



# LHC computing and the GRID

**Евгений Мартынов**

Институт теоретической физики им. Н.Н. Боголюбова  
Национальная академия наук Украины  
ALICE Collaboration



## План

- LHC, эксперименты, данные, требования к вычислительным ресурсам
- Высокопроизводительные вычисления
  - Суперкомпьютеры
  - Распределенные вычисления
- Грид-технологии:
  - идея и история развития
  - Грид-приложения
- Грид в ЦЕРНе: WLCG
- Грид в Украине: UNG



WLCG  
Worldwide LHC Computing Grid

# LHC, эксперименты

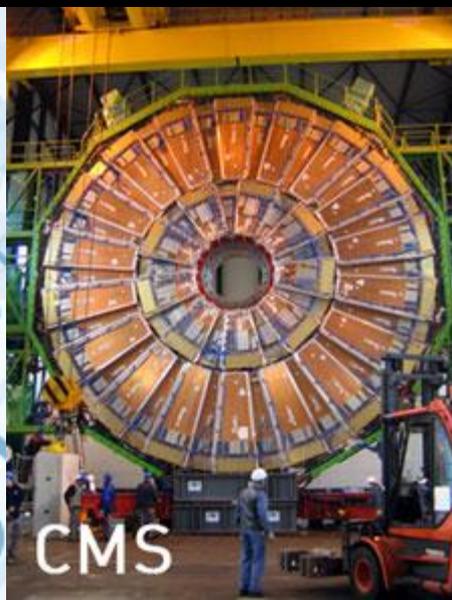


**Size:** 25 x 46 meters  
**Weight:** 7500 tons

**Size:** 16 x 21 meters  
**Weight:** 12500 tons

**Size:** 16 x 26 meters  
**Weight:** 10000 tons

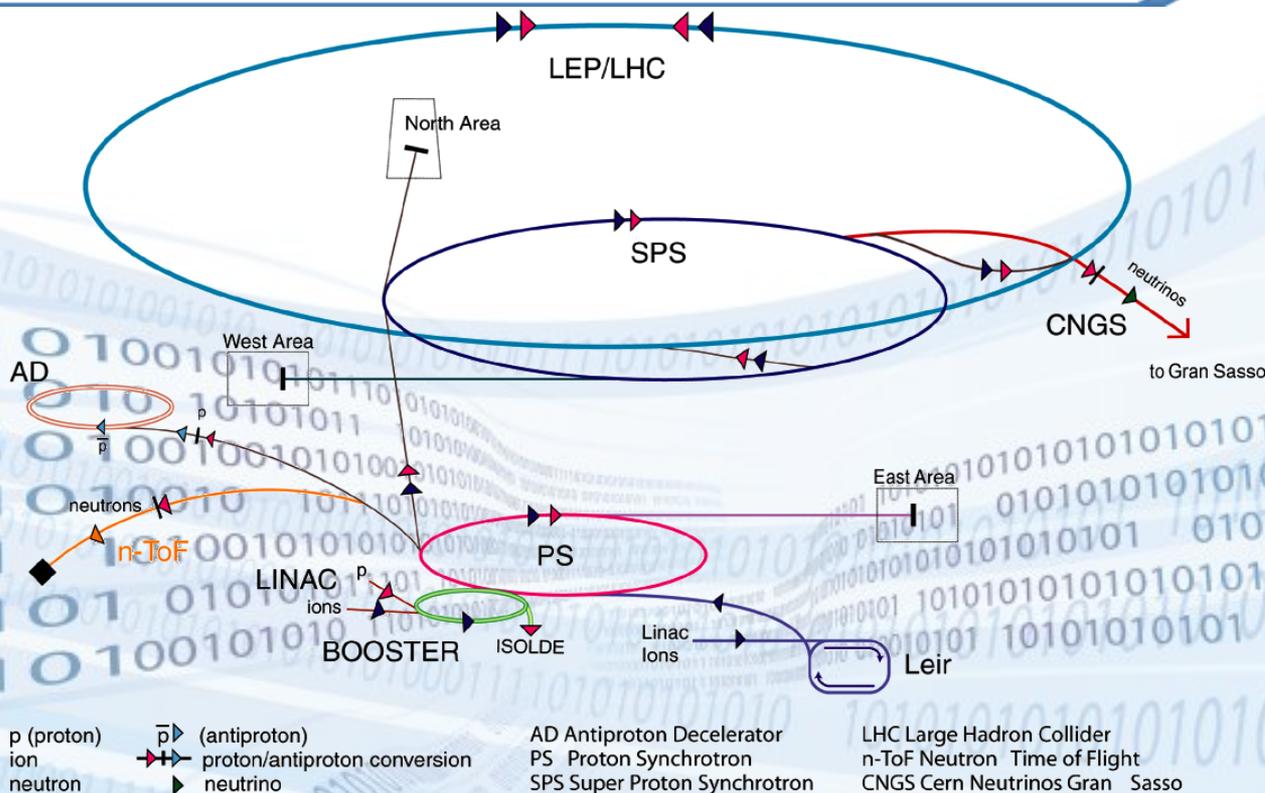
**Size:** 15 x 20 meters  
**Weight:** 4500 tons





# LHC

Worldwide LHC Computing Grid



No. of bunches per proton beam	2808
No. of protons per bunch (at start)	$1.1 \times 10^{11}$
Number of turns per second	11 245
Number of collisions per second	600 million



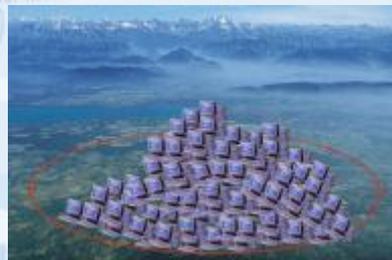
# Данные, вычислительные ресурсы

**В шести экспериментах производится 15-20 петабайт (15-20 млн. гигабайт) данных в год. Их нужно сохранить и проанализировать.**

Данные только одного эксперимента	ежегодно	Вес DVD дисков
Первичные данные	~ 4 PB	14000 кг
Физические данные	0.2 PB	784 кг

**Для всех экспериментов – около 4 млн. DVD дисков или около 20 млн. CD дисков**

**Для обработки и анализа этих данных требуется около 100000 персональных компьютеров (конец прошлого – начало нынешнего столетий)**

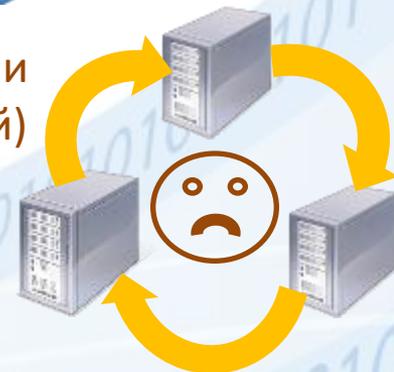


**Компьютерный центр ЦЕРНа даже после существенной модернизации последних лет может обеспечить только ~40% ресурсов для хранения данных и около 30% требуемых вычислительных ресурсов.**



# Высокопроизводительные вычисления

Идея соединить компьютеры для более эффективного решения той или иной задачи (создание системы высокопроизводительных вычислений) очевидна и существует с тех пор, как появились компьютеры. Однако реализация оказалась более трудной, чем представлялось вначале.



Два направления развития высокопроизводительных вычислений:

## локальные системы, распределенные системы

### ❑ Локальные системы:

Суперкомпьютеры (дорого в целом, дорогая техн. поддержка)



Roadrunner



Jaguar



Pleades

Кластеры.

- Shared memory systems
- Distributed memory systems

Первый Beowulf cluster построен в 1994. Так назывался один из Linux-кластеров в NASA. Особенностью такого кластера также является **масштабируемость**, то есть возможность увеличения количества узлов системы с пропорциональным увеличением производительности.

**!!!** Кластеры могут быть гомогенными и гетерогенными



WLCG  
Worldwide LHC Computing Grid

# Высокопроизводительные вычисления



PRESENTED BY  
UNIVERSITY OF  
MANNHEIM

ICL  
INNOVATIVE  
COMPUTING LABORATORY  
UNIVERSITY OF TENNESSEE

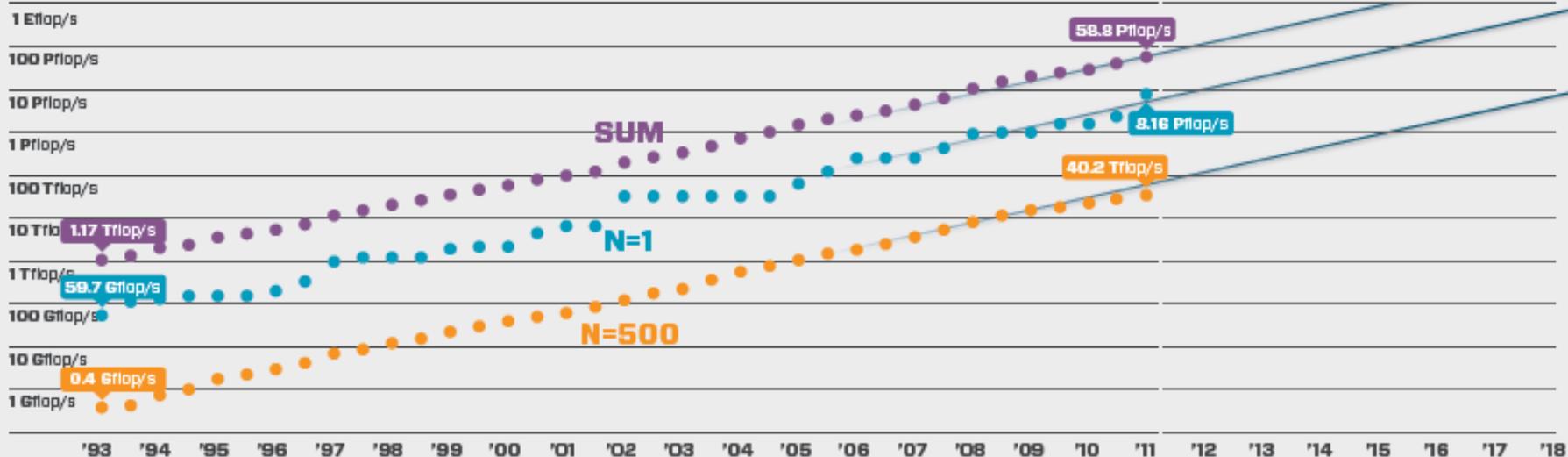
BERKELEY LAB  
Lawrence Berkeley  
National Laboratory

FIND OUT MORE AT  
[www.top500.org](http://www.top500.org)

	NAME/MANUFACTURER/COMPUTER	LOCATION	COUNTRY	CORES	R <sub>max</sub> P flop/s
1	<b>K Computer</b> SPARC64 VIII fx 2.0GHz, Tofu interconnect	RIKEN	Japan	548,352	8.16
2	<b>Tianhe-1A</b> 6-core Intel X5670 2.93 GHz + Nvidia M2050 GPU w/custom interconnect	NUDT/NSCC/Tianjin	China	186,368	2.56
3	<b>Jaguar</b> Cray XT-5 6-core AMD 2.6 GHz w/custom interconnect	DOE/SC/ORNL	USA	224,162	1.76
4	<b>Nebulae</b> Dawning TC3600 Blade Intel X5650 2.67 GHz, NVidia Tesla C2050 GPU w/ Iband	NSCS	China	120,640	1.27
5	<b>Tsubame 2.0</b> HP Proliant SL390s G7 nodes (Xeon X5670 2.93GHz), NVIDIA Tesla M2050 GPU w/Iband	TITech	Japan	73,278	1.19

## PERFORMANCE DEVELOPMENT

## PROJECTED





## □ Распределенные системы

*Распределенные вычисления = Метакомпьютинг => Грид*

**Главная задача (и трудность) – научить систему удаленных друг от друга компьютеров работать согласованно**

### **90-е годы:**

Аргоннская Национальная Лаборатория - первые успешные попытки.

Подготовлены условия для реализации идеи

- наличие мощных компьютеров и кластеров,
- программные методы параллельных вычислений,
- высокая пропускная способность интернет-каналов

Предшественники грида – проекты

**FAFNER** (Factoring via Network-Enabled Recursion, распределенные методы в задачах шифрования) и

**I-WAY** (Information Wide Area Year, объединение суперкомпьютеров и кластеров)

I. Foster, C. Kesselman, S. Tuecke - Globus Toolkit – системные средства для распределенных вычислений



## 1997 – год рождения грида

Конференция в АНЛ;  
«Построение вычислительного грида».

Сформулированы основные определения,  
компоненты и требования к грид-системе.



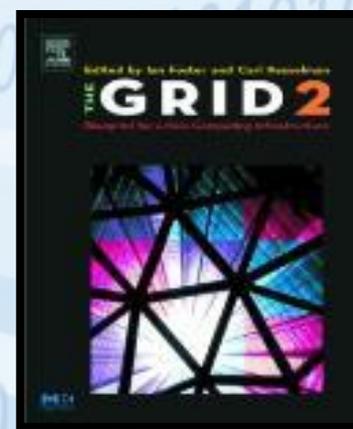
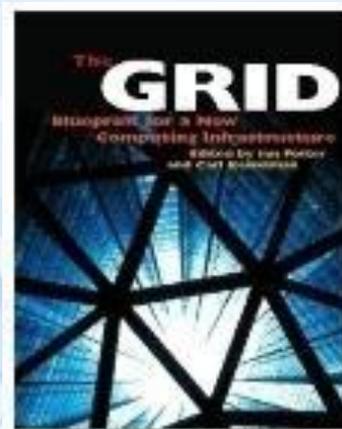
I. Foster



C. Kesselman

1998 - I. Foster, C. Kesselman: “The Grid: Blueprint  
for a New Computing Infrastructure”

Книга получила название «Библия Грида»





# Грид-технологии

I. Foster, C. Kesselman:

« Грид – это **пространственно распределенная** операционная среда с гибким, безопасным и скоординированным разделением ресурсов для выполнения приложений в динамически образующихся **виртуальных организациях**.

ВО – это добровольное объединение институтов лабораторий, согласных выделить часть своих вычислительных ресурсов в общее пользование, получая при этом доступ к ресурсам других участников ВО. Это есть основа реальной работы в гриде

## World Wide Web

Интернет обеспечивает прямое обращение к информации, хранящейся в миллионах географически различных мест



## Grid

Инфраструктура, которая обеспечивает прямой доступ к вычислительным мощностям и ресурсам хранения данных, распределённых по всему Земному шару.

( ≈ вычислительный Интернет, но не только)





# Грид-технологии

Электрические сети

Аналогия

Грид

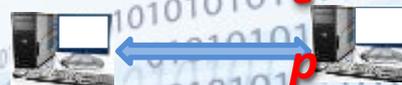


Электростанции

Пользователь

Линии связи

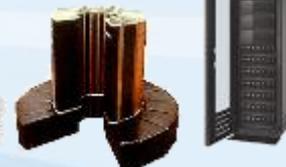
Ресурсы



Управляющий узел



Вычислительные ресурсы



Хранилища данных

П  
р  
о  
г  
р  
а  
м  
н  
о  
е

Г  
р  
и  
д

о  
б  
е  
с  
п  
е  
ч  
е  
н  
и  
е



# Грид-технологии. Идея

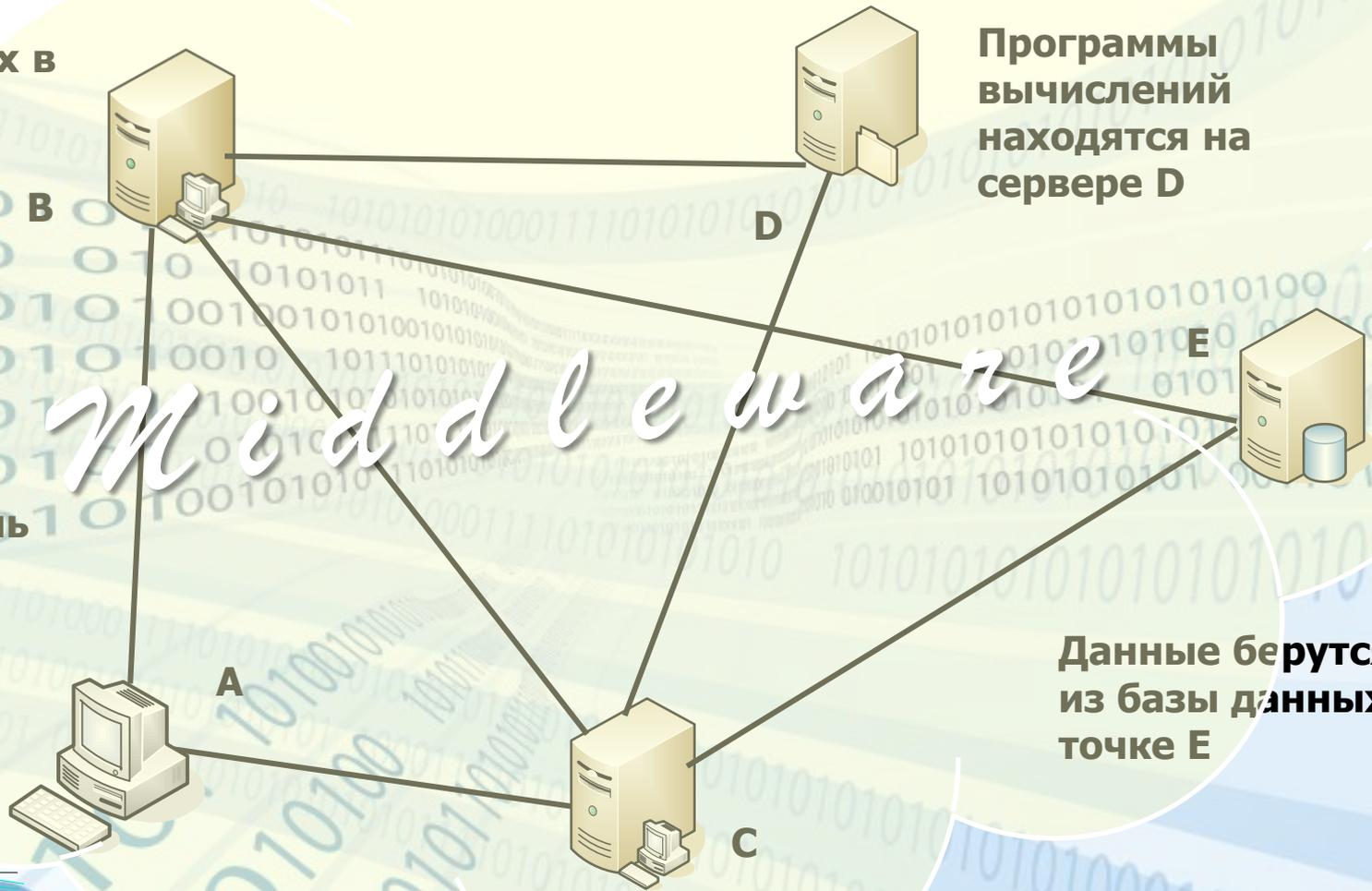
Задача  
решается на  
компьютерах в  
точках В и С

Программы  
вычислений  
находятся на  
сервере D

Middleware

Пользователь  
направляет  
задачу из  
точки А

Данные берутся  
из базы данных в  
точке Е





# Грид-технологии. Основные элементы

## Три компонента грид-инфраструктуры:

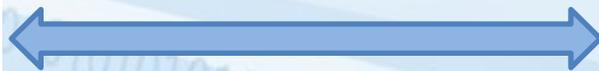
Вычислительные ресурсы (OS – Linux)  
(кластеры, хранилища данных)



Middleware



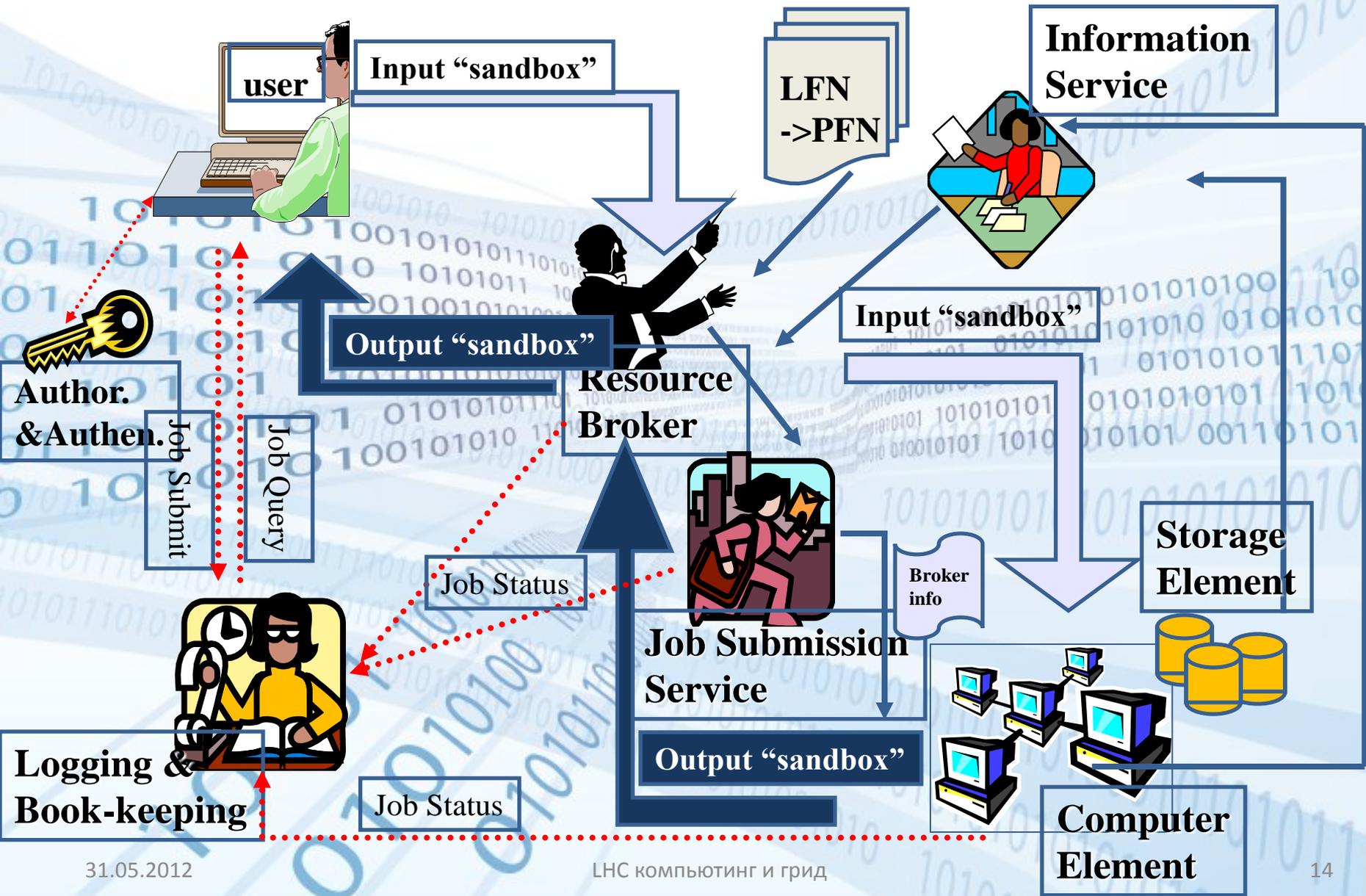
Каналы связи



Чем более быстрые каналы связи, тем более эффективно работает грид

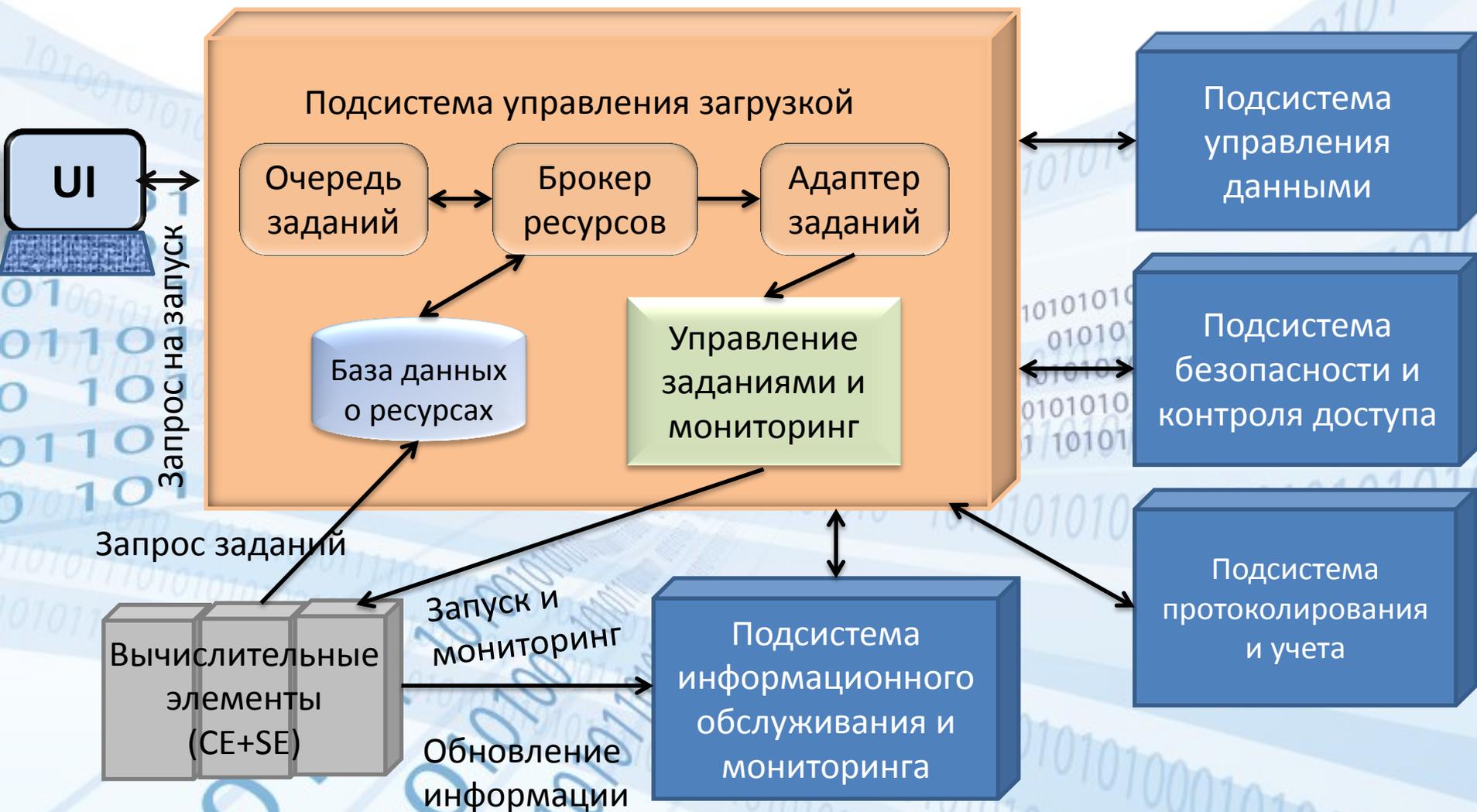


# Как «устроен» грид?





# Общая структура системы управления заданием в грид-среде





# Возможности грид-технологий

**Решение вычислительных задач, требующих беспрецедентных вычислительных ресурсов**

**Решение задач, требующих перебора большого количества вариантов (а следовательно, большого времени вычислений)**

**Ставятся задачи, которые ранее не было смысла формулировать и решать из-за недостатка ресурсов или нереального времени решения**

**Обработка и анализ сколь угодно больших объемов данных**

**Обработка и анализ данных графического характера высоко разрешения (статические изображения, аудио-, видео-потoki)**

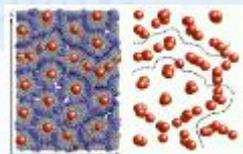
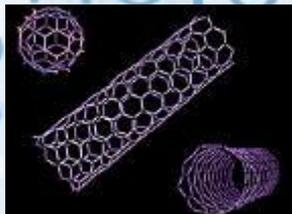
**Онлайновые задачи мониторинга и управления процессами.**

***Необходимые ресурсы обеспечиваются «простым объединением» географически распределенных элементов (кластеров, компьютеров, баз данных)***



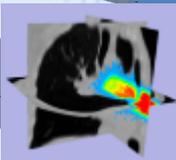
Worldwide LHC Computing Grid

# Где применяется грид?

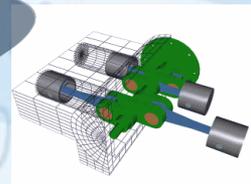


Solid state physics

Earth Sciences Environment



Engineering New materials electronics



Economics



Medicine



Astrophysics Astronomy Radioastronomy

Molecular and cell biology

High Energy Physics

## Grid applications



Worldwide LHC Computing Grid

# Примеры использования грида



Open eHealth



## Medical Grid

**Medical Grid** создается как широкомасштабная грид-инфраструктура, которая объединит распределенные по всему миру и связанные грид-сетью базы медицинских данных, системы визуализации для анализа медицинских изображений.

## Health-e-Child

Цель – создание грид-инфраструктуры для европейской педиатрии



**Результат использования (2006 г.)  
грид-технологий в исследовании H5N1**

**WISDOM** (Wide In Silico Docking On Malaria) – расчёты, требующие значительных компьютерных ресурсов, для поиска лекарств от заболеваний



**Имеется множество других примеров**

RESULTS ALREADY ACHIVED	
Number of docked compounds	2,5 million
Duration of the experience	6 weeks
Estimated duration on 1 PC	105 years
Number of computers	2000
Number of countries giving computers	17
Volume of data produced	600 GB



WLCG  
Worldwide LHC Computing Grid

# Международные грид-проекты и организации

## WLCG (Worldwide LHC Computing Grid, ЦЕРН)

<http://lcg.web.cern.ch/LCG/>



WLCG  
Worldwide LHC Computing Grid

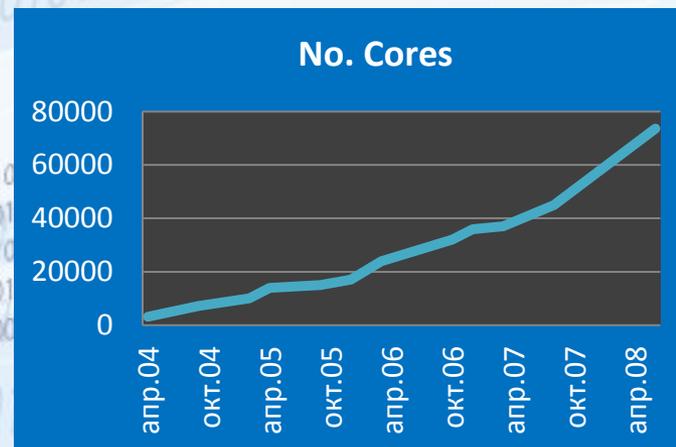
## EGEE (Enabling Grids for E-science)

50 countries, 94 partners

<http://www.eu-egee.org/>



Более 15 приложений в EGEE



## EGI (European Grid Infrastructure)

38 National Grid Initiatives

<http://www.egi.eu/>



European Grid Infrastructure  
Towards a sustainable grid infrastructure

и многие другие...



Грид-технологии - это богатые возможности для принципиально новой организации научных исследований и международной кооперации ученых.

Виртуальные грид-организации в сфере научных исследований – это, по сути, международные научные институты и лаборатории. Они не имеют юридического статуса, но могут и уже очень эффективно решают научные проблемы.

Взаимодействие между учеными не только на индивидуальном уровне (e-мэйл, интернет), но и резко усиливаются коллективная компонента сотрудничества, междисциплинарные связи.

Грид позволяет более комплексно решать научные проблемы, быстрее доводить найденные решения до практических приложений.

Все более глубокое содержание приобретает термин E-наука.



**Компьютерные вычисления, вычисления, требующие все более значительных вычислительных ресурсов становятся неотъемлемой частью, а иногда и основным методом научных исследований разработки научно-технических приложений.**



**Е-наука – новый термин, появившийся около 10 лет назад, применяется к тем наукам, в которых многие результаты невозможно получить, не используя беспрецедентные вычислительные ресурсы и новые информационные технологии (например, грид или суперкомпьютеры).**



На первый план выходят глобальные проекты, направленные на создание европейской инфраструктуры, которая охватывает не только все страны, но и все стороны жизни общества, от науки и индустрии до образования, медицины, общественного управления (e-Government), гуманитарных аспектов, с широким использованием высокоскоростных каналов обмена информацией, новейших информационных технологий, включая грид и облачные (cloud) приложения.



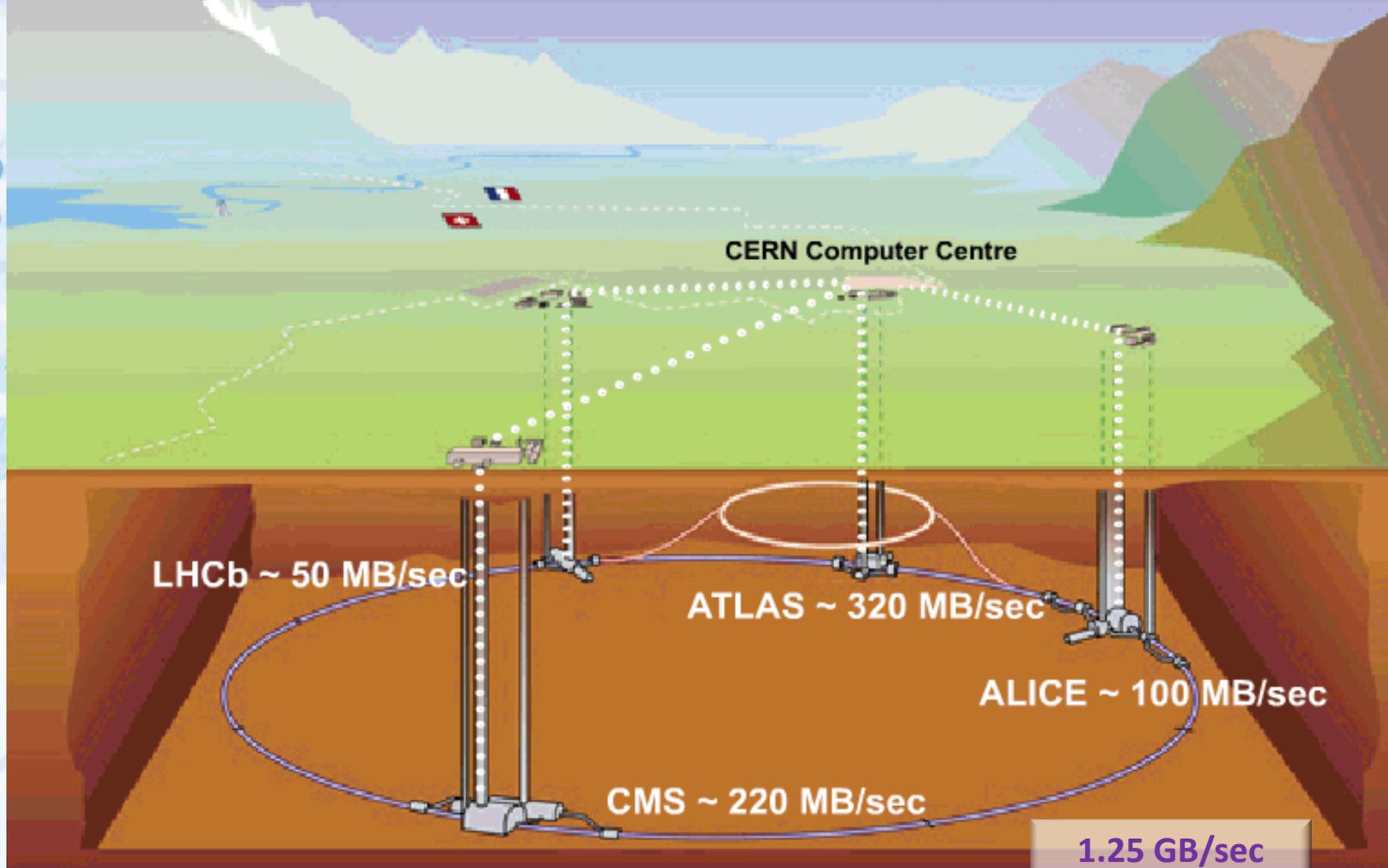
ESFRI

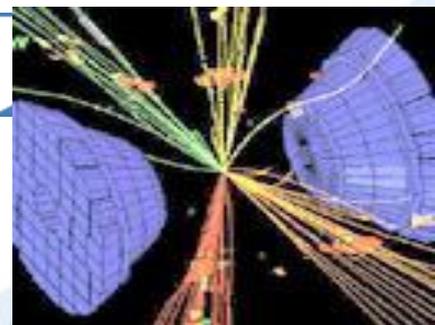
European Strategy Forum  
on Research Infrastructures



e-solutions

LHC: съем и первичный (триггер) отбор экспериментальной информации;  
решение проблемы обработки данных – LHC-грид (WLCG)





**Селективность:  
1 из  $10^{13}$**

*Подобно поиску  
одного человека  
среди тысячи  
мировых  
популяций!*

*Или поиску одной  
иголки в 20  
миллионах стогах  
сена!*

- Столкновение частиц = **событие (an event)**

- Цель - отследить и описать все рождённые частицы и таким образом полностью **реконструировать** процесс.

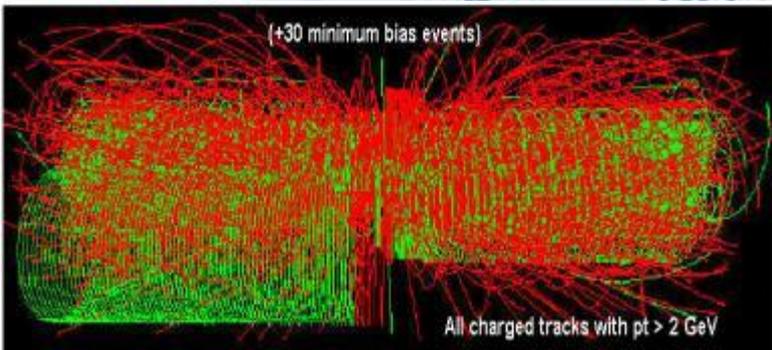
- Существование особых видов или форм событий (“special shapes”) среди все треков является указанием на наличие интересных взаимодействий.

*Один из путей найти **Higgs boson**: искать характерный распад с образованием 4-х лептонов*

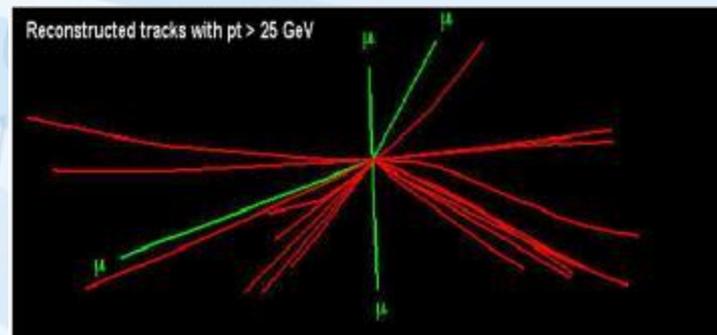
Начиная с такого события,

нужно выделить

«характерное, искомое» событие

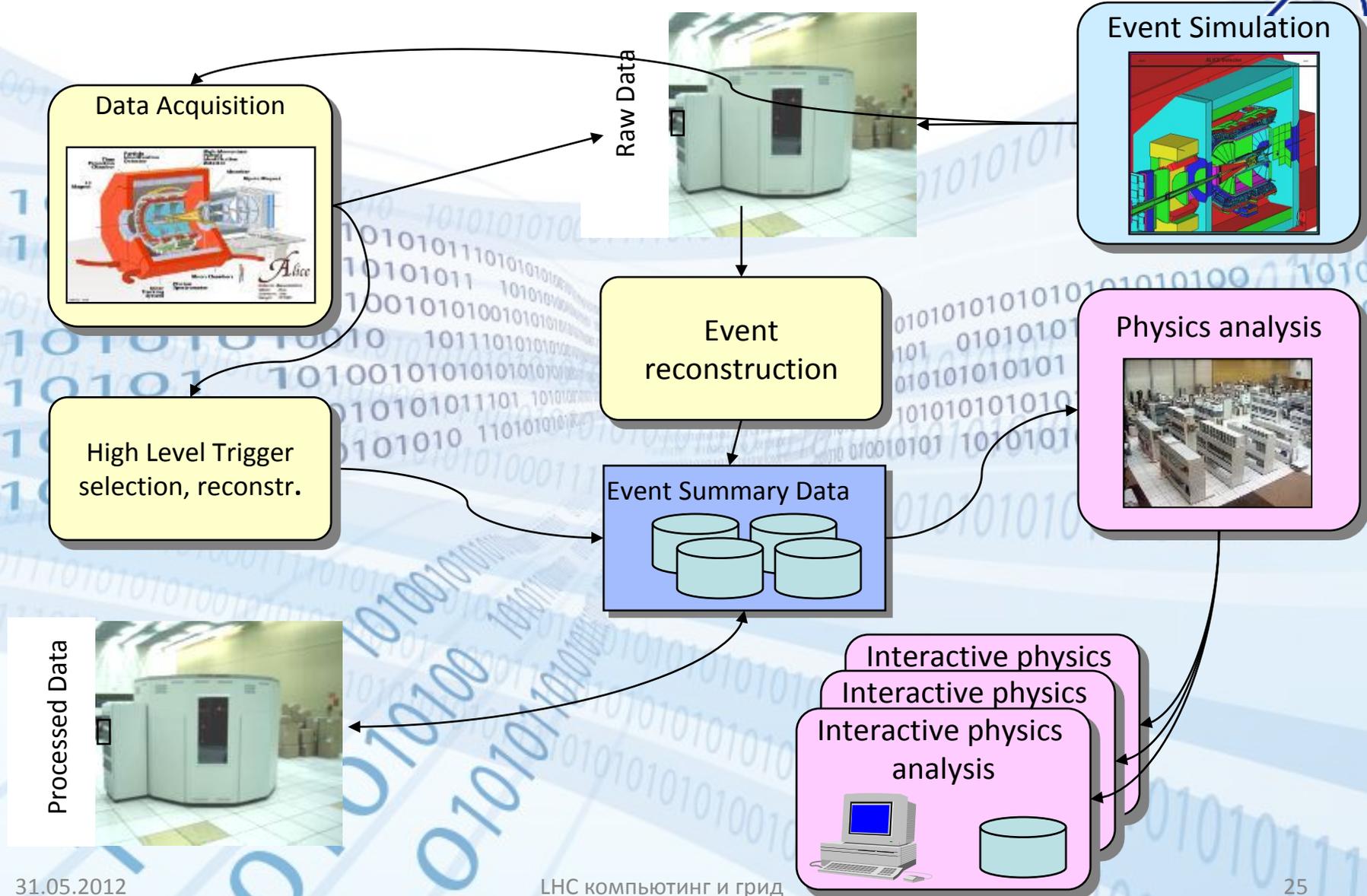


LHC компьютеринг и грид





# Поток экспериментальных данных





WLCG  
Worldwide LHC Computing Grid

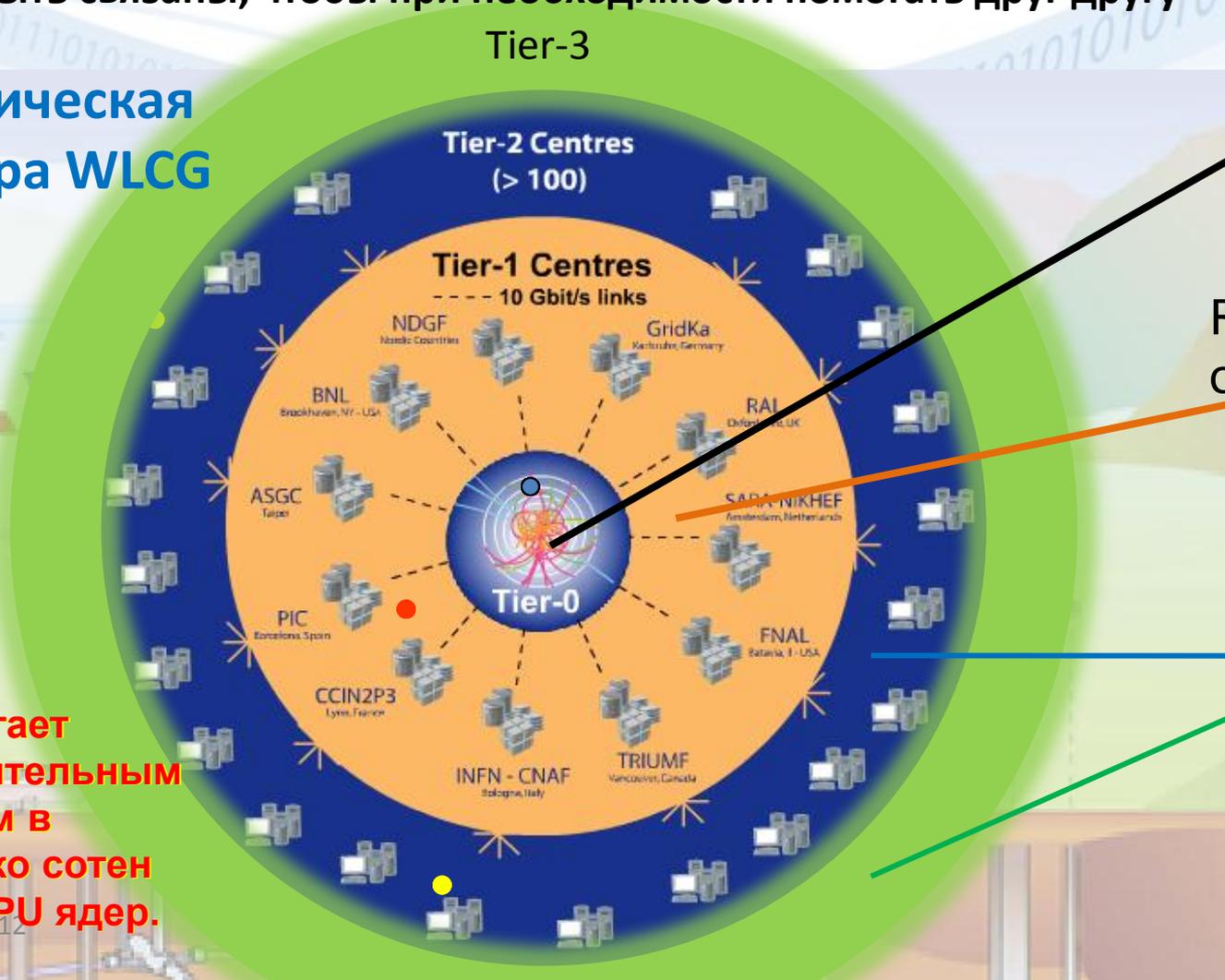
# Грид в ЦЕРНе: WLCG



<http://lcg.web.cern.ch/LCG/Default.htm>

Каждый из экспериментов имеет свою грид-инфраструктуру, однако, они должны быть связаны, чтобы при необходимости помогать друг другу

## Иерархическая структура WLCG



Raw data

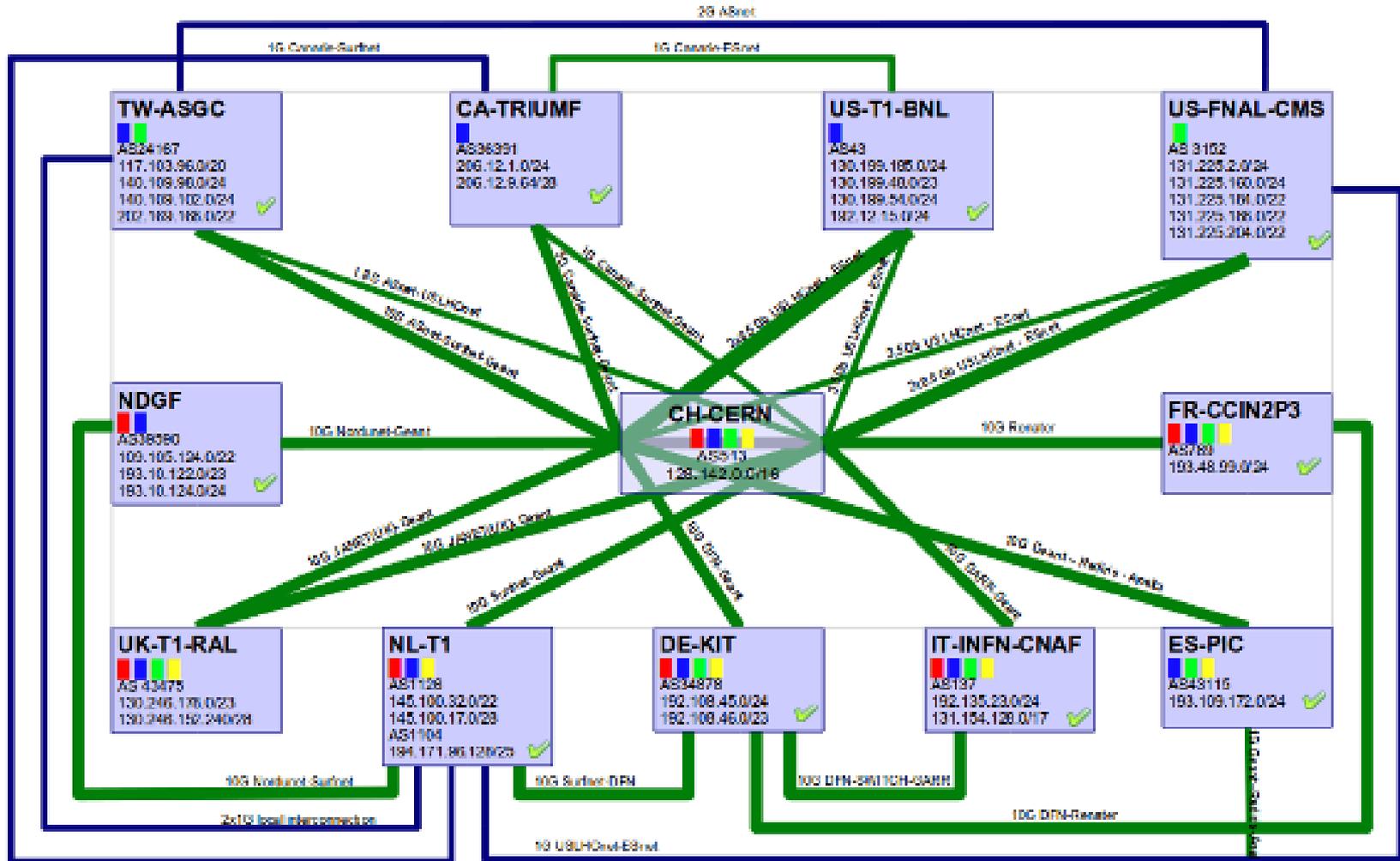
Reconstruction of events

Physical analysis

Сейчас WLCG располагает вычислительным ресурсом в несколько сотен тысяч CPU ядер.

31.05.2012

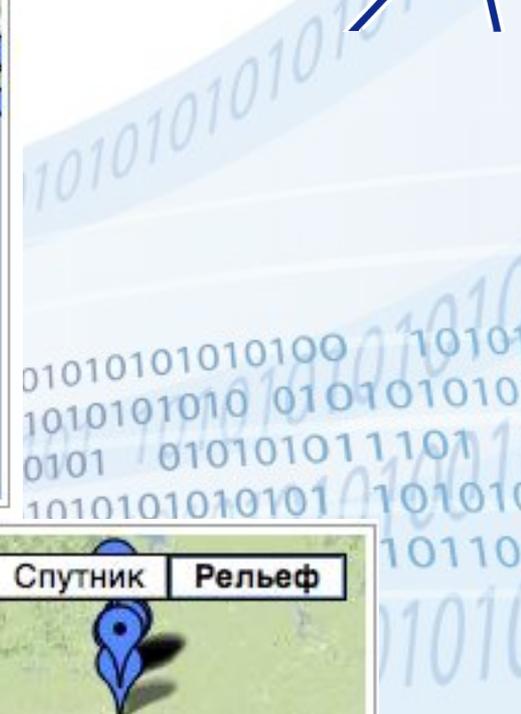
## LHCOPN



— T0-T1 and T1-T1 traffic  
— T1-T1 traffic only  
— Not deployed yet  
(thick) > 10Gbps  
(thin) = 100Gbps  
■ = Alice ■ = Atlas  
■ = CMS ■ = LHCb  
✓ = internet backup available  
 p2p prefix: 192.18.188.0/24  
 edward.mactell@cern.ch 20100918



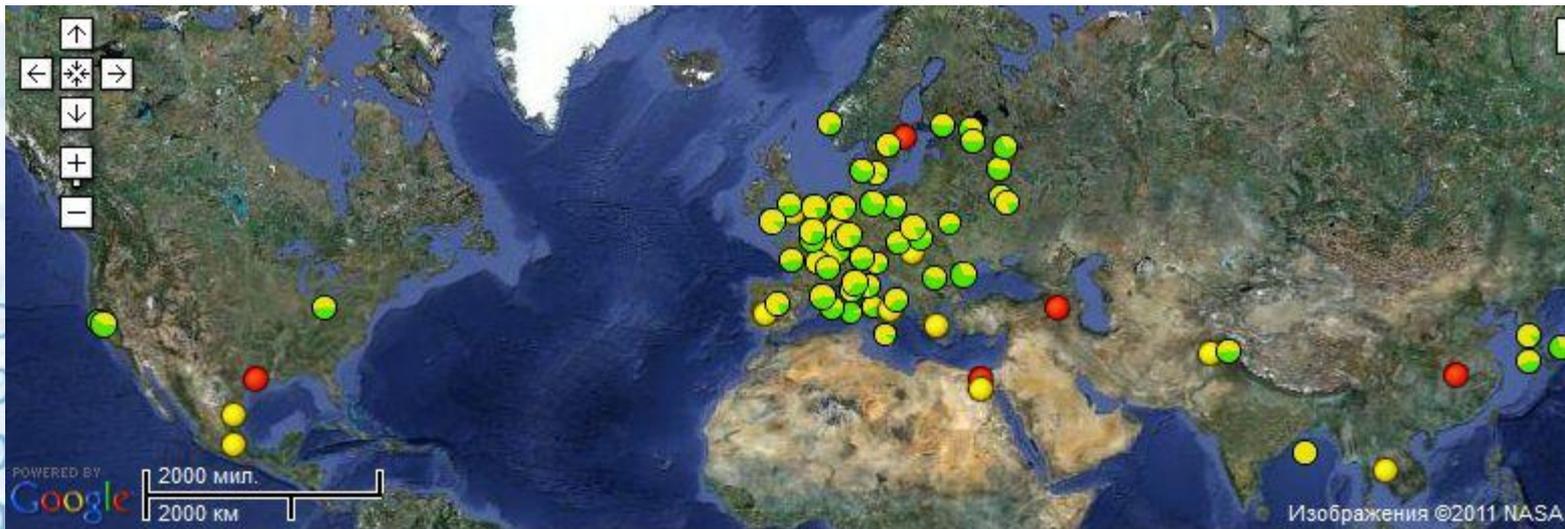
# Monitoring WLCG





Worldwide LHC Computing Grid

# Грид в эксперименте ALICE



**5 Tier-1 сайтов**  
**71 Tier-2 сайт**  
**CPU - ~50000**  
**SE - 13PB**

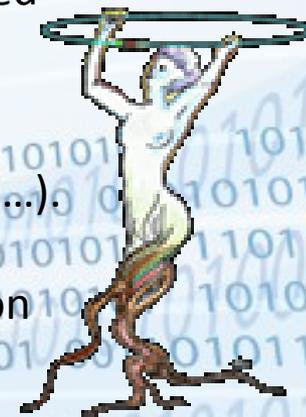
## – средства обработки и анализа экспериментальных данных в эксперименте ALICE

**AliEn** is a lightweight Open Source Grid Framework built around other Open Source components using the combination of a Web Service and Distributed Agent Model.



**ROOT** is an object-oriented framework aimed at solving the data analysis challenges of high-energy physics (☺ визуализация, вычисления, MINUIT, ...).

**AliRoot** is the name ALICE Off-line framework for simulation, reconstruction and analysis. It uses the **ROOT** system as a foundation on which the framework and all applications are built.



ROOT

*Programming inside a **framework** is a little like living in a city. Plumbing, electricity, telephone, and transportation are services provided by the city. In your house, you have interfaces to the services such as light switches, electrical outlets, and telephones. The details, for example, the routing algorithm of the phone switching system, are transparent to you as the user. You do not care; you are only interested in using the phone to communicate with your collaborators to solve your domain specific problems.*

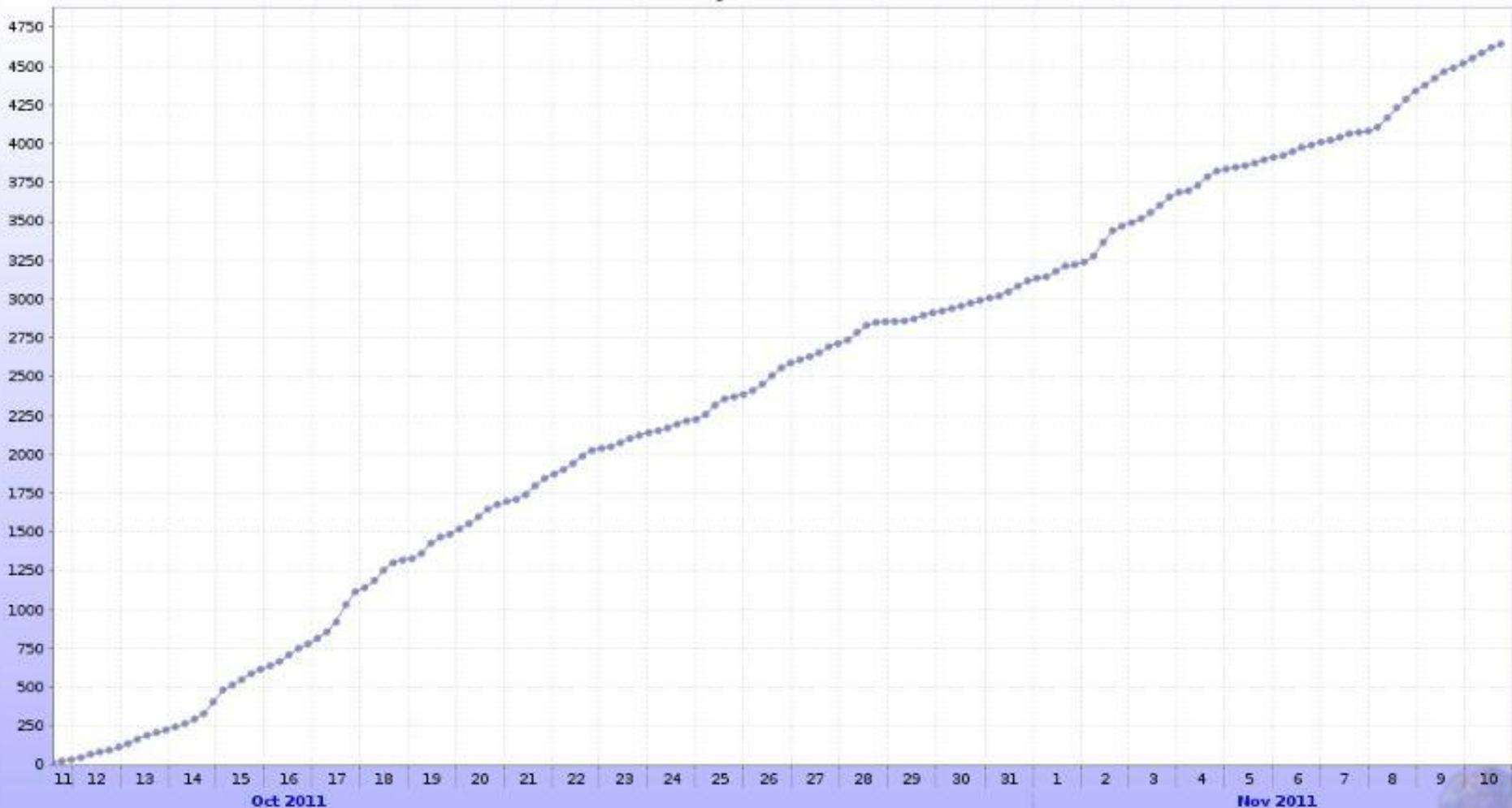
С 11 октября по 10 ноября 2011г.

Interval selection: last month or « 2011-10-11 10:00 - 2011-11-10 20:00 » Plot

Annotations

What is this about?

## Done jobs in BITP



С 11 октября по 10 ноября 2011г.

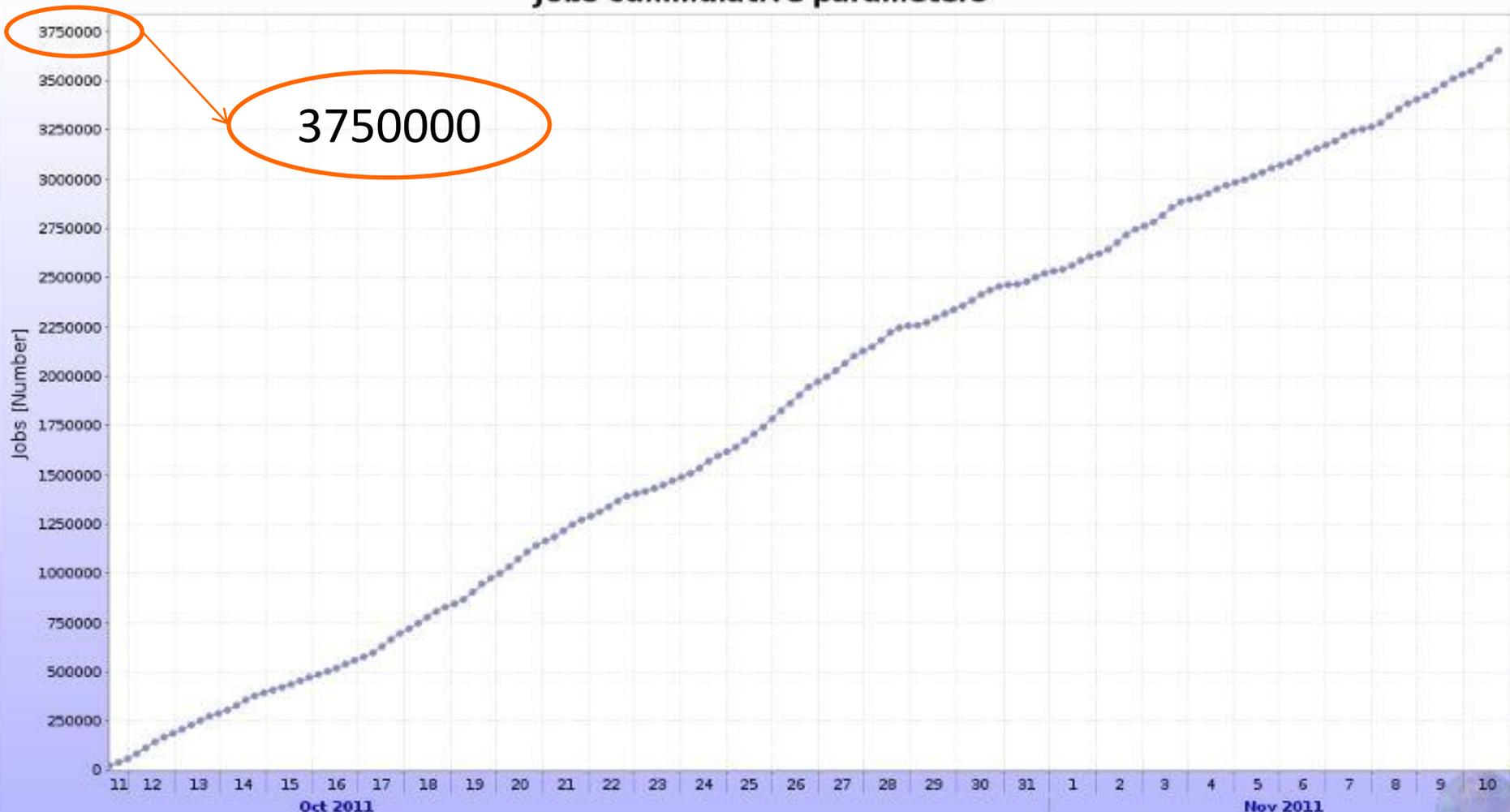
Interval selection: last month or 2011-10-11 10:00 - 2011-11-10 20:00

Plot

Annotations

What is this about?

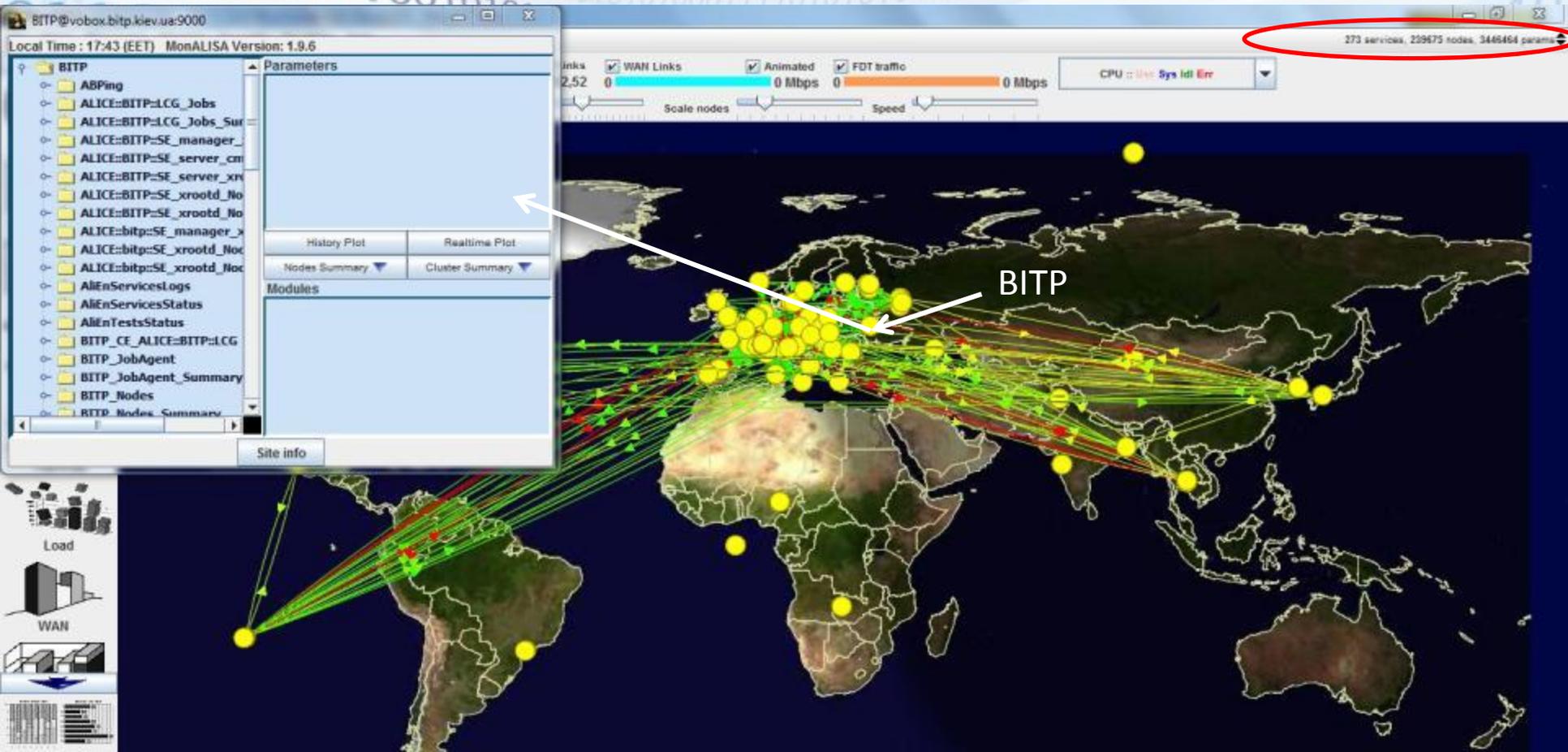
## Jobs cumulative parameters



Done

# Мониторинг всемирного грида

Мониторинг осуществляется из Калифорнийского технологического института. Видны только грид-сайты, на которых установлено и работает соответствующее программное обеспечение.



## Грид в Украине, история создания

- 2004 - 2005** Первые грид кластеры. Харьковский физико-технический институт (часть Российского грида, CMS эксперимент). Грид-кластер в ИТФ и КНУ, AliEnGrid для эксперимента ALICE)
- 2006** - Проект “Внедрение грид-технологий и строительство кластеров в НАН Украины”.
- 5 кластеров в институтах НАН Украины.
  - Украина подписывает соглашение об участии в WLCG
- 2007** - Грид-сегмент (*6 институтов НАН Украины и КНУ в Киеве, один институт во Львове*)
- Украина становится ассоциированным членом EGEE
  - Украина становится членом EGI-DS
- 2008** - Распределенная грид-инфраструктура НАН Украины, Украинский Академический Грид (*22 института НАН Украины + КНУ, КПИ*) в Киеве, Харькове, Днепропетровске, Сумах, Львове.
- 2009** - Государственная целевая научно-техническая программа внедрения и использования грид-технологий на 2009-2013 годы. Бюджет – 300 млн. грн.

# Грид в Украине, состояние

BITP, Institute for theoretical physics  
 ICYB, Institute of cybernetics  
 IFBG, Institute of food biotechnology and genomics  
 IMAG, Institute of magnetism  
 IMATH, Institute of mathematics математики  
 IMBG, Institute of molecular biology and genetics  
 IMP, Institute of metal physics  
 IMMPS, Institute of mathematical machine and system problem  
 Inparcom, PA «Electronmash»  
 IOP, Institute of physics  
 ISoftS, Institute of software systems  
 KMA, Kyiv-Mogila Academy  
 KNU, Kyiv national university  
 KPI, Kyiv polytechnic institute  
 MAO, Main astronomic observatory  
 PIMEE, Institute for Modelling in Energy Engineering  
 SRI, Space research institute

ILTPE, Institute for low temperature physics and engineering  
 ISMA, Institute of scintillating materials  
 IRE, Institute of radiophysics and electronics  
 KIPT, Kharkiv institute for physics and technology  
 RI, Radioastronimics Institute

IAP, Institute of applied physics

IGTM, Institute of geotechnical mechanics

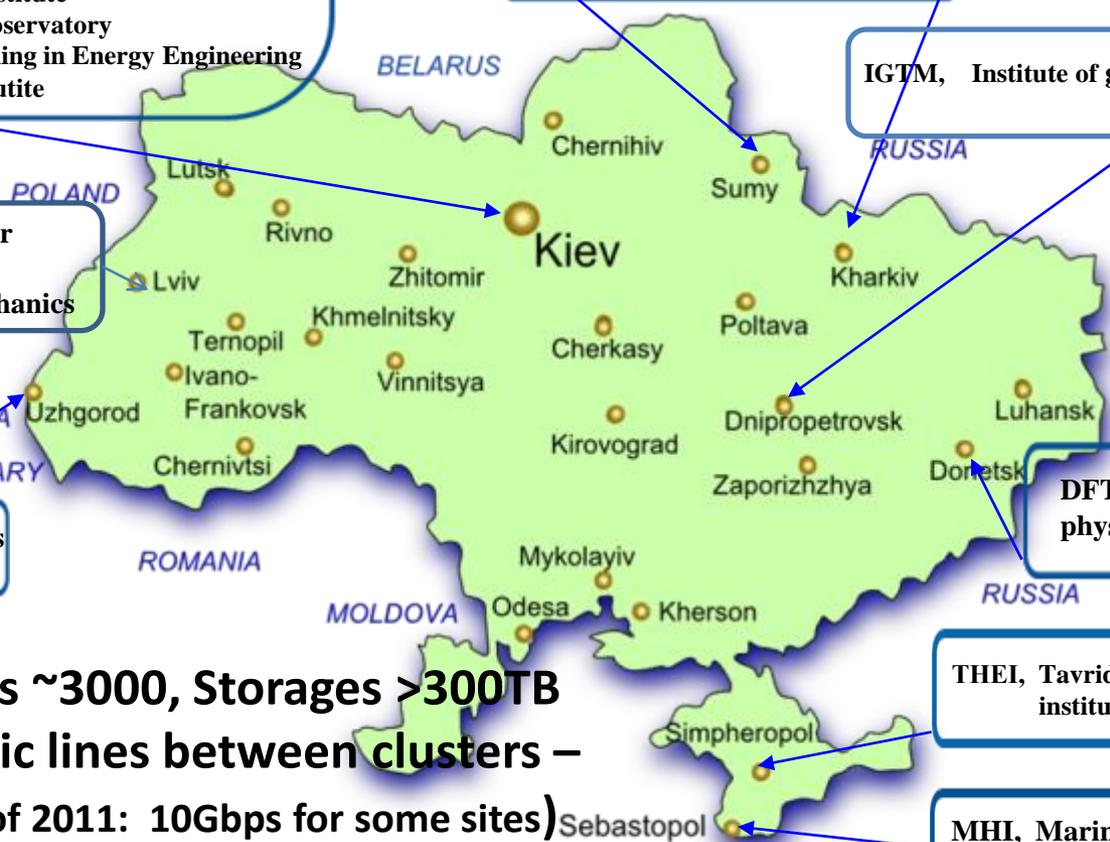
ICMP, Institute of condensed matter physics  
 IPM, Institute of physics and mechanics

IEP, Institute of electron physics

DFTI, Donetsk institute for physics and engineering

THEI, Tavrida humanitarian ecological institute

MHI, Marine hydrophysic institute



**Resources: CPU cores ~3000, Storages >300TB**  
**Capacity of fiber optic lines between clusters – 300-1000Mbps (end of 2011: 10Gbps for some sites)**

31.05.2012

# Грид в Украине

## оптоволоконные каналы передачи данных



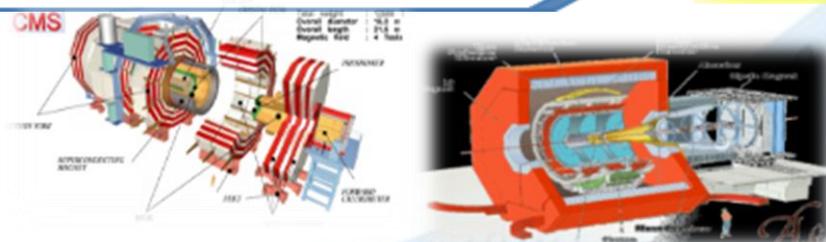
# Грид в Украине, приложения

## Физика высоких энергий

Обработка и анализ данных LHC.

CMS эксперимент: **KIPT**

ALICE эксперимент: **BITP, KNU, KPI, ISMA**



## Астрофизика и астрономия

MAO, BITP, KNU



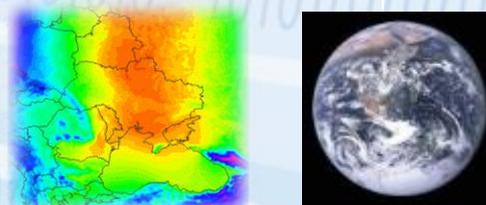
## Молекулярная и клеточная биология, медицина

ILPTE, ICYB, IMBG, IFBG, KNU, ISMA, IPMMS



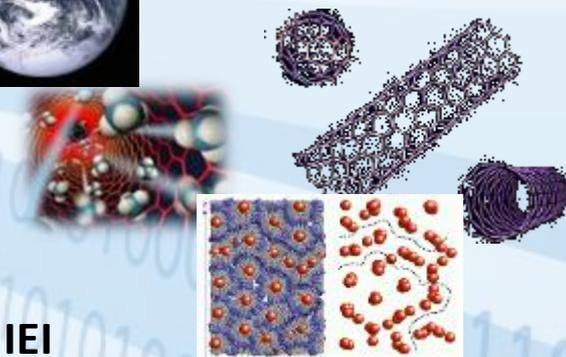
## Геология, гидрофизика, мониторинг окружающей среды

SRI, ICYB, IPMMS, MHI, IGP



## Нанотехнологии, новые материалы, радиоэлектроника

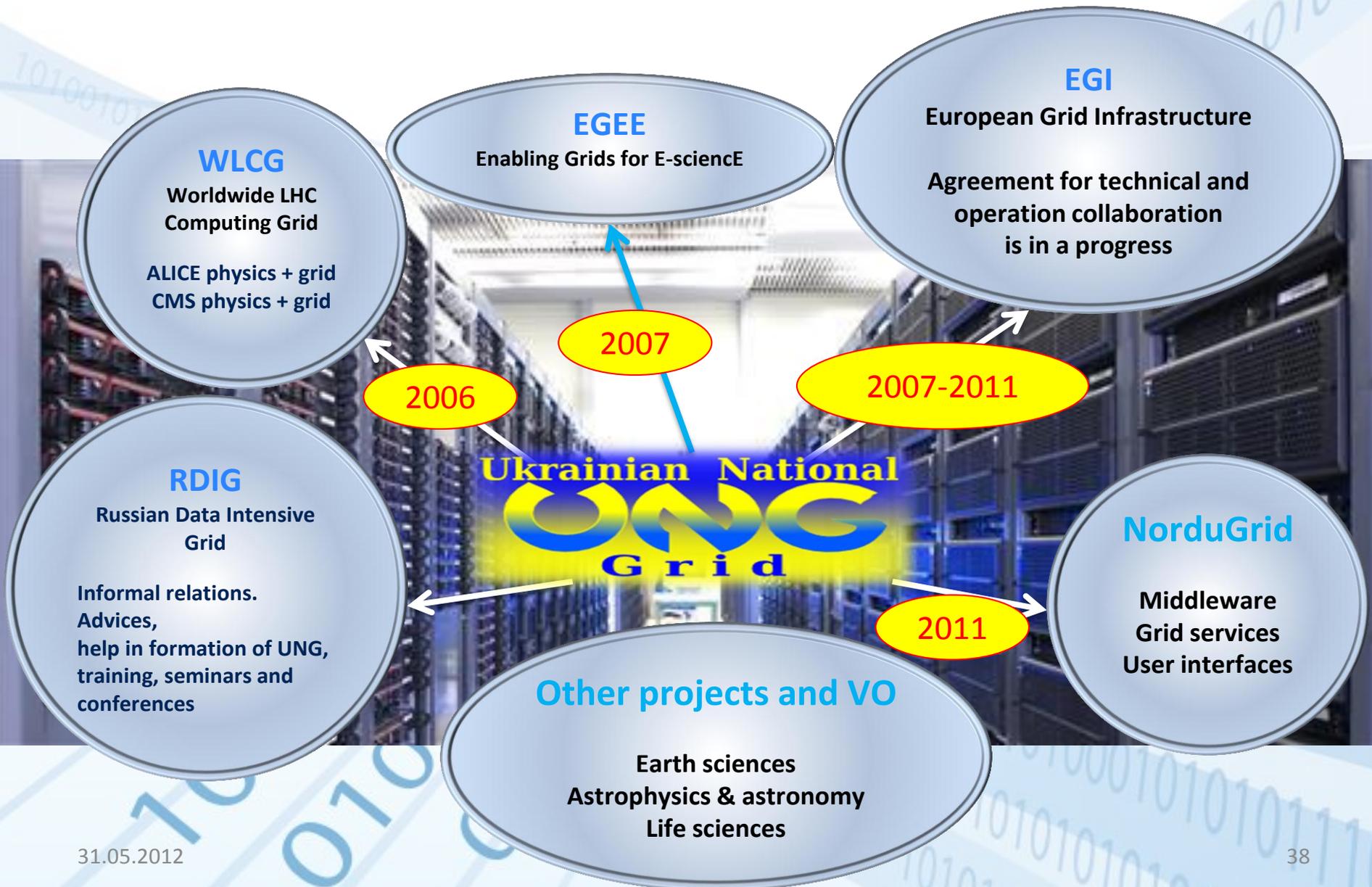
IMP, ILPTE, ICYB, IRE, ICP, DonPTI



## Грид-разработки, экономика

BITP, KNU, KPI, ILPTE, ICYB, ISS, IEI

# Грид в Украине, международное сотрудничество





<http://lcg.web.cern.ch/LCG/>

<http://www.egi.eu>

[http://en.wikipedia.org/wiki/Grid\\_computing](http://en.wikipedia.org/wiki/Grid_computing)

<http://ru.wikipedia.org/wiki/Грид>

<http://www.gridclub.ru/>

<http://grid.nas.gov.ua/>

<http://infrastructure.kiev.ua/ua/>



Спасибо за внимание!