

Введение в EGEE проект

Ю.Ф.Рябов



Название проекта



- EGEE- Enabling Grigs for E-scienceE

Развертывание инфраструктуры
Грид для науки

Содержание

- Основные цели и задачи проекта
- Функциональность проекта EGEE
- Основные направления работ в проекте
- EGEE партнеры
- Роль ЦЕРНа в EGEE
- LCG : LHC computing GRID
- Россия в проекте EGEE
- О консорциуме RDIG-EGEE
- ПИЯФ в EGEE проекте



Основные цели и задачи проекта EGEE



- Сохранить научно-технические результаты предыдущих GRID-проектов
- Сохранить технические и кадровые ресурсы от предыдущих EU программ
- Обеспечить создание высокопроизводительной продукционной GRID инфраструктуры, ее поддержку и развитие
- Гармонизировать национальные и интернациональные активности в развитии GRID технологии в рамках этого проекта
- Привлекать пользователей из различных сфер деятельности и обеспечить им высокий уровень обучения и поддержки

Функциональность проекта EGEE



- **Упрощённый доступ.**
EGEE уменьшит издержки, связанные с разнообразием не связанных между собой систем учёта пользователей. Пользователи смогут объединяться в виртуальные организации с доступом к grid – инфраструктуре, содержащей нужные каждому пользователю рабочие ресурсы.
- **Выполнение вычислений по требованию.**
Эффективно распределяя ресурсы, grid-технологии значительно сокращает время ожидания доступа к ним.
- **Географически распределённый доступ.**
Инфраструктура будет доступна отовсюду, где обеспечена хороший доступ к сети. Ресурсы становятся широко доступными.
- **Исключительно большой объём ресурсов.**
Благодаря согласованности ресурсов и пользовательских групп, в прикладной работе в рамках EGEE будут доступны ресурсы таких объёмов, какие не может предоставить ни один компьютерный центр.
- **Совместное использование программного обеспечения и данных.**
Благодаря единой структуре вычислительных средств, в EGEE пользователям будет легко совместно использовать программное обеспечение и базы данных и разрабатывать программное обеспечение.
- **Высокий уровень поддержки приложений.**
Компетентность всех участников EGEE обеспечит тщательную, всестороннюю поддержку всех основных приложений.

EGEE-II: Expertise & Resources

- More than 90 partners
 - 32 countries
 - 12 federations
- Major and national Grid projects in Europe, USA, Asia



+ 27 countries through related projects:

- BalticGrid
- SEE-GRID
- EUMedGrid
- EUChinaGrid
- EELA



Проект EGEE (Enable Grid for E-science)



○ Страны - поставщики ресурсов

● Страны, готовые вступить

EGEE/LCG-2 грид:

160 центров, 36 стран

>15,000 процессоров,

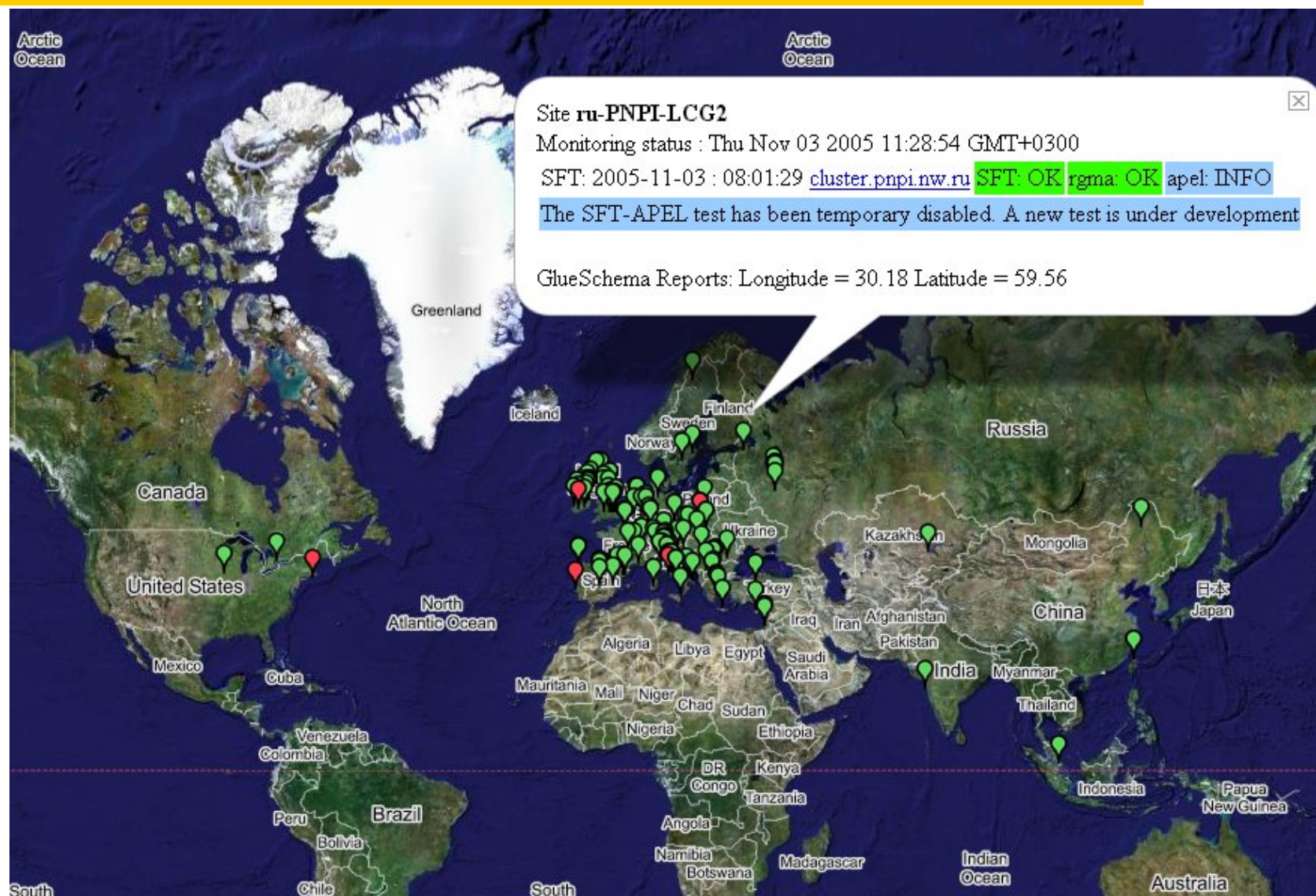
~5 PB памяти

Другие национальные и региональные гриды:

~60 центров, ~6,000 процессоров

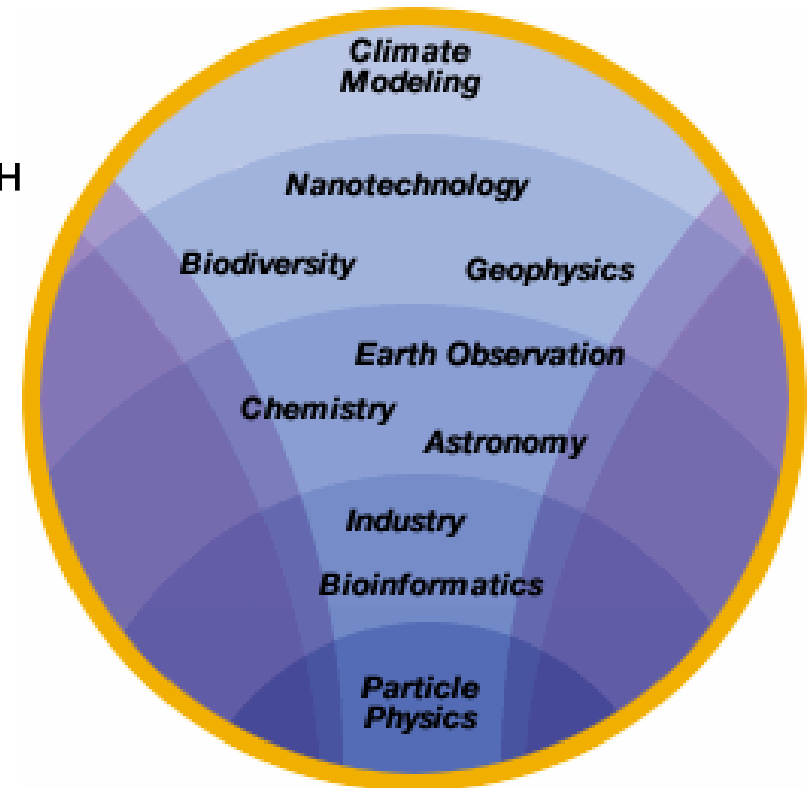
страна	центры	страна	центры	страна	центры
Австрия	2	Индия	1	Россия	10
Бельгия	1	Израиль	2	Сингапур	1
Болгария	4	Италия	25	Словакия	3
Канада	6	Япония	1	Словения	1
Китай	1	Корея	1	Испания	13
Хорватия	1	Нидерланды	2	Швеция	2
Кипр	1	Македония	1	Швейцария	2
Чехия	2	Пакистан	2	Тайвань	4
Франция	8	Польша	4	Турция	1
Германия	8	Португалия	1	Великобритания и Ирландия	35
Греция	6	Пуэрто Рико	1	США	3
Венгрия	1	Румыния	1	Югославия	1

Мониторинг узлов EGEE



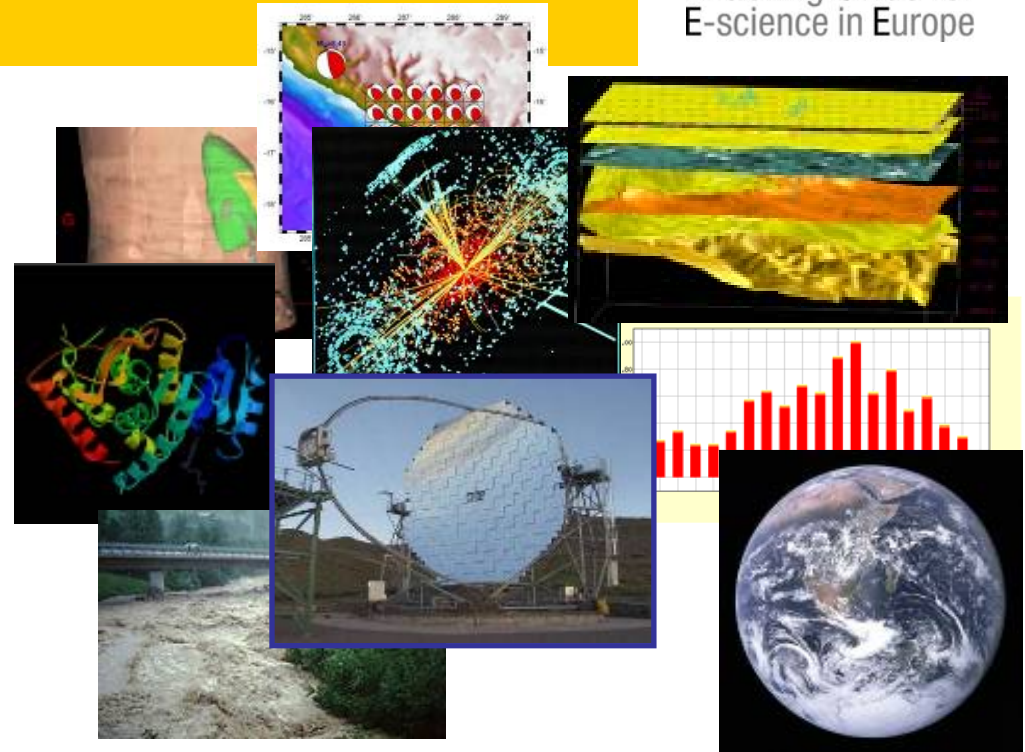
Использование результатов

- Масштаб применения р-тов EGEE: наука(открыто для индустриальных и социоэкономических направлений)
- Главный критерий успеха проекта: кол-во пользует. и кол-во обл. примен 5000 пользователей (3000 после 2-ух лет) мин.,более 10 областей применений
- Две области выбраны для пилотного использования р-тов проекта:
ФИЗИКА
БИОИНФОРМАТИКА



Applications

- Many applications from a growing number of domains
 - Astrophysics
 - Computational Chemistry
 - Earth Sciences
 - Financial Simulation
 - Fusion
 - Geophysics
 - High Energy Physics
 - Life Sciences
 - Multimedia
 - Material Sciences
 - ...



Applications have moved from testing to routine and daily usage ~80-90% efficiency

Активности в EGEE



NA2 –распространение информации

WEB сайты,е-почта,презентации,визиты,конференции ,публикации,...

NA3 – Обучение

Подготовка обучающих материалов. Проведение курсов

NA4 – применение и поддержка Grid-структуры в приложениях

Поддержка ФВЭ и Биологии в пилотном использовании. Привлечение пользователей.Определение общих интерфейсов и инструментария..

SA1 – Поддержка функционирования Европейского Grid

Мониторинг и управление Grid структурой. Распространение Grid ПО и включение новых выч.ресурсов.Поддержка пользователей.

SA2 – Обеспечение сетевыми ресурсами

JRA - Отладка и тестирование ПО промежуточного уровня

Роль ЦЕРНа в проекте EGEE

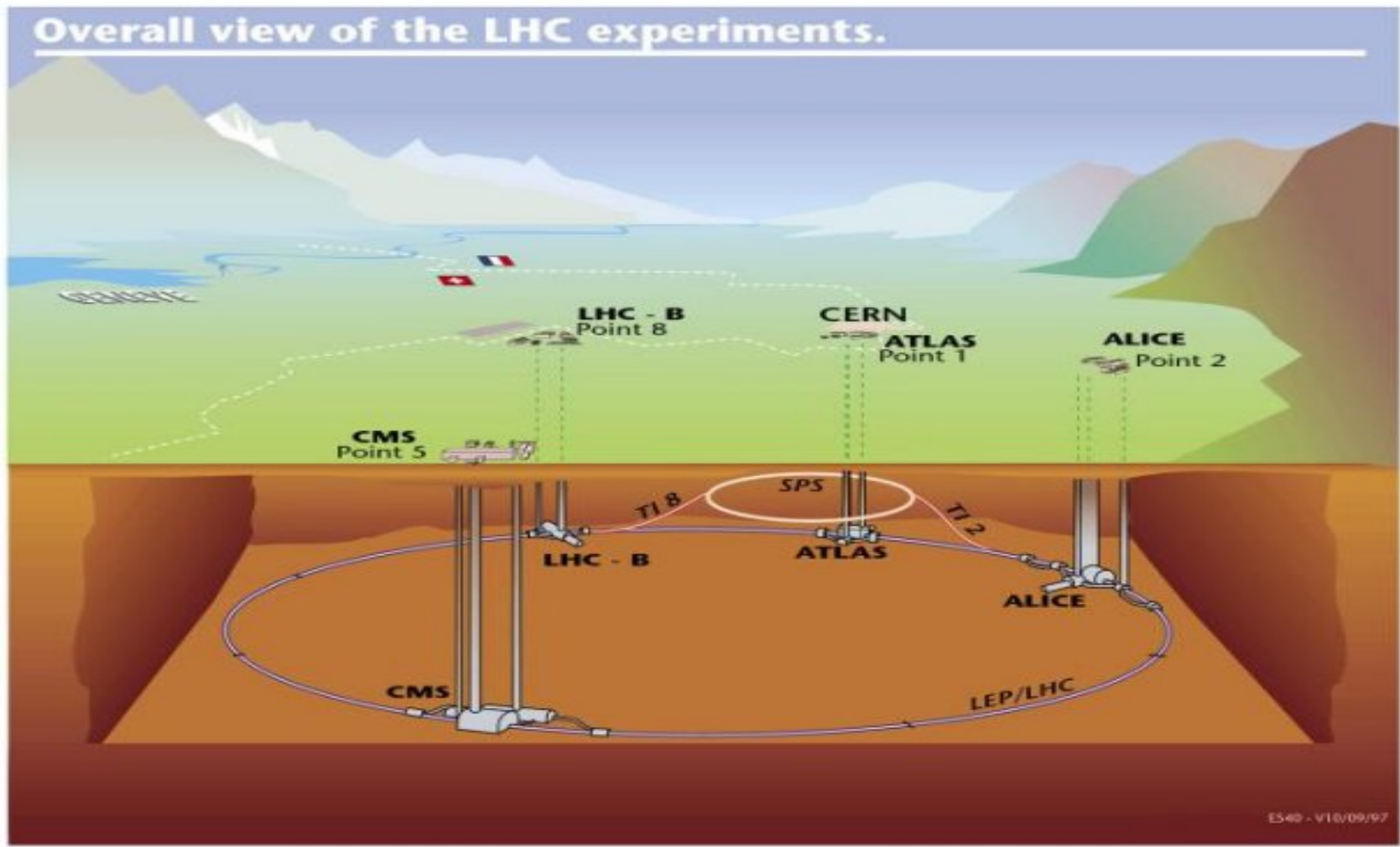


Эксперименты на LHC предполагают беспрецедентное использование компьютеров

LHC Computing GRID (LCG проект) и GRID технологии являются ответственностью ЦЕРНа

По этим принципам ЦЕРН является головным партнером в EGEE проекте

Большой адронный коллайдер



Данные БАК/ЛНЦ

1 Мегабайт (1MB)
Цифровая фотография
Enabling Grids for
1 Гигабайт (1GB)
in Europe
= 1000MB

DVD фильм

1 Терабайт (1TB)
= 1000GB

Объём всех книг,
изданных за год в мире

1 Петабайт (1PB)
= 1000TB

Производит за год один
эксперимент БАК

1 Экзобайт (1EB)
= 1000 PB

Объём информации,
которую производит за
год всё человечество

- 40 000 000 столкновений в секунду
- После фильтрации, остаётся только 100 событий в секунду
- Объём оцифрованных данных для одного события ~ 1Мб
- За год необходимо записать 10^{10} событий = 10 Петабайт данных в год

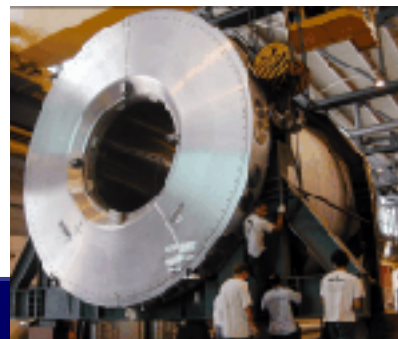
CMS



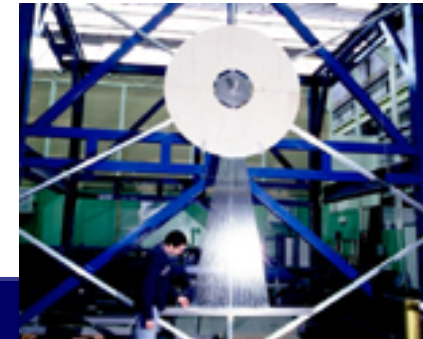
LHCb



ATLAS



ALICE



Необходимость ГРИД для LHC

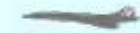
- Объём получаемых данных LHC соответствует 20 миллионам записанных CD дисков в год. *Где их хранить?*
- Анализ данных LHC потребует вычислительных мощностей, эквивалентных мощности 100000 самых современных процессоров. *Где их взять?*
- Ресурсы ЦЕРН уже сейчас составляют более 1000 2-х процессорных ПК и 1 Пб памяти на дисках и на лентах. *Но этого мало!!!*

**Выход – объединение
вычислительных ресурсов
физиков всего мира**



**Воздушный шар
(30 Км)**

**Столпа CD дисков с
данными БАК за 1
год (~ 20 Км)**



**Конкорд
(15 Км)**

**Монблан
(4.8 Км)**



LHC Computing Grid Project (LCG)

- **Основной задачей проекта LCG является создание глобальной инфраструктуры региональных центров для обработки, хранения и анализа данных физических экспериментов LHC.**
- **Новейшие технологии GRID являются основой построения этой инфраструктуры.**

Проект LCG осуществляется в две фазы.

- **1 фаза (2002-2005 гг.) - создание прототипа и разработка проекта системы (LCG TDR).**
- **2 фаза (2005-2007 гг.) - создание инфраструктуры LCG, готовой к обработке, хранению и анализу данных на момент начала работы ускорителя в 2007 году.**

Какое отношение имеет EGEE к существующим гридам и исследовательским сетям?

Инфраструктура EGEE будет построена на исследовательской сети Европейского союза (ЕС) GEANT и в ней будет использован опыт, накопленный в таких проектах, как DataGrid, Datatag, Crossgrid, поддержанных ЕС, в национальных проектах, например, e-Science (Британия), INFN Grid, Nordugrid и американском Trillium. Она обеспечит возможность совместной работы с другими системами во всём мире, включая США и Азию, что будет способствовать установлению всемирной GRID-инфраструктуры.

Участие России в проекте EGEE



- В целях обеспечения полномасштабного участия России в этом проекте был образован 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”).

Russian Data Intensive GRID (RDIG) Consortium *EGEE Federation*



- **ИФВЭ** (Институт физики высоких энергий, Протвино)
- **ИМПБ** (Институт математических проблем биологии, Пущино)
- **ИТЭФ** (Институт теоретической и экспериментальной физики)
- **ОИЯИ** (Объединенный институт ядерных исследований, Дубна)
- **ИШМ** (Институт прикладной математики)
- **ПИЯФ** (Петербургский институт ядерной физики, Гатчина)
- **ИАЭ** (Институт атомной энергии)
- **НИИЯФ МГУ** (Научно-исследовательский институт ядерной физики МГУ)

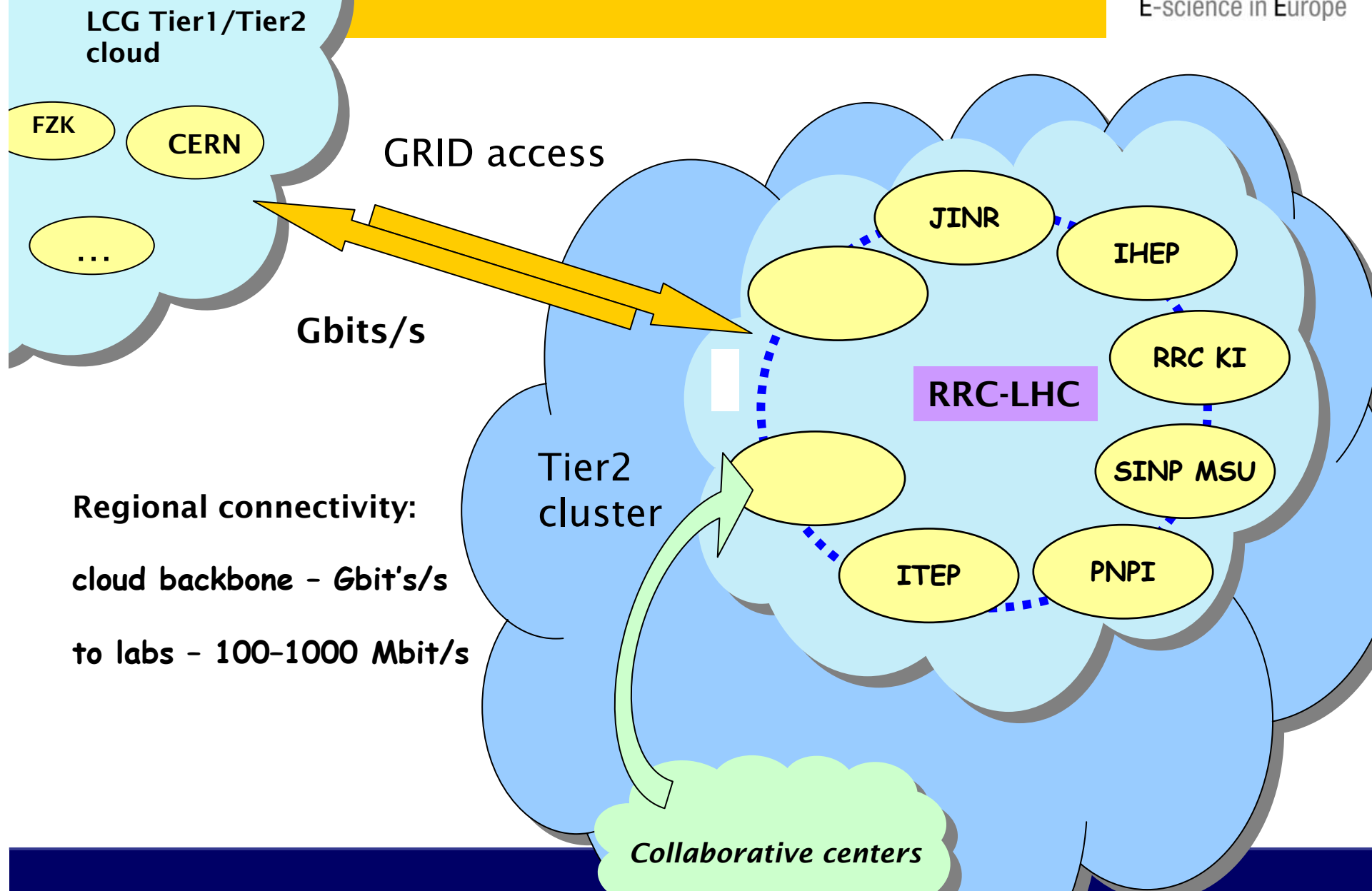
Участники РДИГ



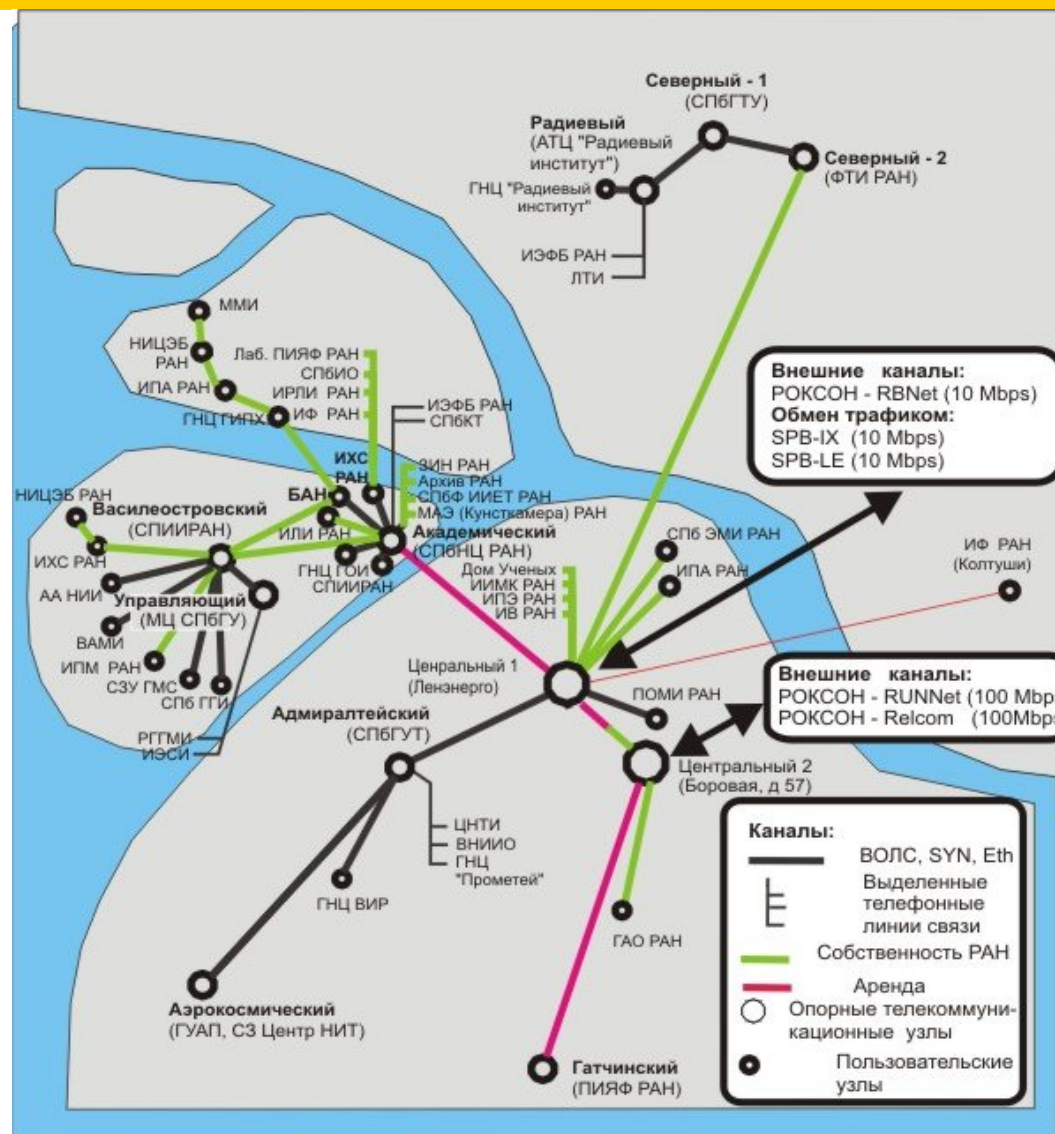
Первоначально – 8 институтов (**3 академических**). В 2005 году к числу ресурсных центров РДИГ добавились:

- **Геофизический центр РАН (Москва)**
- **Новгородский государственный университет им. Ярослава Мудрого (Новгород)**
- **Институт ядерных исследований РАН (Троицк)**

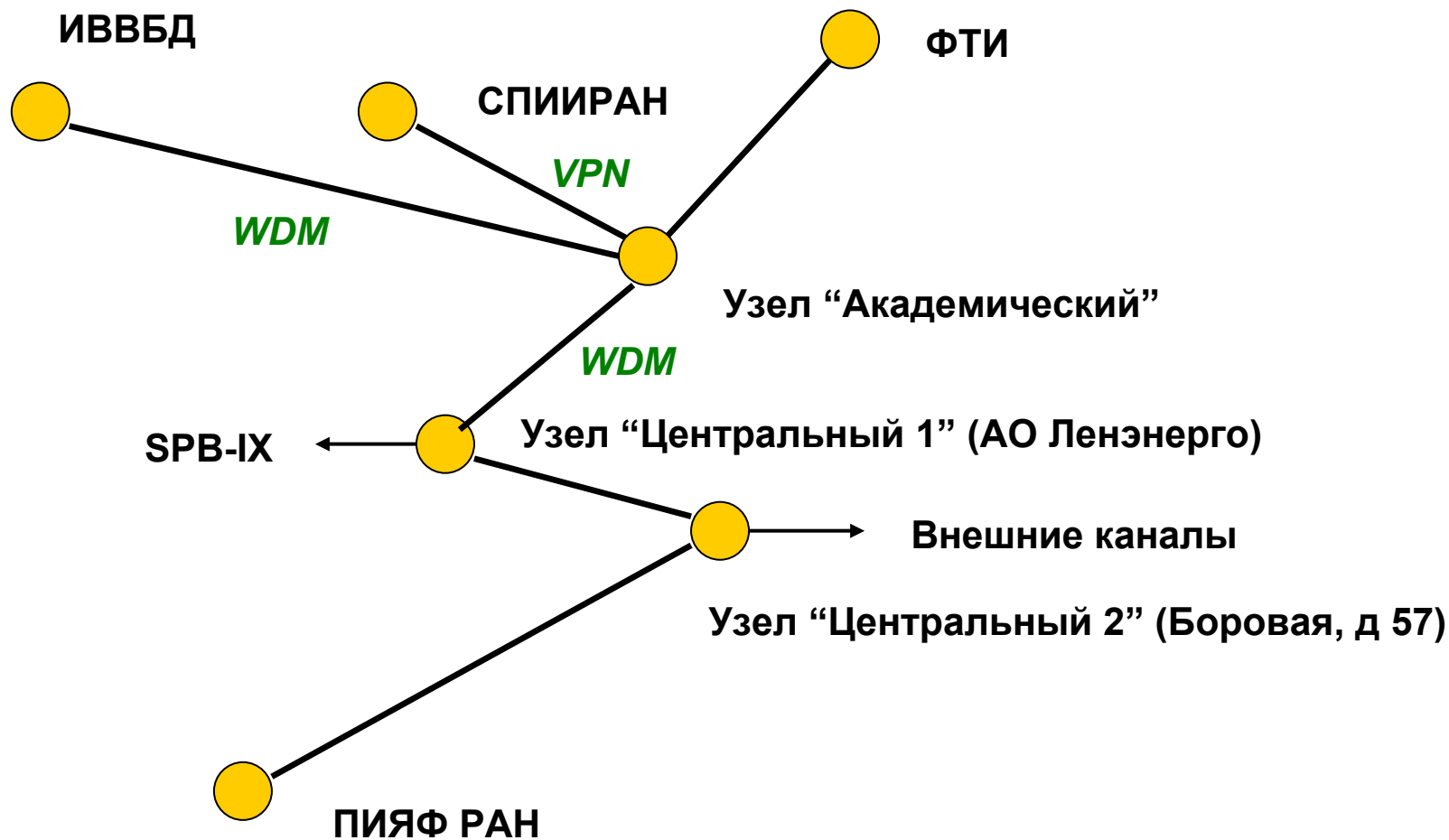
Российский региональный центр: *the DataGrid cloud*



Академическая сеть РОКСОН



Транспортная среда для организации распределённых вычислений



Направление работ по проекту в ПИЯФ РАН



NA2 - разработка и поддержка сайта (EGEE PNPI.nw.ru),
распространение Web инструментария, организация
презентаций в Санкт-Петербургском регионе

NA3(EGEEII) - проведение курсов, включая применение в ФВЭ,
подготовка
материалов для обучения

NA4 - подготовка для использования GRID технологии в LHC
экспериментах (симуляционные сеансы)

SA1(EGEEII) - поддержка функционирования и управление
региональным центром (ПИЯФ)

Создание вычислительной структуры для VO NW_RU

Кластерные вычислительные ресурсы РАН в Санкт-Петербурге

Санкт-Петербургский филиал МСЦ РАН (площадка ФТИ)	2xP-III 1,2 GHz –12 шт. 2xAMD Athlon XP 2400 – 7 шт. (38 CPU)
Санкт-Петербургский институт информатики и автоматизации РАН (площадка на Менделеевской линии В.О. И в Гавани)	2xP-IV 2.8 GHz – 10 шт. (20 CPU)
Петербургский институт ядерной физики им. Б.П.Константинова РАН (Гатчина)	P-IV 3,0 GHz – 30 шт. 2xXEON 2,8 GHz – 2 шт. 2x P-III-850 MHz – 5 шт. (44 CPU)

Useful References:

- **GLOBAL GRID FORUM:** [//www.gridforum.org](http://www.gridforum.org)
- **European GRID FORUM:** [//www.egrid.org](http://www.egrid.org)
- **PPDG:** [//www.ppdg.org](http://www.ppdg.org)
- **IVDGL:** [//www.ivdgl.org](http://www.ivdgl.org)
- **EU DATAGRID:** [//www.eu-datagrid.org](http://www.eu-datagrid.org)
- **DATATAG:** [//www.datatag.org](http://www.datatag.org)
- **GLOBUS:** [//www.globus.org](http://www.globus.org)
- **GriPhyn:** [//www.griphyn.org](http://www.griphyn.org)
- **CrossGRID:** [//www.crossgrid.org](http://www.crossgrid.org)
- **Condor:** [//www.cs.wisc.edu/condor](http://www.cs.wisc.edu/condor)
- **TERAGRID:** [//www.teragrid.org](http://www.teragrid.org)
- **EGEE:** [//www.eu-egee.org](http://www.eu-egee.org)
- **LCG:** [//lcg.web.cern.ch/LCG/](http://lcg.web.cern.ch/LCG/)
- **EGEE-RDIG:** [//www.egee-rdig.ru](http://www.egee-rdig.ru) - представлены материалы (слайды этих и других курсов)



If "The Grid"
vision leads us
here...

... then where are
we now?