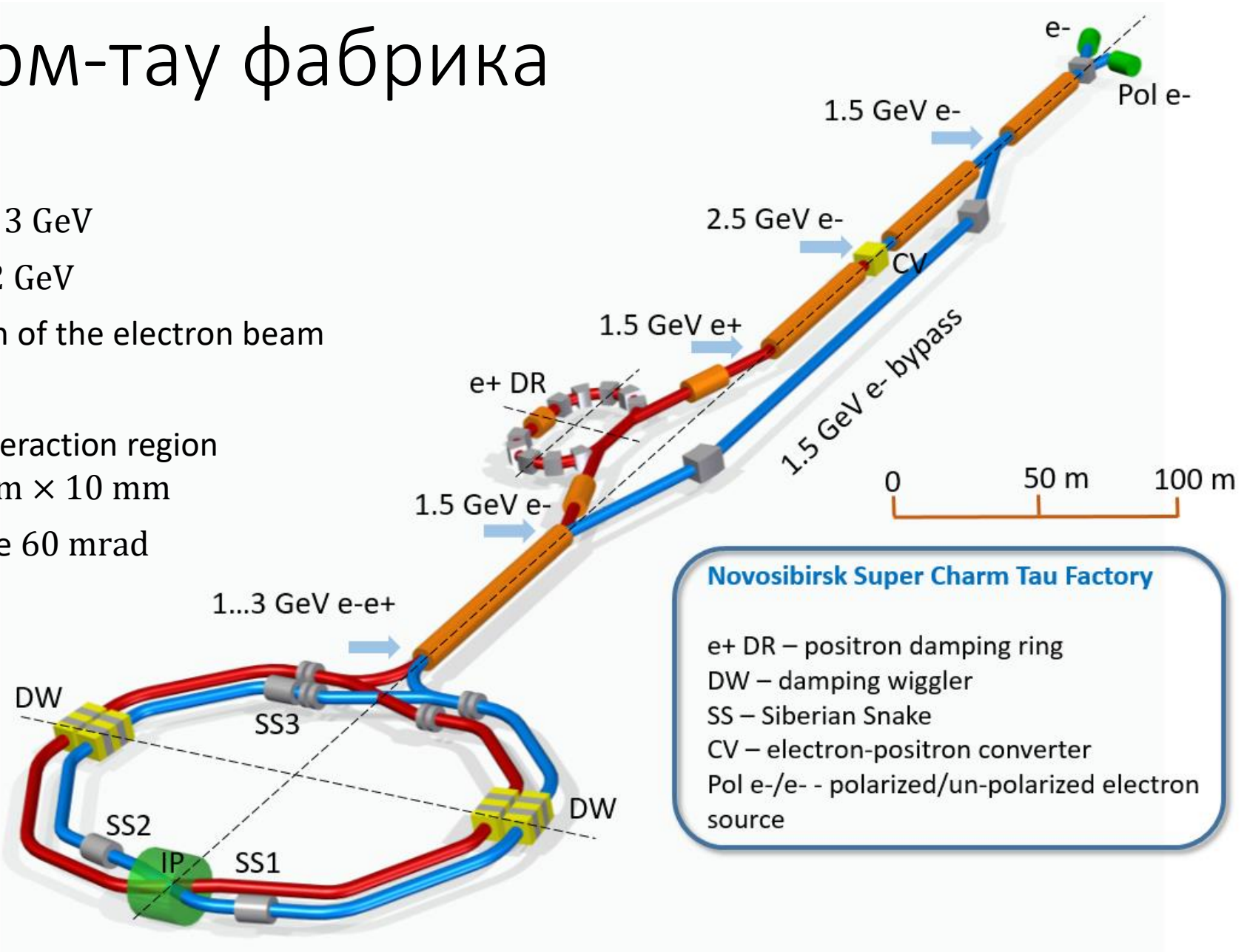




Действующие коллайдеры

Супер чарм-тау фабрика

- Beam energy: from 1 to 3 GeV
- $\mathcal{L} = 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ @ 2 GeV
- Longitudinal polarization of the electron beam
- Crab-waist collisions
 - Beam size in the interaction region
 $20 \mu\text{m} \times 0.2 \mu\text{m} \times 10 \text{mm}$
 - Beam crossing angle 60 mrad



ctd.inp.nsk.su

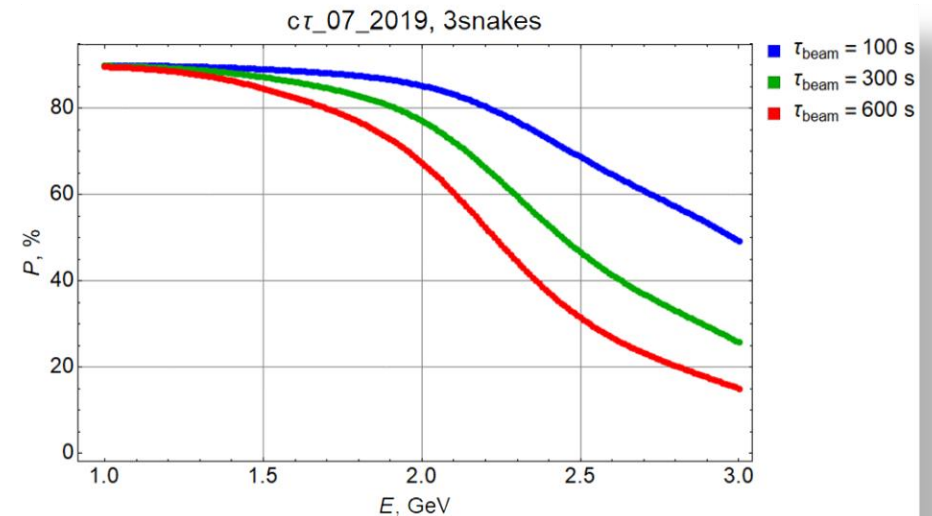
Параметры коллайдера

Circumference	478.092 m				
2θ	60 mrad				
β_x^*/β_y^*	50 mm / 0.5 mm				
F_{RF}	349.9 MHz				
E_{beam} (GeV)	1*	1	1.5	2	3
I (A)	1	1	2.2	2.2	2
N_{bunch}	500	500	490	420	290
ε_x (nm)	11.3	16.3	8.8	7	10.9
L_{peak} ($\text{cm}^{-2}\text{s}^{-1} \times 10^{35}$)	0.21	0.14	0.8	1.3	1.1

* With two $B_w = 3.5$ T wigglers that suppress intrabeam scattering

More details can be found here: <https://c-tau.ru/indico/event/3/contributions/200/>

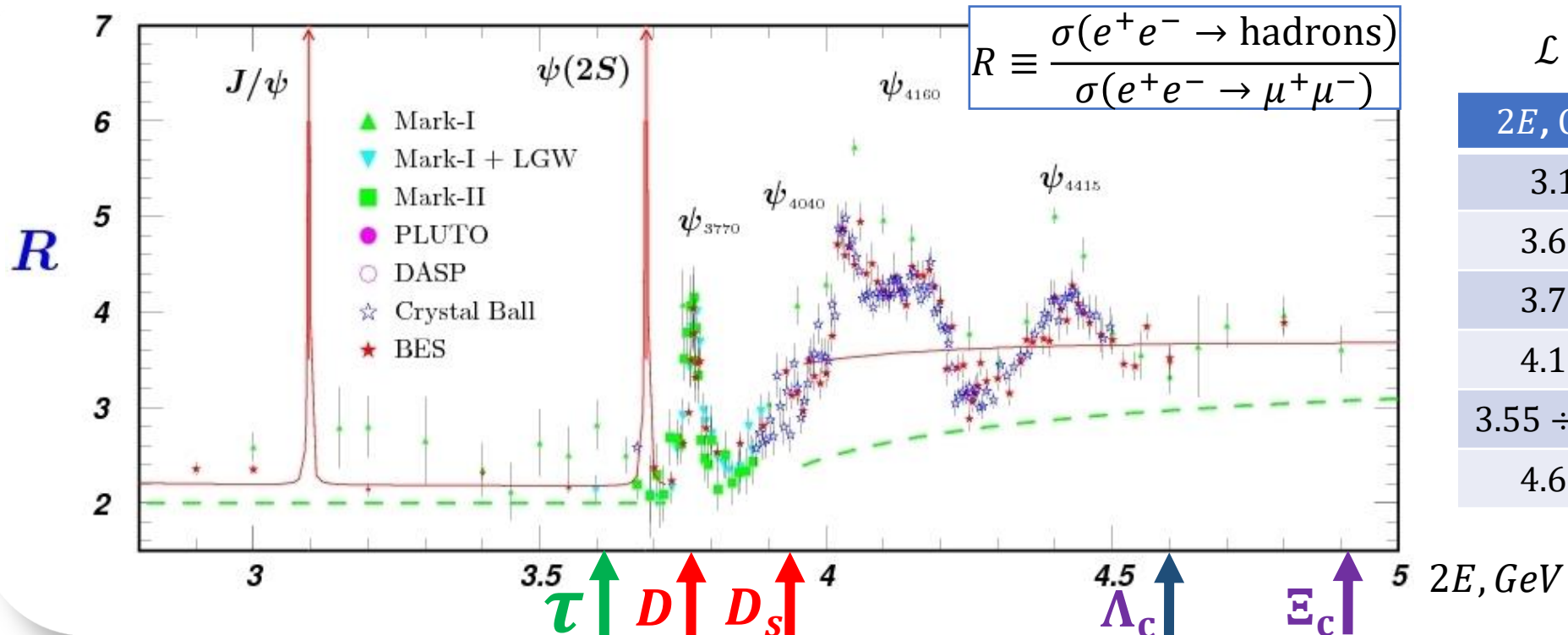
Electron beam polarization with 3 Siberian Snakes



Что производит наша фабрика

\sqrt{s} from 2 GeV to 6 GeV

The production of non-relativistic particles near the threshold makes it possible to study them in detail



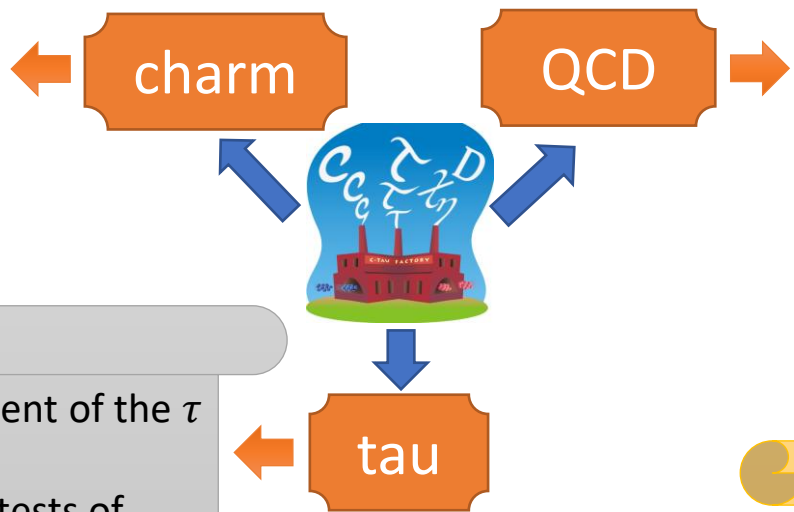
$$\mathcal{L} = 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$$

$2E, \text{ GeV}$	Annual yield
3.1	$10^{12} J/\psi$
3.69	$10^{11} \psi(2S)$
3.77	$10^9 D\bar{D}$
4.17	$10^8 D_s\bar{D}_s$
3.55 ÷ 4.3	$10^{10} \tau\tau$
4.65	$10^8 \Lambda_c^+\Lambda_c^-$

Физическая программа

- ✓ Measurement of the strong phases of D decay amplitudes
- ✓ Measurement of absolute branching fractions
- ✓ Searches for rare and forbidden decays of the charm quark
- ✓ CP violation in charm
- ✓ ...

Input for B meson studies at LHCb and Belle II



- ✓ Physics of highly-excited quarkonium
- ✓ Molecular states
- ✓ Baryon interaction at threshold
- ✓ Search for glueballs in decays of J/ψ and $\psi(2S)$
- ✓ ...

Test of the electroweak sector of the SM

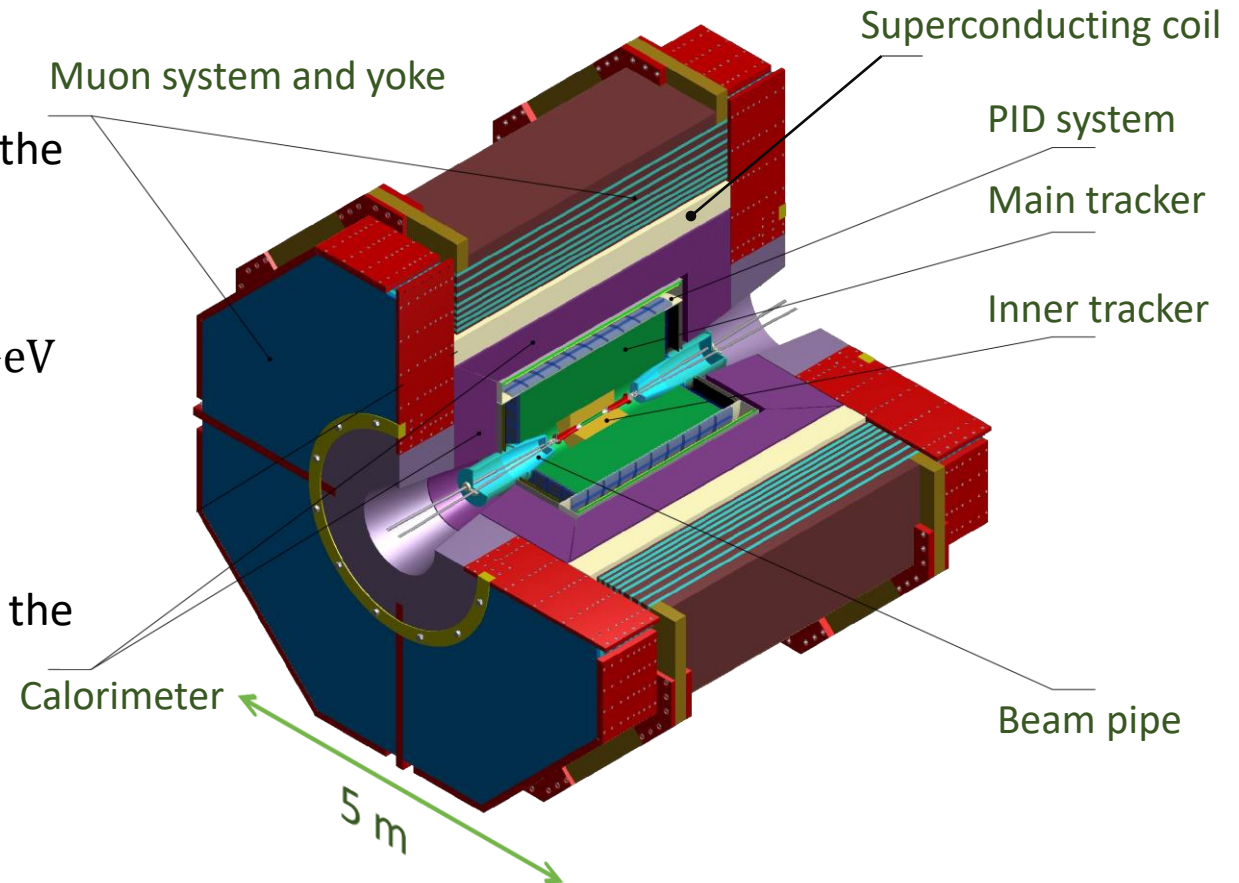
- ✓ Precision measurement of the τ lepton properties
- ✓ Michel parameters, tests of lepton universality
- ✓ Precision measurement of hadronic τ decays
- ✓ Search for CP and T violation in τ decays
- ✓ ...

QCD, α_s , V_{us} . Test of the electroweak model, searches for non-standard contributions

Концепция детектора

- Physics requirements

- **Momentum resolution:** $\approx 0.3\%$
- **CP symmetry and hermeticity** $\approx 95\%$ of the full solid angle
- **Soft track** detection with $p_t \gtrsim 50$ MeV
- Excellent **$\mu/\pi/K/p$ separation** up to 1.5 GeV
 - dE/dx in tracking system
 - Cherenkov light detector for PID
 - The importance of μ/π separation
- Good **π^0/γ separation** and **γ detection** in the energy range from 10 MeV to 3000 MeV
 - Good energy resolution in calorimeter
 - Fast calorimeter ($\sigma_t < 1$ ns)
- **DAQ rate** ~ 300 kHz @ J/ψ peak

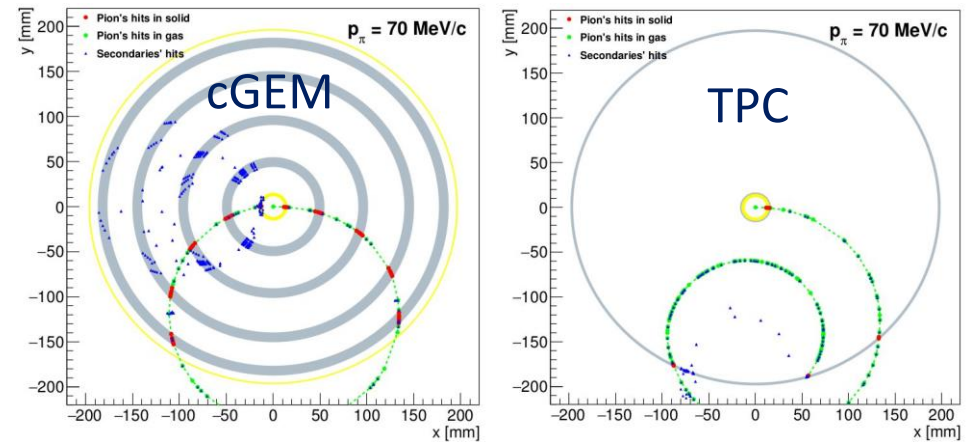


Трековая система

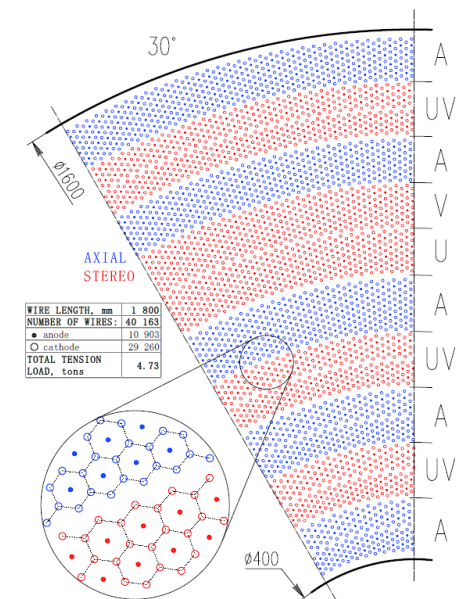
- Восстановление треков заряженных частиц по ионизации в газе
- Измерение импульса
- Информация о типе частицы по величине ионизации

12

Внутренний трекер (две опции)

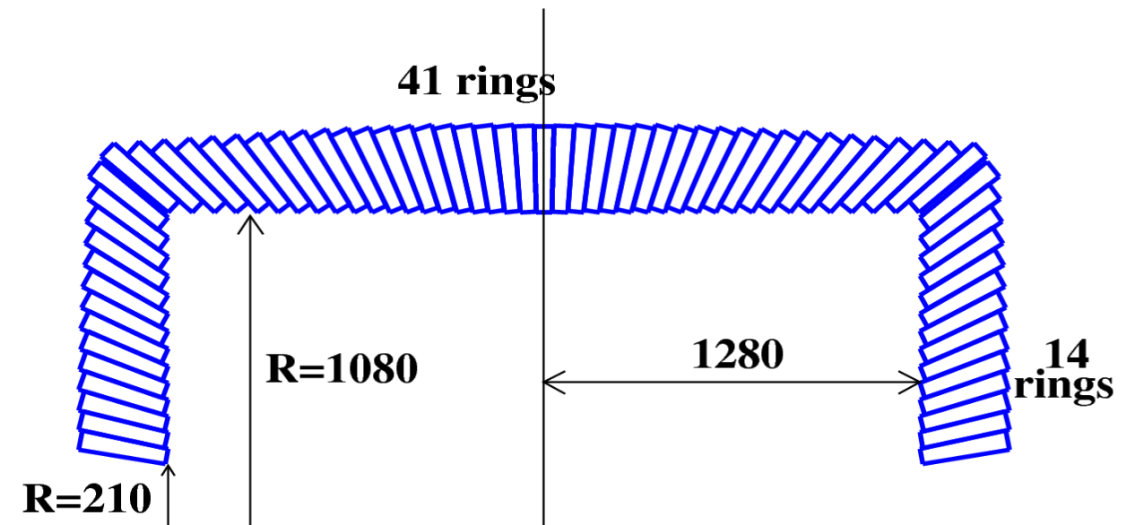
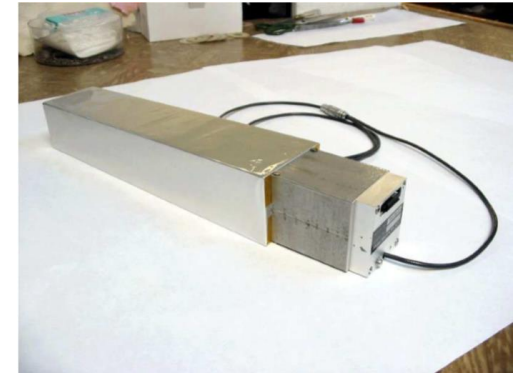
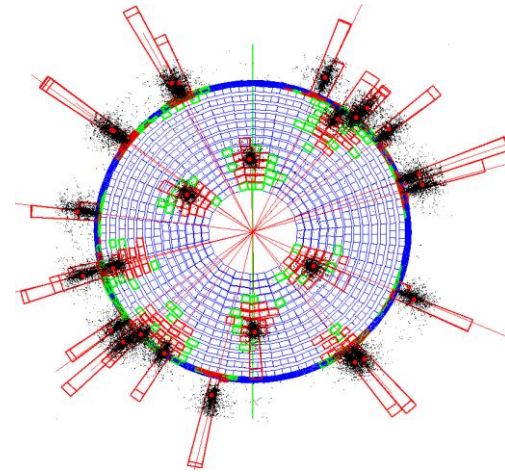


Центральный трекер: дрейфовая камера



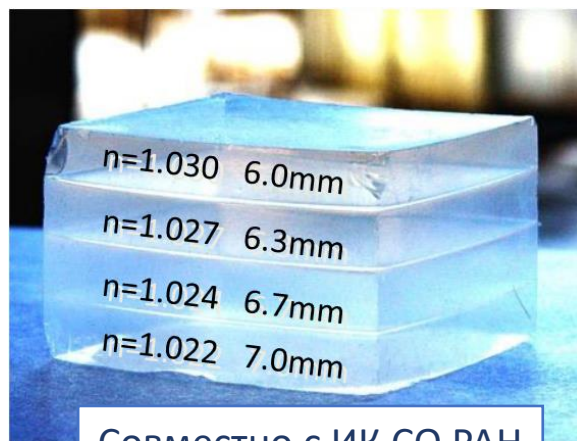
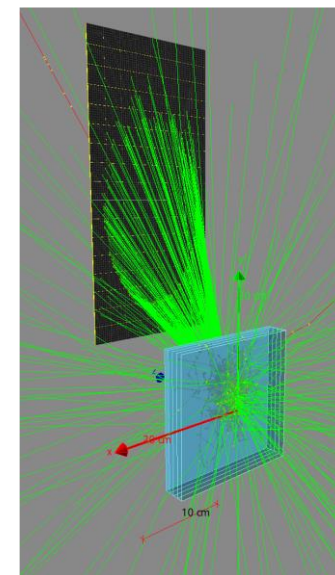
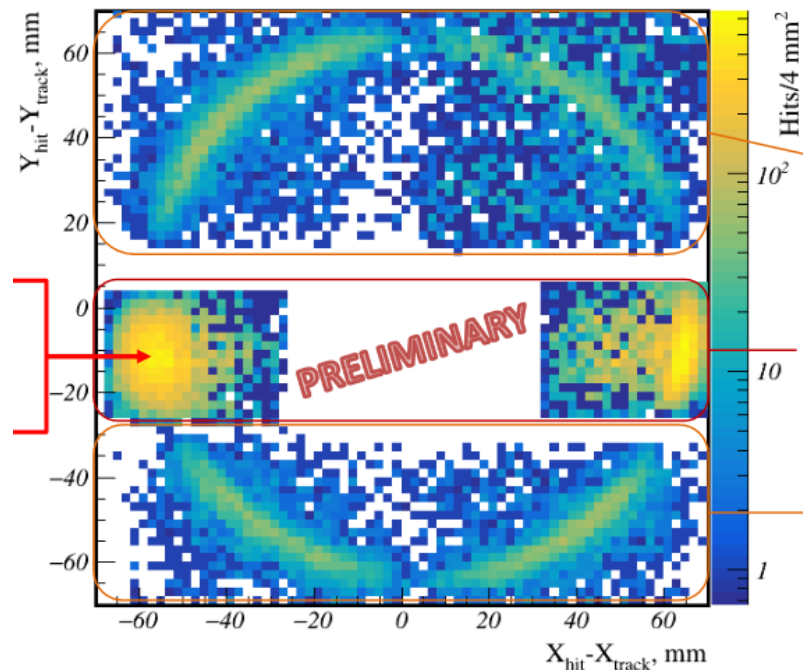
Электромагнитный калориметр

- Кристаллы CsI в форме усеченной пирамиды, размер $5 \times 5 \times 30$ см
 - 5248 штук в цилиндрической части
 - 2176 штук в боковых частях
- Вес 36 тонн
- Самая дорогая часть детектора
 - Около 50 М\$, включая электронику

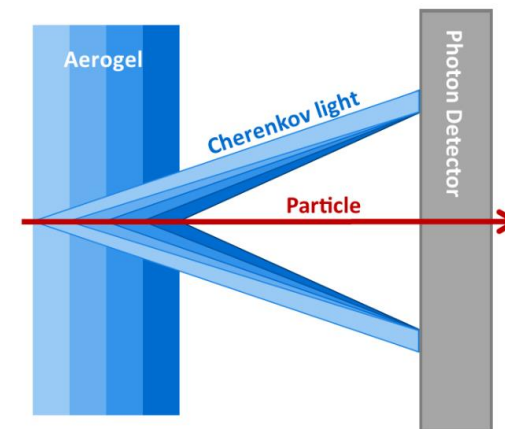


Система идентификации

- ФАРИЧ: четырехслойный фокусирующий детектор черенковского излучения
- Считывание с помощью матрицы пикселей
 - $O(10^6)$ каналов электроники
- Технология позволяет получить рекордные характеристики идентификации заряженных частиц
- Система сложна в реализации
 - Производство качественного аэрогеля в промышленных масштабах
 - Большое количество каналов
 - Большая частота шумовых срабатываний
- Альтернативная опция: DIRC
 - Поддерживается университетом Гиссена



Совместно с ИК СО РАН

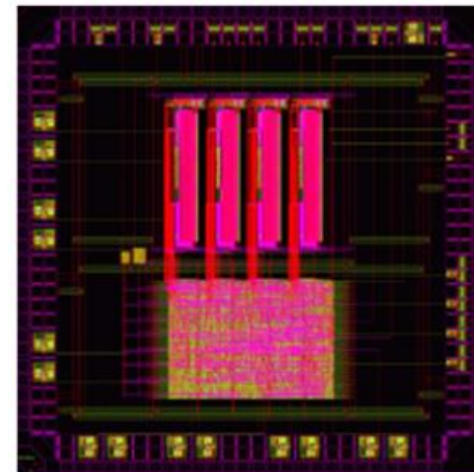


Subdetector parameters

Detector	Channel quantity	Channel size	Power consumption	Trigger participation	Type of information
VD – GEM – Timepix	56 K 10 M	2 x 2 mm ² 150 x 150 μm ²	1 kW	Possibly	Amplitude (A), Time (T), Position (P)
DC	7.1 K	12 x 20 mm ²	2 kW	Yes	A, T, P
FARICH	1 M	4.5 x 4.5 x 50 mm ³	100 kW (21 m ²)	No	T, P
Calorimeter	7.5 K	(40 -70) ² mm ²	2 kW	Yes	A, T, P
Mu	4 - 44 K	∅ 20 mm	1 – 9 kW	Yes	T, P

- Analog-to-digit conversions **inside** the detector
- Data transfer over the **optical links**
- Wide using of **ASIC** (application-specific integrated circuit)
- **Water** cooling of the Front-End electronics

ATLAS LAr FEB SCA



слайд
Алексея Талышева

Parameters of signals and FrontEnd electronics

Detector	Channel load	Duration of signals	Sampling rate	Bit quantity of ADC (TDC)	Time measure precision
VD	33 kHz	Rising edge 25 ns	from 20 MHz to 80 MHz	10 -12 bit ADC	1 ns
DC	50 kHz	Rising edge 100-200 ns, Falling edge – 300 ns	50 MHz (ordinary mode) 500-1000 MHz (cluster mode)	10 bit ADC	1 ns
FARICH	1 MHz	5 – 10 ns	TDC	6 bits (for 8 ns) TDC	200 ps
Calorimeter	15 kHz	30 ns for clear CsI	40 – 50 MHz	18 bit ADC	1 ns
Mu	100 kHz	30 – 200 ns	TDC	11-12 bit TDC	60 ps

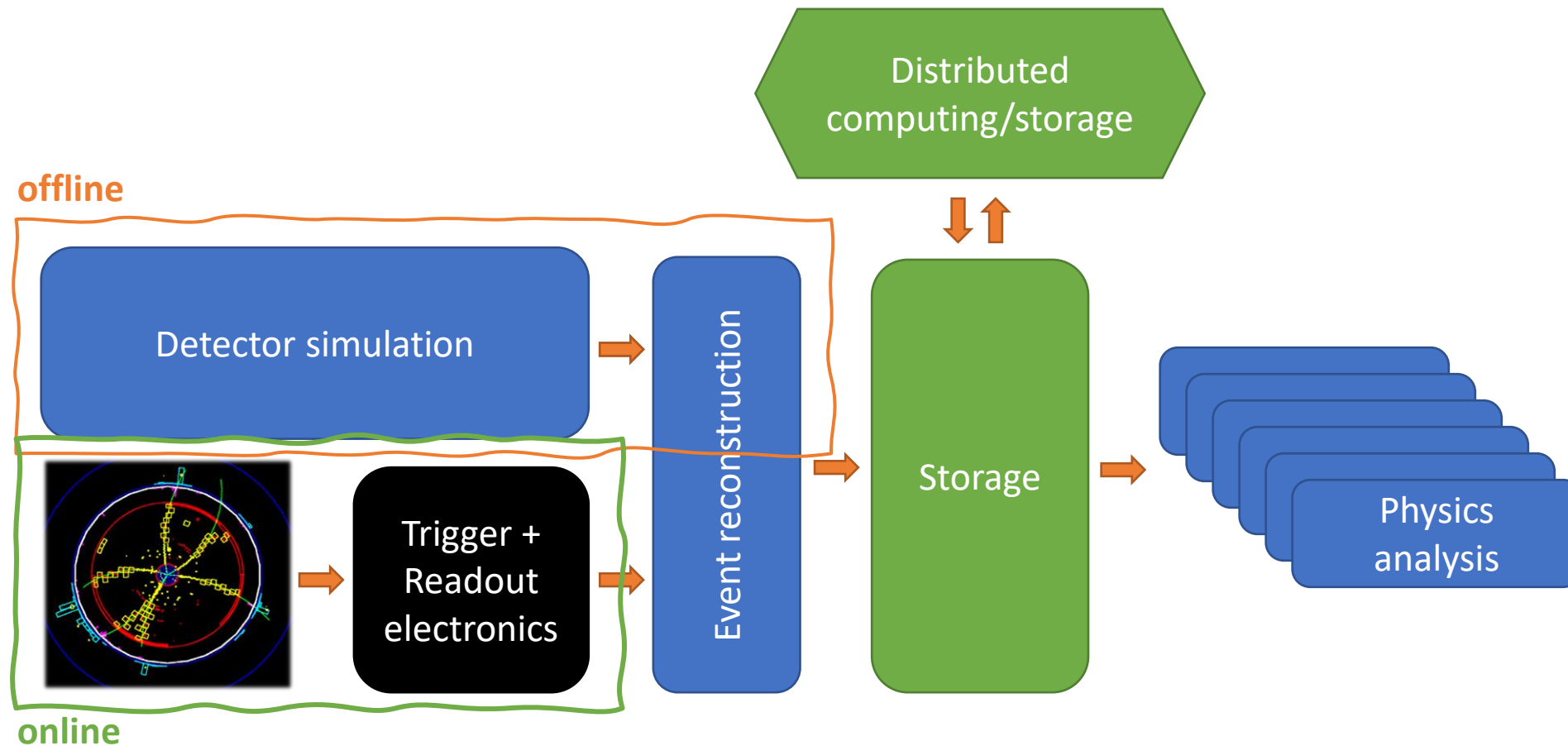
Rates in J/ψ at luminosity $10^{35} \text{ sm}^{-2}\text{s}^{-1}$:

- events 260 kHz
- Bhabha 90 kHz
- cosmic 2 kHz
- Background - several tens of kHz

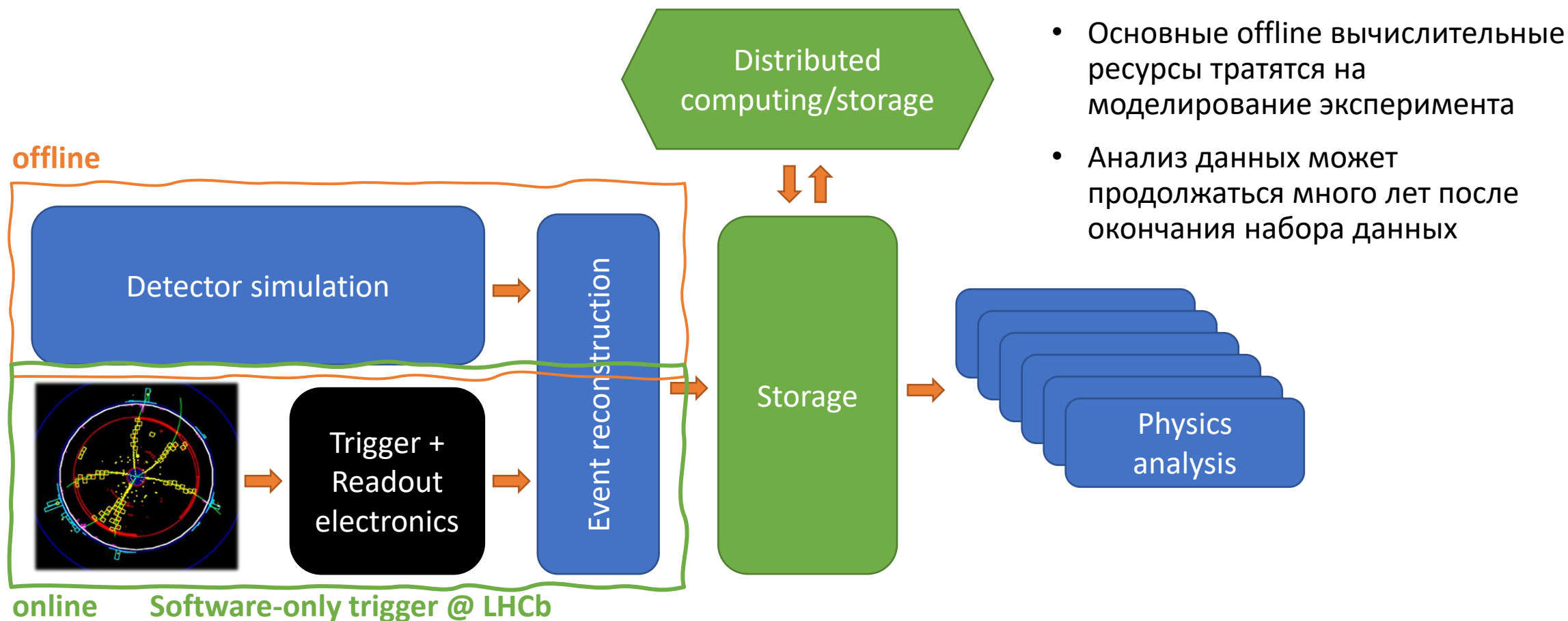
- The maximum readout rate up to **400 kHz**
- The event size **30 KB**

слайд
Алексея Талышева

Поток данных в коллайдерных экспериментах



Поток данных в коллайдерных экспериментах



- Основные offline вычислительные ресурсы тратятся на моделирование эксперимента
- Анализ данных может продолжаться много лет после окончания набора данных

Загрузка и объемы данных

Основные процессы			
	J/ψ	$\psi(2S)$	$\psi(3770)$
E (МэВ)	3097	3686	3770
σ (нбн)	1400	370	≈ 6
f (кГц)	110	34	0.6
Частота фоновых процессов (кГц)			
Космика	≈ 2		
Адронные события	19	17	16
Vhabha	90	80	80

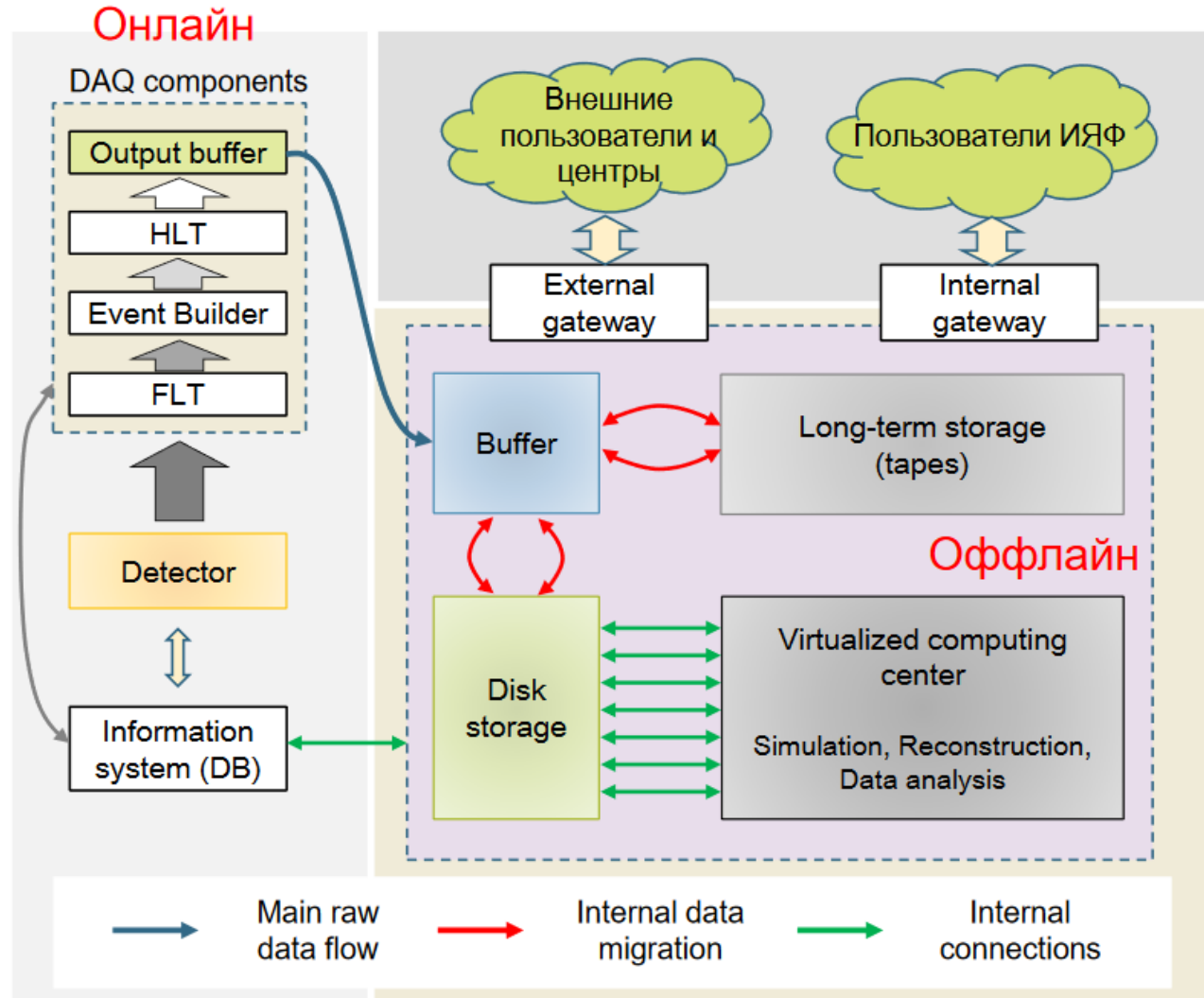


Максимальная частота
триггера 300 кГц

	BESIII	SCT	Belle II
Интеграл светимости (1/аб)	≈ 0.02	10	50
Число событий (10^{10})	≈ 4	200	10
Размер события (кБ)	12	50	300
Объем сырых данных (ПБ)		100	200
Объем обработанных данных (ПБ)		10	80

- Характерные параметры системы хранения и вычислительного кластера понятны
- Подобные системы существуют

Общая структура



Основные параметры

Максимальный поток данных на входе:

10 ГБ/с

Общий объем системы хранения

240 ПБ

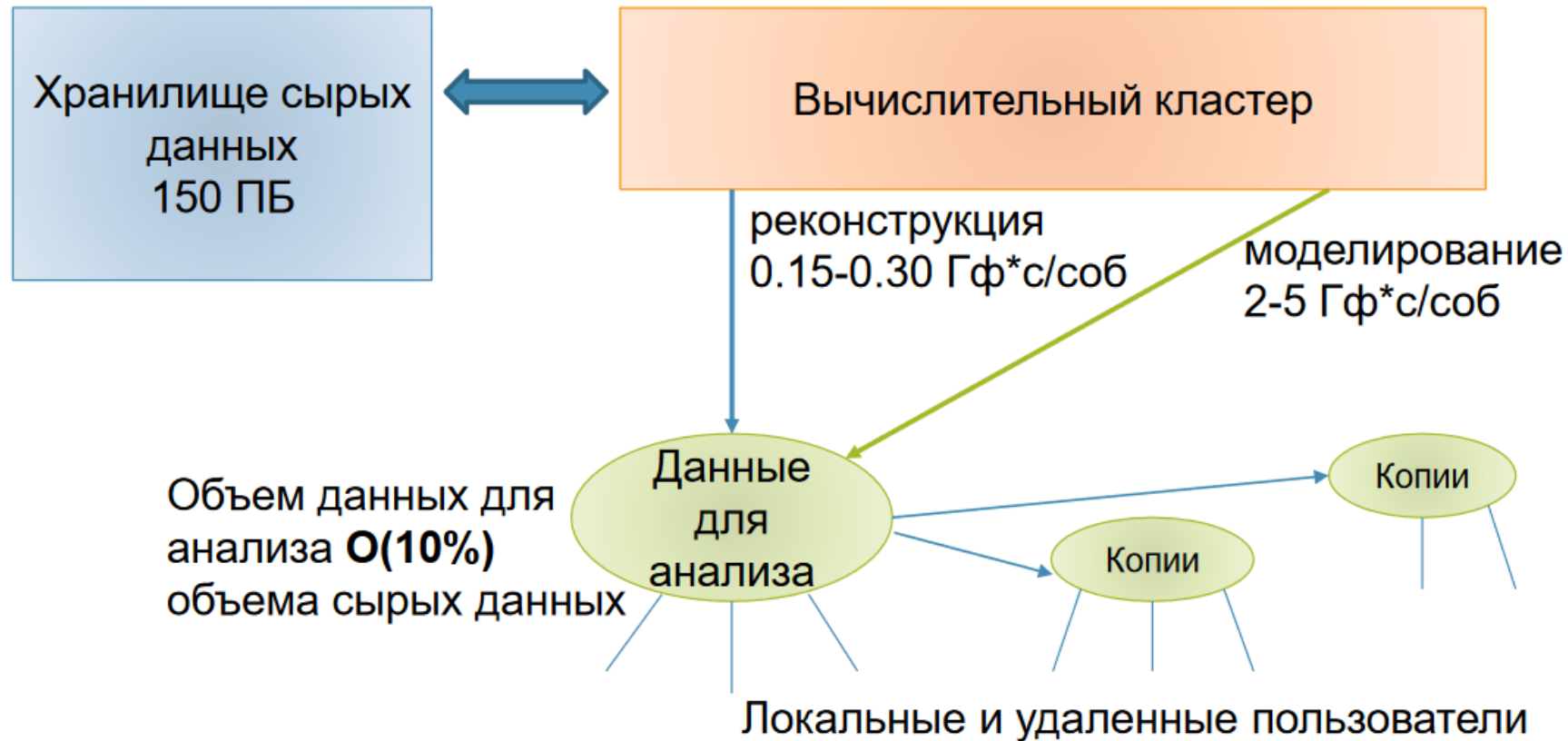
Вычислительная мощность

**0.6 Пфлопс
(1 МВт)**

Основные компоненты локальной вычислительной инфраструктуры

- **Вычислительный кластер «онлайн»**
 - Триггер высокого уровня
 - Высокая коннеktivность 100-1000 Гб/с
 - Возможно использование специализированных архитектур
- **Вычислительный кластер «оффлайн»**
 - Универсальная архитектура
 - Виртуализация
 - Вычислительная мощность ~0.6 Пфлопс
 - Тесная интеграция с системой хранения
- **Система хранения на 240 ПБ**
 - Включает дисковую и ленточную подсистемы

Типичный сценарий



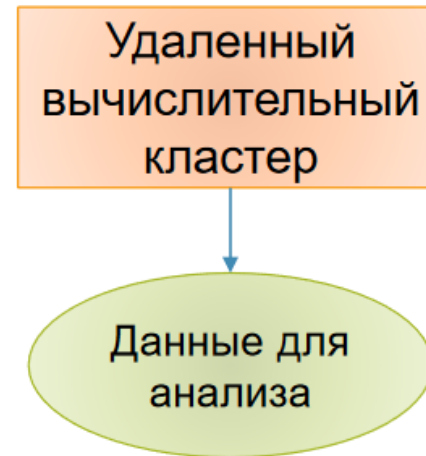
- Основной объем занимают сырые данные
- Основные вычислительные ресурсы тратятся на моделирование данных.
Использование внешних ресурсов для моделирования позволит снизить требования к локальным ресурсам

Сценарии использования географически удаленных компьютерных центров

Анализ данных удаленными группами



Моделирование данных на удаленных ресурсах



Создание удаленной копии сырых данных



коннеktivность 1-10 Гб/с

коннеktivность 40-100 Гб/с

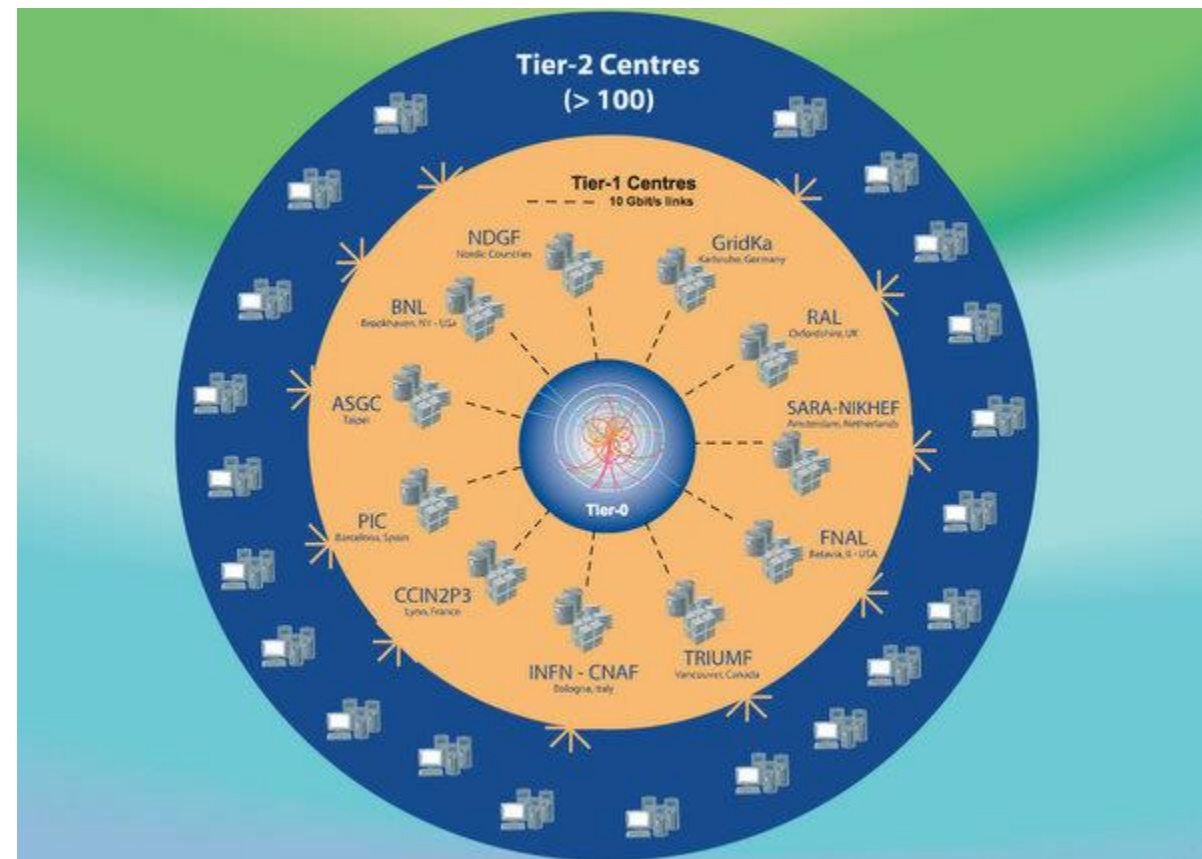
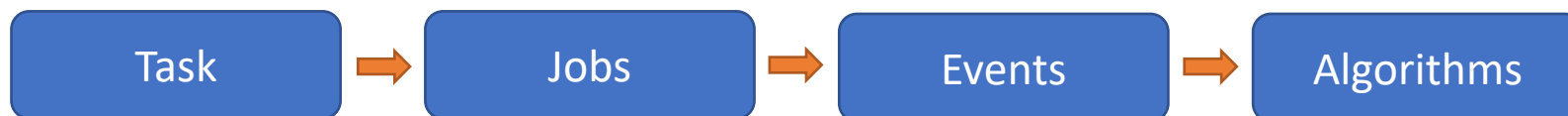
Заключение

- Для проведения экспериментов на Супер $s - \tau$ фабрике необходимо создать локальную вычислительную инфраструктуру, вычислительной мощностью 0.6 Пфлопс и способной хранить 240 ПБ.
 - В РФ есть центры такой производительности
 - Сейчас в ИЯФ есть $\sim 1\%$ вычислительной мощности
- Наиболее вычислительно емкая задача моделирования эксперимента может эффективно решаться на удаленных компьютерных ресурсах
- Создание второго центра хранения данных значительно повысило бы отказоустойчивость системы; для создания такого центра требуется организация выделенных линий связи с пропускной способностью 40-100 Гб/с

LHC Computing

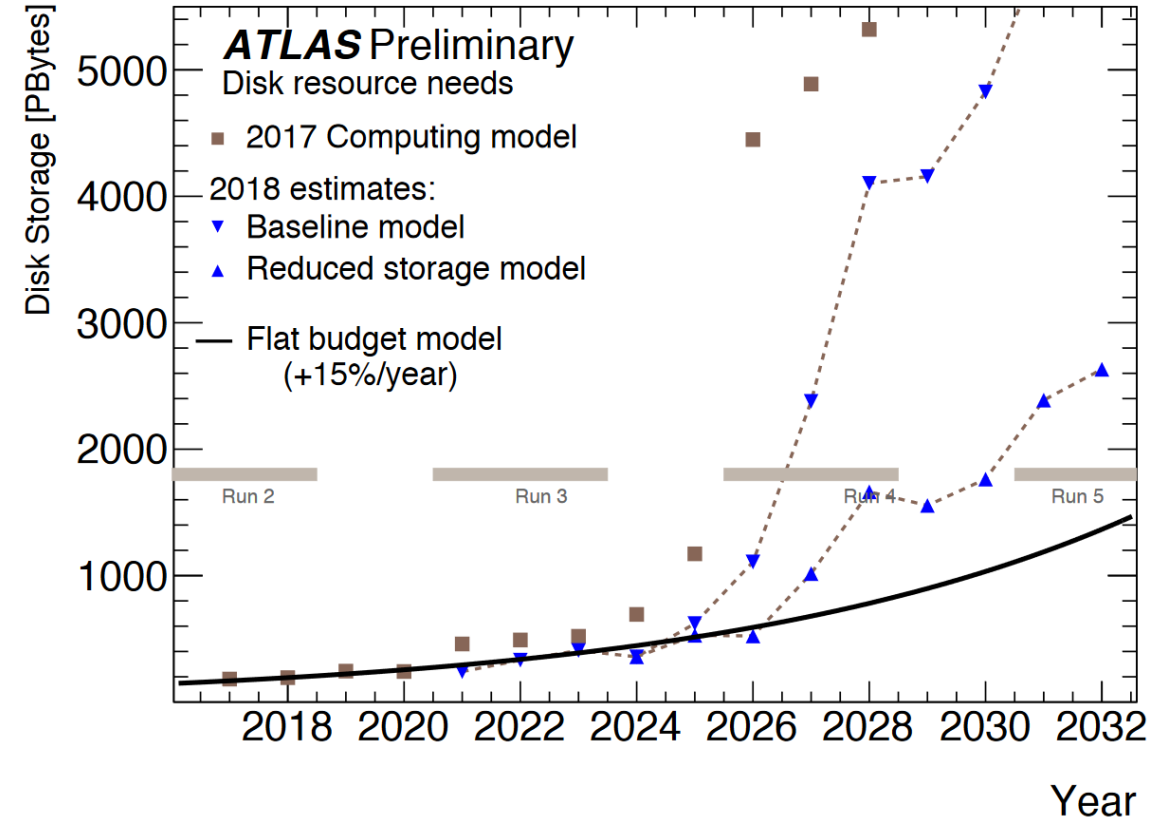
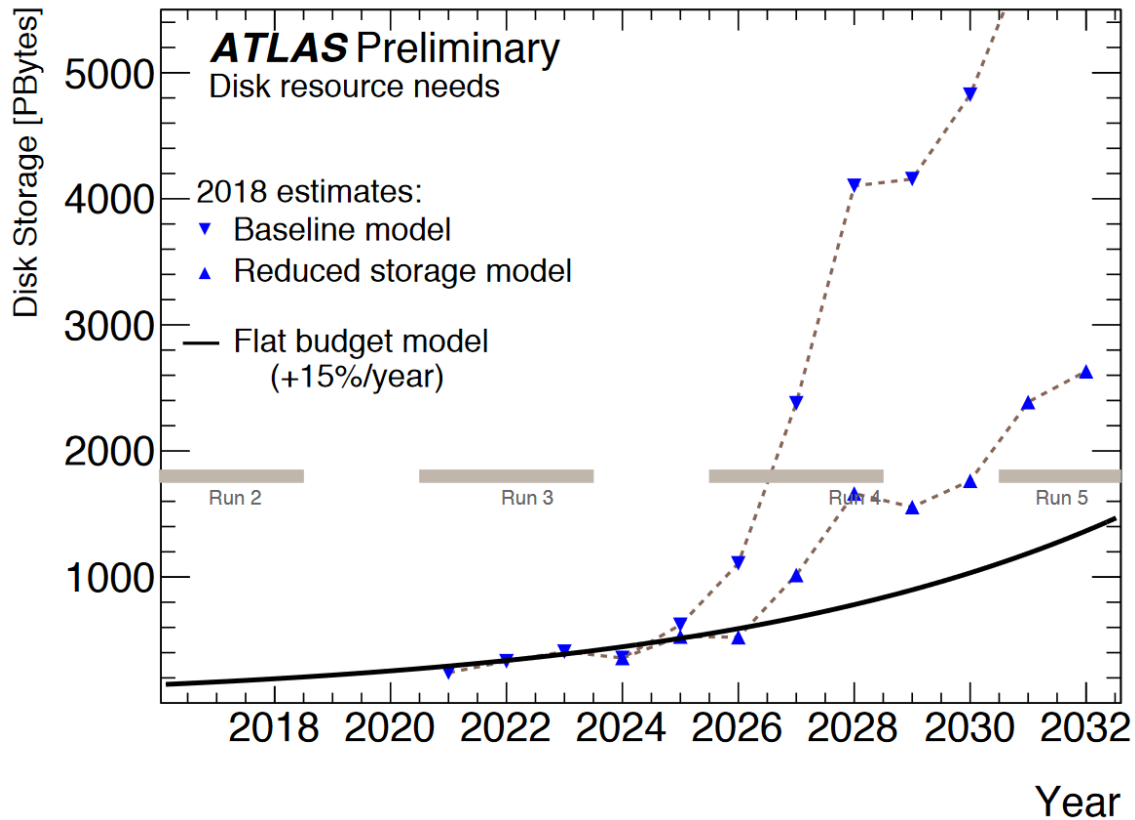
- 167 sites in 42 countries
- $\sim 1M$ CPU cores
- $\sim EB$ of storage
- $> 2M$ jobs per day
- 100 PB of data transfers per year via 10-100 Gb links

Характерная особенность: задания естественным образом распараллеливаются на нескольких уровнях:



LHC computing: future projections

Совещание по обновлению евростратегии в области физики частиц (2019)



HEP software

- “Bad software is extremely expensive!” and “good and clever software allows much more physics to fit in the budget”
- Увеличивающаяся сложность экспериментов больше не позволяет разрабатывать достаточно хорошее ПО независимо для каждого эксперимента. Одновременно увеличиваются частота и сложность событий
- HEP Software Foundation (<https://github.com/HSF>)
 - Стандартизация ПО для детекторов частиц
 - Переход от сложных иерархических структура данных к POD
 - Применение методов машинного обучения



[ROOT: data analysis framework](#)



[Geant4 simulation toolkit](#)

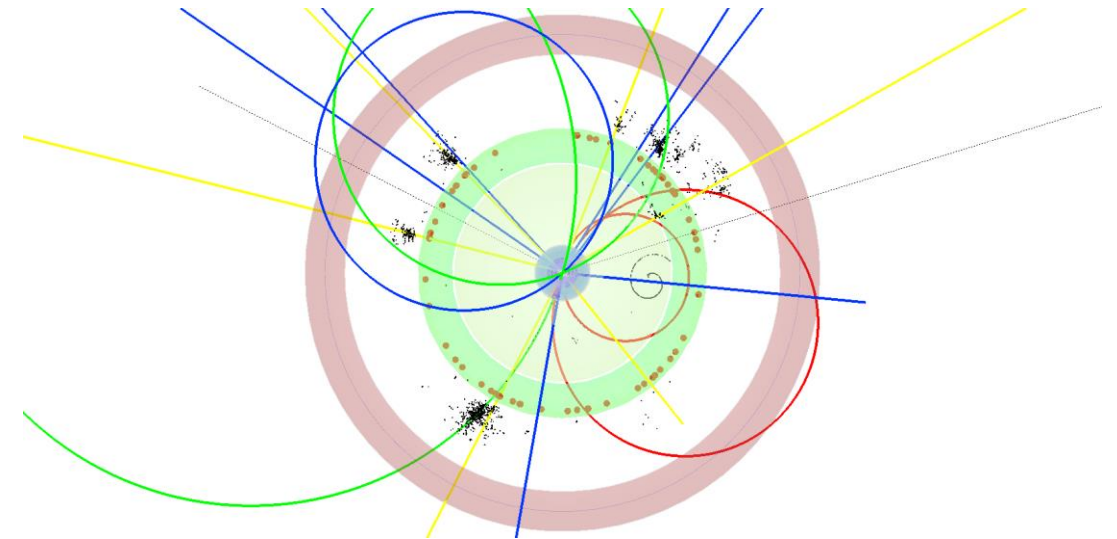
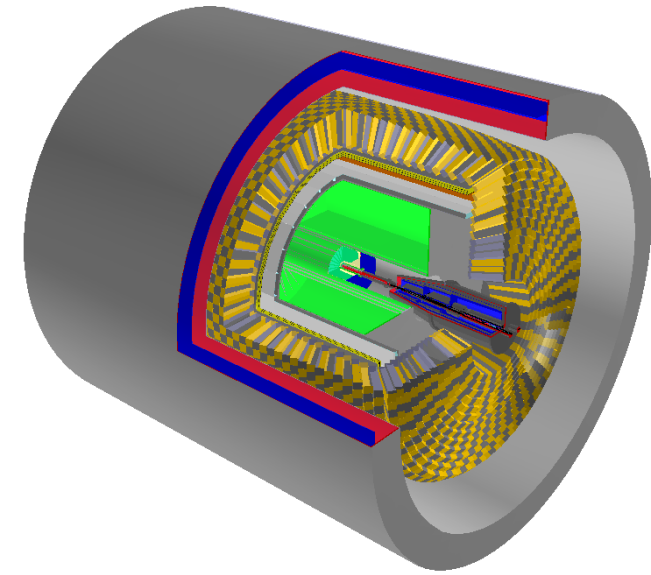
Специализированное ПО
для Супер С-тау фабрики



Стандартное ПО

Моделирование детектора SCT

- Система сборки ПО. Основана на **Athena**
 - Автоматизация через GitLab CI
 - Сервер для разработки ПО
- Фреймворк «Аврора» для программных модулей
 - Основан на **Gaudi**
 - Разработка на C++
 - python-интерфейс для пользователей
- Геометрия детектора описана: **DD4Hep**
- Моделирование взаимодействия частиц с веществом работает: **Geant4**
- Моделирование оцифровки сигналов и разработка алгоритмов реконструкции событий – в процессе
- Есть предварительная версия модели данных: сейчас **PODIO**, в перспективе **EDM4Hep**



РНФ 19-72-20114: 2019-2022

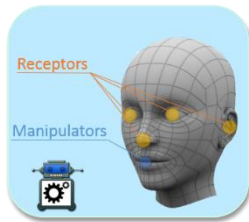
Руководитель гранта, д.ф.-м.н. Логашенко Иван Борисович

- Разработка системы моделирования, обработки и хранения данных установки класса мегасайенс «Супер С-тау фабрика»
- Объект инфраструктуры: ССКЦ
- Исполнители: ИЯФ, ИВМиМГ, ИВТ
- Задачи:
 1. Создание программного обеспечения для моделирования и реконструкции данных экспериментов
 2. Создание системы по автоматизации процессов обработки экстремально большой по объему информации, производимой в экспериментах на СЧТФ
 3. Развитие методов анализа данных в области физики высоких энергий, основанных на методах машинного обучения и байесовского вывода
 4. Разработка модели вычислительной инфраструктуры СЧТФ и проведение имитационного моделирования вычислительной инфраструктуры СЧТФ

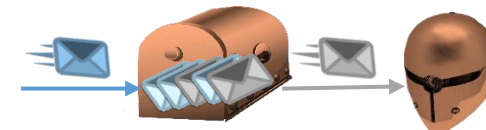
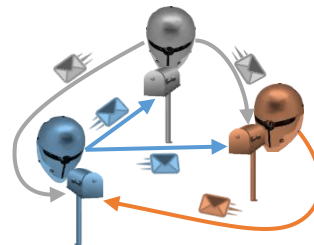
Общая структура имитационного моделирования вычислительной инфраструктуры

Моделирование организовано на основе мультиагентного подхода – вся система описывается как множество взаимодействующих агентов.

Агенты находятся в общей внешней среде, и имеют заданный набор поведений



Агенты взаимодействуют между собой посредством обмена сообщениями. В процессе обмена сообщениями поведение может изменяться

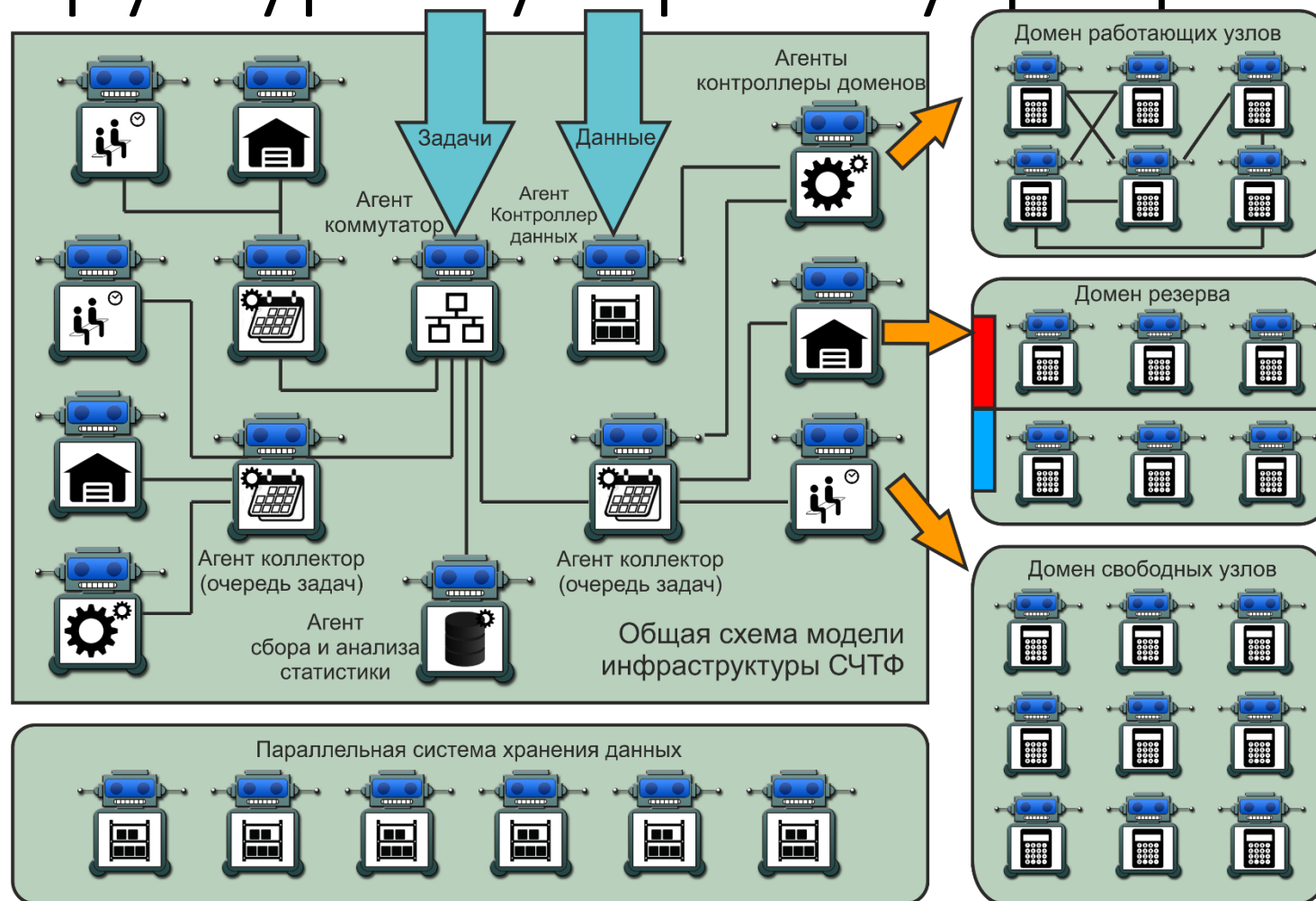


Этапы работы:

- ✓ Выделить основные составные модули вычислительной инфраструктуры, выполняющие однотипные действия;
- ✓ Определить параметры, алгоритмы функционирования и логическую структуру взаимодействия данных модулей.
- ✓ Для каждого модуля создать класс агентов и определить набор поведений в соответствии с параметрами и алгоритмом функционирования.
- ✓ Систематизировать и определить параметры сообщений, которыми будут обмениваться созданные агенты. Для каждого сообщения создать обработчик в соответствующем классе агента.
- В соответствии с характеристиками оборудования кластеров НКС-1П и НКС-30Т создать необходимое число экземпляров классов агентов и задать их параметры. Провести апробацию модели на тестовом наборе входных данных.
- Запуск модельного эксперимента, в процессе которого подобрать характеристики оборудования, необходимые для функционирования СЧТФ.

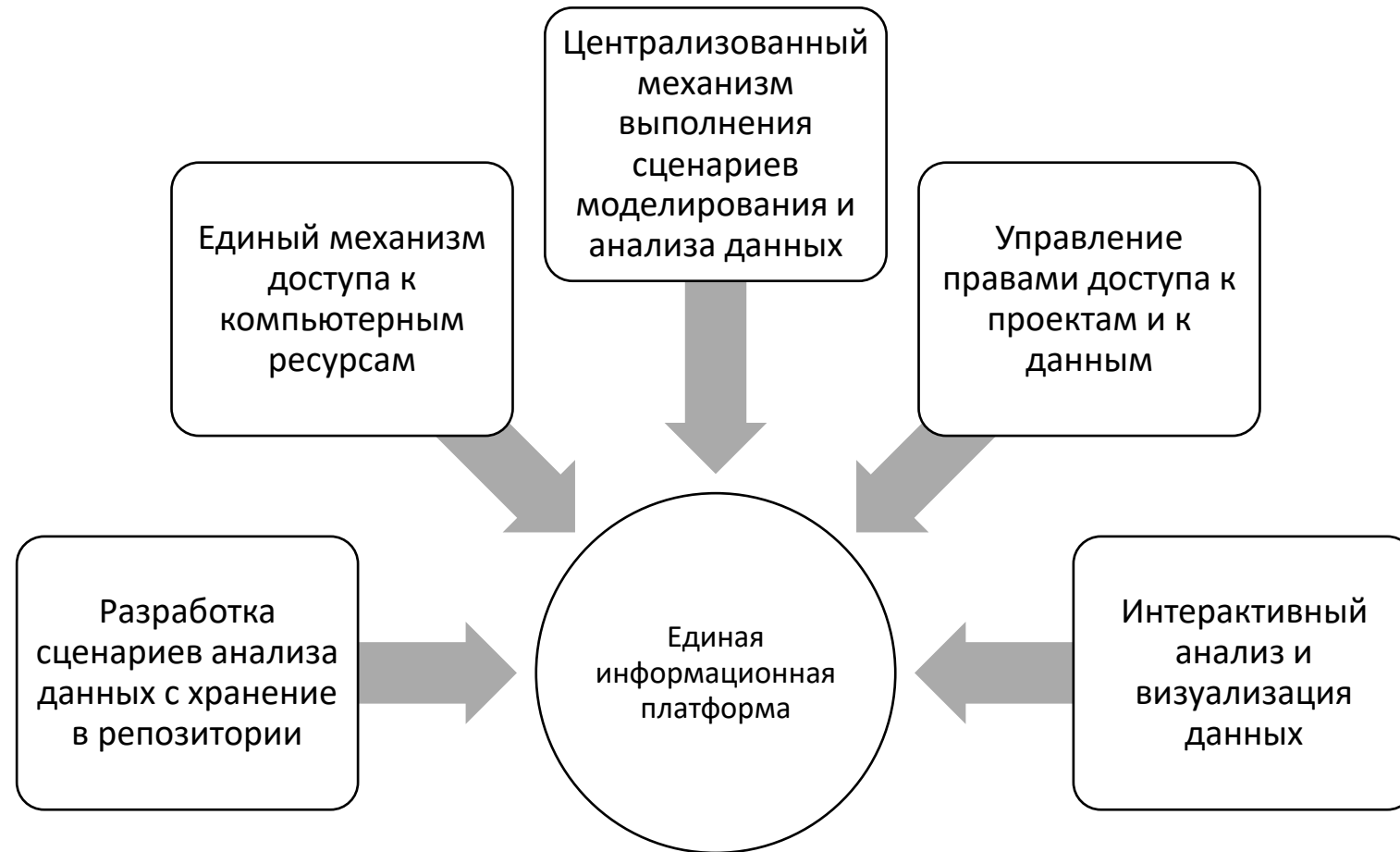
И. Черных
Д. Винс

Структура модели вычислительной инфраструктуры Супер С-тау фабрики



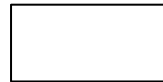
И. Черных
Д. Винс

Задачи и проектные возможности единой информационной платформы

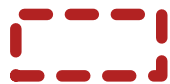


Ф. Колпаков
И. Киселев

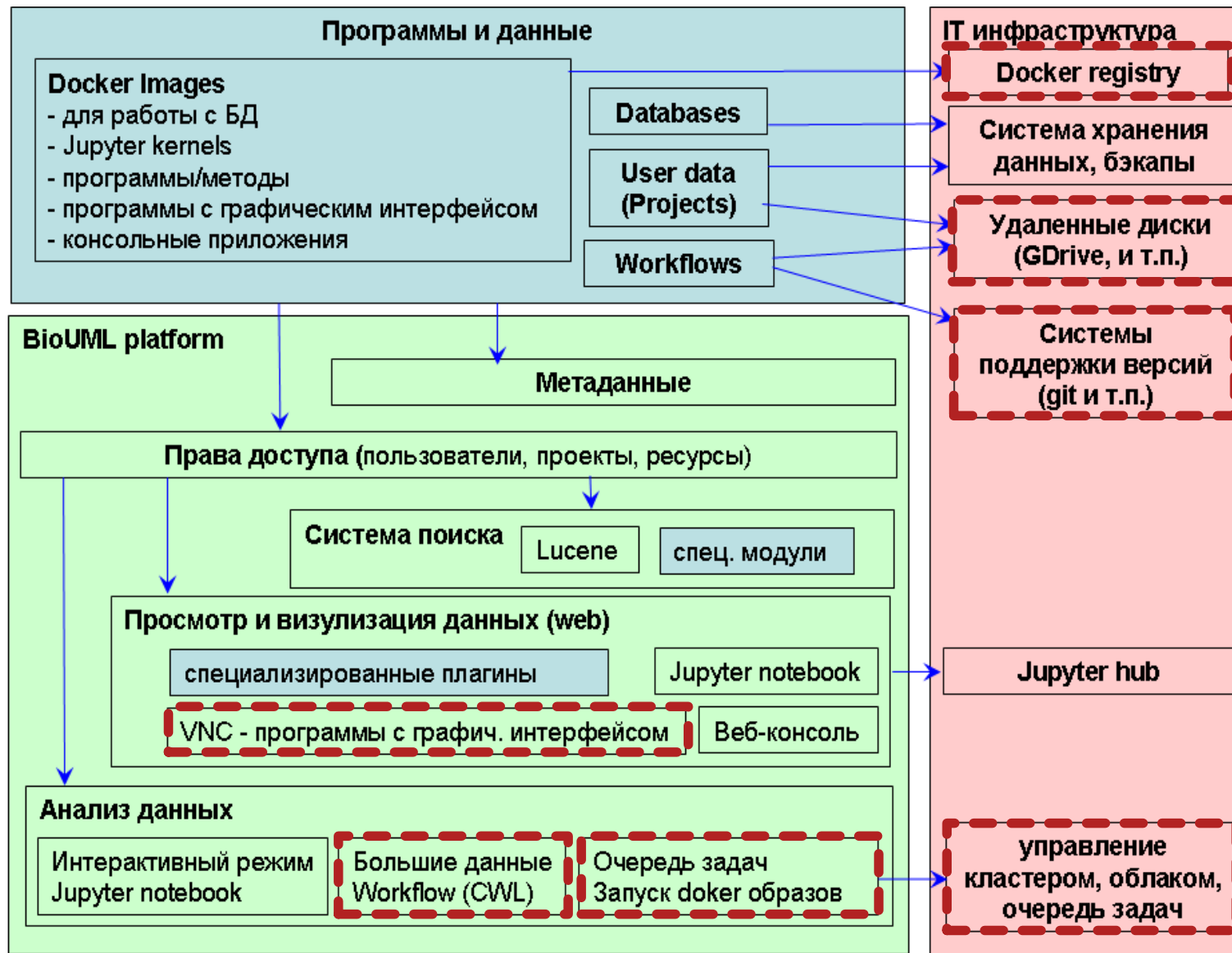
Общая схема единой информационной платформы



Готовые блоки



Планирующиеся блоки

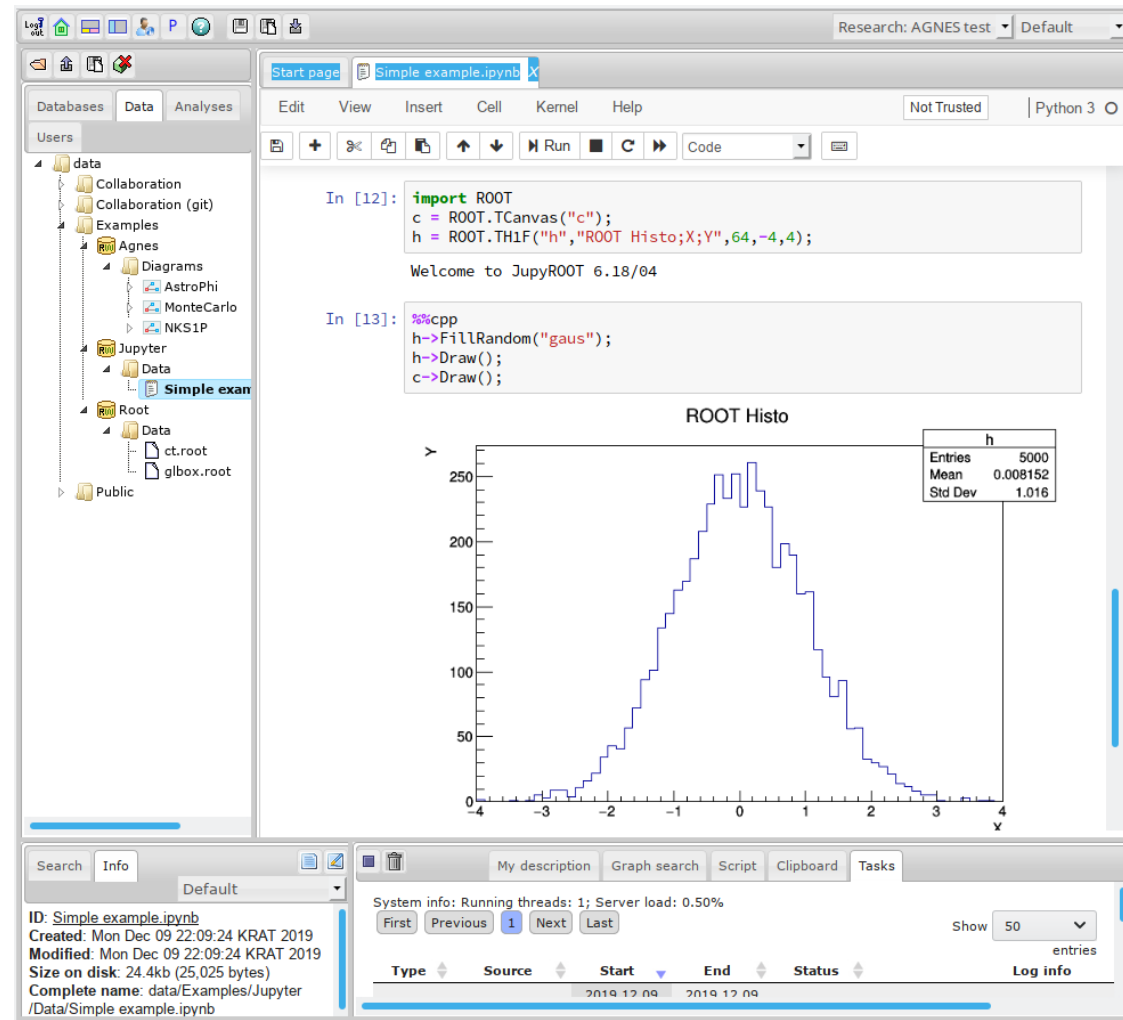


Ф. Колпаков
И. Киселев

Jupyter notebook и ROOT в единой информационной платформе

В платформу были интегрированы инструменты Jupyter notebook и ROOT. В графическом интерфейсе платформы пользователь может создавать и загружать документы jupyter, позволяющие выполнять скрипты на различных языках.

Благодаря интеграции ROOT, пользователь также может работать с ROOT-файлами и выполнять команды ROOT.



The screenshot displays the JupyterROOT web interface. On the left, a file browser shows a directory structure with folders like 'data', 'Jupyter', and 'Root'. The main area contains a Jupyter notebook with two code cells. The first cell imports ROOT and creates a canvas and histogram. The second cell uses C++ macros to fill and draw a Gaussian histogram. Below the code, a plot titled 'ROOT Histo' shows a Gaussian distribution with a mean of 0.008152 and a standard deviation of 1.016. The plot includes a table with the following statistics:

h	
Entries	5000
Mean	0.008152
Std Dev	1.016

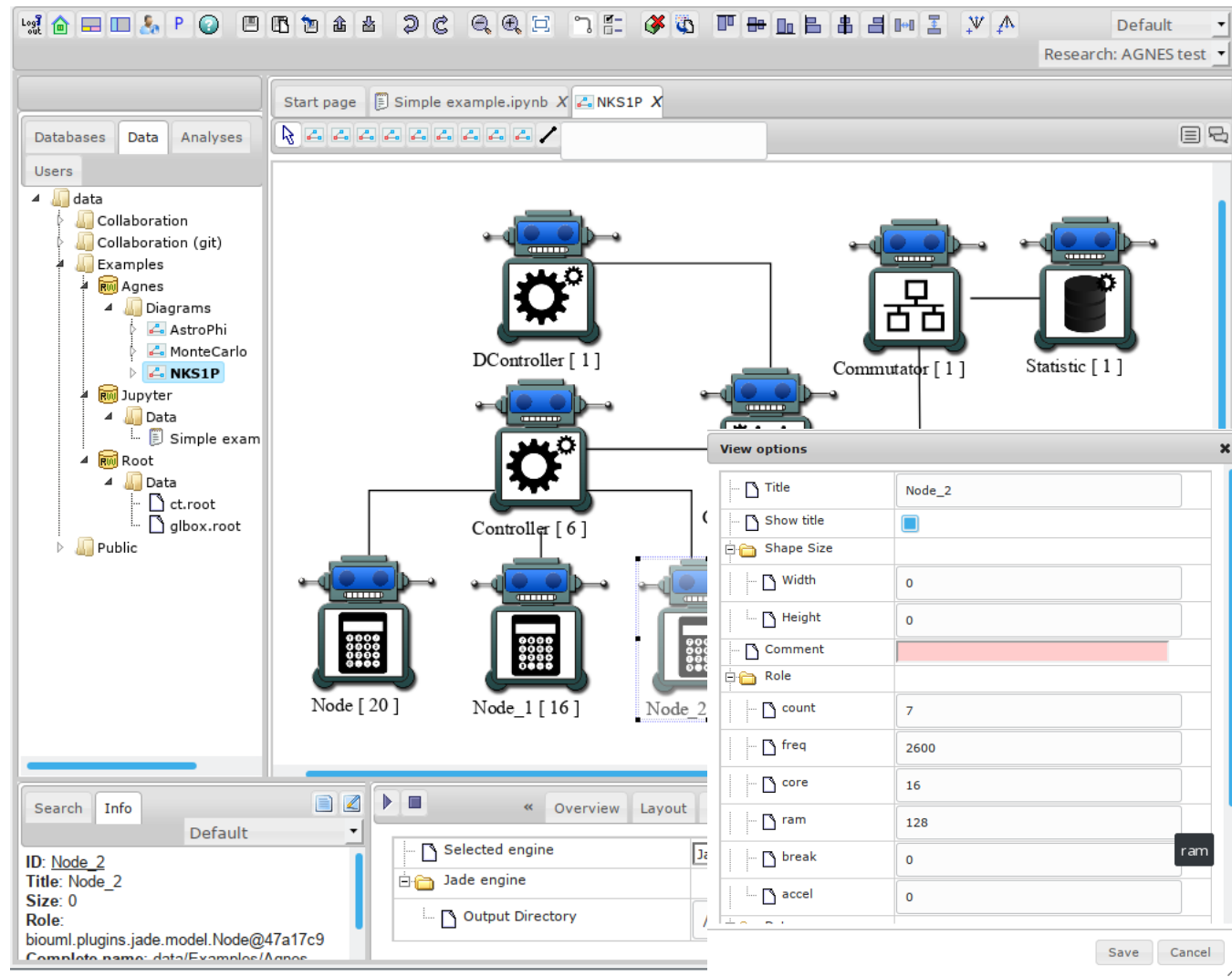
At the bottom of the interface, there is a search bar, system information (running threads: 1, server load: 0.50%), and a log table with columns for Type, Source, Start, End, and Status.

Ф. Колпаков
И. Киселев

Графическое представление имитационной модели в единой информационной платформе

Пользователь может создавать имитационные модели графически в виде диаграмм. Каждый элемент на диаграмме соответствует группе агентов с одинаковыми свойствами.

Пользователь может добавлять группы агентов на диаграмму, редактировать их свойства и запускать численные расчеты из вкладки Simulation.



Заключение

Эксперимент на Супер чарм-тау фабрике подразумевает наличие современной вычислительной инфраструктуры

- Максимальное использование существующих решений
- Решение интересных технических задач
- Улучшение IT-инфраструктуры Академгородка в целом

Backup

Crab-waist collision

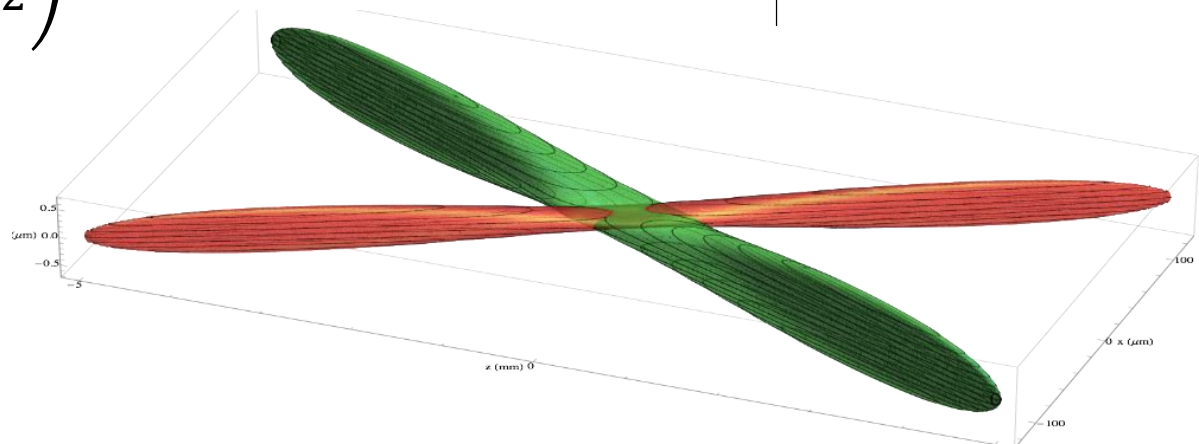
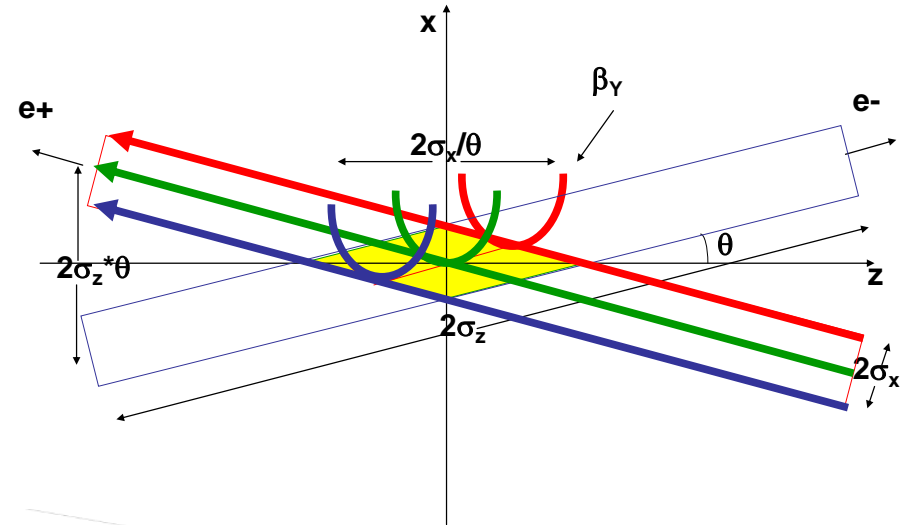
1. *P.Raimondi, 2nd Workshop on Super B-Factory, March 2006.*
2. *P.Raimondi, D.Shatilov, M.Zobov, "Beam-Beam Issues for Colliding Schemes with Large Piwinski Angle and Crabbed Waist", LNF-07-003-IR, Feb. 2007, e-Print: physics/0702033.*

$$L = \frac{\gamma}{2er_e} I \frac{\xi_y}{\beta_y^*} \left(1 + \frac{\sigma_y}{\sigma_x \sqrt{1 + \phi^2}} \right)$$

$$\phi = \frac{\sigma_z}{\sigma_x} \tan \frac{\theta}{2} \simeq 60 \text{ mrad}$$

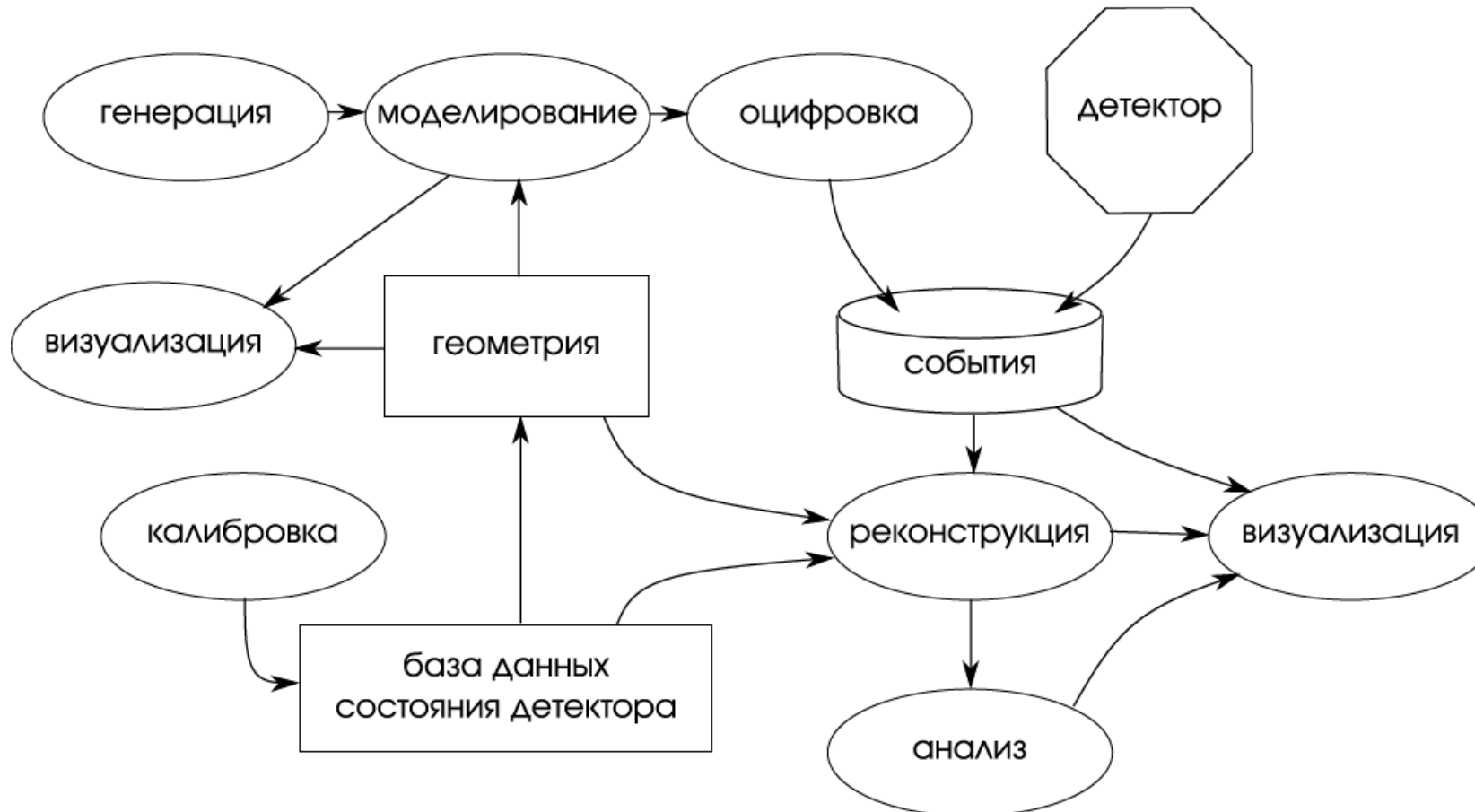
$$\beta_y^* = 0.2 \text{ mm (6-10 mm, B factories)}$$

$$\xi_y = 0.15 \text{ (0.05, B factories)}$$

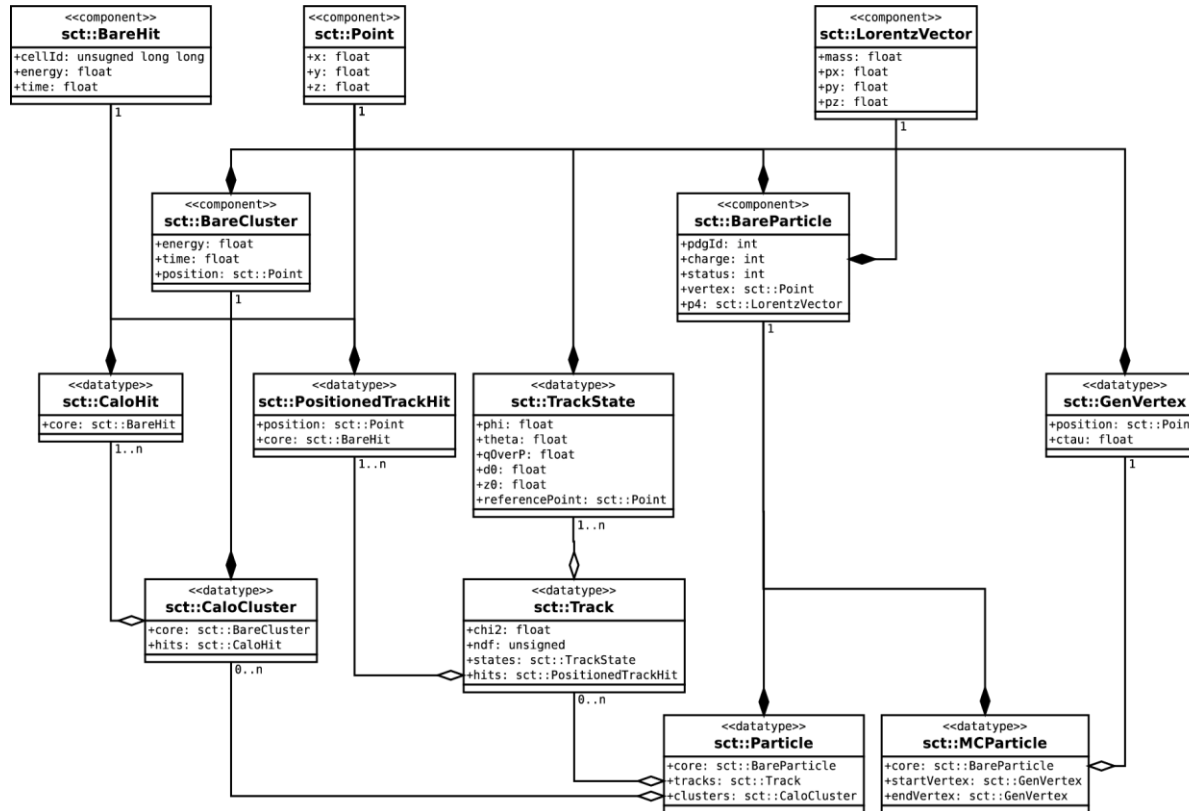


100x luminosity boost without shortening bunches and increasing beam current

Основные элементы программного обеспечения для моделирования и анализа данных



Модель данных



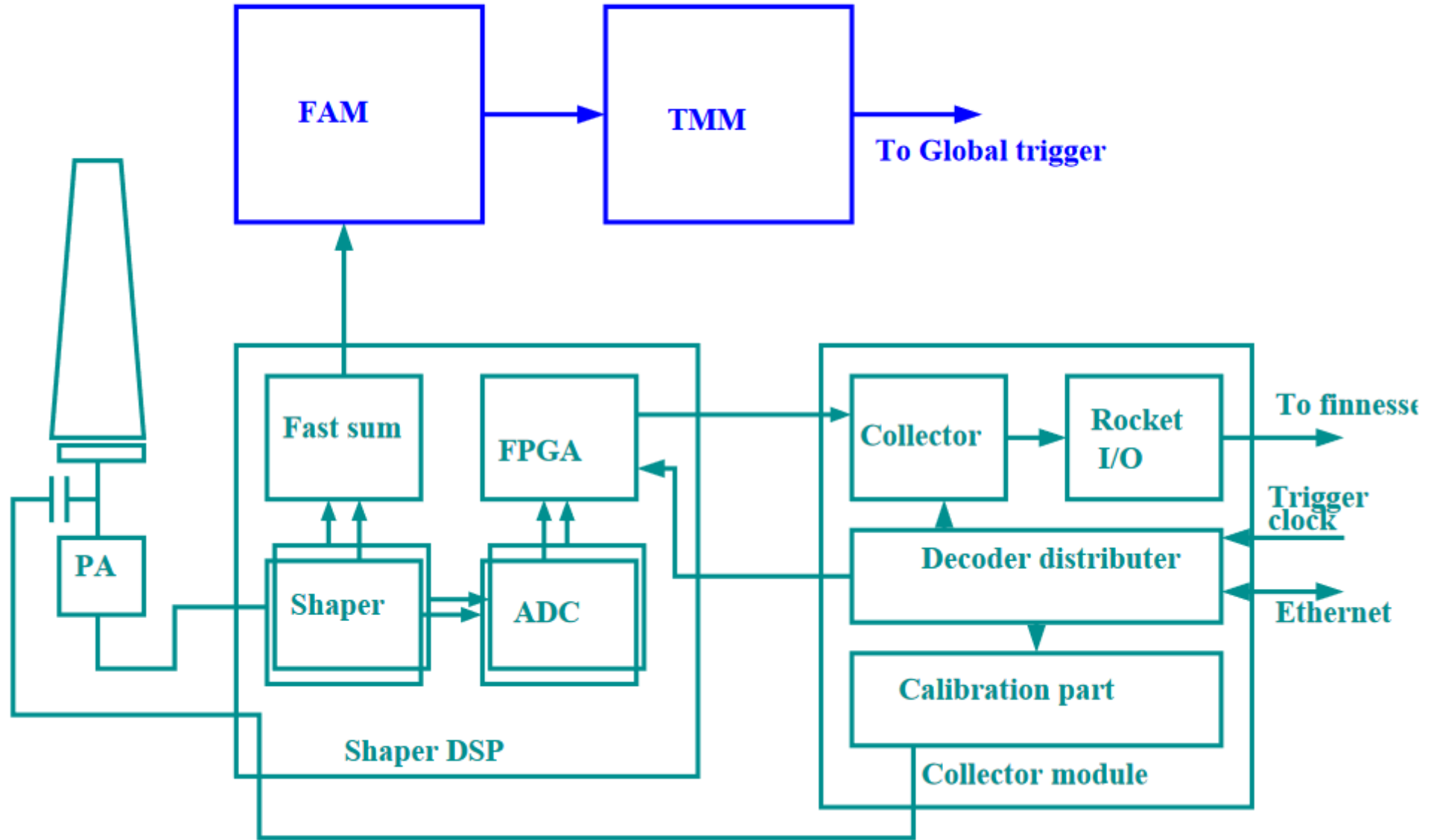
- «datatype» — типы данных, могут быть сохранены в файл для обмена между задачами
- «component» — базовые компоненты типов данных
- Для каждого типа данных поддерживается возможность создания коллекций
- Весь набор C++ классов генерируется из единого описания при помощи библиотеки PODIO
- Предоставляется интерфейс доступа к данным в Python

Collaboration

- Working groups
 - Inner tracker
 - Drift chamber
 - PID
 - Calorimeter
 - Muon system
 - Magnet
 - Physics and simulations
 - Computing
 - DAQ and trigger
 - Beam background
 - Engineering
- International advisory committee
- Dedicated international workshops
 - May 2018, BINP
 - December 2018, Orsay
 - September 2019, Moscow
 - **Fall 2020 in China**
- Monthly online meetings with colleagues working on **the HIEPA project**

WGs are open for international participation





Entangled $D^0\bar{D}^0$ state as a powerful tool

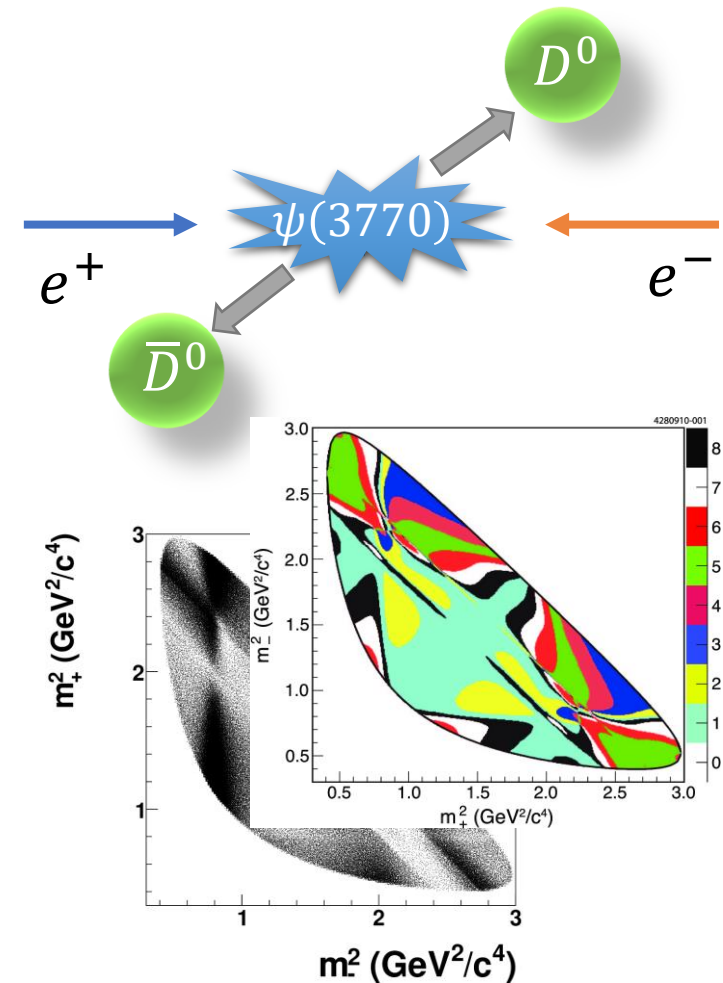
... a powerful tool to study charm and QCD

- Entangled $D^0\bar{D}^0$ pairs are produced at threshold
 - The unique phenomenology
- Measuring charm mixing and CPV in charm
 - **In a unique way**: quantum correlations in action! [1, 2]
 - Precision comparable to Belle II
 - The best environment for the decays with final-state neutrals
- Measurements of the D^0 decay amplitude phases
 - $\delta_{K\pi}, \delta_{K\pi\pi^0}, \dots$ [3]
 - Model-independent Dalitz analysis of $D^0 \rightarrow K_S^0\pi^+\pi^-$
 - Essential **input for the CKM phase γ** measurements at Belle II and LHCb

[1] V.V. slides at [HIEPA2018](#)

[2] [PRD 82 \(2010\) 034033](#)

[3] [PRD 73 \(2007\) 034024](#)



Experiments with polarization

1. Left-right $e^+e^- \rightarrow J/\psi$ cross section asymmetry [1]

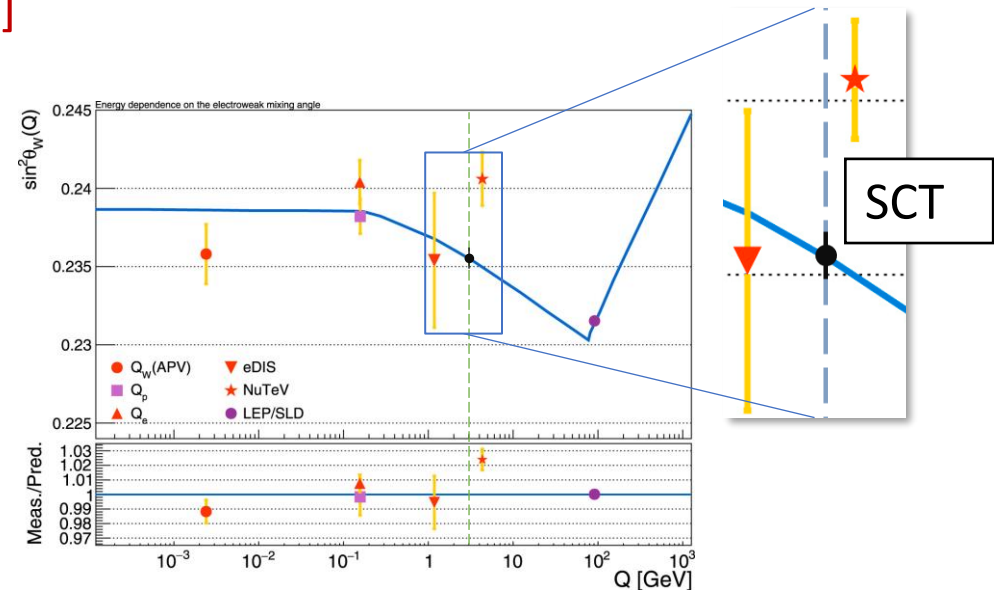
- Measuring the Weinberg angle at 0.3%
- Testing weak interaction of charm
- Sensitivity to heavy Z'

2. Baryons (light and charmed)

- Formfactors: calibration of lattice QCD calculations
- CP and T violation in decays

3. Tau physics

- Lorentz structure of lepton decays
 - No need of spin-spin correlations of $\tau^+\tau^-$. Great simplification of the analysis procedure \rightarrow better systematics and statistics
- CP violation in τ production: EDM
 - Current experimental limit: $|d_\tau| \lesssim 10^{-17} e \cdot cm$
 - With 10^{10} tau pairs at SCT: $\sigma(d_\tau) \sim 10^{-20} e \cdot cm$ [2]
- CP violation in decays, e.g.: $\tau \rightarrow K\pi\nu_\tau$ [3]



[1] [JHEP 2020 \(2020\) 76](#)

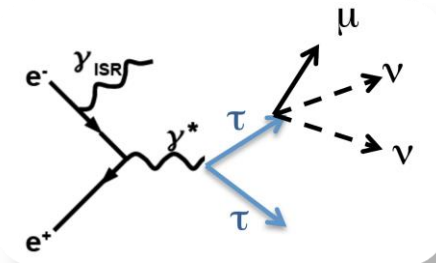
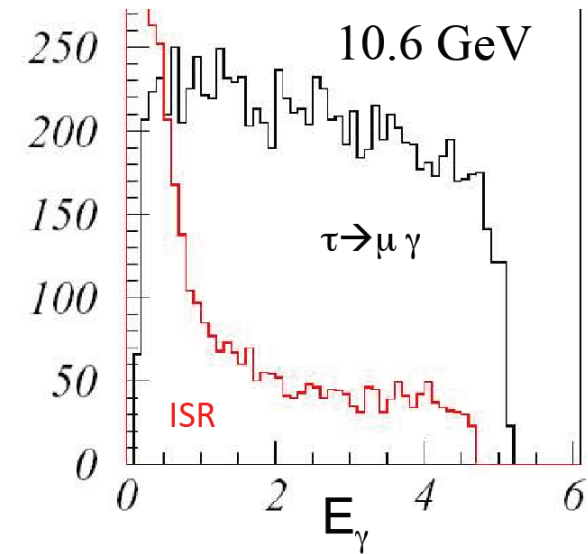
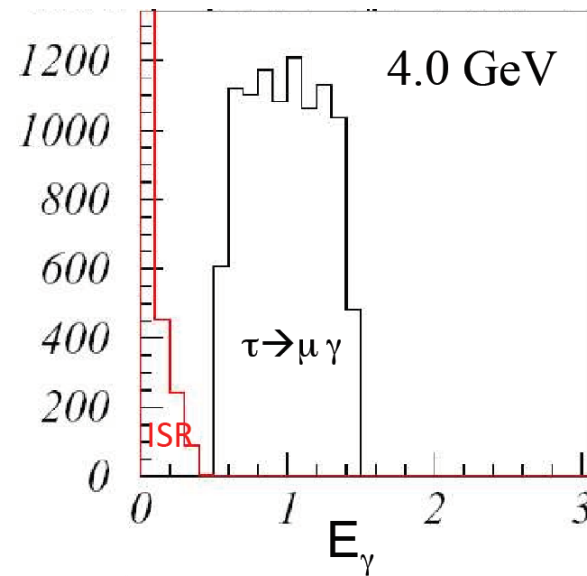
[2] [PRD 51 \(1995\) 5996](#)

[3] [PLB 437 \(1998\) 191](#)

Experiments with τ leptons

- Controlling systematic uncertainties is crucial for τ measurements.
- Threshold kinematics
 - Additional kinematic constraints for τ decays reconstruction
 - Easy background sample (data below threshold)
 - Monochromatic particles from two-body decays
 - Ideal for $\tau \rightarrow \mu\gamma$ search
- Precision weak interaction test with polarized τ

ISR photon background for $\tau \rightarrow \mu\gamma$: B - vs. $c\tau$ -factory
[Bondar, Bobrov (BINP)]



See A. Pich talk [\[pdf\]](#)

From «nuclear chemistry» to nuclear physics of quarkonium

- We have entered the era of «nuclear chemistry»
 - Various hadronic molecular states observed and continue to appear
 - No systematic description yet
- An exciting laboratory for QCD
 - QCD beyond perturbation theory
 - Not a new physics, but **new phenomena interesting by themselves**
- We should come back to nuclear **physics** at some point

