

Консорциум
Российский ГРИД для интенсивных операций с данными проектов мегасайенс (РДИГ-М).

Российский ГРИД сегмент WLCG РДИГ в 2006 г вошел в состав **научной коллаборации WLCG** в качестве региональной федерации. К первому рабочему сеансу БАК (2009-2013 гг) были созданы центры уровня Тьер2, обеспечивающие анализ и моделирование данных. В 2013 году по инициативе директора НИЦ «Курчатовский Институт» М.В. Ковальчука в России были созданы центры уровня Тьер1, обеспечивающие хранение, обработку и анализ данных.

РДИГ меморандум

1 сентября 2003 г.

В целях обеспечения полномасштабного участия России в проекте создания глобальной (паневропейской) компьютерной инфраструктуры типа ГРИД - EGEE (*Enabling Grids for E-science in Europe*), принятом к финансированию летом 2003 года Европейской Комиссией (6-я Рамочная программа, проект FP6-2002-Infrastructures-2 № 508833), и реализации отечественных приоритетов в этом международном проекте, нижеподписавшиеся Институты, участвующие в данном проекте, образовали консорциум **РДИГ (Российский ГРИД для интенсивных операций с данными - Russian Data Intensive GRID, RDIG)** для эффективного выполнения работ по проекту и развитию в России инфраструктуры EGEE, с вовлечением на следующих этапах проекта других организаций из различных областей науки, образования и промышленности.

Консорциум РДИГ, согласно принятой в проекте EGEE структуре, входит в проект в качестве региональной федерации «Россия» («Russia»), наравне с ЦЕРН (Европейской лабораторией по физике частиц), выполняющей координирующие функции по проекту, и другими 8-ю национальными и региональными федерациями (Франция, Италия, Великобритания/Ирландия, Германия/Швейцария, Скандинавские страны, консорциумы Центрально-Европейских стран, Юго-восточной Европы и Юго-западной Европы).

Консорциум РДИГ принял решение обратиться к Министерству промышленности, науки и технологий Российской Федерации с просьбой о выполнении координирующих функций по участию России в проекте EGEE, а также к Минатому РФ и РАН с предложением принять участие в поддержке соответствующих работ по проекту.

Для координации участия России в проекте EGEE создается **Координационный Совет РДИГ**, в который входят руководители участвующих институтов, представитель Минпромнауки России, а также представители других заинтересованных ведомств. Для оперативного управления работами по проекту создается **Управляющий Совет РДИГ**, на основе паритетного представительства участвующих институтов и заинтересованных ведомств. На этап подготовки контракта по проекту (сентябрь-ноябрь 2003 года) функции Управляющего Совета РДИГ выполняет инициативная группа, участвовавшая в подготовке проекта.

Научный директор ГНЦ ИТЭФ
чл.-корр. РАН М.В. Данилов

Директор ГНЦ ИФВЭ
профессор Н.Е. Тюрин

Директор ИМПБ РАН
профессор В.Д. Лажно

Зам. директора по науке ИПМ
им.М.В.Келдыша РАН
профессор Д.А. Коржин

Директор НИИЯФ МГУ
профессор М.И. Панаенок

Директор ОИЯИ
академик В.Г. Кадышевский

Директор ПИЯФ РАН
чл.-корр. РАН В.А.Назаренко

Президент РНЦ КИ
академик Е.П. Велихов

WORLDWIDE LHC COMPUTING GRID COLLABORATION

The European Organization for Nuclear Research (CERN)

and

Federal Agency for Science and Innovation (FASI)

declare that they agree on this Memorandum of Understanding for collaboration in the deployment and exploitation of the Worldwide LHC Computing Grid.

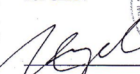
Done in Geneva

Done in Moscow

on _____

on 3 July 2007

For CERN

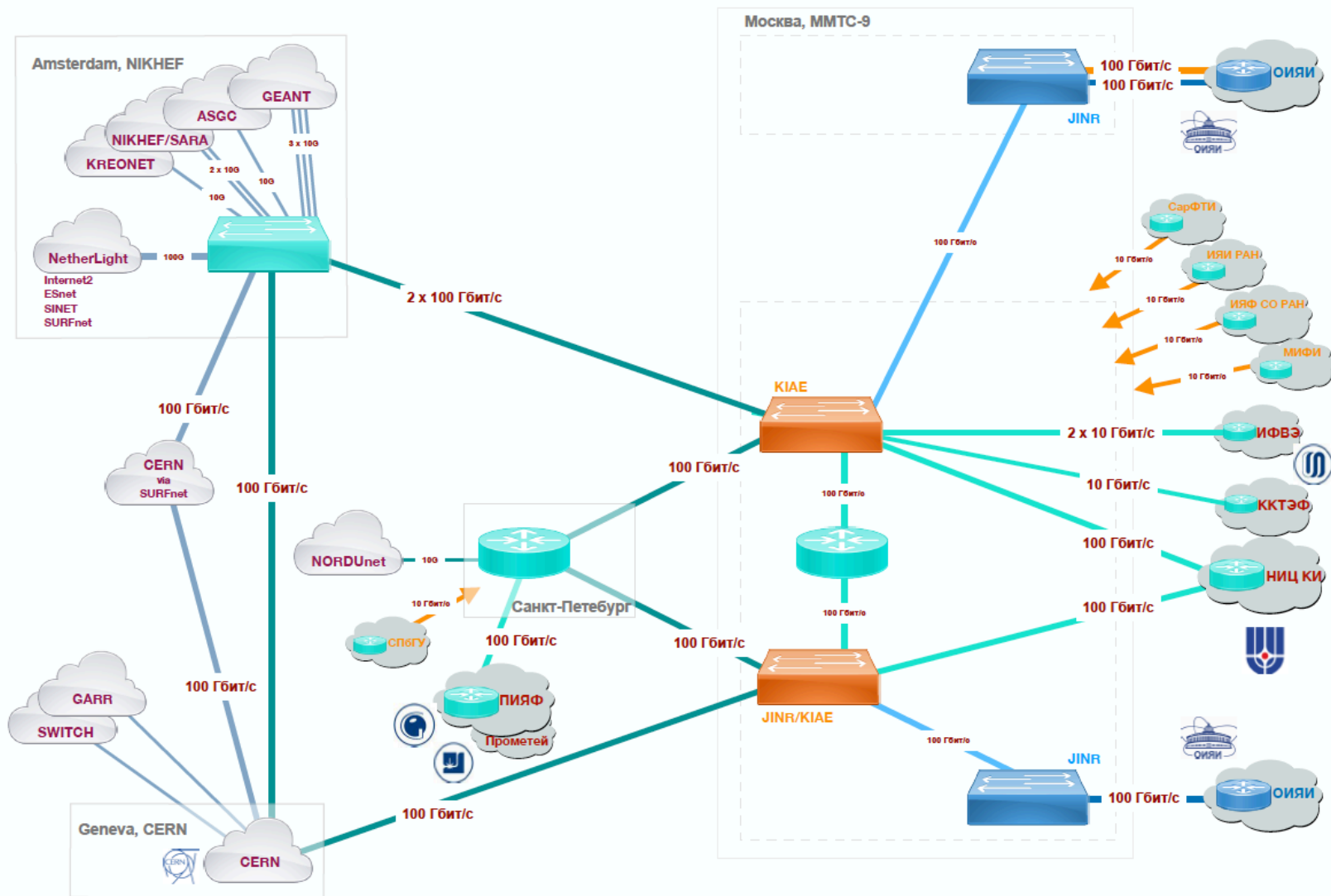

 Jos Engelen
 Chief Scientific Officer

For FASI


 Sergei Mazurek
 Chief

- Консорциум Российских институтов Российский ГРИД для интенсивных операций с данными (РДИГ) был создан 2003 году для участия в распределенной обработке данных экспериментов на Большом адронном коллайдере, являясь Российским сегментом международной научной инфраструктуры класса мегасайенс.
- Ресурсы РДИГ включают более 50 тыс. вычислительных ядер, 50 петабайт (ПБ) дисковых и 70 ПБ ленточных массивов (более 90% ресурсов предоставляют НИЦ «Курчатовский институт» и ОИЯИ).
- За год ресурсами РДИГ обрабатываются десятки миллионов задач, передается и обрабатывается более 150 ПБ данных.

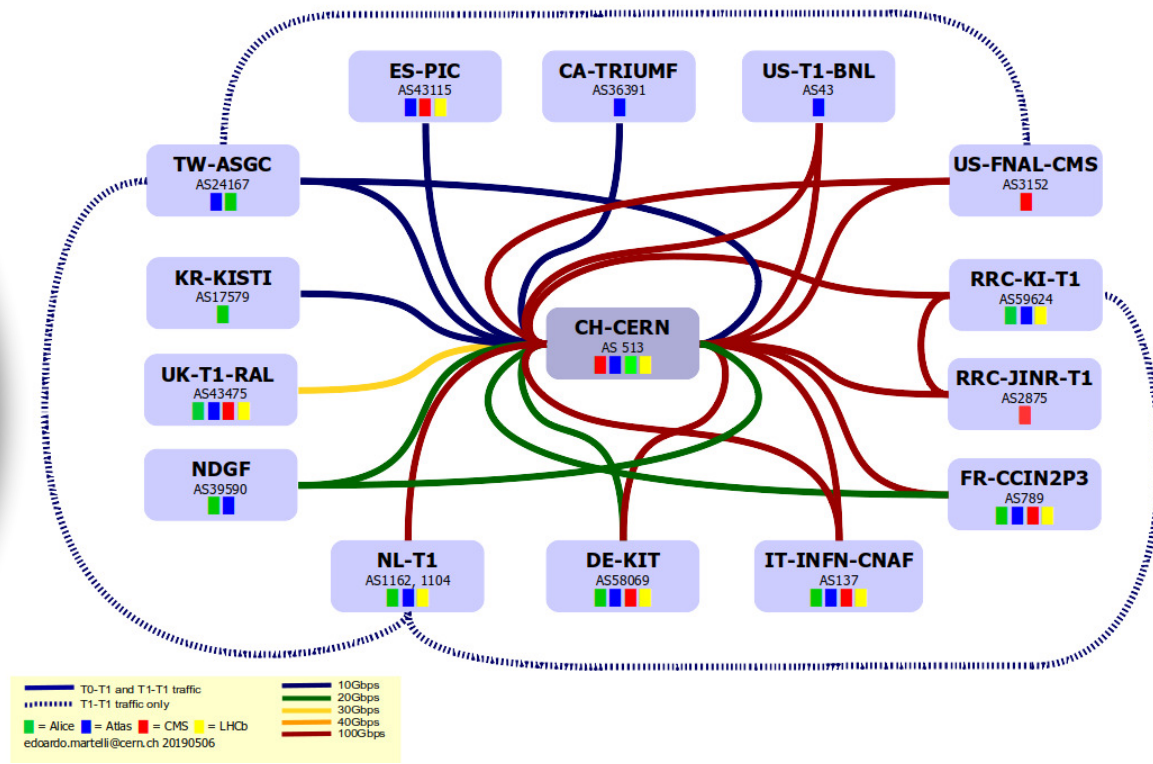
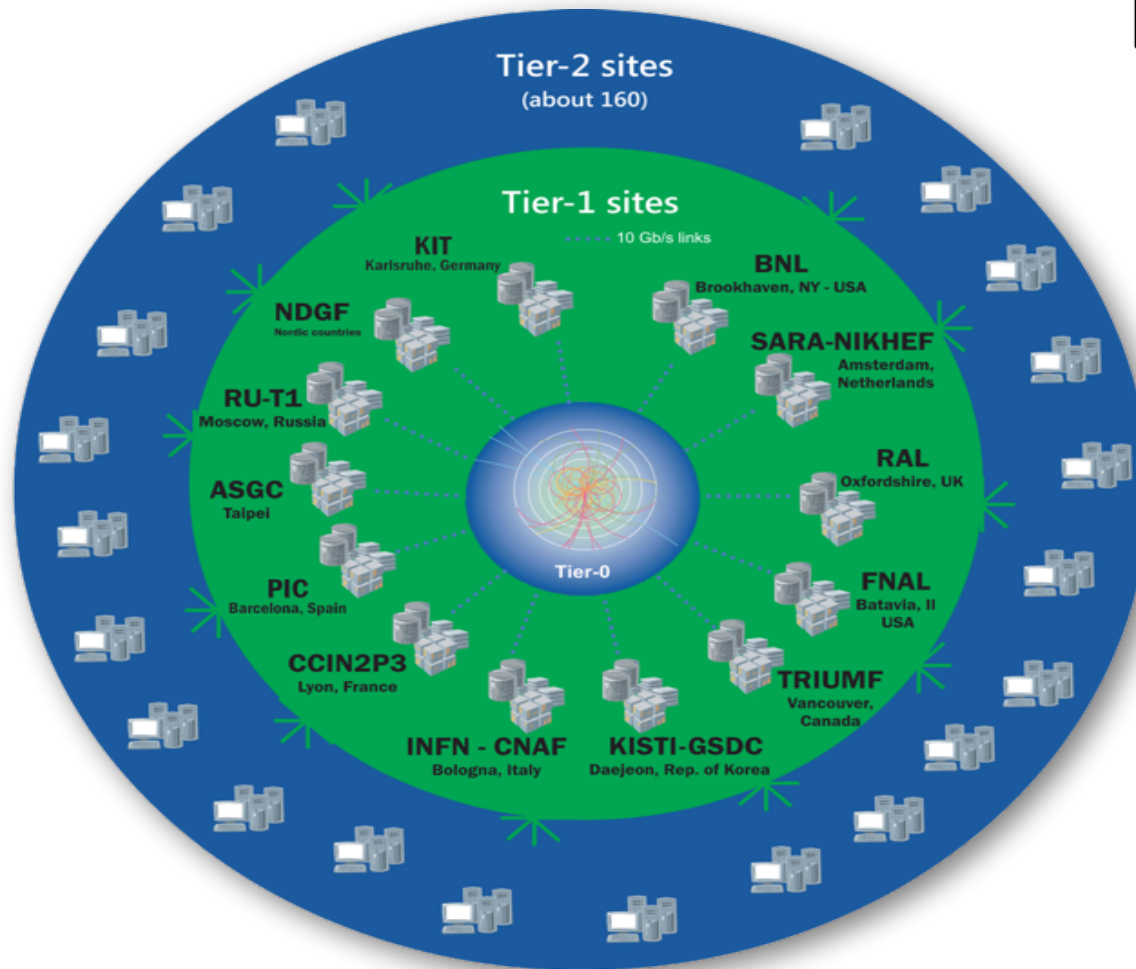
Russian Data-Intensive Grid (RDIG) * Магистральная инфраструктура * 2024 (План)



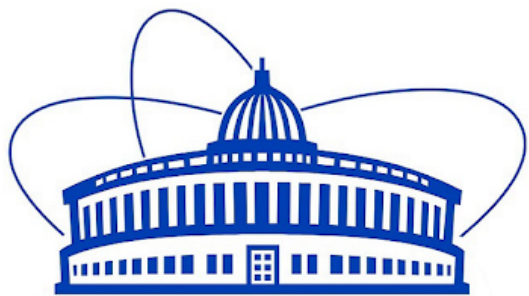
- Администраторы РДИГ - НИЦ «Курчатовский институт» и ОИЯИ



LHCOPN



— T0-T1 and T1-T1 traffic
 T1-T1 traffic only
 ■ = Alice ■ = Atlas ■ = CMS ■ = LHCb
 edoardo.martelli@cern.ch 20190506



Объединенный институт
ядерных исследований



ИСП

РАН

Для создания и обеспечения эффективной работы национальной научной инфо-коммуникационной инфраструктуры для интенсивных операций с данными проектов мегасайенс, включая сервис-ориентированную цифровую среду с использованием высокопроизводительной облачной платформы, сформирован консорциум в составе НИЦ «Курчатовский институт», ОИЯИ и ИСП РАН.



Стороны обязуются объединить усилия в целях обеспечения развития исследований и инфокоммуникационной инфраструктуры как коллаборативной сервис-ориентированной цифровой среды в области интенсивных операций с большими данными, включая суперкомпьютерное моделирование, параллельную обработку и интеллектуальный анализ с использованием методов и технологий искусственного интеллекта, являющихся основой проведения фундаментальных, прикладных исследований и разработок.

СОГЛАШЕНИЕ об образовании научно-технологического консорциума

Федеральное государственное бюджетное учреждение «Национальный исследовательский центр «Курчатовский институт» (далее – НИЦ «Курчатовский институт») в лице президента Ковальчука Михаила Валентиновича, действующего на основании Устава,

Федеральное государственное бюджетное учреждение науки Институт системного программирования им. В.П. Иванникова Российской академии наук (далее – ИСП РАН) в лице директора Аветисяна Арутюна Ишхановича, действующего на основании Устава,

Международная межправительственная научно-исследовательская организация Объединенный институт ядерных исследований (далее – ОИЯИ) в лице директора Трубникова Григория Владимировича, действующего на основании Устава,

именуемые в дальнейшем «Стороны», заключили настоящее Соглашение о нижеследующем:

1. Предмет соглашения

1.1. Стороны договорились об образовании научно-технологического консорциума Российский ГРИД для интенсивных операций с данными проектов мегасайенс (РДИГ-М) (далее – Консорциум) путем заключения настоящего Соглашения о создании Консорциума.

1.2. Стороны обязуются объединить усилия в целях обеспечения развития исследований и инфокоммуникационной инфраструктуры как коллаборативной сервис-ориентированной цифровой среды в области интенсивных операций с большими данными, включая суперкомпьютерное моделирование, параллельную обработку и интеллектуальный анализ с использованием методов и технологий искусственного интеллекта, являющихся основой проведения фундаментальных, прикладных исследований и разработок.

2. Общие положения

- 2.1. Консорциум образуется без ограничений срока действия.
- 2.2. Консорциум не является юридическим лицом.

7.1. В ходе реализации деятельности Консорциума все участники руководствуются принципом доброй воли. В случае возникновения разногласий последние урегулируются путем переговоров. В случае невозможности достичь согласия споры решаются в соответствии с законодательством Российской Федерации.

7.2. Каждый участник самостоятельно принимает решение о формах участия в деятельности Консорциума.

7.3. Настоящее Соглашение вступает в силу с момента подписания.

8. Адреса и подписи Участников Консорциума

Федеральное государственное бюджетное учреждение «Национальный исследовательский центр «Курчатовский институт» (НИЦ «Курчатовский институт») Адрес: 123182 Россия, Москва, пл. Академика Курчатова

Президент



М.В. Ковальчук
№ СОГ-1/р
от 05.02.2024

Федеральное государственное бюджетное учреждение науки Институт системного программирования им. В.П. Иванникова Российской академии наук (ИСП РАН) Адрес: 109004, г. Москва, ул. Александра Солженицына, д. 25

Директор

А.И. Аветисян

Международная межправительственная организация Объединенный институт ядерных исследований, Адрес: 141980 Россия, Московская обл, г. Дубна, ул. Жолио-Кюри, 6

Директор

Г.В. Трубников

Задачами Консорциума являются:

создание передовых технологий в области высокопроизводительных вычислений, анализа больших данных и искусственного интеллекта;

создание и эксплуатация национальной инфраструктуры для передачи, обработки, хранения и анализа научных «больших данных»;

создание коллаборативной сервис-ориентированной цифровой среды науки и образования;

обеспечение и координация научных исследований в соответствии с направлениями деятельности Консорциума;

получение конкурентоспособных на мировом масштабе научных результатов.

Направлениями деятельности Консорциума являются:

разработка и внедрение суперкомпьютерных (включая системы с экстремальным параллелизмом), сетевых, ГРИД и облачных технологий, а также технологий федеративного хранения, интеллектуальной обработки и анализа сверхбольших массивов и потоков научных данных;

создание информационно-вычислительных комплексов и систем обработки и хранения данных проектов «мегасайенс»;

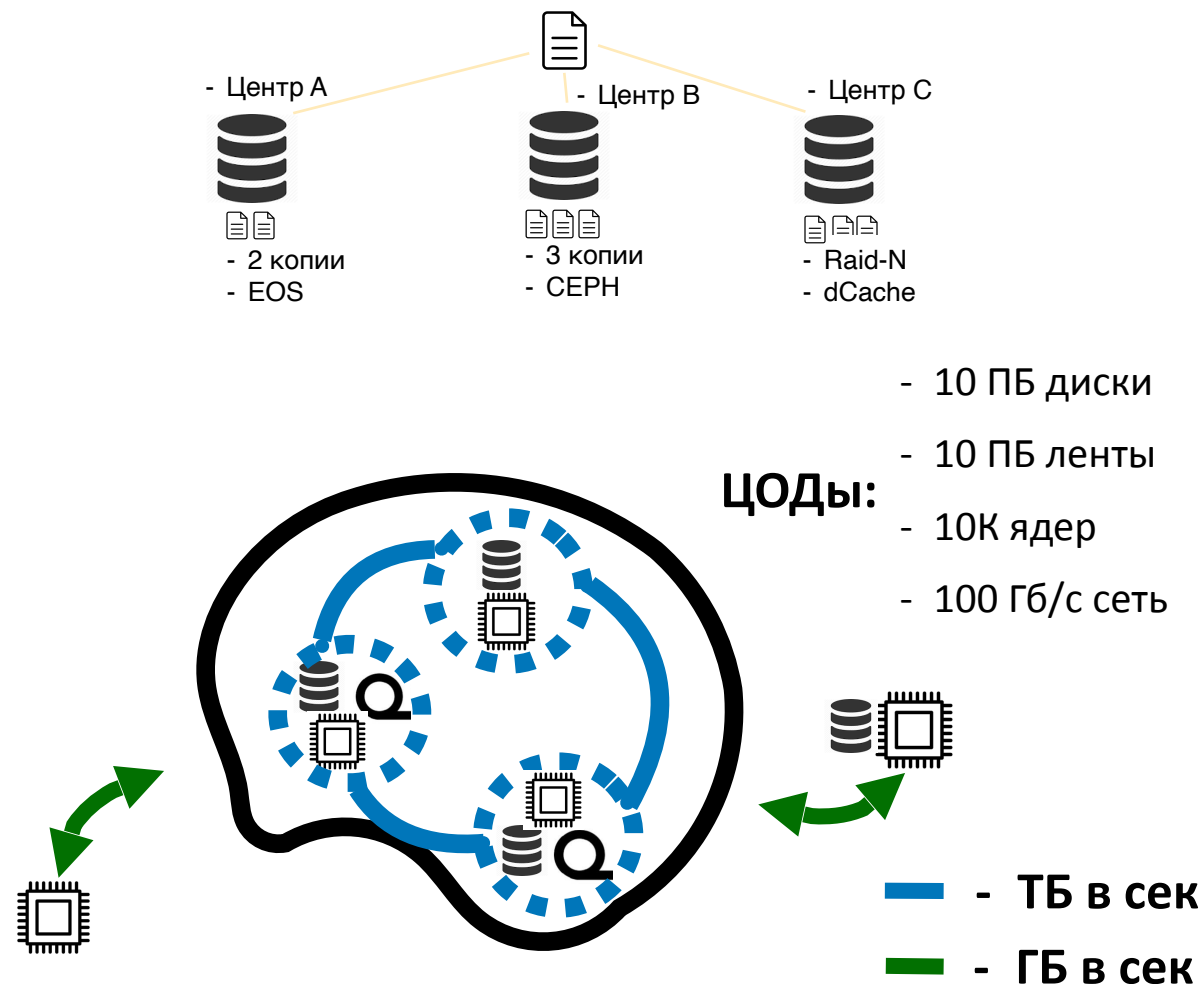
разработка оригинальных алгоритмов обработки и анализа данных, в том числе основанных на использовании искусственного интеллекта, методов квантовых вычислений, технологий информационной безопасности;

создание доверенных аппаратно-программных комплексов.

- РДИГ-М должен обеспечить инфраструктуру для действующих и создаваемых отечественных исследовательских установок мегасайенс – ускорительных комплексов НИКА (Дубна) и У-70 (Протвино), реакторного комплекса ПИК (Гатчина), модернизируемых и вновь создаваемых источников синхротронного излучения КИСИ-Курчатов (Москва), ТНК (Зеленоград), СКИФ (Новосибирск), РИФ (остров Русский), СИЛА и нейтронного источника Омега (Протвино). Прогнозируемый рост объемов передаваемых данных к 2030 г:
- в экспериментах НИКА до 150 ПБ/год,
- в экспериментах синхротронно-нейтронных исследований до 100 ПБ/год.



- Создание на базе НИКС и РДИГ национальной научной инфо-коммуникационной инфраструктуры для интенсивных операций с данными проектов мегасайенс
- Создание единой цифровой платформы для хранения, обработки и анализа экспериментальных данных синхротронно-нейтронных исследований (СНИ) объемом в сотни ПетаБайт с Терабитным сетевым ядром
- Увеличение вычислительных мощностей и объема систем хранения данных для предоставления облачных сервисов для решения задач в ИИ, машинного обучения, больших данных
- Интеграционные мероприятия а рамках сотрудничества с национальными научно-образовательными сетями стран ЕАЭС, СНГ, ШОС, БРИКС.



- ЕДИНАЯ СЕРВИСНАЯ ПЛАТФОРМА НА ОСНОВЕ СТЕКА ТЕХНОЛОГИЙ ИСП РАН

- Организации реального сектора экономики

- Экспертные сообщества, Государственные академии

- Заинтересованные ФОИВ, ведомства

- Университеты, образовательные и научные центры

- Международные участники

- Пространство исследователя :: Образовательная среда :: Web-лаборатории

- Прикладные сервисы

- TALISMAN.Биография



- VisOntia



- TALISMAN.Поток



- SciNoon



- Анализ социальных сетей
- Обработка текстов
- Образовательные платформы
- Представление знаний
- Интеллектуальный поиск
- Обработка текстов

- Инструменты сбора, хранения и анализа больших данных

- TALISMAN



TensorFlow



- Платформа управления ПЖЦ моделей ИИ и наборов данных

- Инструмент работы с большими данными
- Интеграция с внешними и внутренними «поставщиками» данных
- Технологии AI, ML, DL
- AI Ready данные

- Облачная среда

- ASPERITA



- MICHMAN



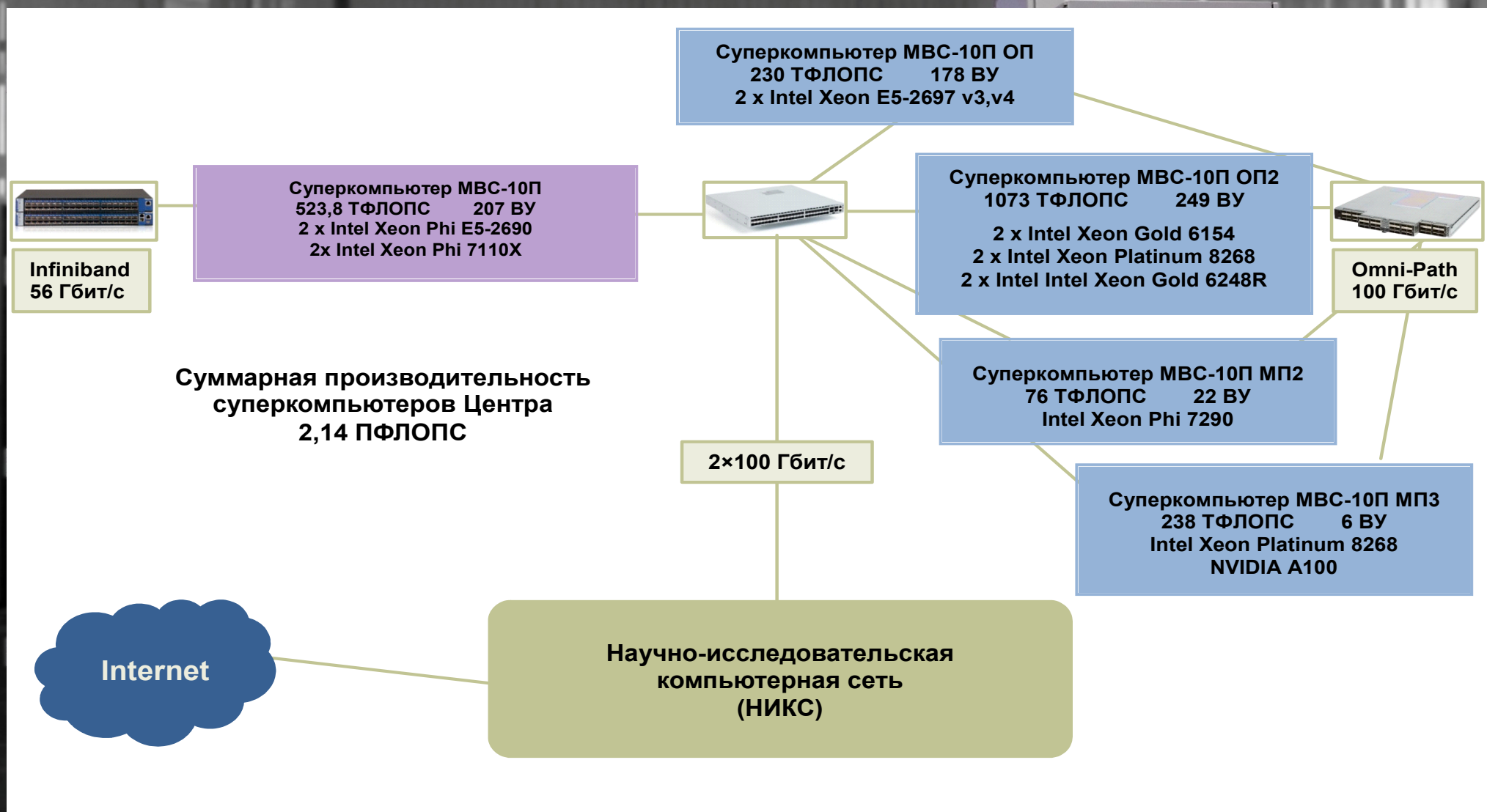
- FANLIGHT



- Миграция legacy-систем в облако
- Инфраструктура как сервис на базе Openstack (IaaS)
- Платформенные сервисы по запросу (PaaS)
- Микросервисная архитектура и контейнеризация

- Аппаратные платформы





Федеральный центр коллективного пользования «Комплекс моделирования и обработки данных исследовательских установок мега-класса»,

2007 г.



HPC1 (34 ТФлопс),
2688 ядер, 40.5 Тбайт
система хранения

2009 г., 2011 г.



HPC2 (123 ТФлопс) 10К ядер
HPC3G (134 ТФлопс) ускорители Nvidia M2070
SMP серверы (512 ГБ RAM)
Параллельная файловая система 120 ТБ

2015 г. 2019 г.



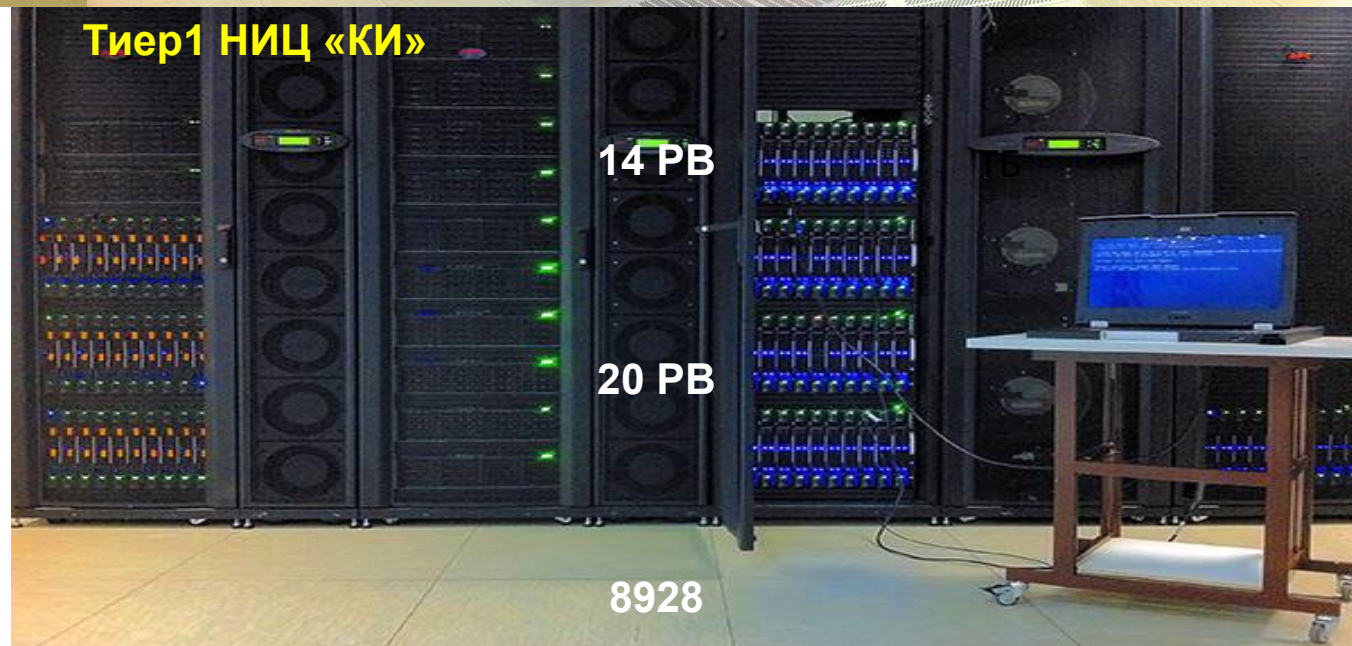
HPC4 (370 ТФлопс) 8К ядер
HPC5 (602 ТФлопс)
ускорители Nvidia K80
(320 шт.)
Параллельная файловая система Lustre 960 ТБ



**Дисковая система
хранения, ПБайт**

**Ленточная система
хранения, ПБайт**

**Вычислительный
массив, ядер**



Теоретическая общая пиковая производительность составляет
364 Тфлопс

Система вычислительных узлов общего назначения: 160
узлов с 28 ядрами и 128 ГБ памяти на каждом (4,5 ГБ на ядро)

Система вычислительных узлов с векторными процессорами:
40 узлов с 68 ядрами (272 виртуальных)

Система вычислительных узлов с большой памятью:
2 узла с 28 ядрами и 1,5 ТБ памяти на каждом
16 узлов с 28 ядрами и 1 ТБ памяти + 1,6 ТБ
энергонезависимой памяти (NVMe) на каждом

Система хранения данных для параллельных вычислений
объёмом 2,9 ПБ Система блочного хранения данных полным
объёмом 2,5 ПБ

Система передачи данных для параллельных вычислений со
скоростью 100 Гбит/с (InfiniBand EDR)

Сеть общего назначения со скоростью от 1 до 10 Гбит/с

Сеть управления со скоростью 1 Гбит/с

Телекоммуникационный узел

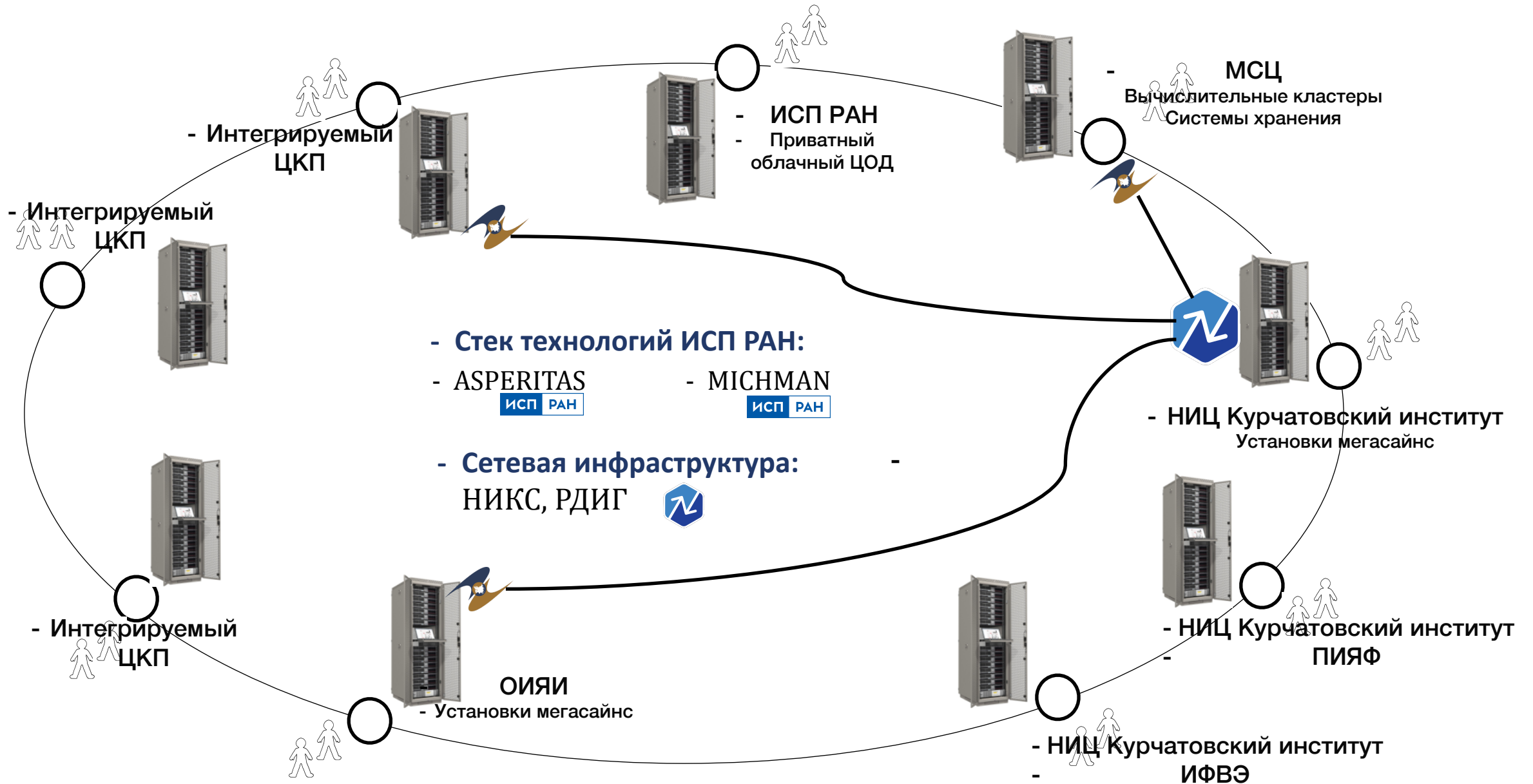


- НИКС – национальная исследовательская компьютерная сеть
- Магистральная инфраструктура от Санкт-Петербурга до Владивостока
- Пропускная способность от 1 до 100 Гбит/с (Москва-СПб)
- Производительность узлов связи от 400 Гбит/с до 2400 Гбит/с
- В 2024 году в контуре НИКС будет более 325 ведущих научных и образовательных организаций, включая операторов ЦКП, УНУ и СКЦ
- 100ПБ передается по сети ежегодно



- Администратор НИКС - НИЦ “Курчатовский институт”- НИИСИ
- (Приказ Минобрнауки России от 12 мая 2021 г. № 375)

- ЕДИНАЯ СЕРВИСНАЯ ПЛАТФОРМА НА БАЗЕ НИКС и РДИГ .



Проект: СТРАТЕГИЧЕСКИЙ ПЛАН развития природоподобных технологий в РФ Этап I (2024 – 2032 годы)

Информационно-коммуникационная платформа НСРПТ с распределенной модульной архитектурой включает следующие модули (вычислительные системы) :

- модули, обеспечивающие разработку прикладных программных продуктов заданного уровня доверия, включая совместную работу программистов, создание и ведение единого фонда моделей, алгоритмов и программ, инструменты сбора и интеллектуального анализа больших данных и управления полным жизненным циклом нейросетевых моделей (включая большие фундаментальные модели) и наборов данных, с использованием коллаборационной цифровой среды и высокопроизводительной облачной платформы;
- модули, обеспечивающие сбор, передачу, обработку, хранение, анализ и моделирование научных данных в области синхротронно-нейтронных исследований, базирующиеся на ГРИД и облачных технологиях для интенсивных операций с большими данными, технологиях суперкомпьютерного моделирования, параллельной обработки и интеллектуального анализа данных с использованием методов ИИ;
- модули, обеспечивающие сбор, хранение, предоставление доступа к биологической информации и её централизованную обработку, и анализ на основе технологий параллельной работы с большими данными, методов и технологий ИИ, с использованием системы «Национальная база генетической информации»;
- модули, обеспечивающие моделирование материалов на уровне связи состава-структуры-свойств, моделирование технологий изготовления и переработки материалов, прогнозирование свойств материалов на микроуровне, создание цифровых двойников изделий, создание и ведение цифрового каталога материалов, с использованием суперкомпьютерных и облачных технологий;
- интеграционную инфраструктуру на базе российского ГРИД для интенсивных операций с данными проектов мегасайенс (РДИГ-М), Федерального центра коллективного пользования научным оборудованием «Комплекс моделирования и обработки данных исследовательских установок мега-класса» (ФЦКП), Национальной исследовательской компьютерной сети нового поколения (НИКС) и Межуниверситетской квантовой сети (МУКС).

Спасибо за внимание!