



ИНСТИТУТ
СИСТЕМНОГО АНАЛИЗА
И УПРАВЛЕНИЯ

Методы и технологии обработки данных в гетерогенных вычислительных средах

Кореньков Владимир Васильевич

**Научный руководитель Лаборатории информационных технологий
имени М.Г. Мещерякова ОИЯИ**

**Зав. кафедрой «Распределенные информационно-вычислительные системы»
Государственного университета «Дубна»**

**Руководитель магистерской программы «Методы и технологии обработки данных в
гетерогенных вычислительных средах» в филиале МГУ в Дубне**

**Осенняя студенческая ИТ-школа,
Дубна, ЛИТ ОИЯИ, 6 октября 2025 г.**

Эволюция ИТ

- менялись концепции
- круг и сложность решаемых задач
- возникал новый технологический набор
- углублялась специализация разработчиков
- сокращалось время ввода новых продуктов и сервисов

первые цифровые платформы
поддержки, работающих в
реальном времени

Первое поколение (1960-е годы) — мейнфреймы

Второе поколение (1970-е годы) — универсальные ЭВМ

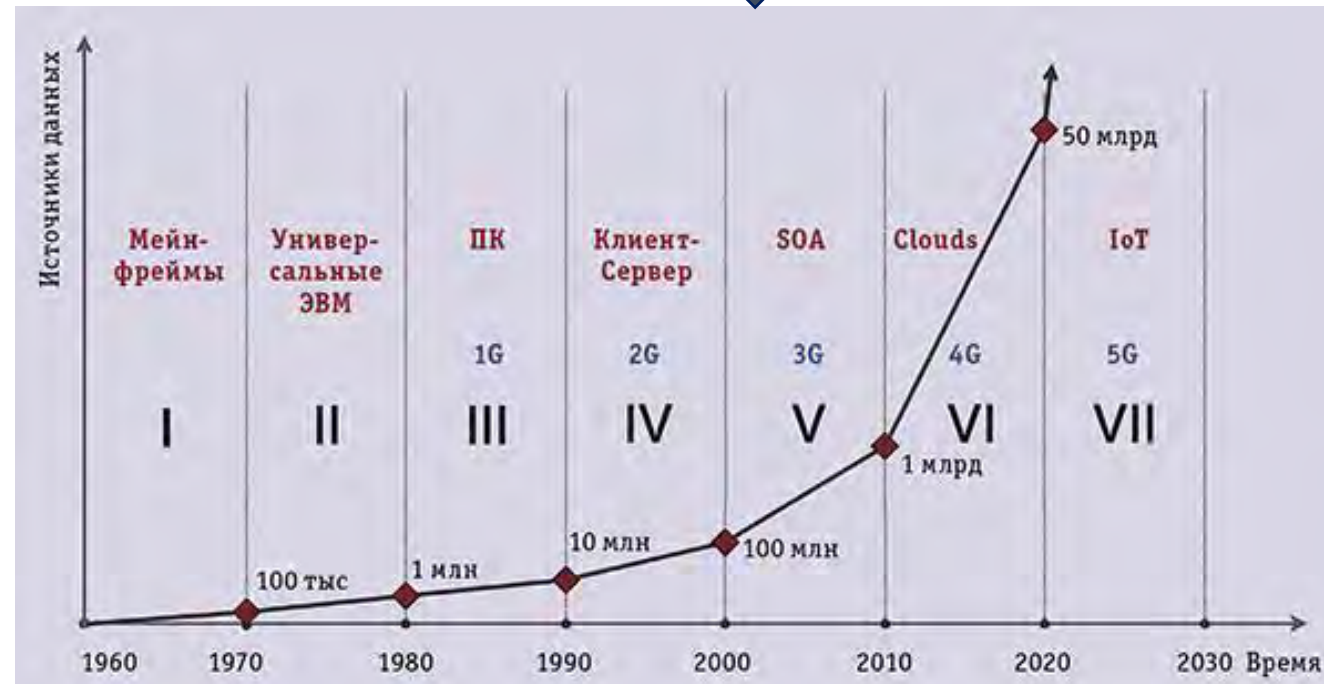
Третье поколение (1980-е годы) — персональные компьютеры

Четвертое поколение (1990-е годы) — клиент-сервер

Пятое поколение (2000-е годы) — сервисная архитектура

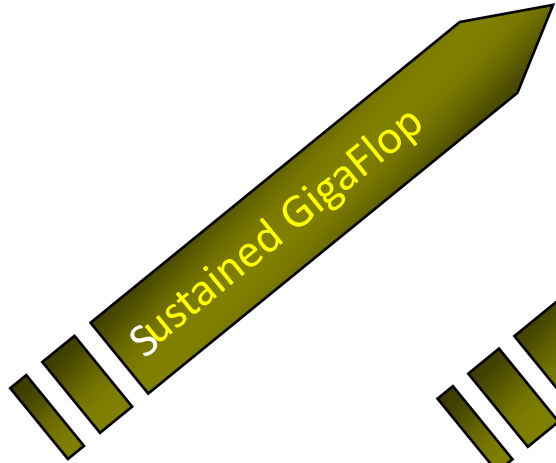
Шестое поколение (2010-е годы) — облака

Седьмое поколение (2020-е годы) — IoT, искусственный интеллект, квантовые вычисления

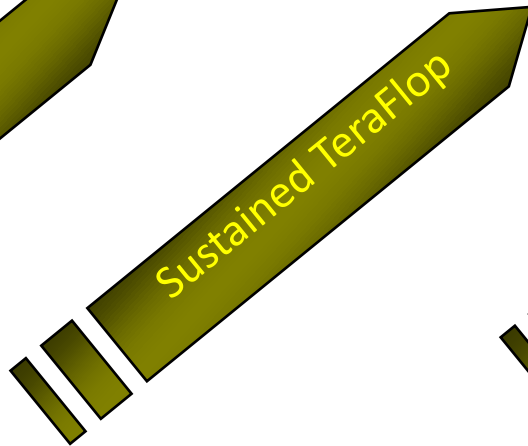


Этапы развития суперкомпьютеров согласно закону Мура (от Gigafllops до Exafllops)

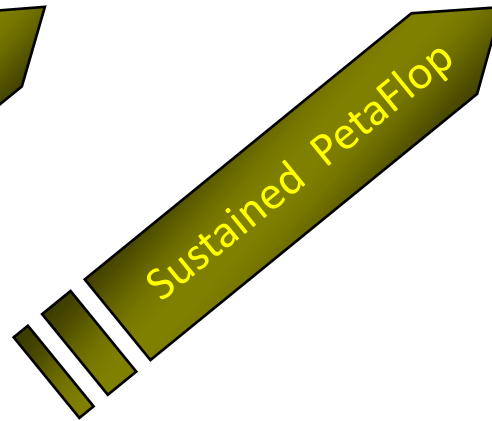
~1987



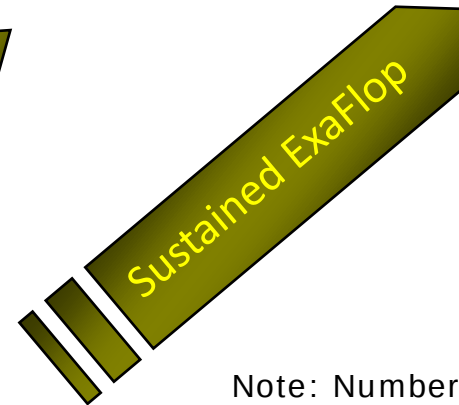
~1997



2008



2022



Note: Numbers are based on Linpack Benchmark.
Dates are approximate.

The El Capitan system at the Lawrence Livermore National Laboratory, No. 1 system on the TOP500.

El Capitan has 11,039,616 cores, 1.742 Exaflop/s, 29,580 MW



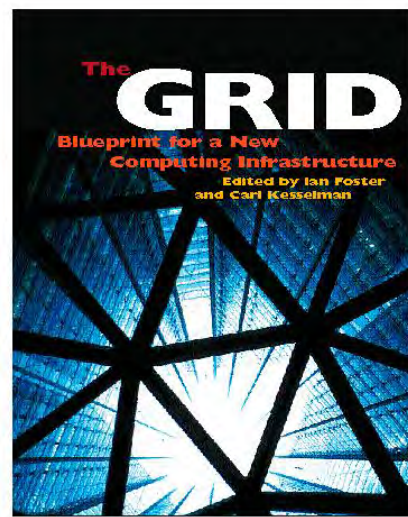
Frontier - HPE Cray EX235a, AMD
Oak Ridge National Laboratory
9,066,176 cores 1,353.00 exaflops/s
24,507 MW



Grids, clouds, fog, edge, supercomputers...

Grids

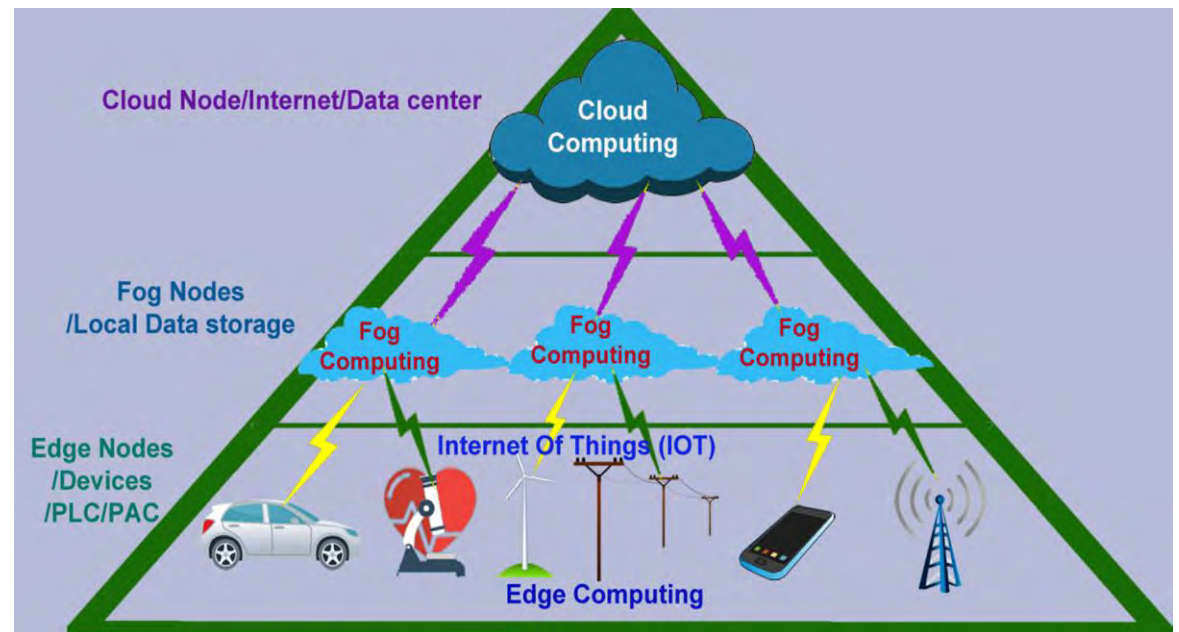
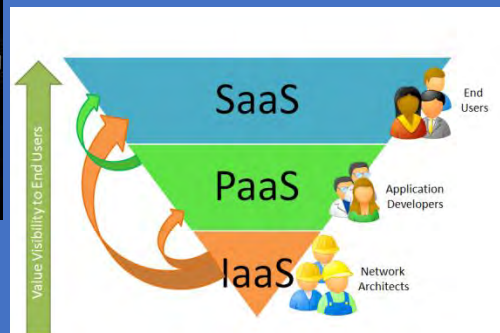
- Collaborative environment
- Distributed resources



Supercomputers



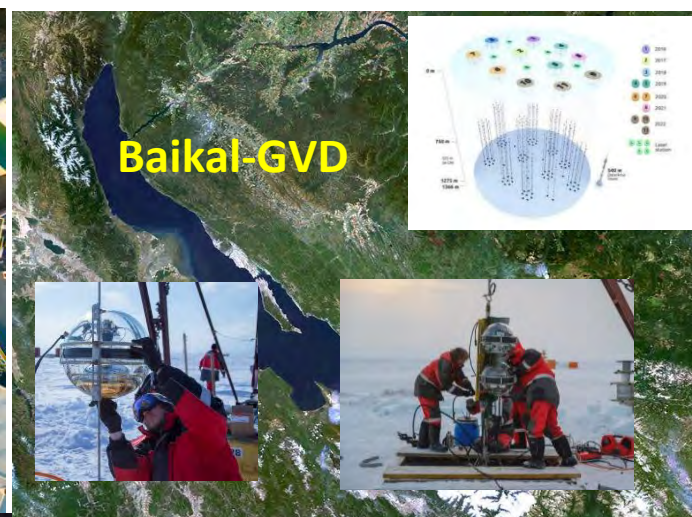
Clouds



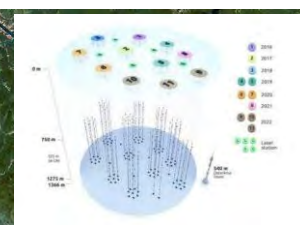
Научно-исследовательская программа ОИЯИ



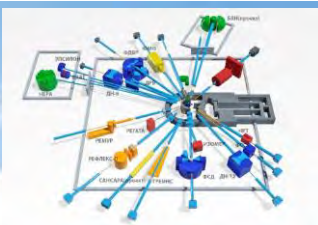
NICA complex



Baikal-GVD



IBR-2M



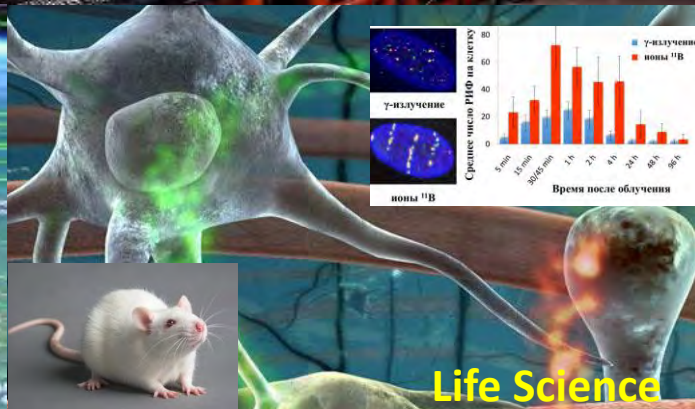
SHEF



Nuclotron



IT & CC



Life Science

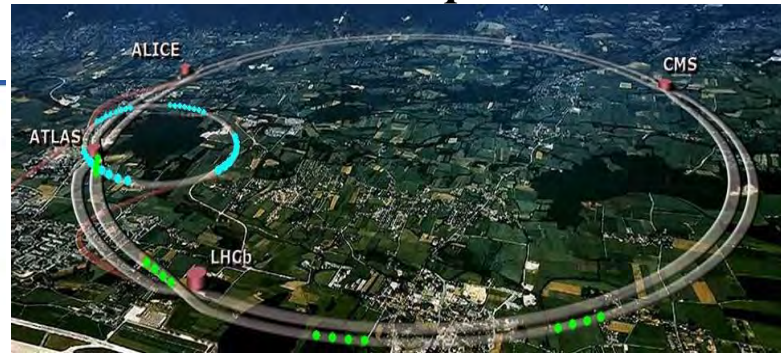


DRIBS-III

HPC+Big Data+Искусственный интеллект



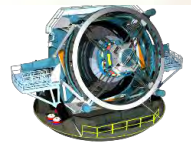
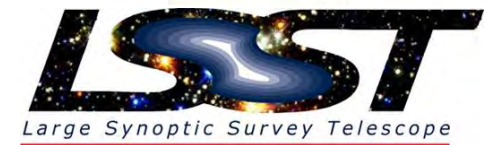
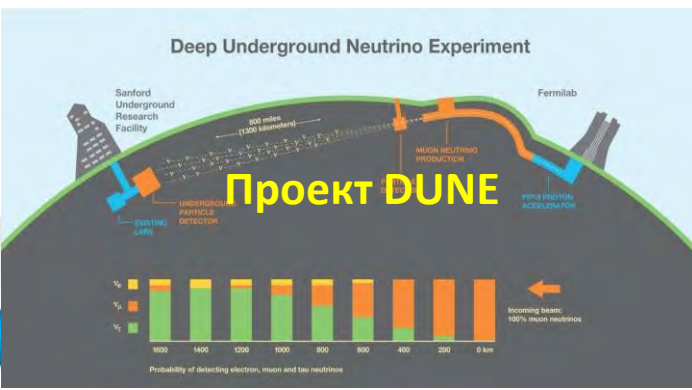
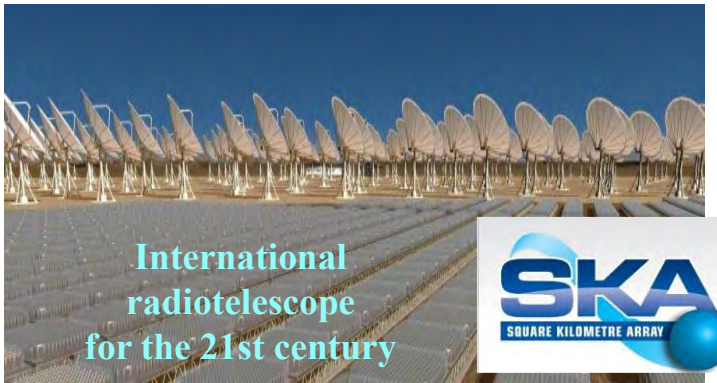
Физика высоких энергий



CERN Large Hadron Collider > 400 PB/Year



Square Kilometer Array radio telescope
> 1 Eb/Year raw data (estimation)



Large Synoptic Survey Telescope
> 10 Pb/Year (estimation)

Концепция Грид

«Грид - это система, которая:

- координирует использование ресурсов при отсутствии централизованного управления этими ресурсами
- использует стандартные, открытые, универсальные протоколы и интерфейсы.
- обеспечивает высококачественное обслуживание»

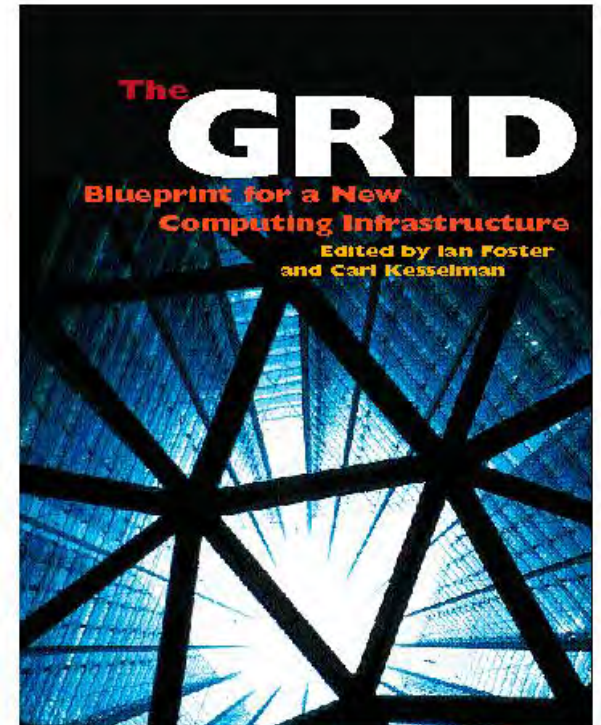
(Ian Foster: "What is the grid? ", 2002 г.)

Модели грид:

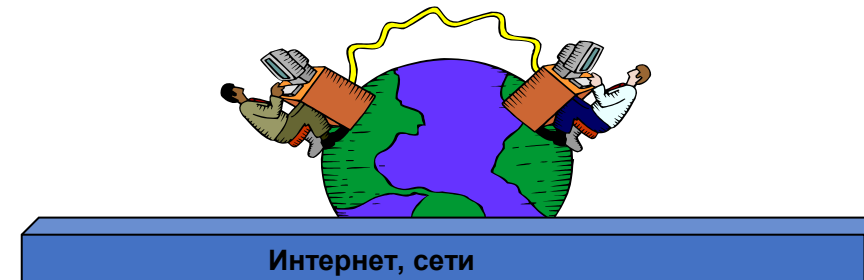
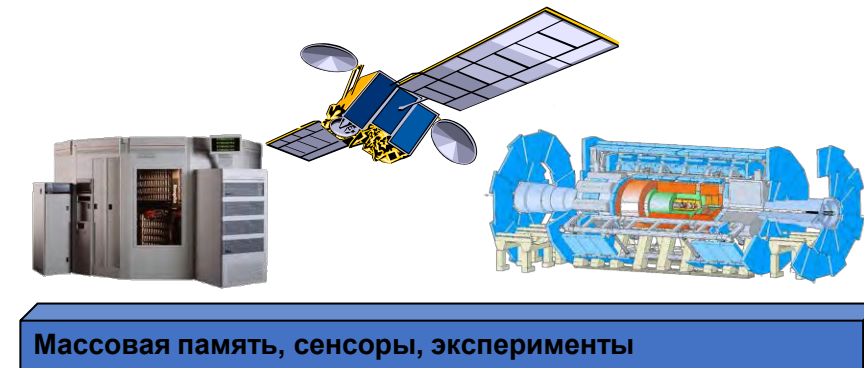
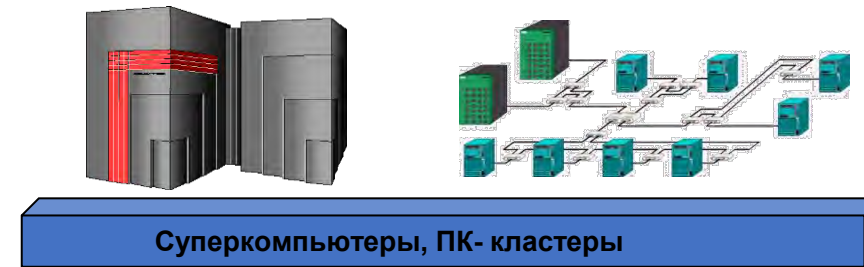
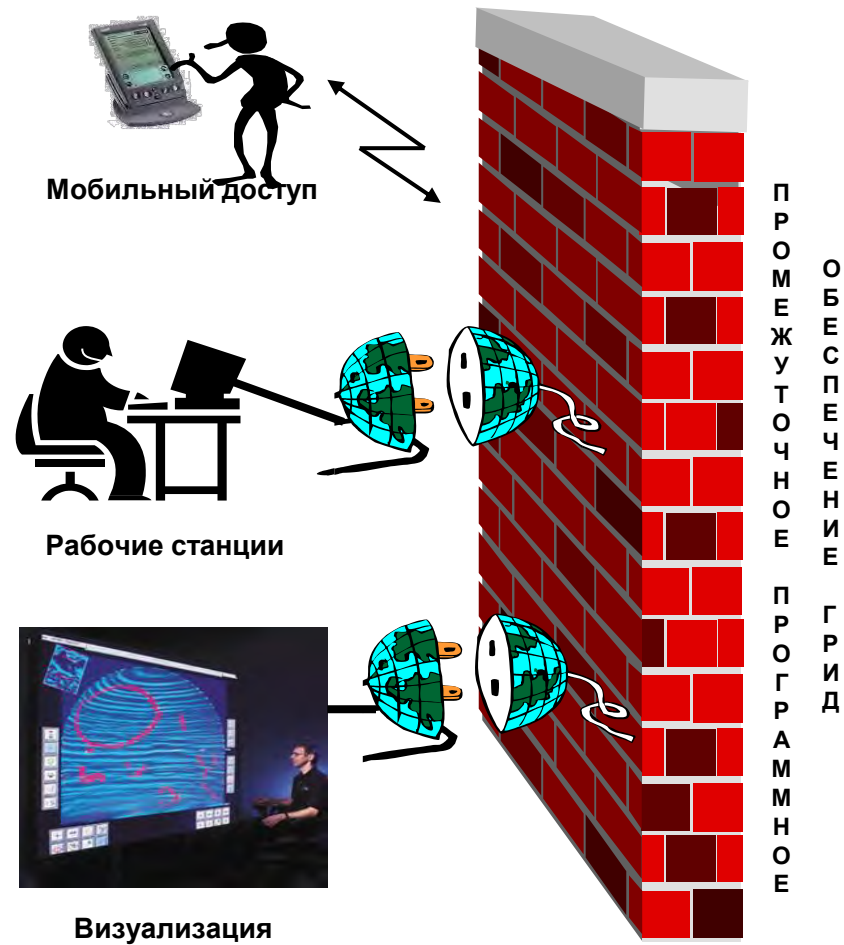
- ❖ Distributed Computing
- ❖ High-Throughput Computing
- ❖ On-Demand Computing
- ❖ Data-Intensive Computing
- ❖ Collaborative Computing

Междисциплинарный характер грид: развиваемые технологии применяются в физике высоких энергий, космофизике, микробиологии, экологии, метеорологии, различных инженерных и бизнес приложениях.

Виртуальные организации (VO)

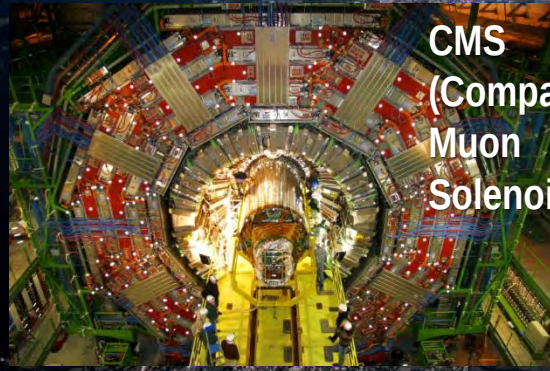


Грид - это средство для совместного использования вычислительных мощностей и хранилищ данных посредством интернета



Large Hadron Collider

Большой адронный коллайдер (БАК), один из крупнейших и по-настоящему глобальных научных проектов, является самым захватывающим поворотным моментом в физике элементарных частиц.



CMS
(Compact
Muon
Solenoid)



LHCb
(Large
Hadron
Collider
beauty
experiment)

Data flow to permanent storage: 4-6 GB/sec

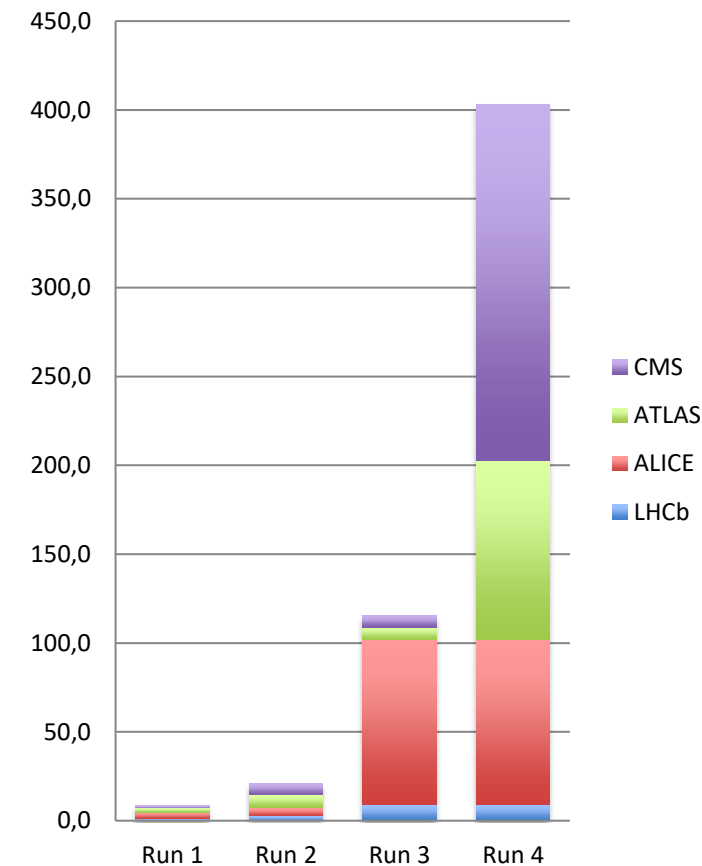
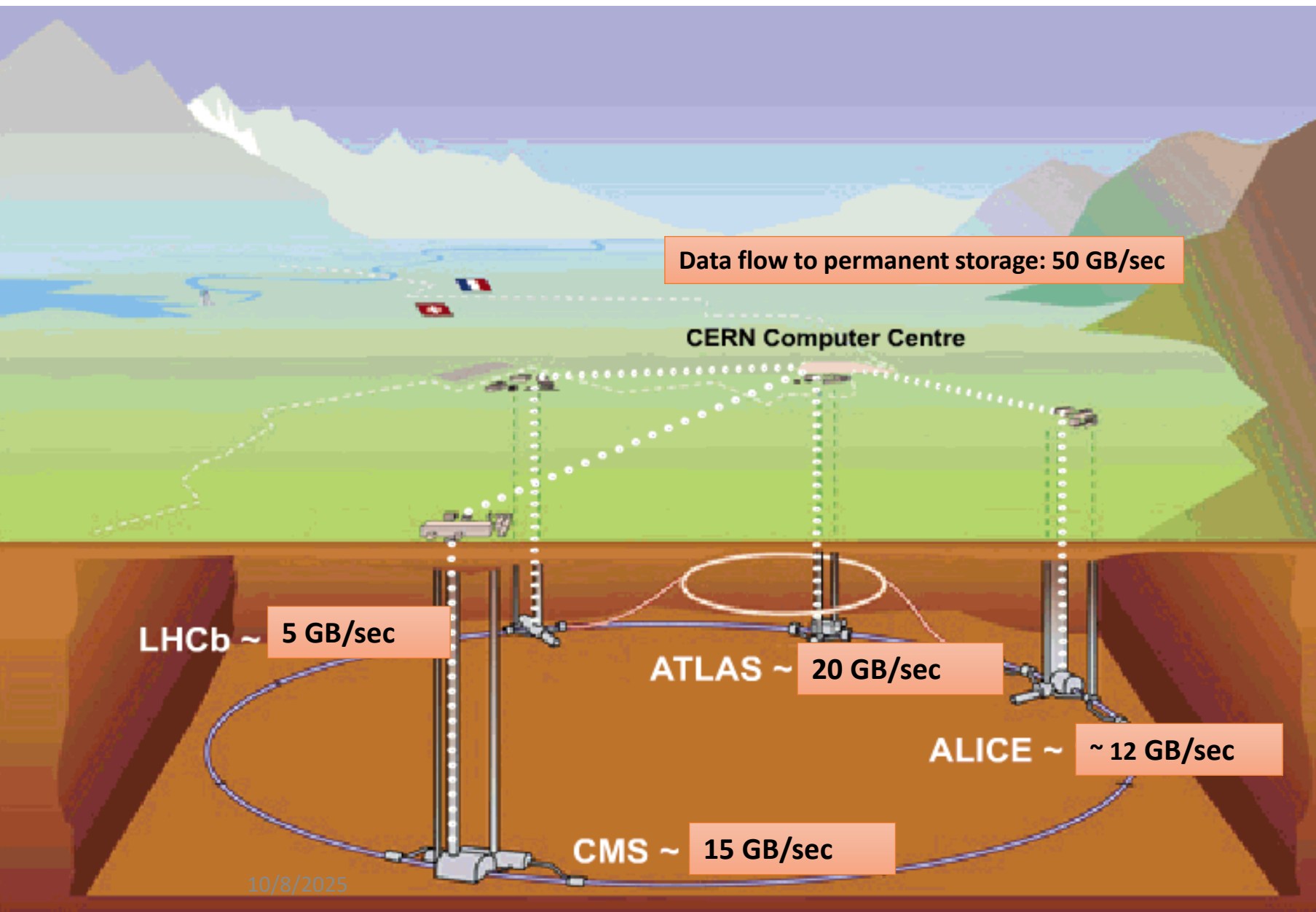


ALICE
(A Large Ion
Collider
Experiment)



ATLAS
(A Toroidal
LHC
ApparatuS)

Сбор и архивирование данных в ЦЕРН



Параметры детектора АТЛАС

Энергия центра масс 14 TeV

Частота столкновений пучков 40 MHz

Светимость :

- начальная: $10^{31} \text{ см}^{-2}\text{с}^{-1}$
- низкая: $2 \cdot 10^{33} \text{ см}^{-2}\text{с}^{-1}$
- целевая: $10^{34} \text{ см}^{-2}\text{с}^{-1}$

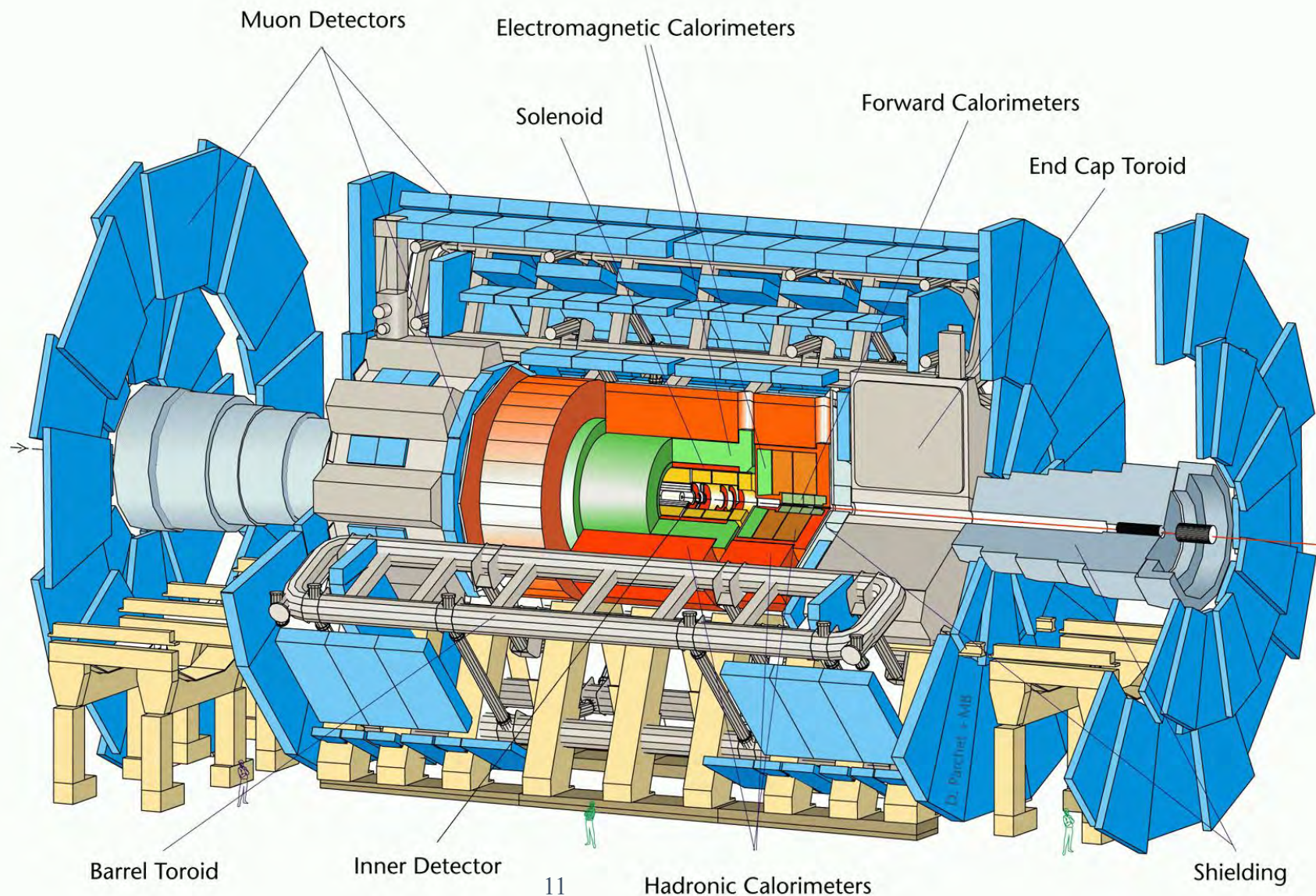
Вес 7000 тонн,

Диаметр 24м,

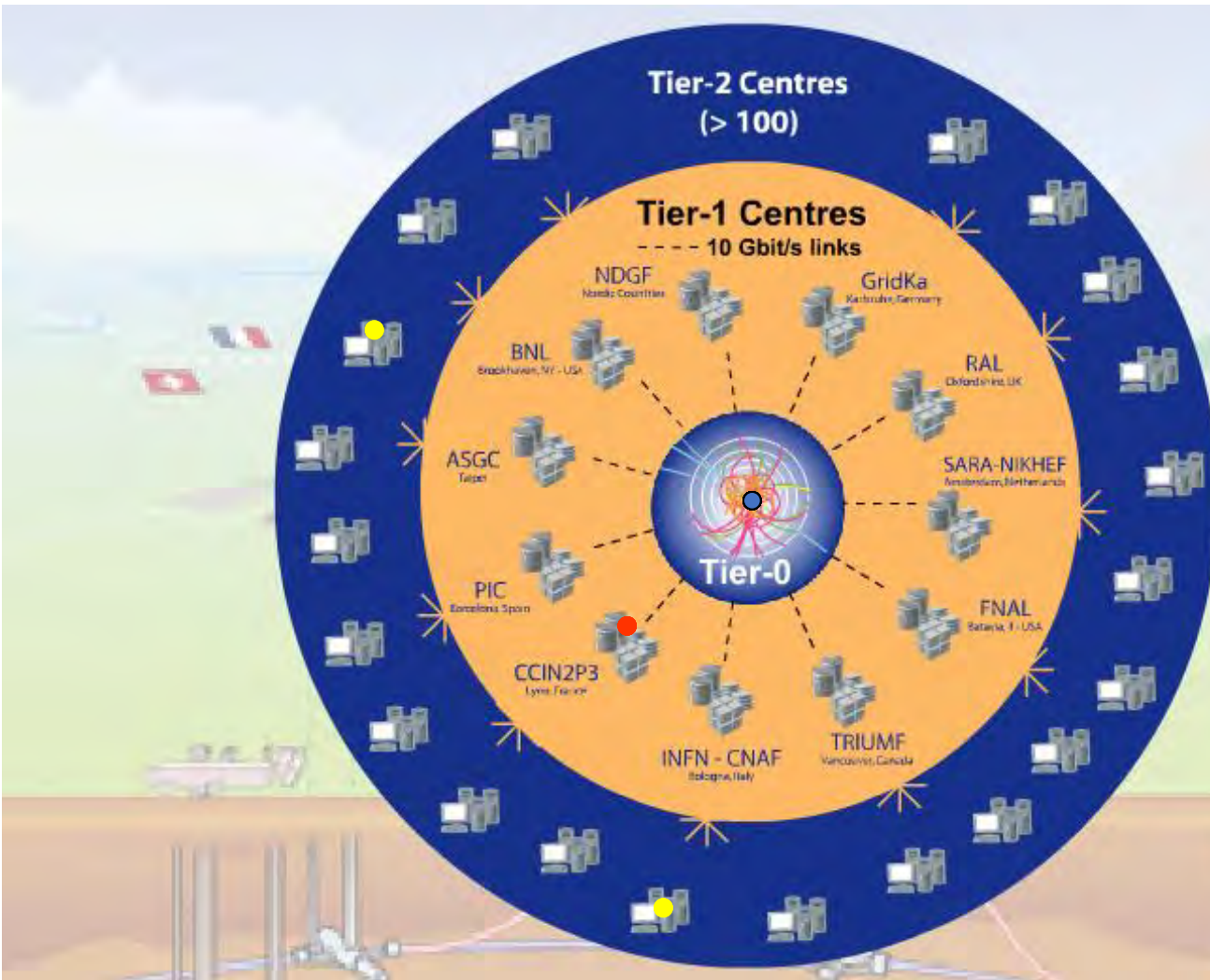
Длина 46м

Количество регистрирующих каналов 140 000 000

D712/mb-26/06/97



Tier Structure of GRID Distributed Computing: Tier-0/Tier-1/Tier-2



Tier-0 (CERN):

- accepts data from the CMS Online Data Acquisition and Trigger System
- archives RAW data
- the first pass of reconstruction and performs Prompt Calibration
- data distribution to Tier-1

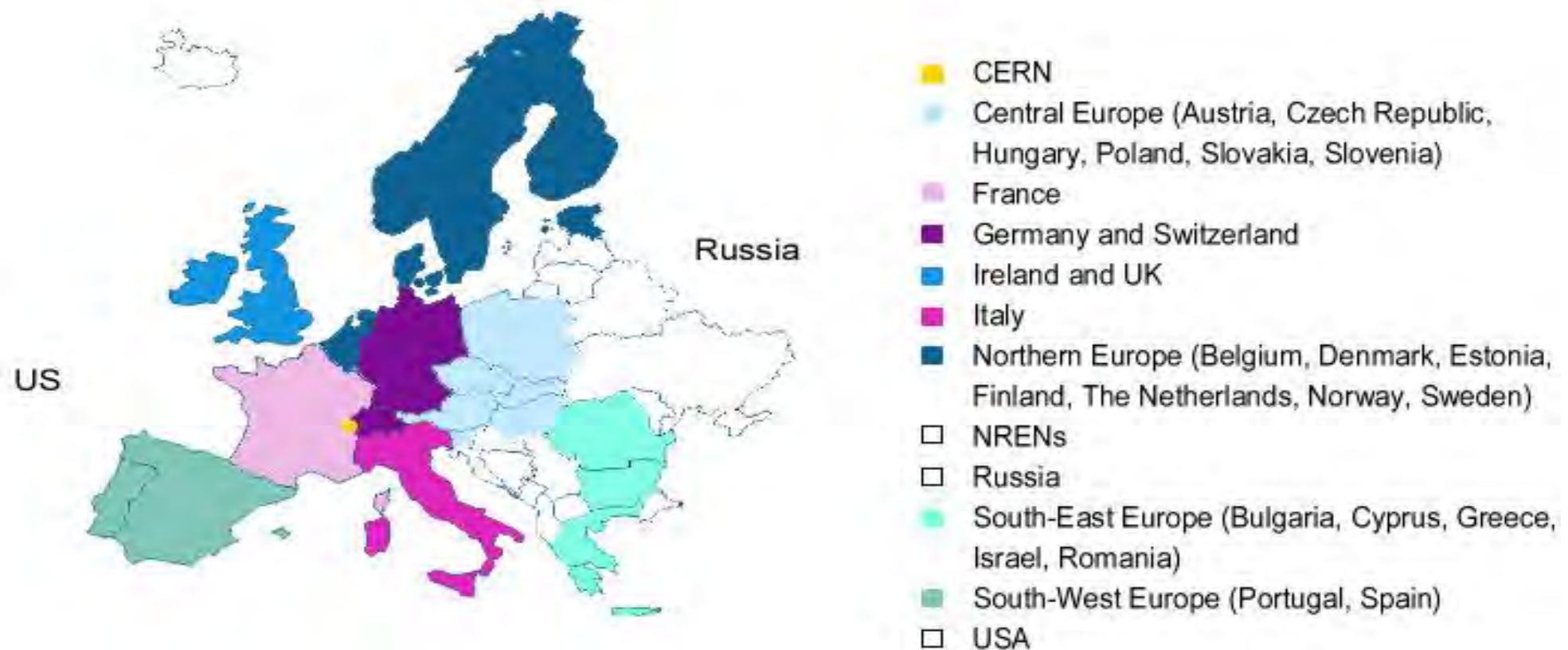
Tier-1 (11 centers):

- receives a data from the Tier-0
- data processing (re-reconstruction, skimming, calibration etc)
- distributes data and MC to the other Tier-1 and Tier-2
- secure storage and redistribution for data and MC

Tier-2 (>200 centers):

- simulation
- user physics analysis

Проект EGEE - Enabling Grids for E-sciencE



Проект EGEE - Enabling Grids for E-sciencE направлен на создание международной инфраструктуры, основанной на технологиях грид. Проект выполняется консорциумом из **70 институтов в 27 странах**, объединенных в региональные гриды.

Russian Data Intensive Grid infrastructure (RDIG)

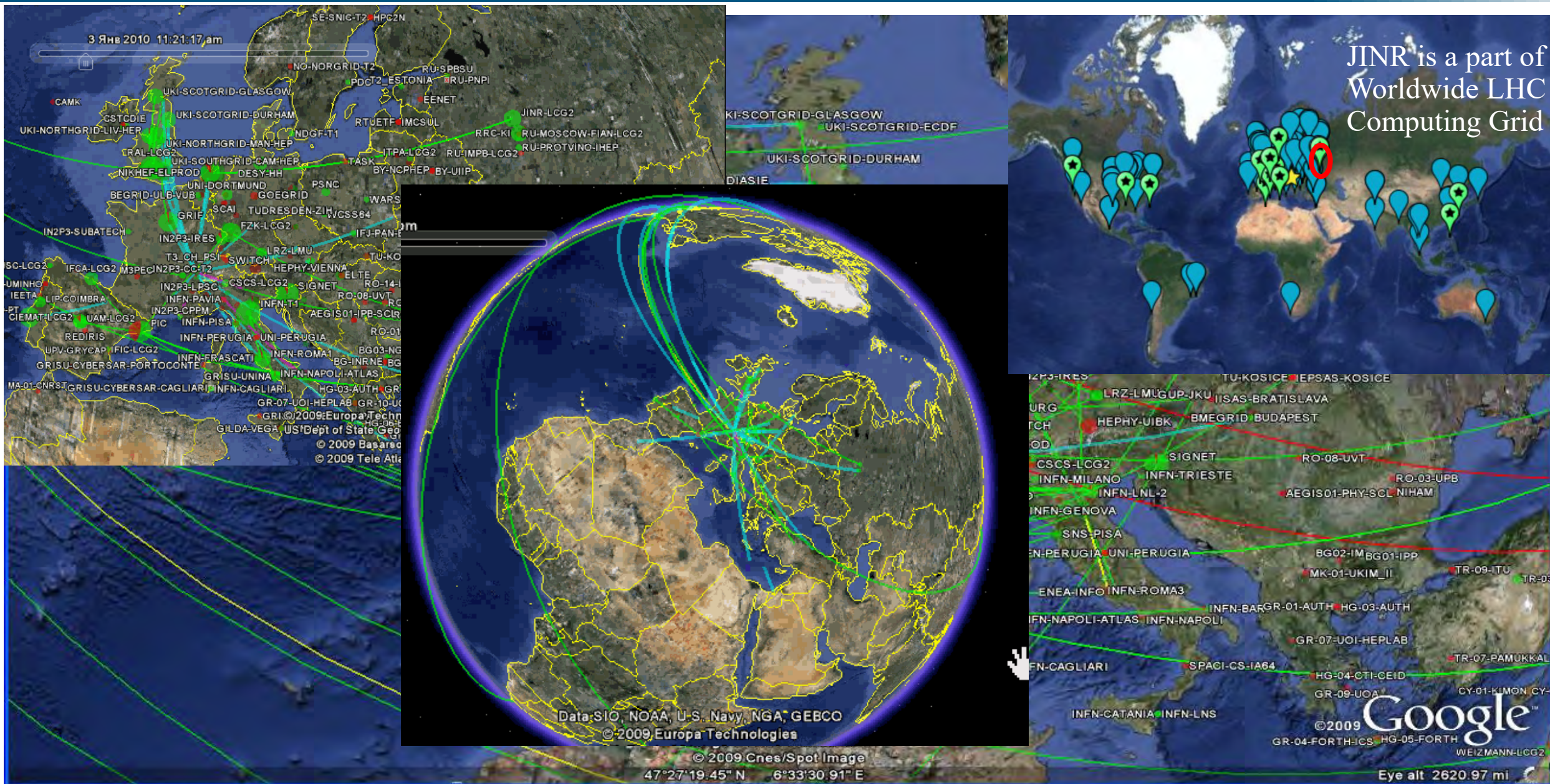
Консорциум Российский ГРИД для интенсивных операций с данными (РДИГ) был создан в 2003 году для активного участия в распределенной обработке данных экспериментов на Большом адронной коллайдере в рамках научной коллаборации WLCG (Worldwide LHC Computing Grid). Созданная инфраструктура RDIG имеет огромное значение для эффективного участия ученых России в научной программе экспериментов на LHC.



Ресурсные центры RDIG:

- ITEP
- JINR-LCG2 (Dubna)
- RRC-KI
- RU-Moscow-KIAM
- RU-Phys-SPbSU
- RU-Protvino-IHEP
- RU-SPbSU
- Ru-Troitsk-INR
- ru-IMPB-LCG2
- ru-Moscow-FIAN
- ru-Moscow-MEPHI
- ru-PNPI-LCG2 (Gatchina)
- ru-Moscow-SINP
- **Kharkov-KIPT (UA)**
- **BY-NCPHEP (Minsk)**
- **UA-KNU**
- **UA-BITP**

The Worldwide LHC Computing Grid (WLCG)



Worldwide LHC Computing Grid (WLCG)

WLCG – Международное сотрудничество по распространению и анализу данных БАК.

Объединяет компьютерные центры по всему миру, предоставляющие вычислительные ресурсы и ресурсы хранения данных, в единую инфраструктуру, доступную всем физикам БАК.

Компьютерные технологии WLCG позволили физикам объявить об открытии бозона Хиггса (Нобелевская премия 2013 года).

Tier0 (CERN):

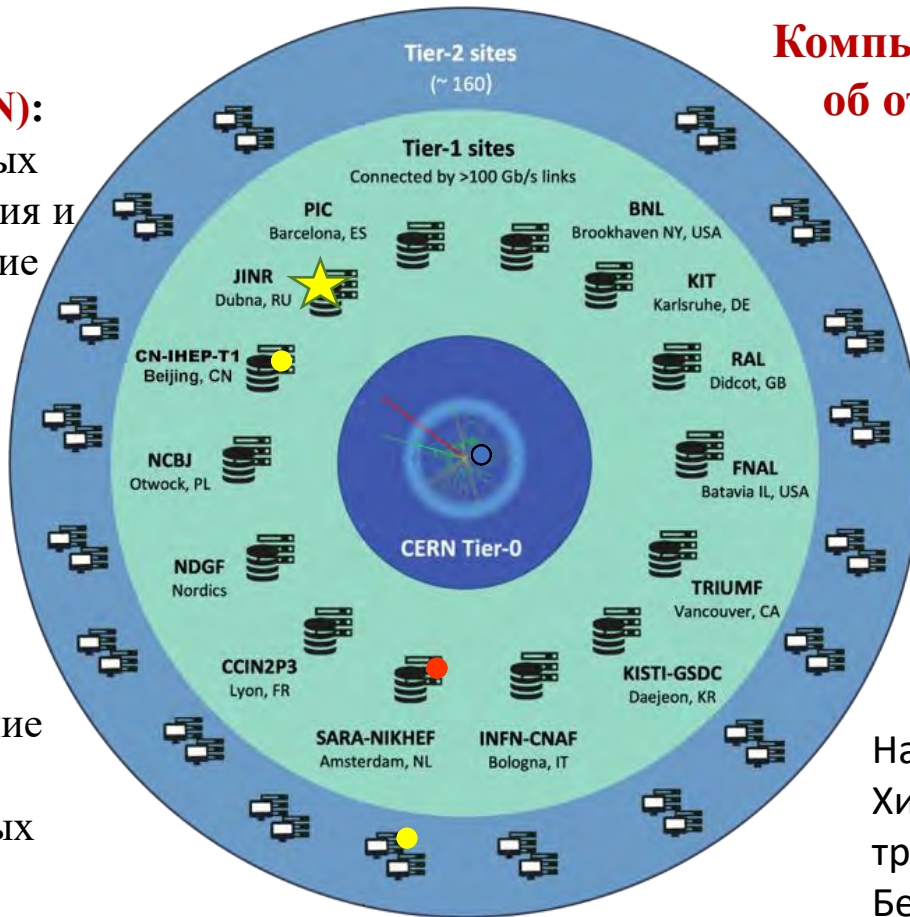
запись данных
реконструкция и
распределение

Tier1:

постоянное
хранение,
повторная
обработка

Tier2:

моделирование
физический
анализ данных



42 страны

170 центров

~2 миллиона ядер

2 EB хранилища

> 2 млн задач/день

100-250 Gb/s связь



На торжестве по поводу получения Нобелевской премии за открытие бозона Хиггса директор ЦЕРНа Рольф Хойер прямо назвал грид-технологии одним из трех столпов успеха (наряду с ускорителем LHC и физическими установками). Без организации грид-инфраструктуры на LHC было бы невозможно обрабатывать и хранить колоссальный объем данных, поступающих с коллайдера, а значит, совершать научные открытия. Сегодня уже ни один крупный проект не осуществим без использования распределенной инфраструктуры для обработки данных.

Миссия проекта WLCG — предоставить глобальные вычислительные ресурсы для хранения, распространения и анализа примерно 400 ПБ данных, ожидаемых каждый год от Большого адронного коллайдера.



M.G. Mesheryakov
(17.09.1910 - 24.05.1994)



N.N. Govorun
(18.03.1930 - 21.07.1989)

ЛИТ (ранее ЛВТА) ОИЯИ была создана в 1966 году.
Главные направления деятельности Лаборатории связаны с развитием сетевой, компьютерной, информационной инфраструктуры, а также с математической и программой поддержкой научной программы ОИЯИ.



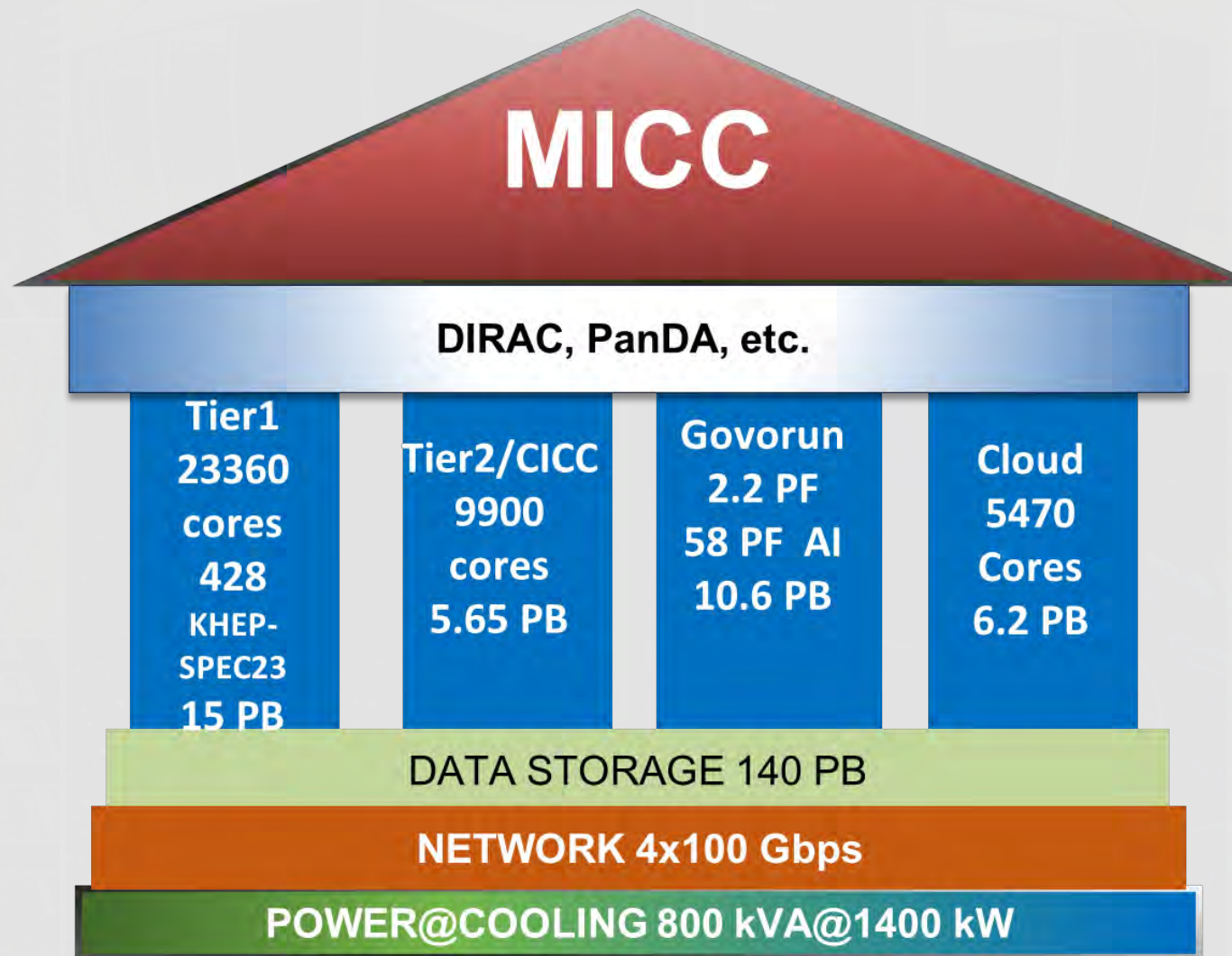
Объединенный институт ядерных исследований

Лаборатория информационных технологий им. М.Г. Мещерякова

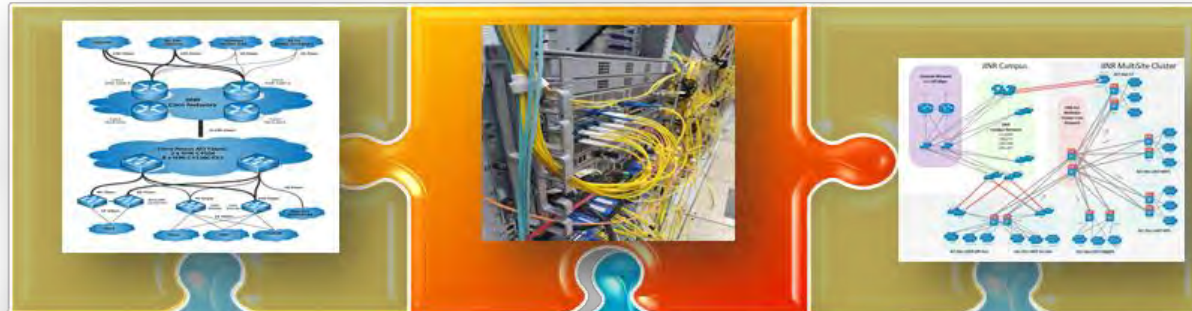


Научная ИТ экосистема:
согласованное развитие взаимосвязанных
ИТ-технологий и вычислительных методов

Многофункциональный информационно-
вычислительный комплекс ОИЯИ



MICC Power @ Cooling @ Network



Wide Area Network 4x100 Gbps
Cluster Backbone 4x100 Gbps
Campus Backbone 2x100 Gbps

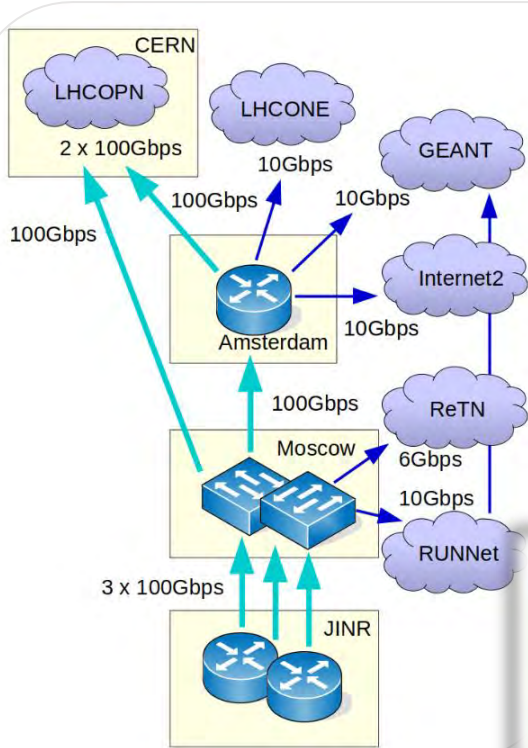


Dry chillers
In-Row systems
Total cooling 1400 kW



Uninterruptible power supplies (UPS) 8x300 kVA
Diesel-generator units (DGU) 2x1500 kVA

Система внешних каналов и локальной сети ОИЯИ



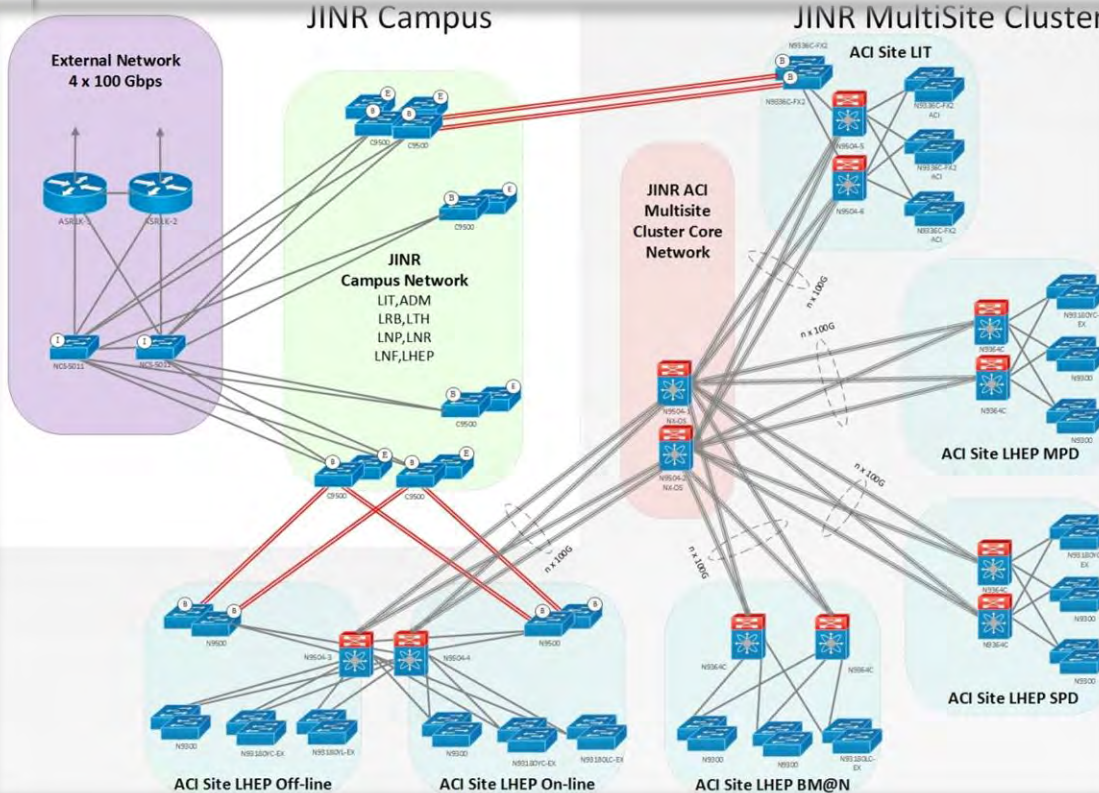
ОИЯИ- Москва 4x100 Гбит/сек

ОИЯИ - ЦЕРН - 100 Гбит/сек и ОИЯИ - Амстердам 100 Гбит/сек для сетей LHCOPN, LHCONE, GEANT

Прямые каналы связи до 100 Гбит/сек для связи с российскими сетями

Локальная вычислительная сеть 2x100 Гбит/сек

Распределенная многоузловая кластерная сеть между ЛИТ и ЛФВЭ 4x100 Гбит/сек



Локальная сеть ОИЯИ:

Пользователей - 5758

из них сотрудников ОИЯИ - 5570

не сотрудников - 188

Сетевых элементов - 13395

IP адресов ipv4 - 22430

IP адресов ipv6 - 1421

Удалённый доступ - 920

Электронные библиотеки - 1163

EDUROAM - 140

Email адреса @jinr.int - 4786

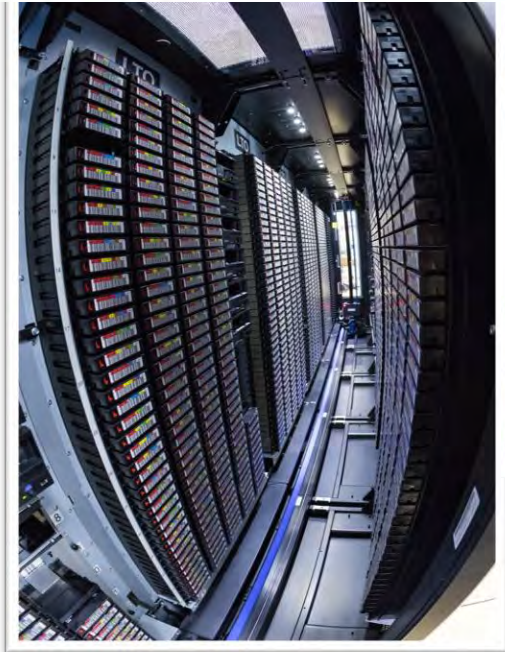
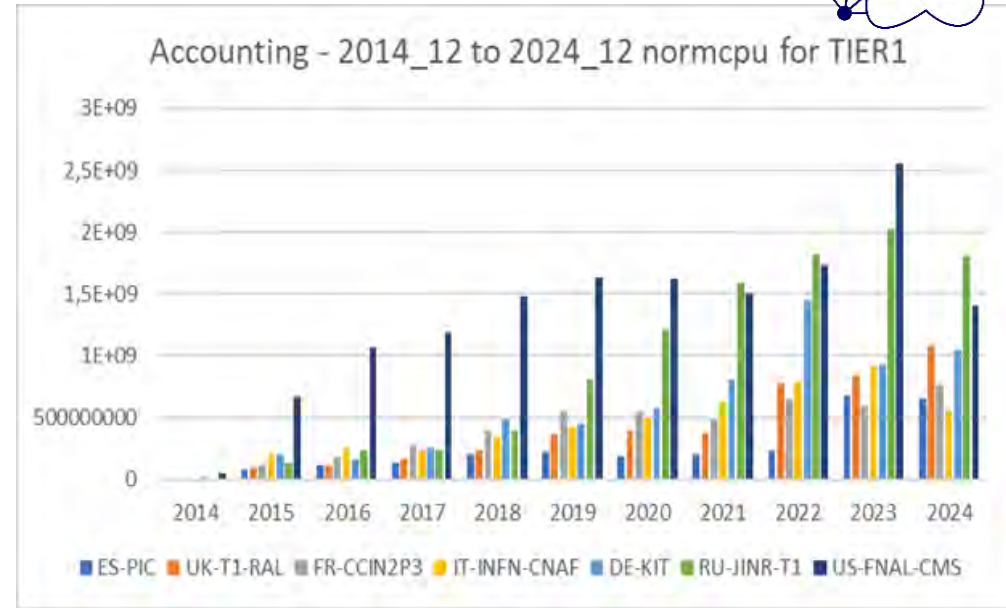
Сетевой трафик в 2024 году

43,45 PB - входящий

27,28 PB - исходящий



Tier1 для проекта CMS в ОИЯИ



- 20500 ядер
- 360 kHS23
- 15.5 ПБ диски
- 100 ПБ лент.робот
- 100% надежность и доступность

Tier1 CMS 2024

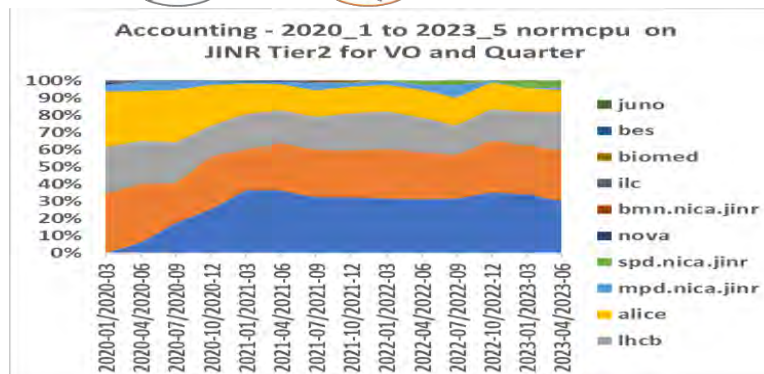
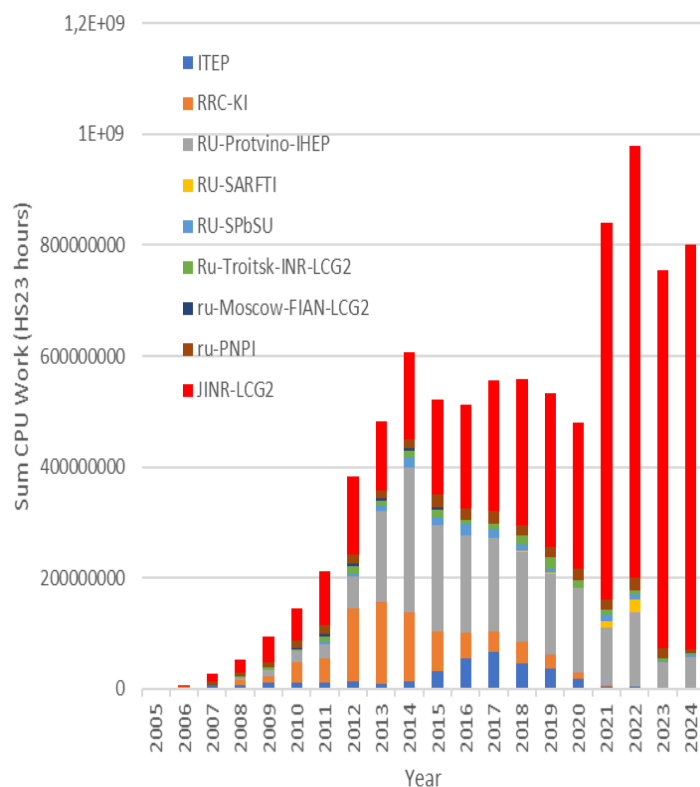
RU-JINR-T1	1,888,913,532	24 %
US-FNAL-CMS	1,665,321,019	21 %
UK-T1-RAL	1,132,299,978	14 %
DE-KIT	1,097,134,332	13 %
FR-CCIN2P3	807,778,212	10 %
ES-PIC	671,291,604	8 %
IT-INFN-CNAF	618,136,609	7 %



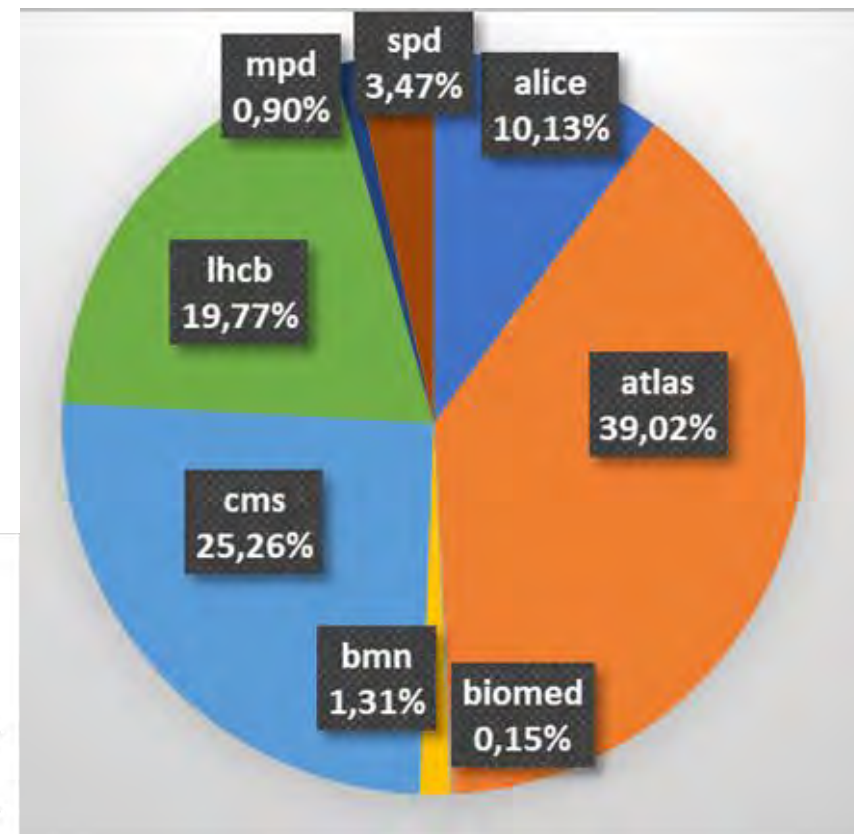
Tier2 at JINR

Tier2 at JINR provides computing power and data storage and access systems for the majority of JINR users and user groups, as well as for users of virtual organizations (VOs) of the grid environment (LHC, NICA, FAIR, etc.).

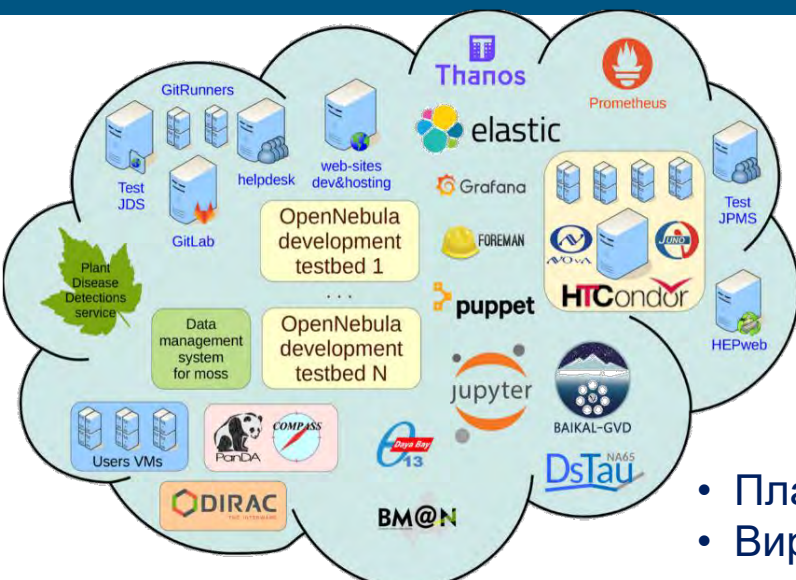
Accounting of RDIG Tier2's and years 2005-2024



Use of the JINR Tier2 site by virtual organizations within grid projects



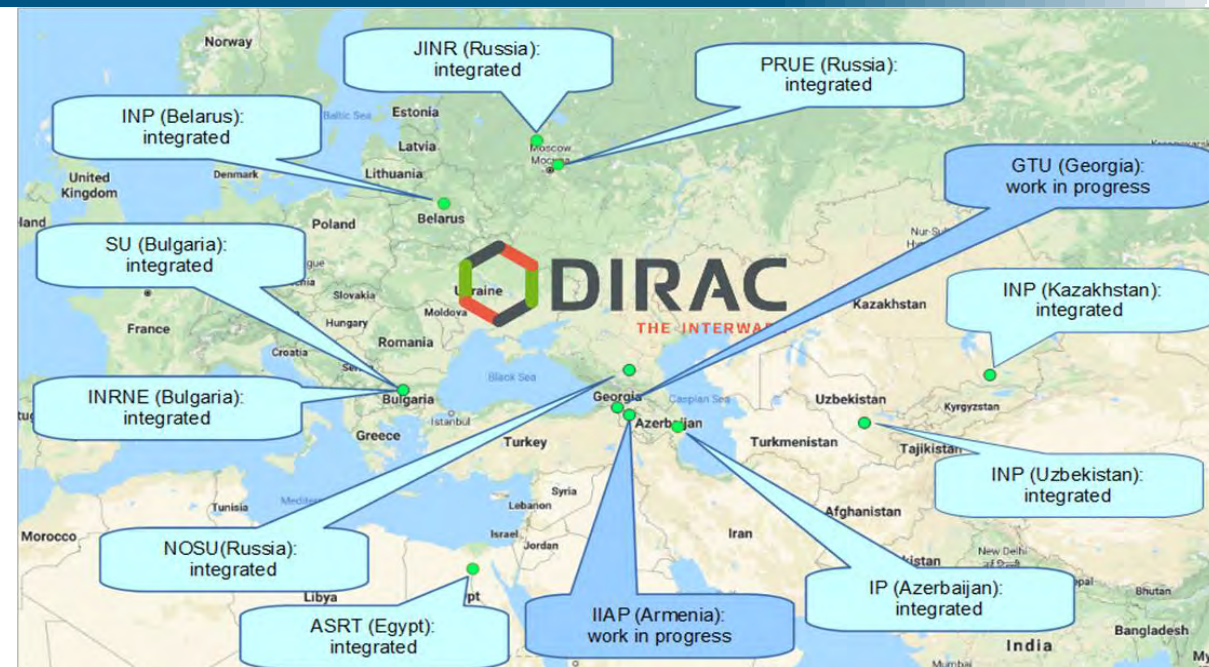
Облачная среда ОИЯИ и стран-участниц ОИЯИ



Распределенная информационно-вычислительная среда (DICE) на базе DIRAC, объединяющая облака организаций государств-участников ОИЯИ.

- Платформа – OpenNebula,
- Виртуализация – KVM

- Хранилище (Local disks, Ceph)
- Ресурсы: ~ 5,152 CPU ядер; 5 PB ceph-хранилище



- Вычислительные ресурсы для нейтринных экспериментов:
- Виртуальные машины для пользователей ОИЯИ
- Испытательные стенды для исследований и разработок в области ИТ
- Сервисы системы обработки данных эксперимента COMPASS
- Система управления данными программы по контролю за загрязнением воздуха Европы (UNECE ICP Vegetation)
- Система диагностирования болезней агрокультур посредством использования современных методов машинного обучения
- Сервис для визуализации данных, Gitlab и некоторые другие.
- Распределенная информационно-вычислительная среда на базе DIRAC (DICE), которая интегрирует облака организаций стран-членов ОИЯИ и т.д.



Суперкомпьютер «Говорун»



- Гиперконвергентная программно-определяемая система
- Многоуровневая система хранения и обработки данных (**1.6 PB + 10 PB + 100 PB**)
- Общая пиковая производительность:
 - **2.2 PFlops** двойной точности
 - **26 PFlops** для задач ИИ
- GPU компонента **NVIDIA V100+A100+H100**
- CPU компонента на **жидкостном охлаждении** РСК «Торнадо»
- Наиболее энергоэффективная система в России (**PUE = 1,06**)
- Скоростью чтения/записи данных **>300 ГБ/с**



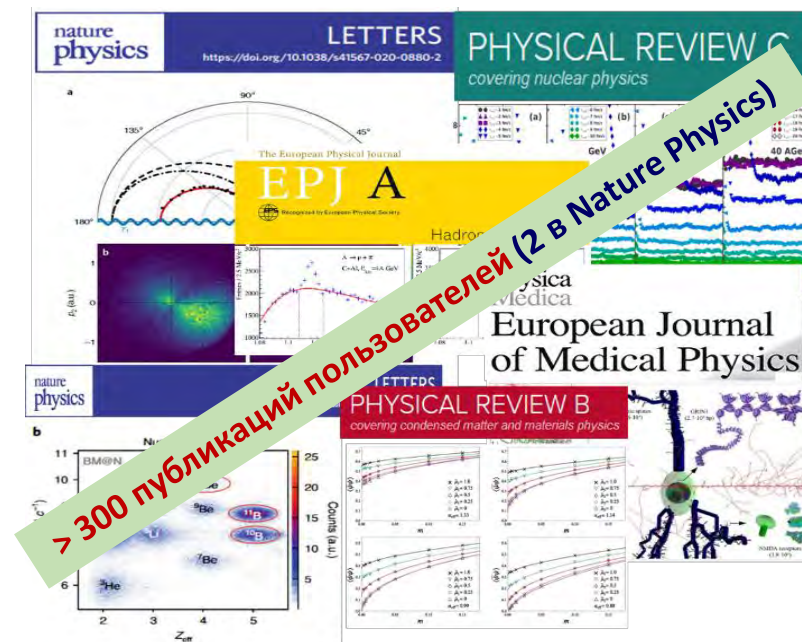
Ключевые проекты, использующие ресурсы СК «Говорун»:

- мегапроект NICA,
- расчеты решеточной квантовой хромодинамики,
- расчеты свойств атомов сверхтяжелых элементов,
- исследования в области радиационной биологии,
- расчеты радиационной безопасности установок ОИЯИ.



Общее количество пользователей: **347** включая из стран-участниц ОИЯИ (Армения, Беларусь, Вьетнам, Египет, ЮАР)

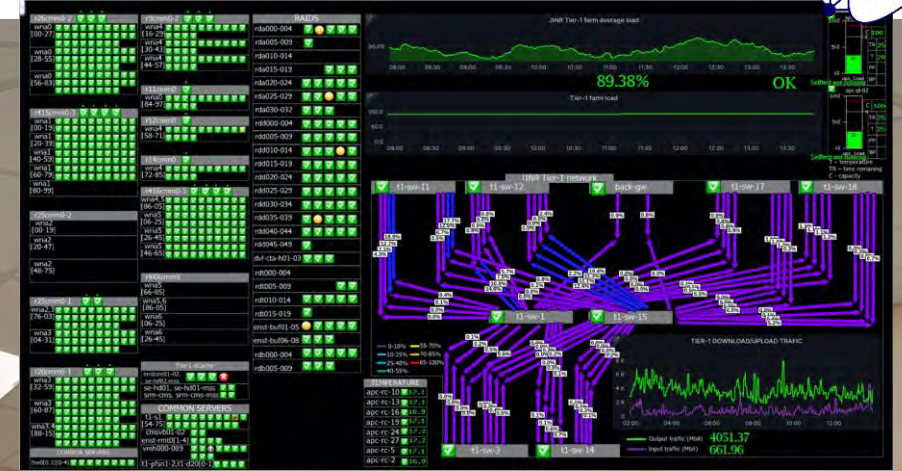
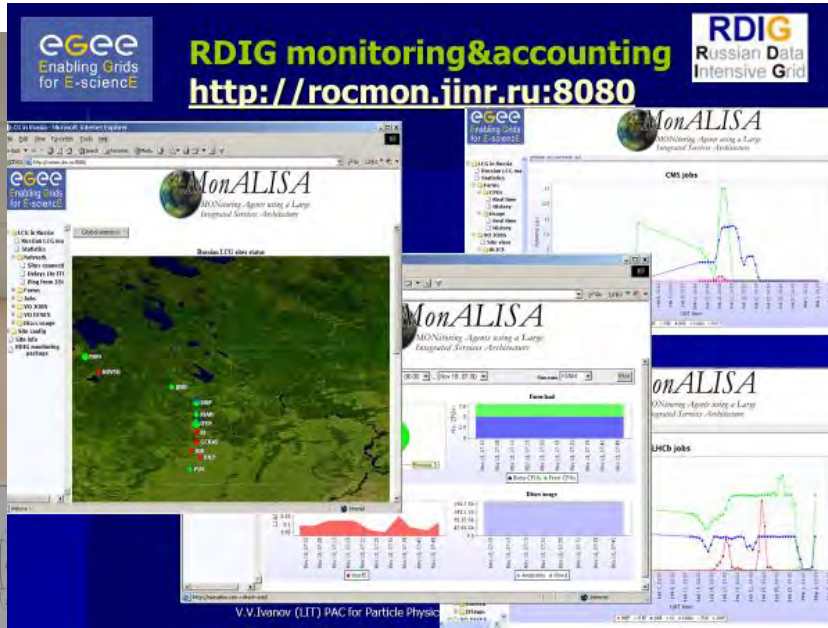
СК «Говорун» включен в единую суперкомпьютерную инфраструктуру на базе Национальной исследовательской компьютерной сети России (НИКС).



Российская премия DC Awards 2020 за «Лучшее ИТ-решение для ЦОДа»



Monitoring



Многоуровневая система хранения данных



- Ограниченное количество данных и краткосрочное хранилище - для хранения самой операционной системы, временных пользовательских файлов
- Распределенная глобальная система AFS – для хранения домашних каталогов пользователей и программного обеспечения
- dCache является традиционным для Grid – для хранения больших объемов данных (в основном для экспериментов на LHC) на среднесрочный период
- EOS распространяется на все ресурсы МИБК – для хранения больших объемов данных на среднесрочный период. В настоящее время EOS используется для хранения BM@N, MPD, SPD, BaikalGVD и др.
- Ленточные роботизированные системы – для хранения больших объемов данных на длительный период (CMS. BM@N, MPD, SPD, JUNO).



На СК “Говорун” разработана и внедрена специальная иерархическая система обработки и хранения данных с программно-определяемой архитектурой.

По скорости доступа к данным различают следующие уровни:

- горячие данные(LUSTRE),
- теплые данные (EOS)
- холодные данные (TAPE)

Ускорительный комплекс NICA



MPD
Многоцелевая
установка



Нуклотрон



Синхротрон-
бустер



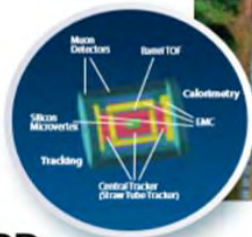
Линейный
ускоритель
легких
ионов



Линейный
ускоритель
тяжелых
ионов



Коллайдер



SPD
Установка
спиновой физики

Инфраструктура
ARIADNA
для прикладных исследований



BM@N
Эксперимент
«Барионная
материя на
нуклотроне»



Распределенная гетерогенная среда на основе платформы DIRAC



Ресурсы ЛФВЭ

cephEOS

EOS

DDC cluster

NICA cluster

DAQ computing farm

Для выполнения расчётов, связанных с массовой генерацией и обработкой данных экспериментов мегасайнс проекта NICA, создана распределённая среда на основе платформы DIRAC Interware.

МИВК ЛИТ

cephCloud

MLIT EOS

MLIT CTA

dCache

JINR Cloud

Tier-1

Tier-2

Governor

Lustre
Ultra-fast storage

Облака

NOSU	INP
IPANAS	ASRT
INP	PRUE
INRNE	SU
	INP

MSC

Polytech

ПОЛИТЕХ
Санкт-Петербургский политехнический университет Петра Великого

НИКС

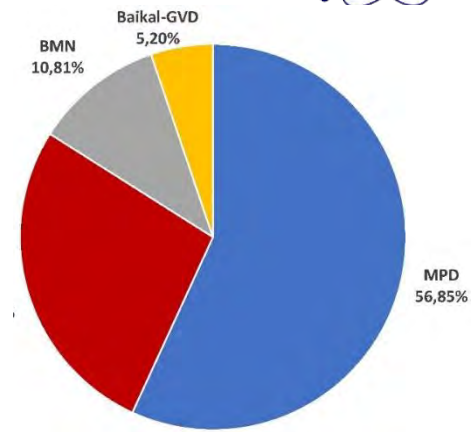
Другие/ресурсы коллабораций

IMDT
MONSIEUR ACADEMY OF SCIENCES
INSTITUTE OF MATHEMATICS AND DIGITAL TECHNOLOGY

MEPHI
МИФИ

UNAM
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

DIRAC используется для решения задач коллабораций всех трех экспериментов на ускорительном комплексе NICA, а также нейтринного телескопа Baikal-GVD



В 2023 году 8-й физический сеанс BM@N стал первым случаем в ОИЯИ, когда вся вычислительная инфраструктура, объединенная платформой DIRAC, использовалась для полной реконструкции необработанных экспериментальных данных.

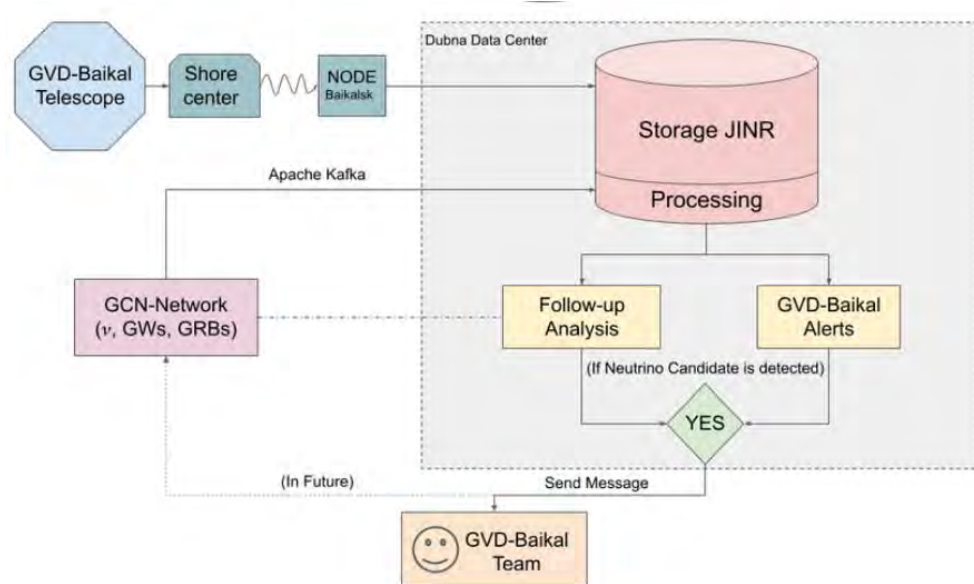
 31 k amount of raw files	 400 TB BM@N Run8 RAW data size	 12 data processing campaigns
 363 k jobs completed	 86 years total computation time	 ~3 PB data transferred

Суммарная статистика использования DIRAC для обработки 8-го физического сеанса BM@N

Neutrino Detector Baikal-GVD (Gigaton Volume Detector)

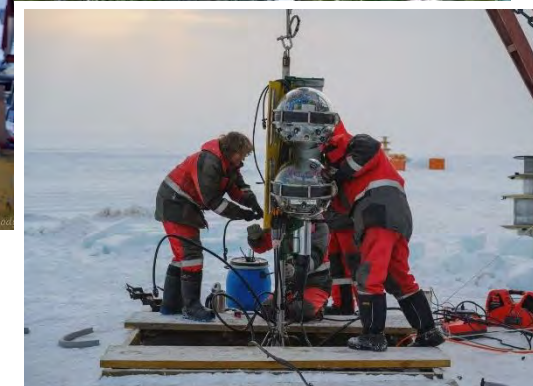
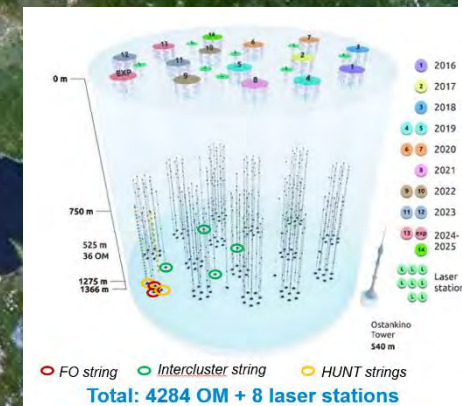


The **Baikal Neutrino Telescope** is a neutrino detector located in Lake Baikal at a depth of about 1300 m, the design capacity is 1 km³. Baikal-GVD is one of three operating large-scale neutrino telescopes in the world and along with the IceCube telescopes at the South Pole and KM3NeT in the Mediterranean, is part of the Global Neutrino Network (GNN). It is designed to register and study ultra-high energy neutrino fluxes from astrophysical sources.



JINR computing recourses:

- 2000 cores,
- ~ 2 PB disk storage,
- 2 PB tape library

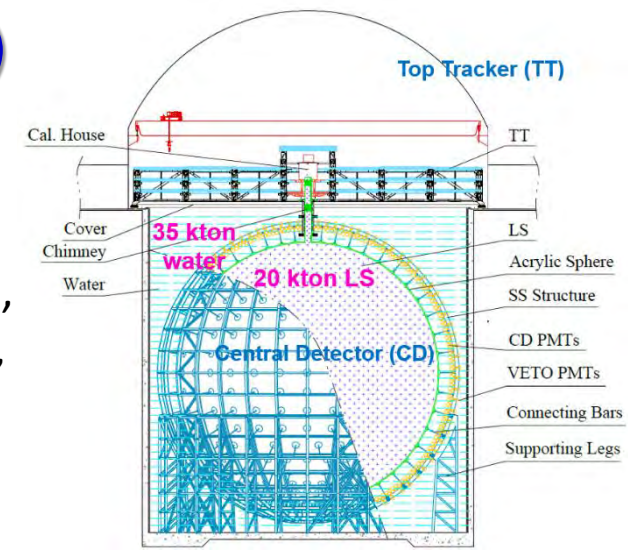


Currently, the total flow of experimental data is within 100 Mbit/s, but a transfer rate of 1 Gbit/s will be required for comfortable operation because of given the growth of the installation over the next three years.



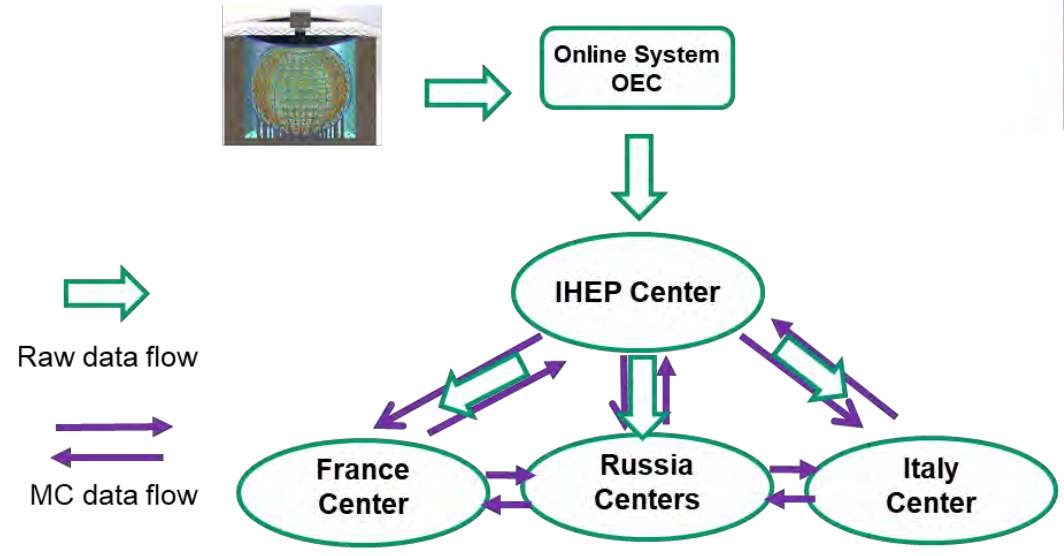
JUNO experiment (Guangzhou, south of China)

JUNO (Jiangmen Underground Neutrino Observatory) is a multipurpose Neutrino Observatory with a rich program in neutrino physics and astrophysics studying neutrinos over a large energy range, including solar, supernova, atmospheric, geo-neutrinos, proton decay, exotic searches.



- ❖ Five data centers joined: IHEP, CC-IN2P3, INFN-CNAF, JINR, MSU (~6000 CPU cores, >10PB disk)
- ❖ Raw data flows from Online to IHEP which then distributes to other centers
- ❖ KUP (Keep up production, 1st reconstruction)* and Calibration run in IHEP
- ❖ MC Simulation, PP(Physics Production, 2nd reconstruction) and Analysis run in data centers

Over the next 5 years, the minimum required data transmission channel capacity for Tier-1 centers is currently estimated at 2.5 Gbps.



raw data, prompt reconstruction, calibration, simulation, analysis

copy of raw data, rereco, simulation, analysis

JINR computing recourses:

- 1700 cores,
- ? PB storage

JUNO Data Storage Volume Forecast by Storage Type (PB) in Tier-1 Data Centers

	2025	2026	2027	2028	2029
Tape	2.79	5.83	11.17	13.96	19.54
Disk	2.98	6.26	7.49	7.49	7.49

Создание консорциума для IT_обеспечения исследовательской инфраструктуры класса «мегасайнс»



- В России реализуется программа масштабных научных проектов, важнейшей частью которых является развитие распределенных гетерогенных компьютерных систем (включая системы с экстремально-массивным параллелизмом) для обработки, хранения, анализа экспериментальных данных, разработка и внедрение эффективных методов, алгоритмов и программного обеспечения для моделирования физических систем, математической обработки и анализа экспериментальных данных, развитие методов машинного обучения, искусственного интеллекта, квантовых вычислений.
- Для решения этой масштабной задачи необходимо развивать распределенную компьютерную инфраструктуру, объединяющую ключевые научные и образовательные институты, участвующие в проектах мегасайнс - РДИГ-М. Созданный консорциум на базе ОИЯИ, НИЦ Курчатовский институт, ИСП РАН должен стать ядром для IT-обеспечения исследовательской инфраструктуры класса «мегасайнс».

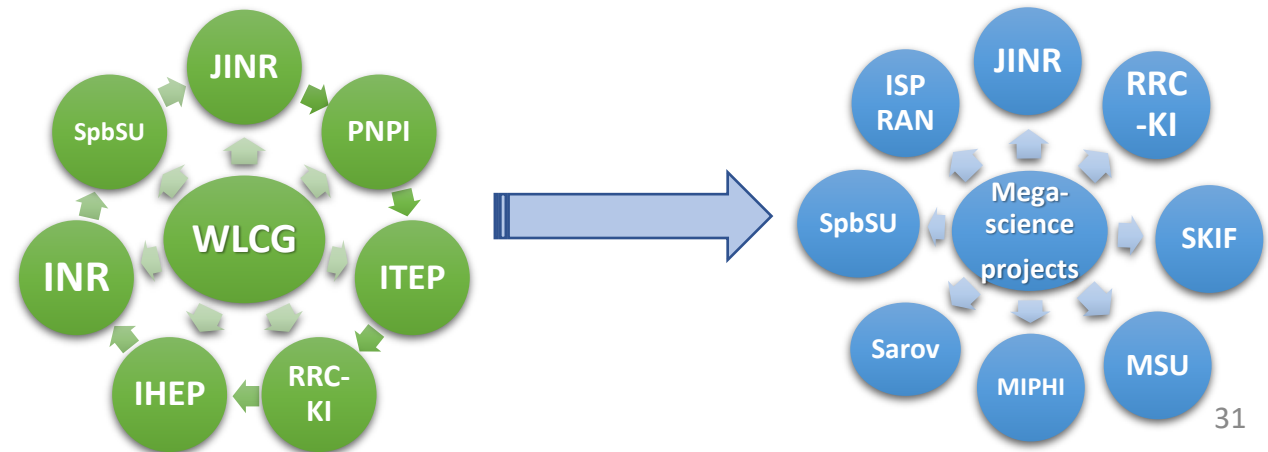


Figure 1: Energy as a function of parameters. (a) 3D surface plot of energy vs. parameters r and T , with corresponding 2D histograms. (b) Crystal structures of BaIrO_3 and Na_2IrO_2 , and a network diagram of the system. (c) Simulation results showing particle distributions and energy spectra. (d) Folding@home simulation results showing protein folding progress and energy landscape.

In 2024 the MLIT staff published
>200 scientific publications,

>100 articles within international collaborations

Numerical modeling of complex physical systems

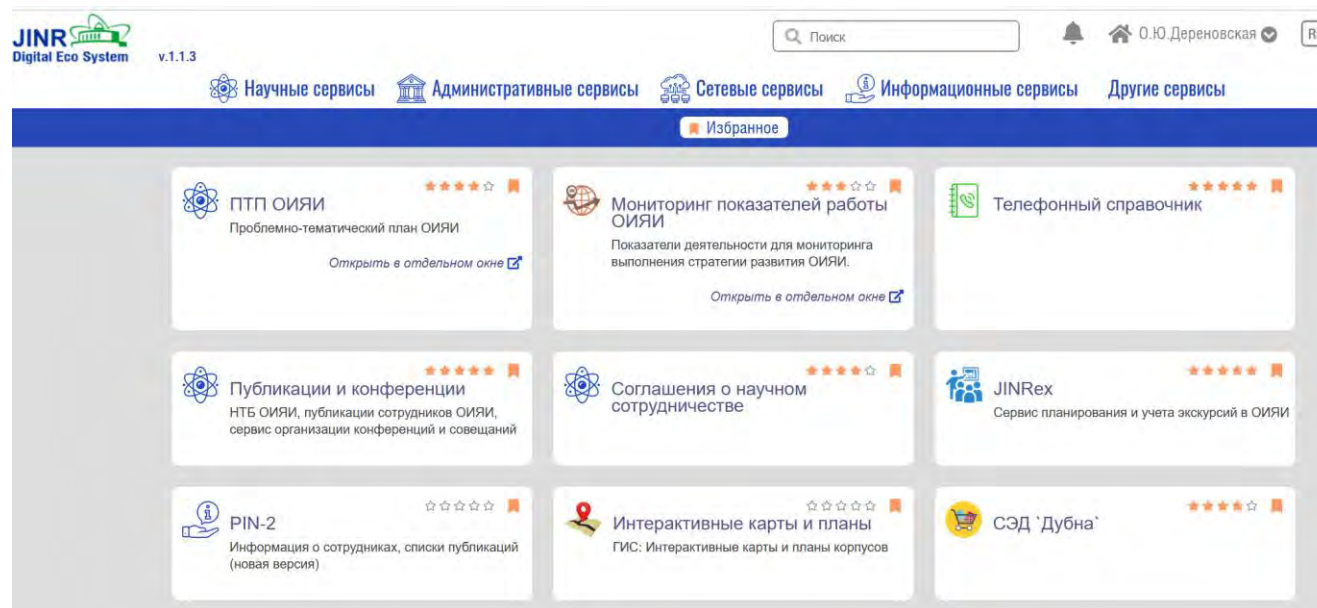
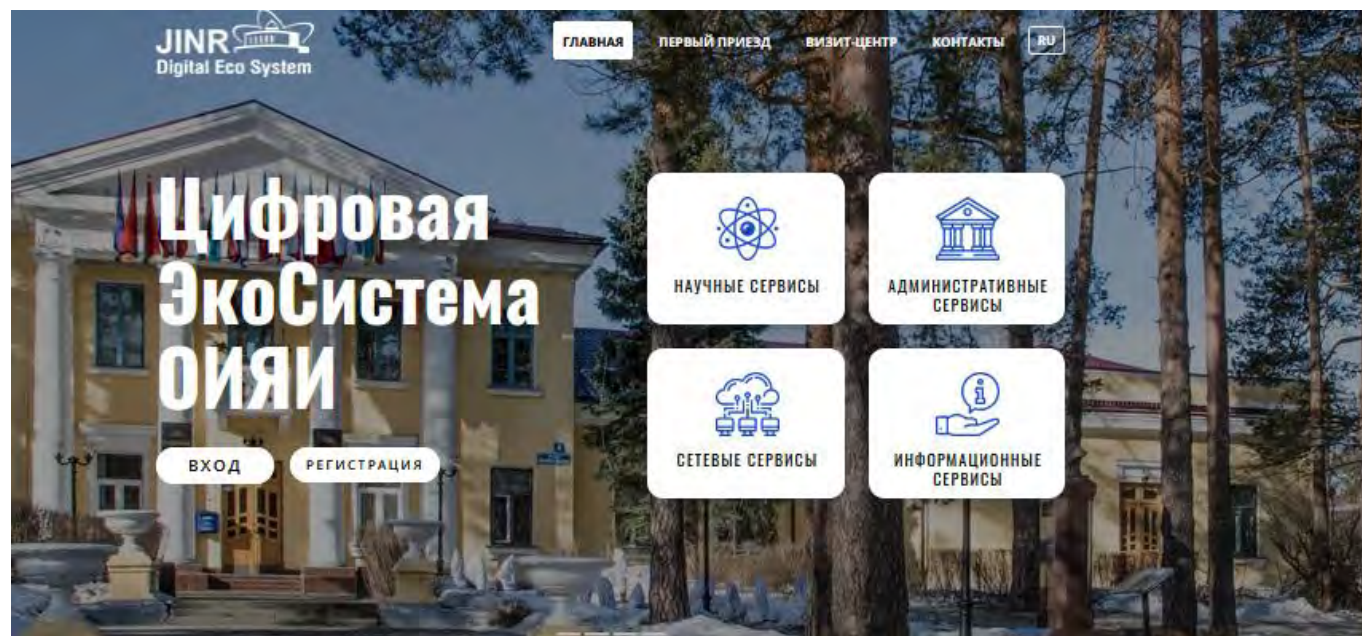
Experimental data processing and analysis

Big Data

Machine and Deep learning

Computer algebra

Quantum computing



Цифровая ЭкоСистема ОИЯИ

единое окно в цифровую среду ОИЯИ

- интеграция существующих и перспективных сервисов поддержки **научной, административной и социальной деятельности**, а также сопровождение инженерной и IT-инфраструктур Института.
- **Интегральный** Личный кабинет сотрудника ОИЯИ
- Уведомления в личном кабинете
- Часть ресурсов доступна для незарегистрированных пользователей
- Актуальность. Информация обновляется **владельцами сервисов** оперативно и регулярно
- Удобный интерфейс администраторов сервисов
- Поддерживает двуязычность: русский и английский языки
- Мобильная версия системы

Основное за 2024 год:

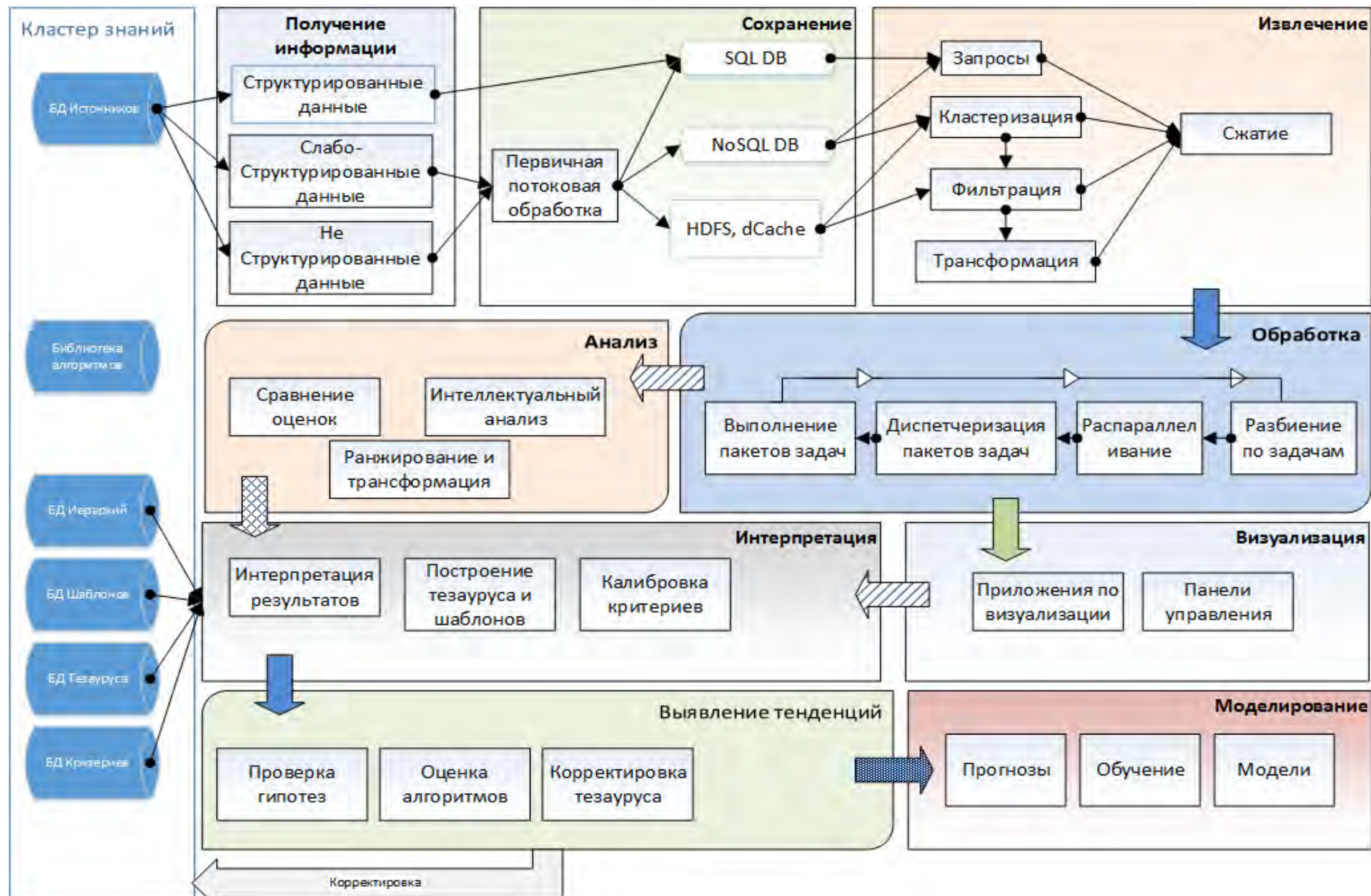
- ПИН-2: глубоко переработан и введён в опытную эксплуатацию.
- Сервисы для совместной работы (календарь, сервис управления документами)
- Репозиторий публикаций сотрудников ОИЯИ: оперативное наполнение (1576 за 2024 г., 8733 с 2020 г.), введён в опытную эксплуатацию.
- Системные сервисы: прототип шины данных.

Цифровая ЭкоСистема ОИЯИ

Планы на 2025 год:

- Ввод в эксплуатацию ПИН-2, интегрированного с ЦЭС и репозиторием публикаций.
- Подготовка к переносу процессов закупочной деятельности в систему документооборота, создаваемую ДРЦС.
- ГИС: средства учёта помещений и резервирования рабочих мест, организация рабочего пространства на базе цифрового двойника здания.
- ИИ-помощник по цифровой экосистеме.
- Система поддержки пользователей ЦЭС полного цикла.
- Среда для хранения и управления данными сервисов ЦЭС.
- **Переход от активности к проекту**

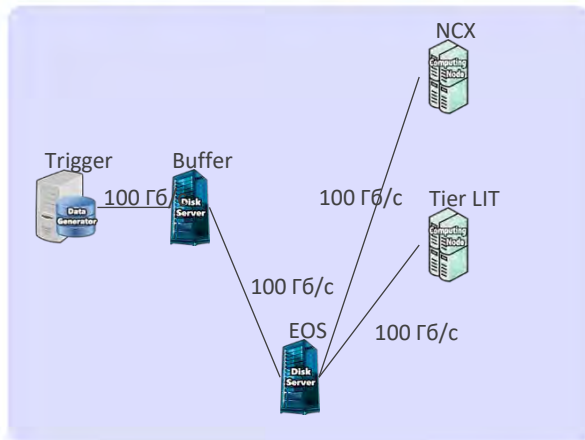
Принципиальная схема системы сбора, хранения и анализа Больших данных



Программный комплекс для создания цифровых двойников распределенных центров сбора, хранения и обработки данных

Цифровой двойник РЦОД

Построение инфраструктуры центра сбора, хранения и обработки данных



Примеры применения для экспериментов комплекса NICA

- Построение цифрового двойника вычислительной инфраструктуры эксперимента VM@N.
- Построение цифрового двойника вычислительной системы онлайн-фильтра данных эксперимента SPD.



Веб-сервис



Результаты эксперимента Test 1

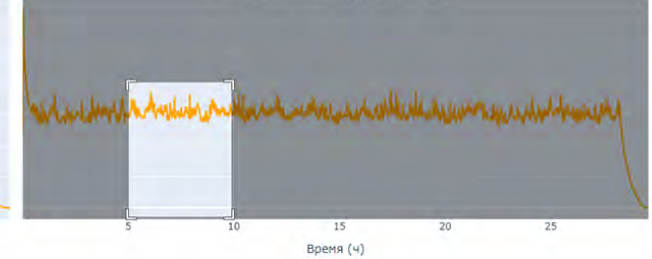
Выберите вкладку для просмотра результатов

Хранилища данных Вычислительные компоненты Каналы связи Очереди задач Распределения файлов

Нагрузка на канал связи compute0



Нагрузка на канал связи compute2



Особенности

✓ Универсальность.

✓ Учитываются важные функциональные параметры распределенных центров:

- характеристики оборудования;
- характеристики потоков данных и задач;
- вероятности сбоев, отказов и изменений в производительности оборудования и других процессов, происходящих в системе.

✓ Результаты работы отличаются от результатов работы существующего распределенного центра не более, чем на 20%.

Development of the system for training and retraining IT specialists



Training courses, master classes and lectures

**MLIT staff and
leading scientists from JINR and its Member States**

**Leading manufacturers of modern computing
architectures and software**

**Parallel
programming
technologies**

OpenMP

MPI



**Tools for debugging and
profiling parallel
applications**



**Work with applied software
packages**

COMSOL
MULTIPHYSICS

Wolfram Mathematica



**Frameworks and
tools for ML/DL tasks**



TensorFlow

NumPy



scikit-learn



**Quantum
algorithms,
quantum
programming and
quantum control**



Активность ЛИТ ОИЯИ в направлении подготовки кадров

Большую роль в подготовке высококвалифицированных кадров играет университет “Дубна”. ЛИТ ОИЯИ осуществляет техническую помощь для этой цели: сетевая инфраструктура, МИВК, программная среда для различных приложений.

На ресурсах ЛИТ организован образовательный процесс, проводится преддипломная практика, реализуются совместные проекты, многие магистры зачисляются в штат ЛИТ по совместительству, все аспиранты, которые работают по тематике ЛИТ зачисляются в штат, лучшие студенты получают стипендию ОИЯИ.

33 сотрудника ЛИТ преподают в университете Дубна,

52 выпускника университета Дубна работают в ЛИТ.

На базе ЛИТ создана базовая кафедра РИВС университета Дубна.

Регулярно организовываются студенческие ИТ-школы, в том числе в рамках крупных конференций (NEC в Варне, в Черногории, ММСР в Словакии, в Ереване, в Дубне, GRID в Дубне).

Различного уровня ИТ-школы были организованы в ОИЯИ, университете Дубна, в СОГУ, в РУДН, в МИФИ, в странах- участницах ОИЯИ.

Кафедра РИВС ИСАУ



Бакалавриат: Программная инженерия:

Магистратура: «Математическое моделирование и анализ данных»:

«Цифровые платформы и аналитика больших данных»

Подготовка высококвалифицированных специалистов, способных развивать цифровые платформы на основе интеграции распределенных, параллельных, гибридных вычислений, систем распределенного хранения огромных массивов данных. Цифровые платформы являются базой для крупных приложений во всех сферах деятельности, включая цифровую экономику, научные проекты класса мегасайенс.

Для реализации этой цели необходимо сочетание фундаментальных знаний в области математических методов и алгоритмов (включая методы глубокого машинного обучения, искусственного интеллекта) с практическими навыками применения современных информационных технологий (программная инженерия, параллельные, гибридные и распределенные вычисления, распределенные хранилища данных (DataLake), аналитика Больших данных, компьютерное зрение, видео-аналитика и другие перспективные технологии).

Среди основных направлений подготовки отдельное внимание уделено развитию модели компьютинга, программной платформы системы сбора, хранения, обработки и анализа данных экспериментов на установках класса мегасайенс (NICA, LHC, JUNO, Baikal-GVD...).



Филиал МГУ в Дубне

Подготовка специалистов в области теоретической и экспериментальной физики высоких энергий, релятивистской ядерной физики на базе ОИЯИ, для прикладных исследований и разработок в медицине, биологии и других областях с использованием ядерно-физических методов и информационных технологий.

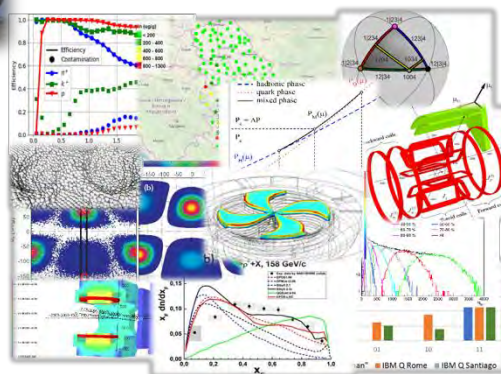
2024 год: Магистерские программы по направлению
03.04.02 Физика

- Физика элементарных частиц
- Фундаментальная и прикладная ядерная физика

16 декабря 2024: Получена лицензия на новую образовательную магистерскую программу по направлению
01.04.02 Прикладная математика и информатика

Сентябрь 2025:
первый набор
на **10 бюджетных мест**

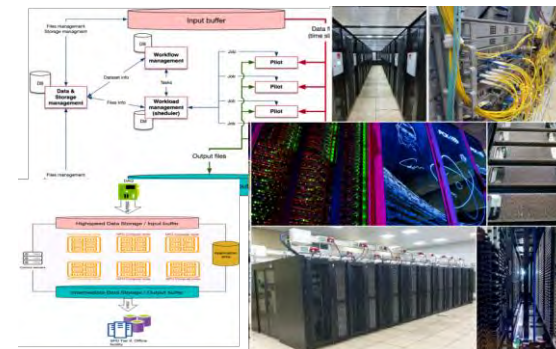
Методы и технологии обработки данных в гетерогенных вычислительных средах



Математическое моделирование, численные методы и комплексы программ



Глубокое машинное обучение и аналитика больших данных



Компьютинг (программные средства и модели) для проектов класса мегасайнс

Образовательная траектория подготовки ИТ-специалистов



JINR School of Information Technology

for Russian speaking students



Involving young specialists in solving tasks that face JINR using high-performance and distributed computing, data analysis methods and algorithms, state-of-the-art information technologies

Stage 1

Autumn School

acquaintance with the directions of JINR scientific research

Stage 2

Spring School

presentation the results of joint work with the Institute's specialists



7-11 октября 2024

58 студентов из Российских университетов



57
студентов



21
университет

Россия

Москва :

- Государственный университет управления
- Национальный исследовательский университет «МЭИ»
- Национальный исследовательский ядерный университет «МИФИ»
- Российский университет дружбы народов им. Патриса Лумумбы
- Российский экономический университет им. Г.В. Плеханова

Санкт-Петербург :

- Санкт-Петербургский государственный университет
- Университет ИТМО

Томск :

- Национальный исследовательский Томский государственный университет
- Томский политехнический университет

Узбекистан

- Ташкентский филиал СПбГУ

- **Воронеж** : Воронежский государственный университет
- **Владикавказ** : Северо-Осетинский государственный университет им. К.Л. Хетагурова
- **Дубна** : Государственный университет «Дубна»
- **Казань** : Казанский федеральный университет
- **Новочеркасск** : Южно-Российский государственный политехнический университет
- **Самара** : Самарский университет
- **Тула** : Тульский государственный университет
- **Челябинск** : Южно-Уральский государственный университет (НИУ)

Казахстан

- Казахский национальный университет им. Аль-Фараби

Беларусь

- Гомельский государственный технический университет им. П.О.Сухого
- Полесский государственный университет

MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS 2024

20–25 Oct 2024
Yerevan, Armenia

More than **150 participants**

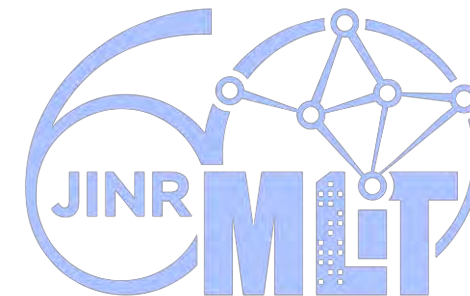
21 Plenary reports **110** Sessional reports

18 Countries: Armenia, Belarus, Bulgaria, Canada, the Czech Republic, Egypt, France, Georgia, Iran, Kazakhstan, Mongolia, New Zealand, Poland, Romania, Slovakia, Tajikistan, Uzbekistan and a large number of Russian research centers and universities.

Conference Topics:

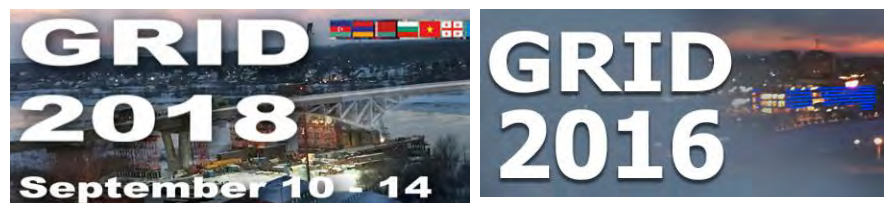
- Mathematical methods and tools for modeling complex physical systems;
- Mathematical methods in life sciences;
- Modern methods for data processing and analysis in Mega-science projects;
- Machine learning and big data analytics;
- Methods of quantum computing and quantum information processing;
- Numerical and analytical calculations in modern mathematical physics;
- Methods and numerical algorithms in high-energy physics.





GRID'2004 – 204 participants from 8 countries;
 GRID'2006 – 173 participants from 8 countries;
 GRID'2008 – 211 participants from 20 countries;
 GRID'2010 – 198 participants from 21 countries;
 GRID'2012 – 243 participants from 22 countries;
 GRID'2014 – 190 participants from 12 countries;
 GRID'2016 – 280 participants from 18 countries;
 GRID'2018 – 260 participants from 16 countries;
 GRID'2021 – 239 participants from 19 countries;
 GRID'2023 – 280 participants from 17 countries;

GRID'2025 – 298 participants from 14 countries
37 plenary and over 135 sessional talks



GRID'2025

7 — 11 July, Dubna



296 participants

37 Plenary reports

127 Sectional reports

16 Countries:

Armenia
Belarus
Bulgaria
CERN
Egypt
France
Georgia
Iran

Kazakhstan
Mexico
Rwanda
Russian Federation
Romania
South Africa
Taiwan
Uzbekistan

Russia was represented by participants from
44 universities and research centers.



Platformix

RSC

IT Cost



1. Distributed Computing Systems, Grid and Cloud Technologies, Storage Systems: architectures, operation, middleware and services

2. High Performance Computing
3. Application software in HTC and HPC
4. Computing for MegaScience Projects
5. Methods and Technologies for Experimental Data Processing

Korenkov V.V. - Chairman
Strizh T.A. - Vice-Chairman
Shmatov S.V. - Vice-Chairman
Derenovskaya O.Ya. - Scientific Secretary

Belyakova N.E.	Podgorny D.V.
Koragin I.I.	Prizhking D.I.
Korotseva T.I.	Ryugovskaya O.V.
Khmelev A.V.	Sokolov I.A.
Makhalik A.A.	Stankov D.B.
Mozhitova Ye.	Stoytsova O.I.
Molozet K.R.	Sviridov A.S.

 Phone: +7 (496) 216-40-19, 216-48-26

Avetisyan A. I. (SP RAS, Moscow)
Bogdanov A. V. SPbSU, St. Petersburg
Chetverushkin B. M. KIAM RAS, Moscow
Dimitrev V. SU, Bulgaria
Ilyin V. A. NRC "Kurchatov Institute", Moscow
Fedorov M. V. ITP RAS, Moscow
Gina El-Feky ASST, Egypt
Kallaev I. A. SSC RAS, Pustovo-Lok
Kalachev I. A. FIC KazanSC RAS, Kazan
Khvedelidze A. M. JINR, Dubna / GTU, Georgia
Zaitsev A. M. JINR, Dubna

Kilin S. V. NIS, Irkutsk
Kiryakov A. K. RNP, Gatchina
Korenkov V. V. JINR, Dubna
Krukov A. P. SINT MSU, Moscow
Lakhno V. D. IMPB RAS, Pushchino
Odoevskaya E. V. HSE, Moscow
Oganov A. R. SkolTech, Moscow
Oganov A. R. MOSEI, Ufa, Ufa Institute

Pospyskin M. A. FRC CSC RAS, Moscow
 Proletarskiy A. V. BMSU, Moscow
 Sahakyan V. G. IAP NAS, Armenia
 Sakhiyev S. K. IIR, Kazakhstan
 Sargsyan S. S. RAU, Armenia
 Shabanov B. M. JSCC, Moscow
 Shchur L. N. HSE, Moscow

Sithole H. NICIS, South Africa
Solfer V. A. SU, Saransk
Sokolov I. A. CSC MSU, Moscow
Sokolinsky L. B. SUStU, Chelyabinsk
Temkin I. O. MSIS, Moscow
Toporkov V. V. NRU MPEI, Moscow
Tran C. T. VINATOM, Vietnam

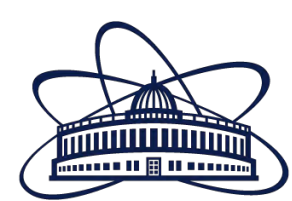
Trubnikov G. V. JINR, Dubna
Tsaregorodtsev A. Yu. CPCM-IMP-CNR, FN
Velikhov V. E. NRC 'Kurchatov instn
Voevodin V. V. SBCC MSU, Moscow
Zhang X. IHEP CAS, China
Zolotarev V. I. pr-5500, St. Petersburg



Тенденции развития распределенных вычислений для масштабных научных проектов

Развитие компьютерных архитектур и их интеграция:

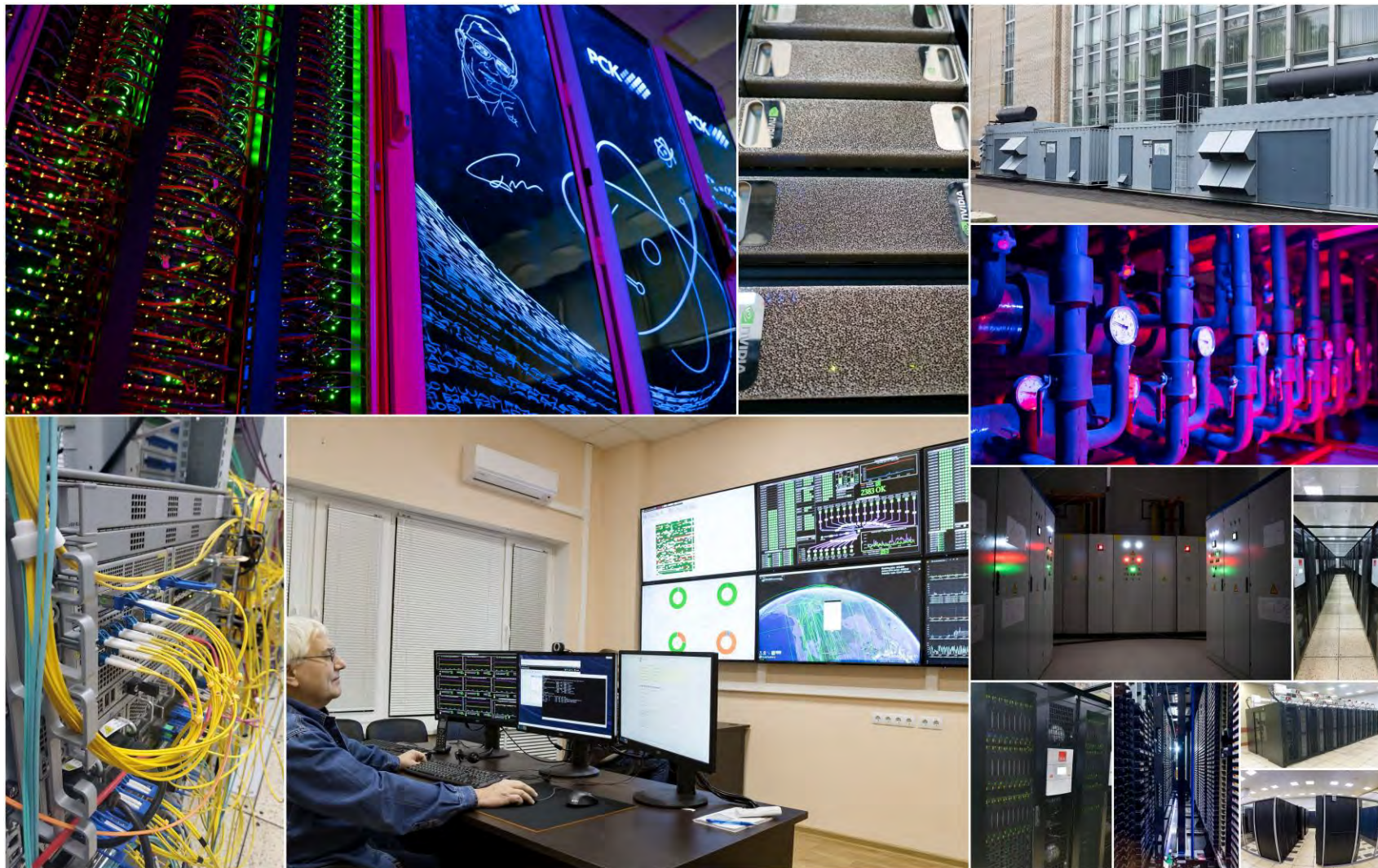
- суперкомпьютеры масштаба Peta-ExaFlop, графические, квантовые, фотонные процессоры как специализированные вычислители вокруг универсальной системы
- распределенные иерархические системы хранения информации масштаба Peta-ExaByte, работающие с общими метаданными для защищенного эффективного доступа и надежного хранения информации
- центры обработки данных на базе облачных технологий, обеспечивающие эффективные сервисы для пользователей и поддержку функционирования систем туманных и граничных вычислений
- интеллектуальные системы управления заданиями в распределенной гетерогенной среде, включающей различные компьютерные архитектуры (суперкомпьютеры, грид, облака, кластеры, серверы, системы добровольных вычислений)
- развитие методов, алгоритмов, платформ, систем аналитики Больших данных, интеллектуального анализа данных, ML/DL, искусственного интеллекта, цифровых сервисов
- создание цифровых двойников экспериментальных установок, распределенных компьютерных комплексов и других сложных объектов
- подготовка высококвалифицированных специалистов в области распределенных вычислений, хранения, обработки, анализа данных для научных проектов



JINR



Спасибо за внимание



<https://lit.jinr.ru>



t.me/MLIT_news