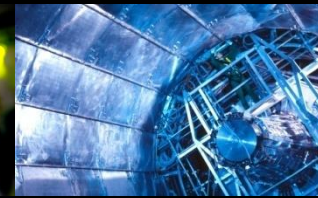
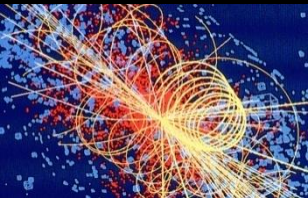


# World LHC Computing Grid

Юлия Андреева  
ЦЕРН

В презентации использованы слайды  
коллег из IT отдела ЦЕРНа

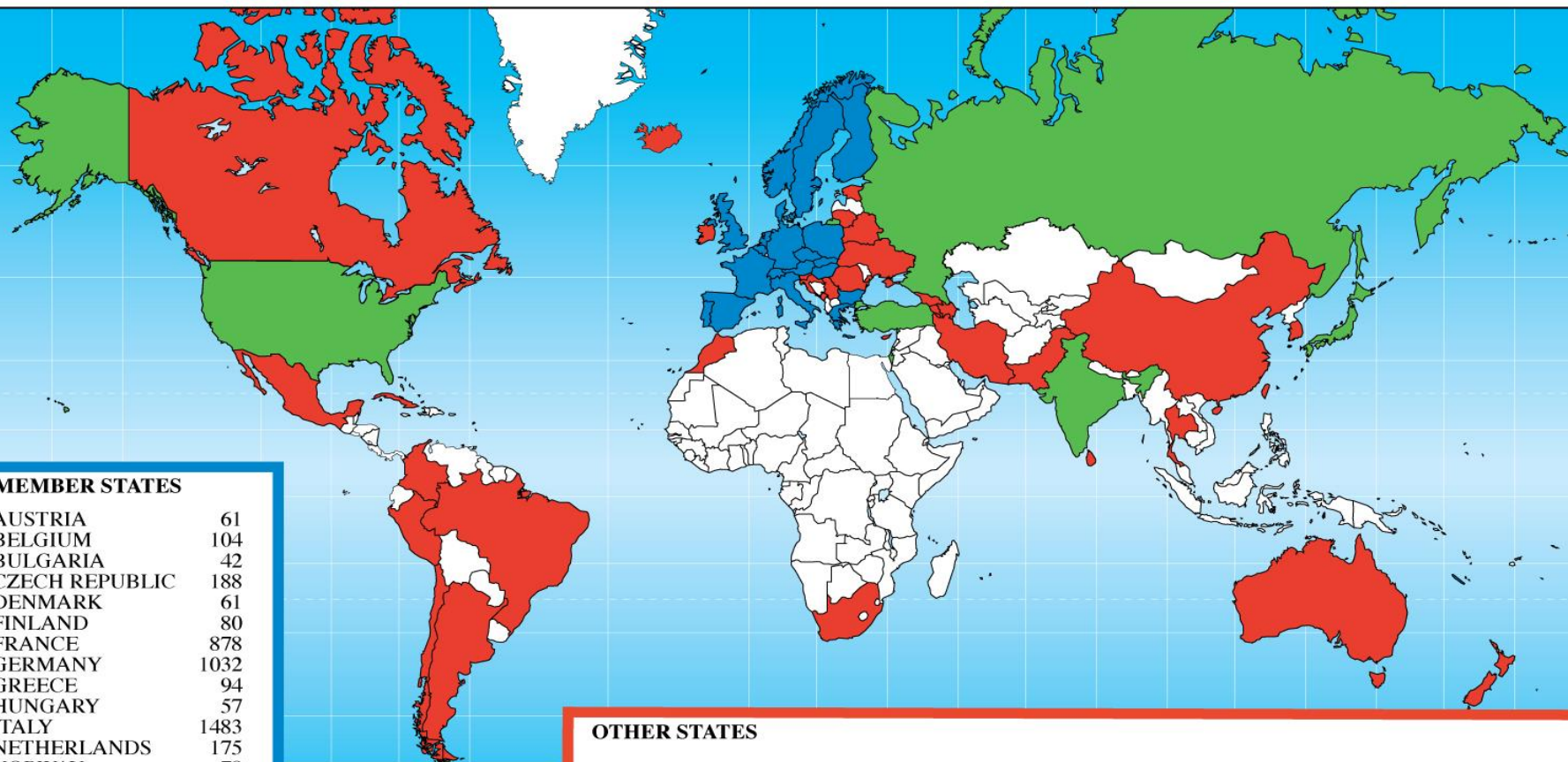


# 20 стран участниц ЦЕРН...



...и огромное сообщество пользователей в разных уголках планеты

## Distribution of All CERN Users by Nation of Institute on 6 January 2009



### MEMBER STATES

AUSTRIA	61
BELGIUM	104
BULGARIA	42
CZECH REPUBLIC	188
DENMARK	61
FINLAND	80
FRANCE	878
GERMANY	1032
GREECE	94
HUNGARY	57
ITALY	1483
NETHERLANDS	175
NORWAY	78
POLAND	174
PORTUGAL	111
SLOVAKIA	49
SPAIN	286
SWEDEN	73
SWITZERLAND	330
UNITED KINGDOM	715

**6071**

### OBSERVER STATES

INDIA	89
ISRAEL	59
JAPAN	200
RUSSIA	883
TURKEY	52
USA	1485

**2768**

### OTHER STATES

ARGENTINA	10	CUBA	3	MONTENEGRO	1	SRI LANKA	1
ARMENIA	15	CYPRUS	6	MOROCCO	5	TAIWAN	42
AUSTRALIA	14	ESTONIA	11	NEW ZEALAND	6	THAILAND	1
AZERBAIJAN	1	GEORGIA	11	PAKISTAN	24	UKRAINE	18
BELARUS	19	ICELAND	1	PERU	1		
BRAZIL	73	IRAN	12	ROMANIA	49		
CANADA	136	IRELAND	12	SERBIA	17		
CHILE	4	KOREA	51	SLOVENIA	16		
CHINA	64	LITHUANIA	5	SOUTH AFRICA	8		
COLOMBIA	11	MEXICO	28				
CROATIA	20						

**696**

# Альтернативы по созданию компьютерной системы

- ЦЕРН не обладает вычислительными мощностями необходимыми для обработки и хранения данных
- Возможные решения:
  - Многкратное увеличение компьютерного центра ЦЕРНе
  - ИЛИ
  - Использование ресурсов институтов участвующих в экспериментах и расположенных в разных уголках земного шара

# Концепция ГРИДа

- Грид является географически распределённой инфраструктурой, объединяющей множество ресурсов разных типов (процессоры, долговременная и оперативная память, хранилища и базы данных, сети), доступ к которым пользователь может получить из любой точки, независимо от места их расположения.
- Идея грид-компьютинга возникла вместе с распространением персональных компьютеров, развитием интернета и технологий пакетной передачи данных на основе оптического волокна а также технологий локальных сетей (Gigabit Ethernet). Полоса пропускания коммуникационных средств стала достаточной, чтобы при необходимости привлечь ресурсы другого компьютера.
- Термин «грид-вычисления» появился в начале 1990-х гг., как метафора о такой же лёгкости доступа к вычислительным ресурсам, как и к электрической сети (англ. power grid) в сборнике под редакцией Яна Фостера и Карла Кессельмана

# Компьютерный ГРИД для ЛНС - ЭТО ...

- Распределенная компьютерная система предоставляющая ресурсы для ЛНС вычислений
- Разработанная, оперируемая и управляемая мировой научной коллаборацией между экспериментами и компьютерными центрами
- Главная идея заключается в использовании компьютерных ресурсов независимо от места их расположения

# Компьютерный ГРИД для ЛНС - почему?

Мы имеем дело распределенными ресурсами (компьютерные мощности, кадры, финансирование)

Участники ЛНС заинтересованы вкладывать деньги в ресурсы в своей стране одновременно решая общую задачу

С технической точки зрения распределенная система более надежна.

# Компьютерный ГРИД для LHC это система созданная для пользователей

Несколько слоев  
матобеспечения решают задачу  
абстагирования пользователя от  
сложности имплементации  
компьютерной системы  
Пользователь видит ГРИД как  
единый вычислительный ресурс





# Сложность имплементации диктуется высокими требованиями

- Объем данных
  - Высокая скорость набора данных помножить на большой объем данных помножить на 4 эксперимента
  - 20 Petabytes новых данных в год
  - Все данные должны архивироваться и храниться десятилетиями
- Вычислительные мощности
  - Сложные события × большое количество событий × тысячи пользователей: около 350,000 компьютерных ядер
- Распределенные ресурсы и финансирование
- Неоднородное матобеспечение
- Федерация ГРИД систем (EGEE, OSG, NorduGrid)
- Большое количество пользователей
  - 10 000 ученых
  - 500 институтов



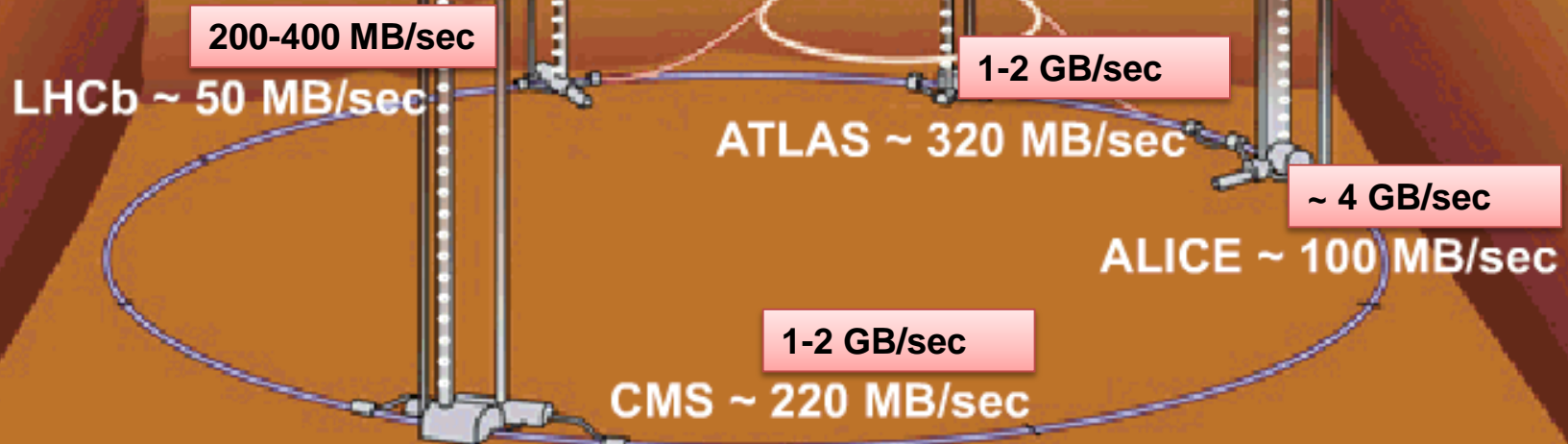
- WLCG инфраструктура работает с 2005 года

# Решаемые задачи

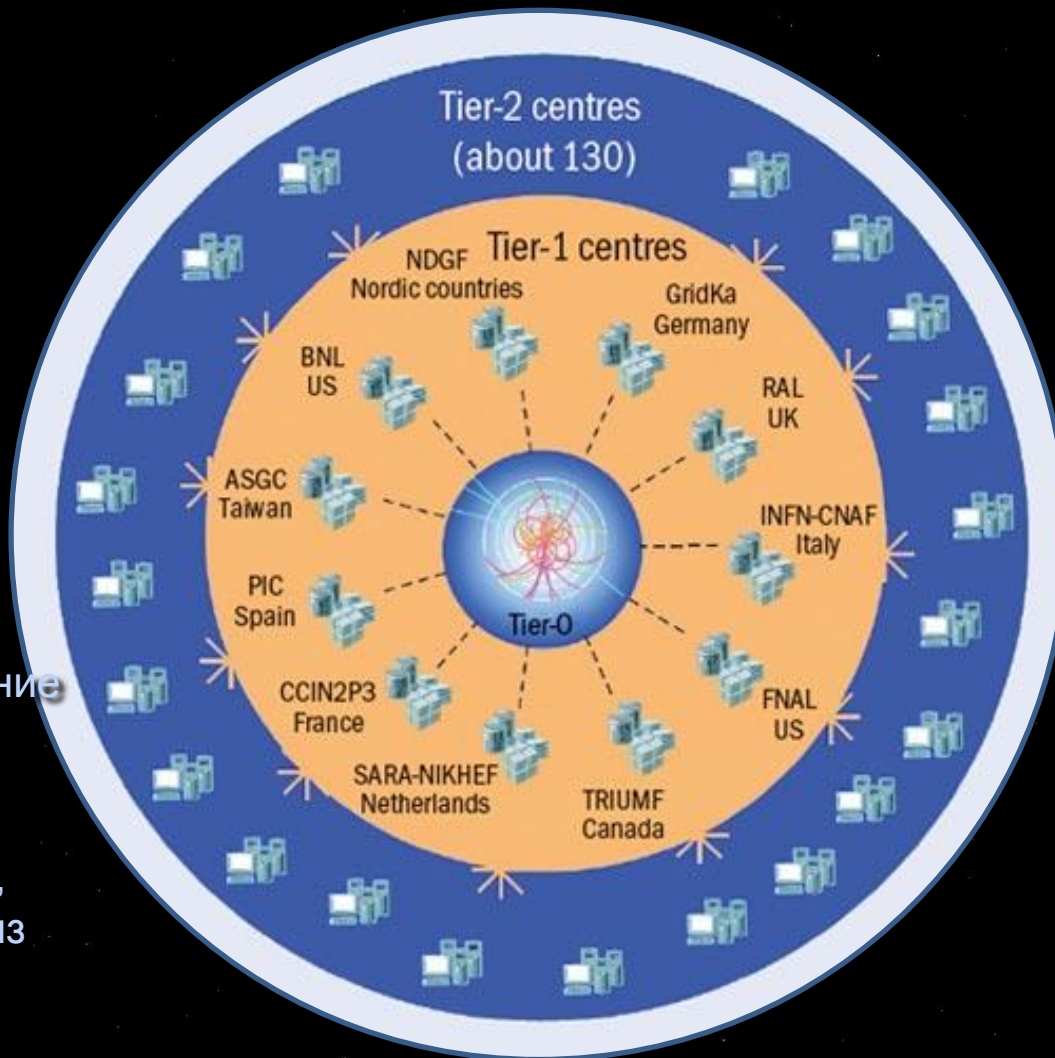
- Передача данных
  - Туда, где они должны храниться и обрабатываться
- “Production” вычисления (реконструкция, симуляция)
- “Analysis” вычисления

Передача данных к пункту хранения и обработки: 4-6 GB/sec

CERN Computer Centre



# The Worldwide LHC Computing Grid



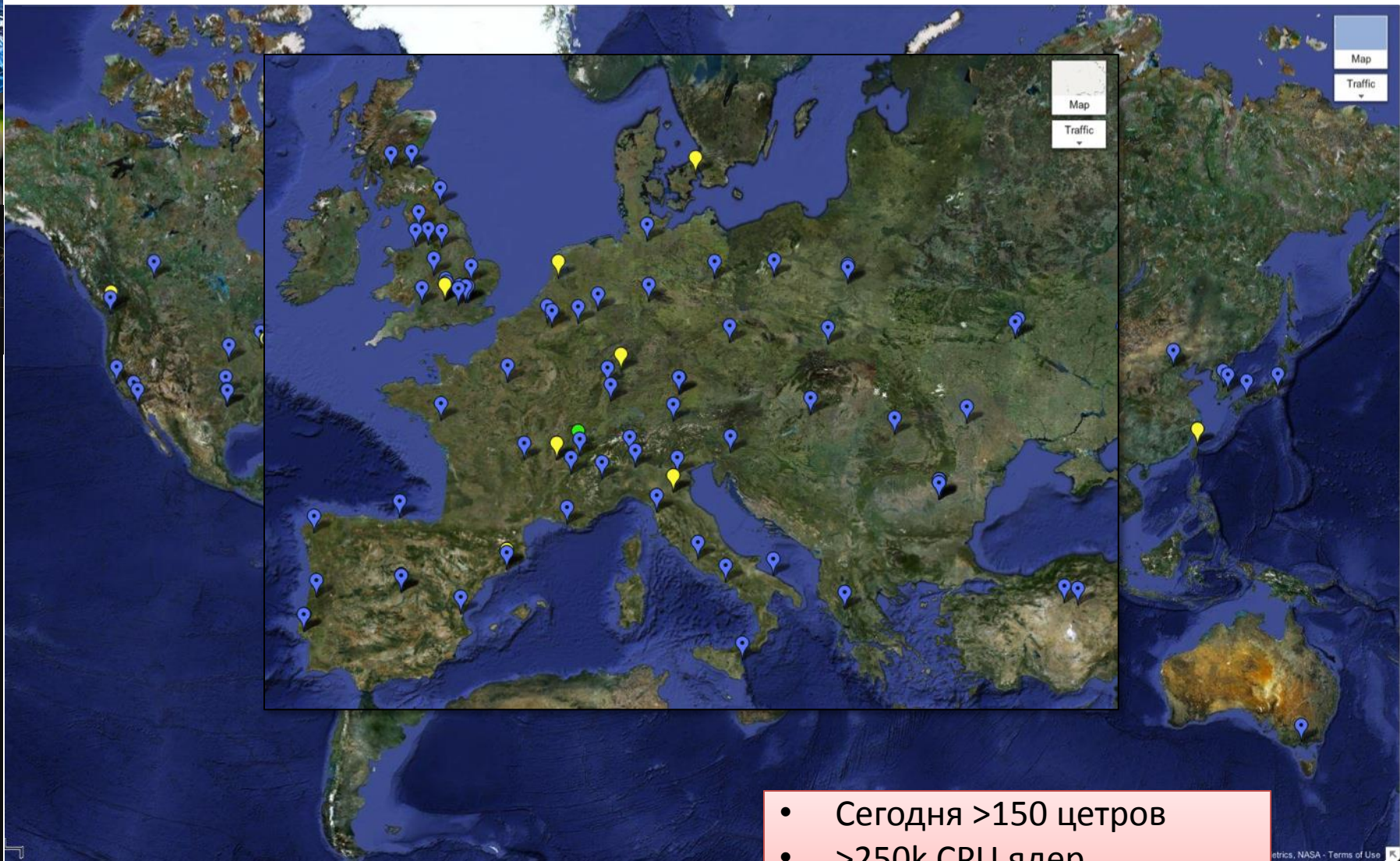
Tier-0 (CERN):  
запись и  
архивирование  
данных,  
реконструкция и  
передача данных

Tier-1: архивирование  
данных, повторная  
обработка, анализ

Tier-2: Симуляция,  
физический анализ

Около 160  
центров,  
35 стран  
~350'000  
вычислительных  
ядер  
200 PB ресурсы  
хранения  
данных  
> 2 миллионов  
задач в день  
10 Gb  
ЛИНИИ СВЯЗИ

# WLCG ГРИД центры



- Сегодня >150 центров
- >250к CPU ядер
- >150 PB дискового пространства



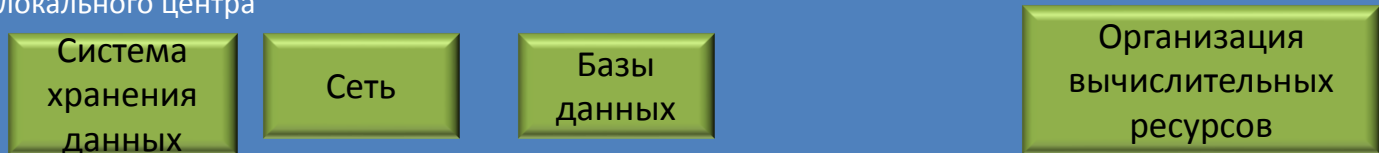
# Матобеспечение

Матобеспечение экспериментов

Матобеспечение ГРИД

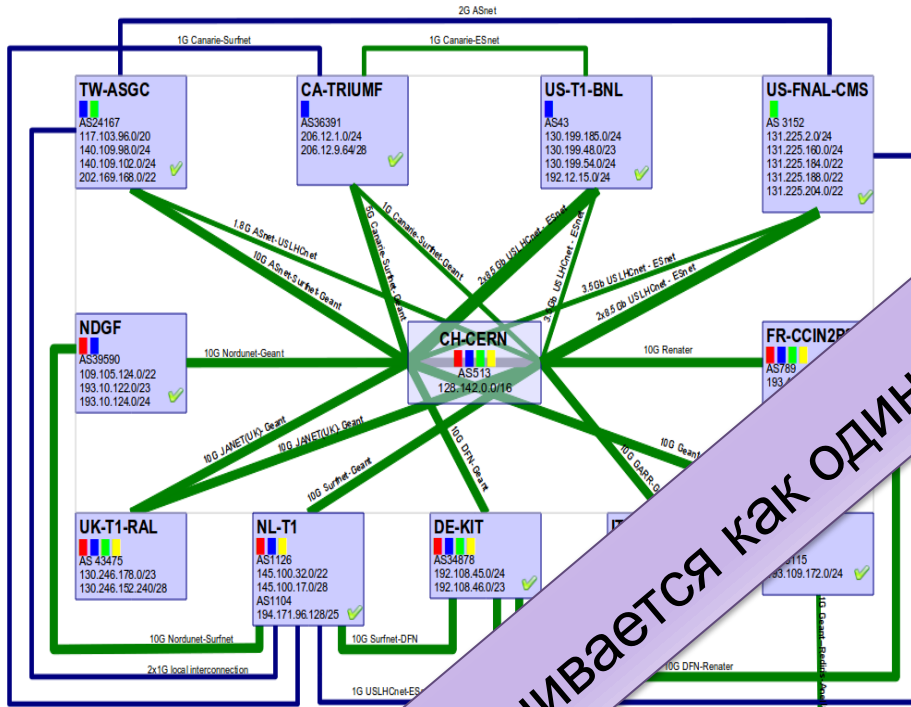


Уровень локального центра



# LHC Networking

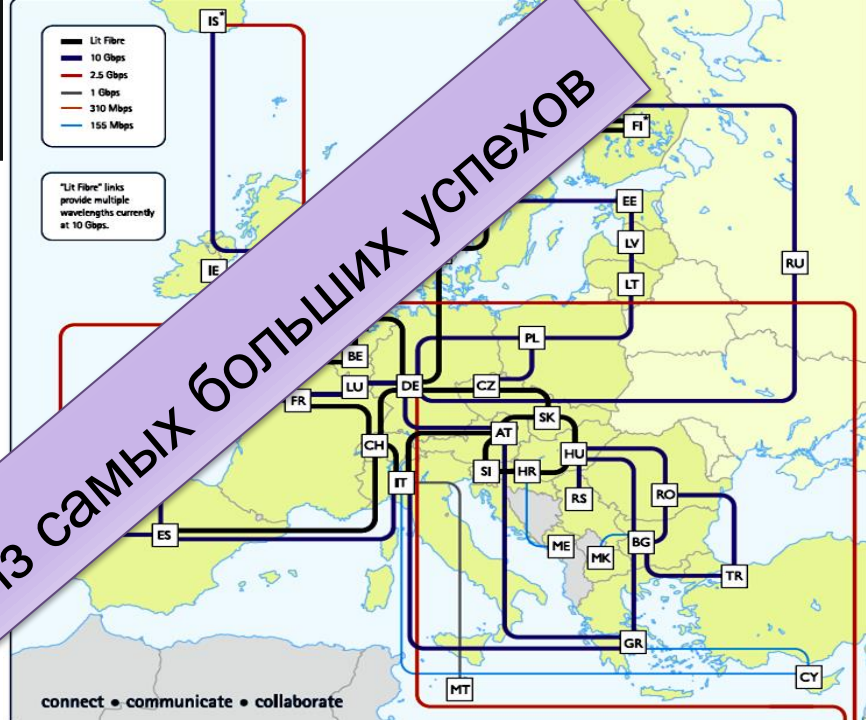
## LHCOPN



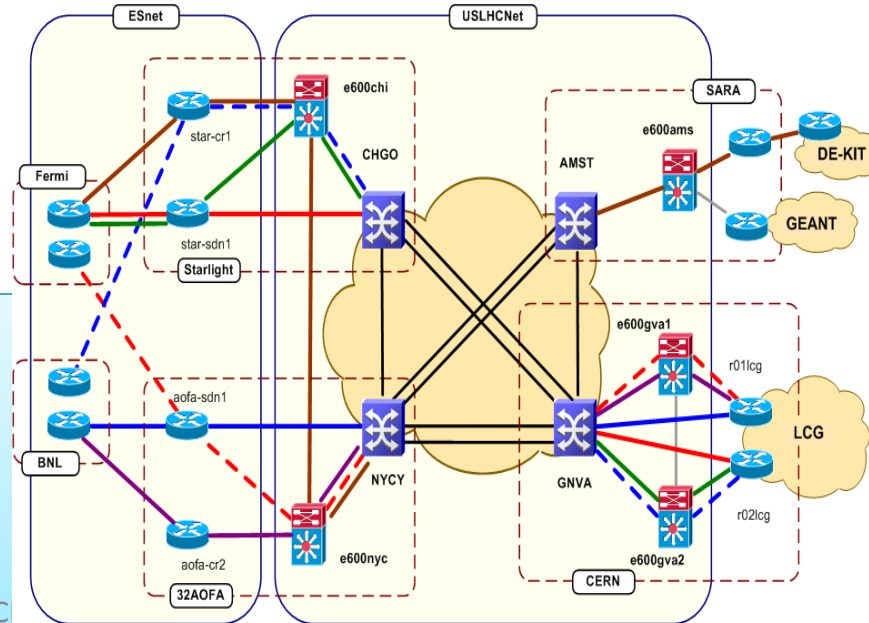
- 10-T1 and T1-T1 traffic
- T1-T1 traffic only
- Not deployed yet
- (thick) >10Gbps
- (thin) <10Gbps
- = Alice = Atlas
- = CMS = LHCb
- ✓ = internet backup available
- pdp prefix: 192.16.166.0/24
- esbarco.marcell@cern.ch 20 100%

• **Re...**  
**N, GEANT, US-LHCNet**  
**NRENs & other national**  
**& international providers**

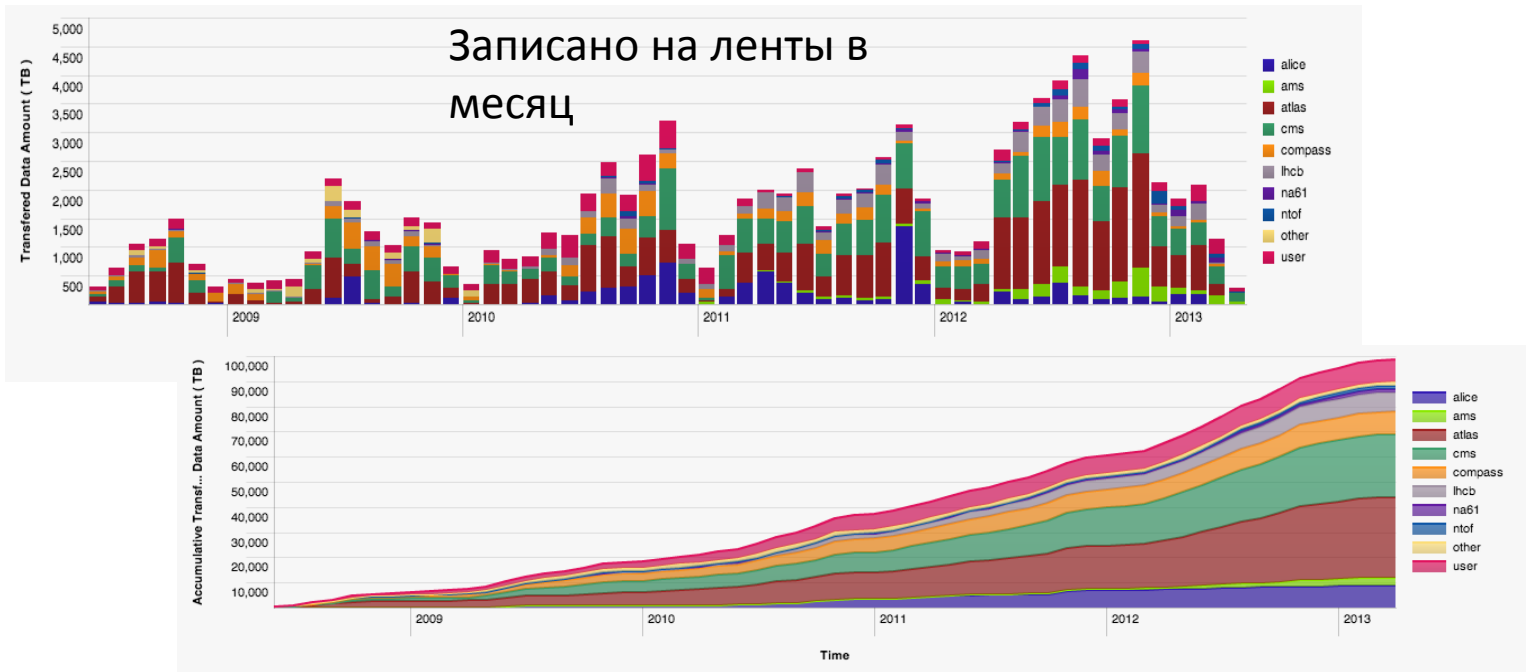
Ian Bird, C



Planned Backbone Topology by the end of 2010. GÉANT is operated by DANTE on behalf of Europe's NRENs.



# Сбор и архивирование данных

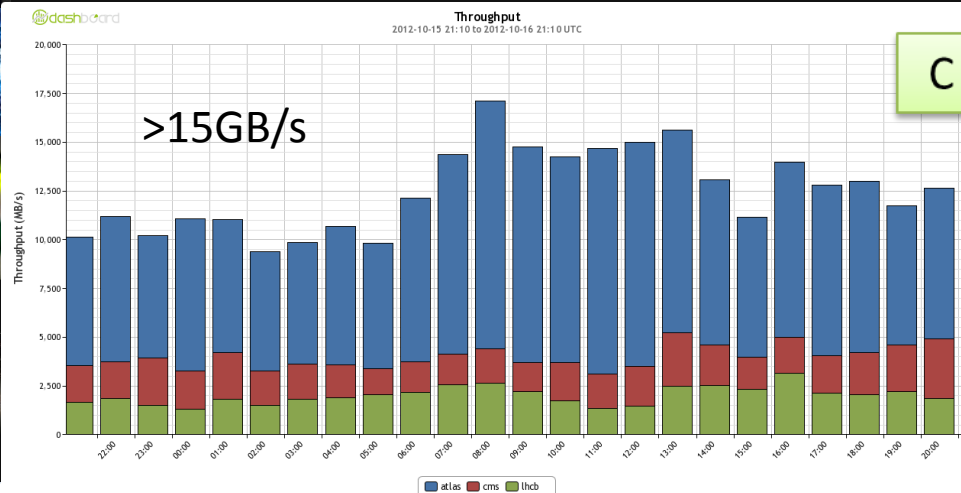


- 27 PB заархивировано на лентах в 2012+13 (p-p and p-Pb)
- Всего заархивировано 100 PB из которых ~70 PB is LHC data
- До 4.6 PB/месяц

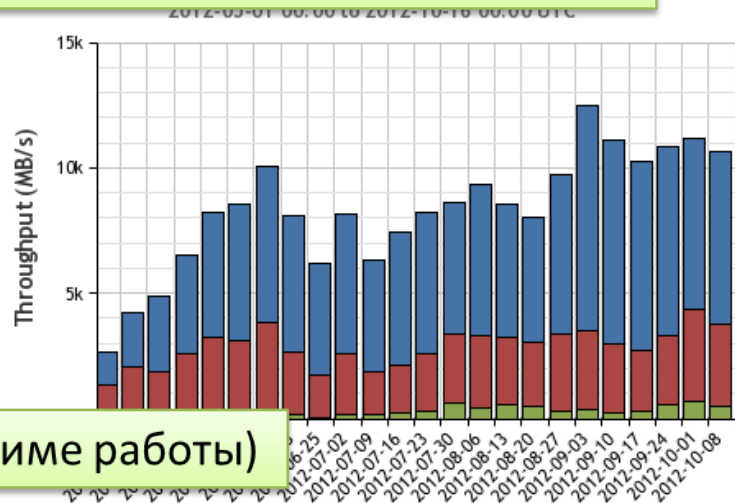




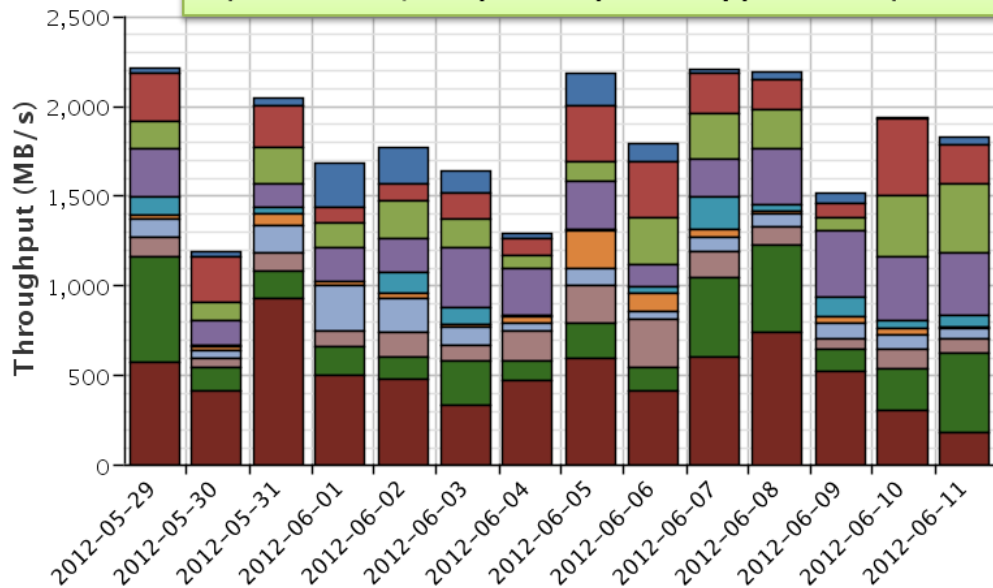
# Передача данных



С мая 2012 до начала октября 2012



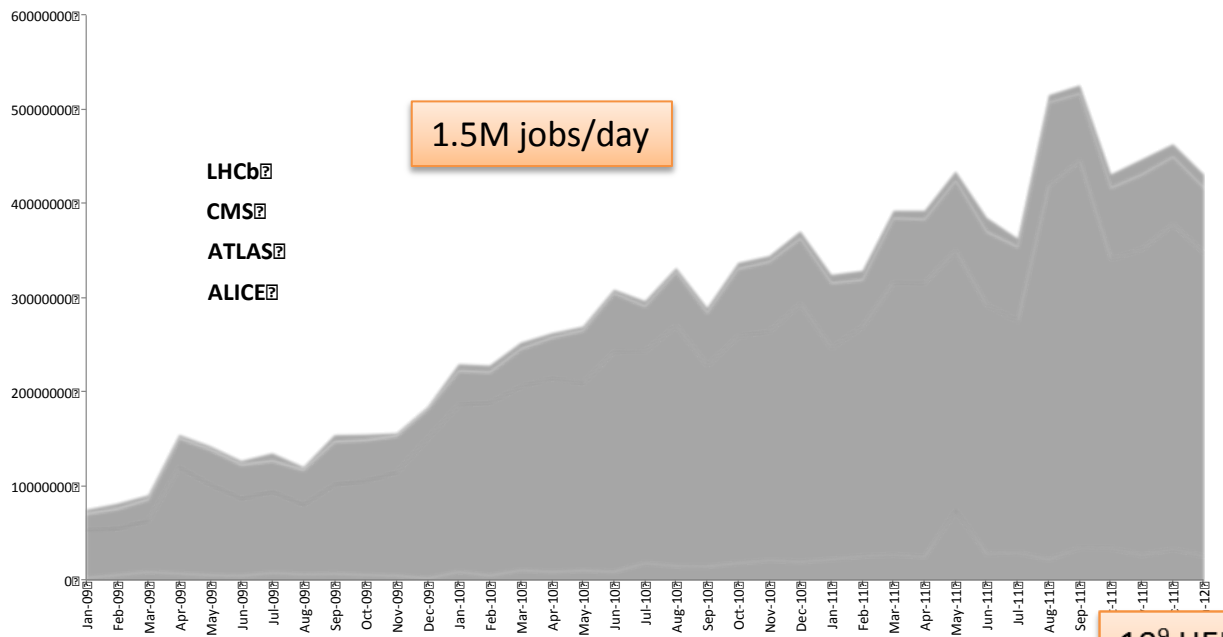
ЦЕРН → центры первого уровня ( LHC в режиме работы)



Планируемая  
скорость:

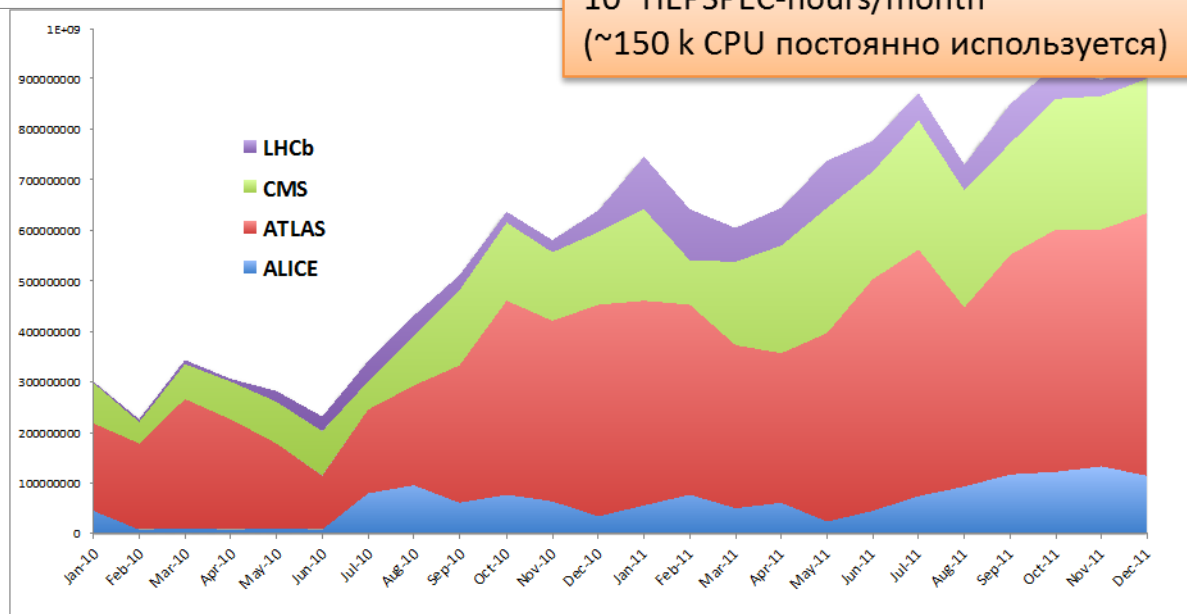
Tier 0 – Tier 1: 650 MB/s

# Использование вычислительных мощностей

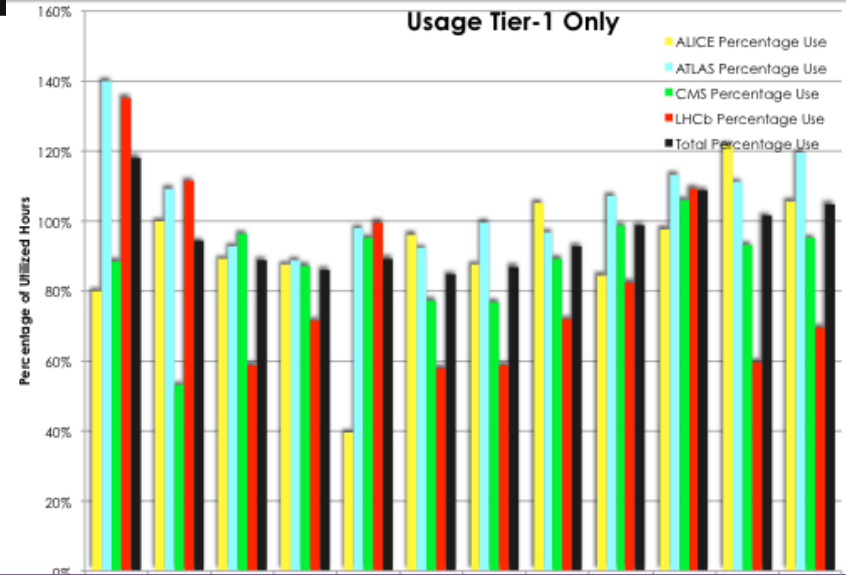
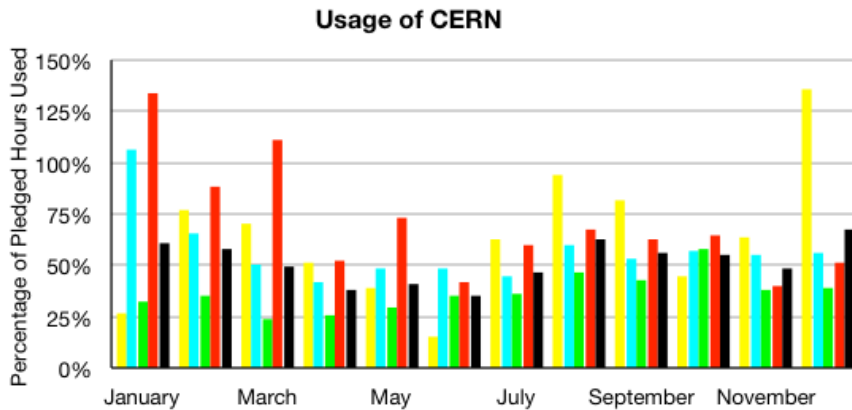


Продолжает расти...

- # задач в день
- CPU

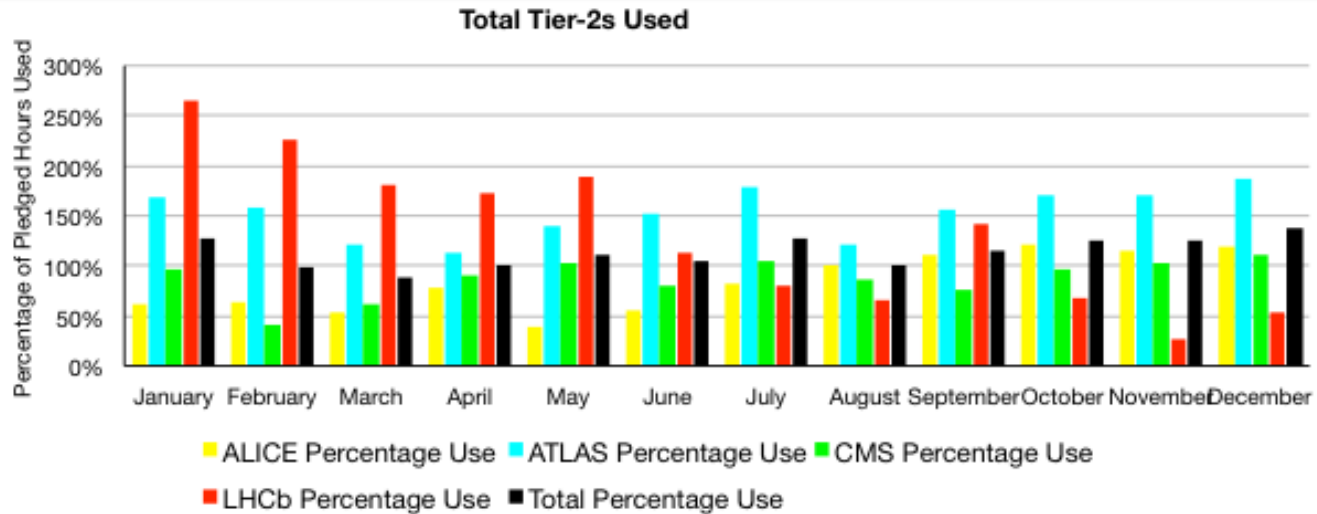


# Использование ресурсов в сравнении с квотами



■ ALICE Percentage Use   
 ■ ATLAS Percentage Use   
 ■ CMS Percentage Use  
■ LHCb Percentage Use   
 ■ Total Percentage Use

**Мы используем все предоставленные ресурсы!**

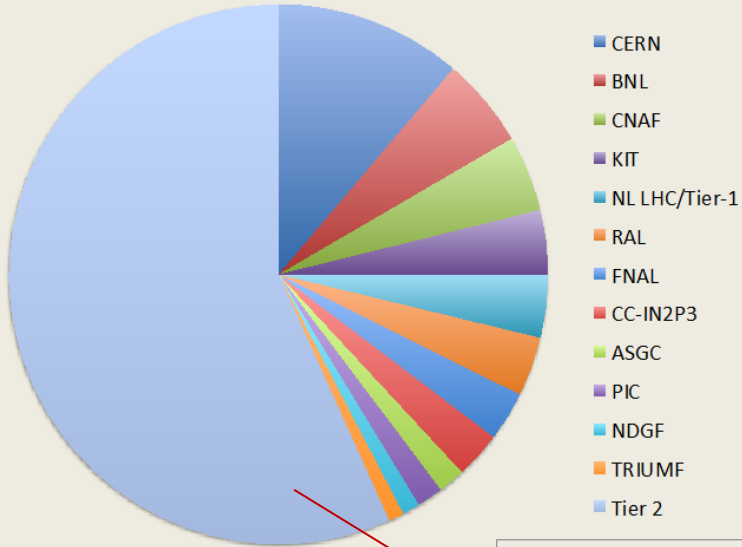


■ ALICE Percentage Use   
 ■ ATLAS Percentage Use   
 ■ CMS Percentage Use  
■ LHCb Percentage Use   
 ■ Total Percentage Use

# CPU – распределение по центрам

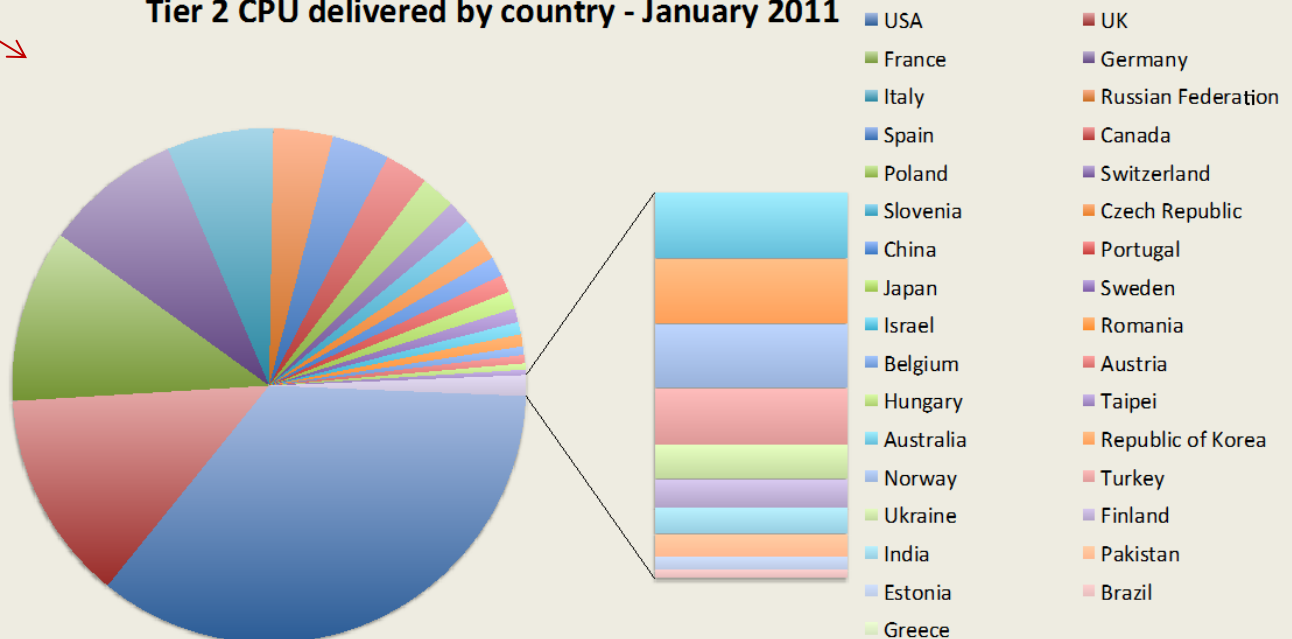
## РАЗНОГО УРОВНЯ

CPU delivered - January 2011



- ГРИД действительно работает
  - Все центры большие и маленькие
- Вносят свой вклад
- И этот вклад востребован!**

Tier 2 CPU delivered by country - January 2011



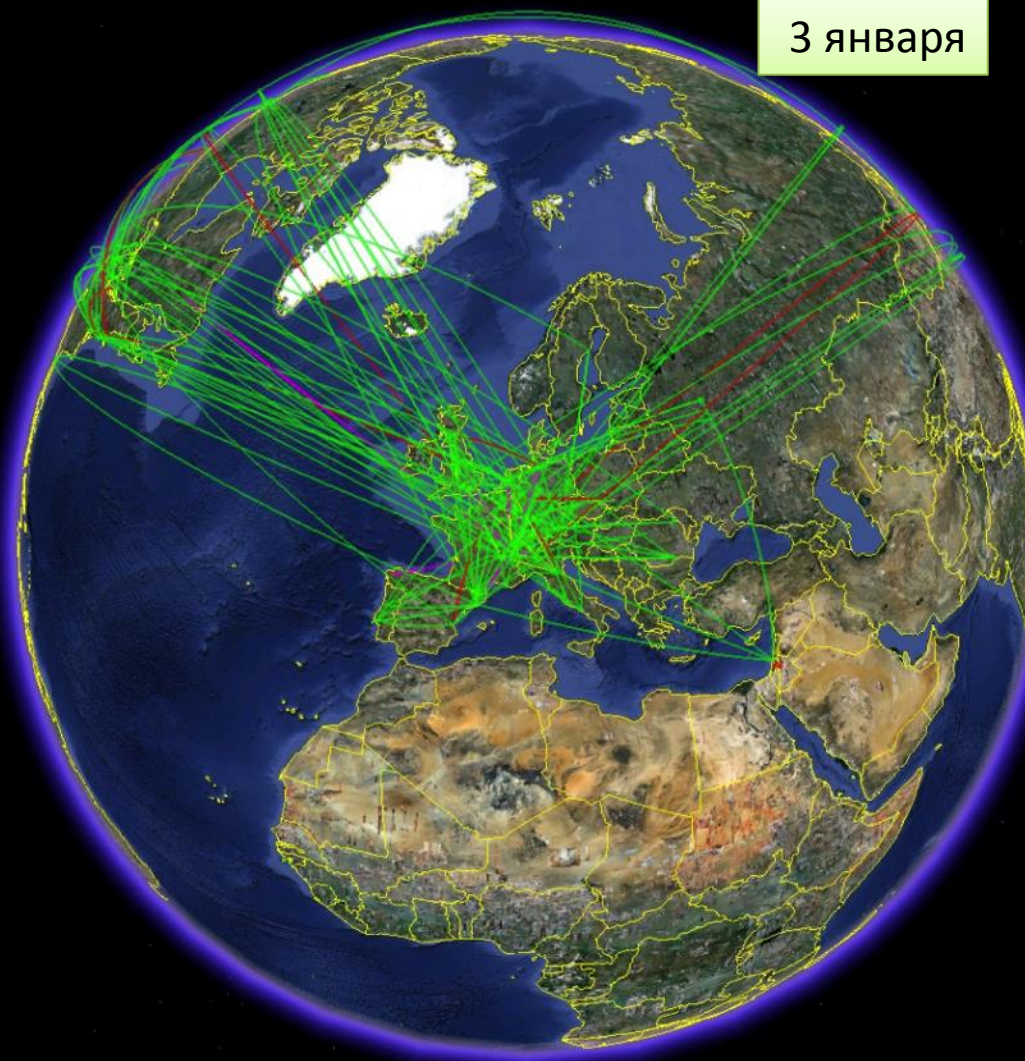
# Вклад Российских центров

- 9 Российских и один Украинский центр предоставляют свои ресурсы для LHC . Эти центры являются частью Российского ГРИДА (Russian Data Intensive GRID)
- Они расположены в Москве, Помосковье, Петербурге, Ленинградской области и на Украине
- Являются центрами второго уровня
- На базе Дубны и Курчатовского института создается центр первого уровня

# WLCG – работа без выходных

Running jobs: 151768.0  
Transfer rate: 4.72 GiB/sec

3 января



© 2011 Europa Technologies  
US Dept of State Geographer  
© 2011 Google  
© 2011 MapLink/Tele Atlas

Google™ earth

Eye alt 18391.55 km

WLCG  
Worldwide LHC Computing Grid



# Как добиться надежной и эффективной работы инфраструктуры

- Добиться надежной и эффективной работы такой большой, сложной и неоднородной инфраструктуры очень трудно
- Компьютерные смены
- Постоянное общение между экспериментами и вычислительными центрами (ежедневные получасовые рабочие собрания)
- Тестирование распределенных центров и сервисов
- Мониторирование

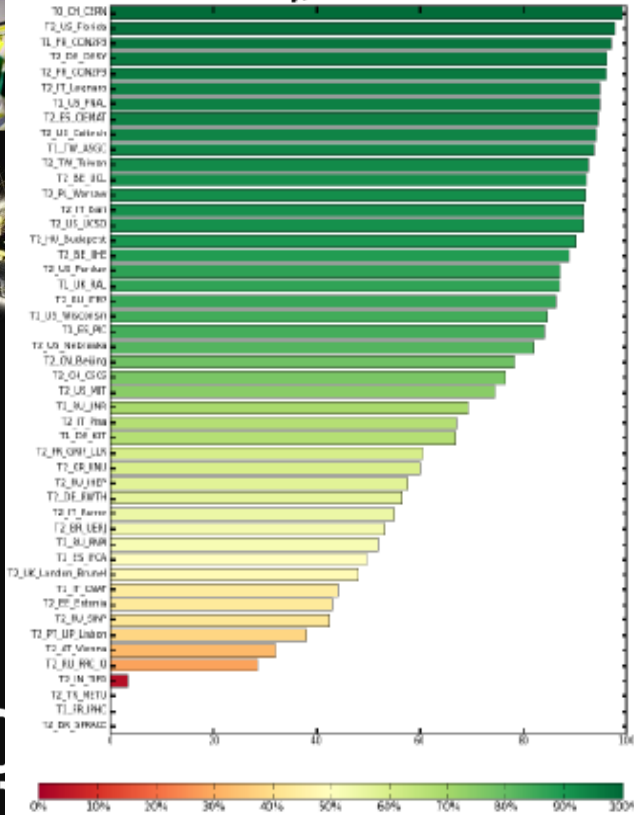
# Тестирование

- Разработана система удаленного тестирования распределенных сервисов и центров
- Тесты проверяющие определенную функциональность регулярно запускаются как ГРИД задачи на все центры инфраструктуры
- Результаты выполнения этих тестов сохраняются в центральной базе данных
- На основании результатов этих тестов вычисляется эффективность работы центра
- Центры не справляющиеся с критическими тестами временно исключаются из информационной системы и становятся невидимыми для системы управления задач

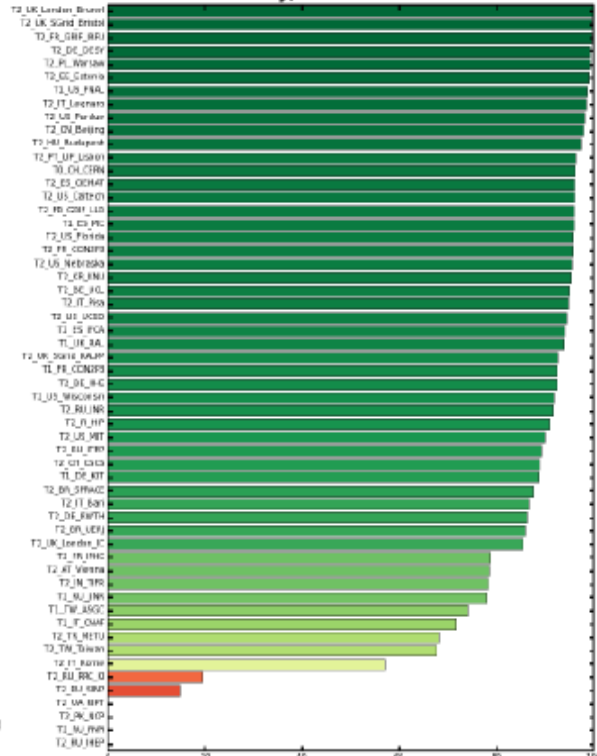


# Видимое улучшение качества работы инфраструктуры

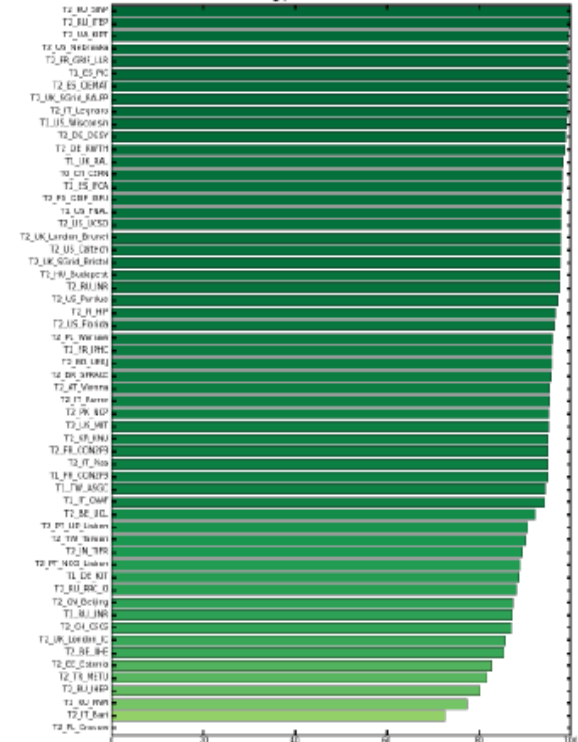
Site Availability, 2008-04-01 - 2008-04-30



Site Availability, 2009-04-01 - 2009-04-30



Site Availability, 2010-04-01 - 2010-04-30



# Мониторирование

- Разработано большое количество систем позволяющих мониторировать состояние и эффективность работы отдельных сервисов, сайтов, а так же эффективность передачи данных и ведения вычислений
- Роль систем мониторинга очень важна, т.к. они позволяют обнаружить проблемы, идентифицировать их причины и принять меры по их устранению

# Эволюция центра 0 уровня

- Начиная с 2005 г стали очевидны проблемы связанные с достижением максимальной нагрузки компьютерного центра, что делало невозможным дальнейшее наращивание **вычислительных мощностей**
  - Ограничение на уровне 2.5 MW
  - В то время как потребность 2020 году оценивается ~10 MW
- **Кроме того наблюдались другие ограничения**
  - Перегрузка систем UPS
  - Существенный недостаток систем электроснабжения в критических ситуациях (дизельные устройства)
  - Отсутствие резервного центра в случае глобальной проблемы с компьютерным центром ЦЕРНа

# Новый центр 0 уровня в Будапеште



**После проведения тендера, новый центр было решено строить в Будапеште**

**Первый этап работ был завершен в конце 2012 года**

**1100m<sup>2</sup> (725m<sup>2</sup>) в существующем здании но с новой инфраструктурой  
Полное покрытие UPS и дизельными устройствами б максимум 2.7MW**

**Новый центр должен быть полностью введен в действие в первом квартале 2014 года**

**Оба центра должны работать как единое целое**





# Виртуальные машины

Виртуальная машина эмулирует работу реального компьютера или отдельных компонентов аппаратного обеспечения (включая BIOS, оперативную память, жёсткий диск и другие периферийные устройства) или некоторой вычислительной среды

Например, Windows можно запускать в виртуальной машине под Linux или наоборот

На одном компьютере может функционировать несколько виртуальных машин => можно имитировать несколько серверов на одном компьютере

Появление виртуальных машин позволяет существенно облегчить поддержку вычислительных кластеров и установку/переустановку на них требуемого окружения, сервисов, т.д.

Появление виртуальных машин дало толчок Cloud computing

# Cloud computing

- Cloud computing (Облачные вычисления) предлагает путь как можно по мере необходимости, возможно временно увеличить вычислительные мощности, не закупая оборудования, не расширяя помещения, не нанимая и обучая персонал т.д.
- **Облачные вычисления представляют собой динамически масштабируемый способ доступа к внешним вычислительным ресурсам в виде сервиса, предоставляемого посредством Интернета**  
При этом пользователю не требуется никаких особых знаний об инфраструктуре "облака" или навыков управления этой "облачной" технологией.
- На сегодняшний день "китами" индустрии облачных вычислений считаются компании Google, Amazon, Microsoft, IBM, Sun, Ubuntu и другие.



# GRID или CLOUD?

- Скорее всего будущее за комбинированным решением: ГРИД состоящий из облаков, где отдельный компьютерный центр входящий в ГРИД структуру может быть технически реализован в виде облака
- В настоящее время ведутся работы по созданию не коммерческих облаков на базе компьютерных ресурсов ЦЕРНа и других центров и адаптации заданий симуляции, реконструкции и физического анализа на коммерческих и не коммерческих облаках

Global Effort → Global Success

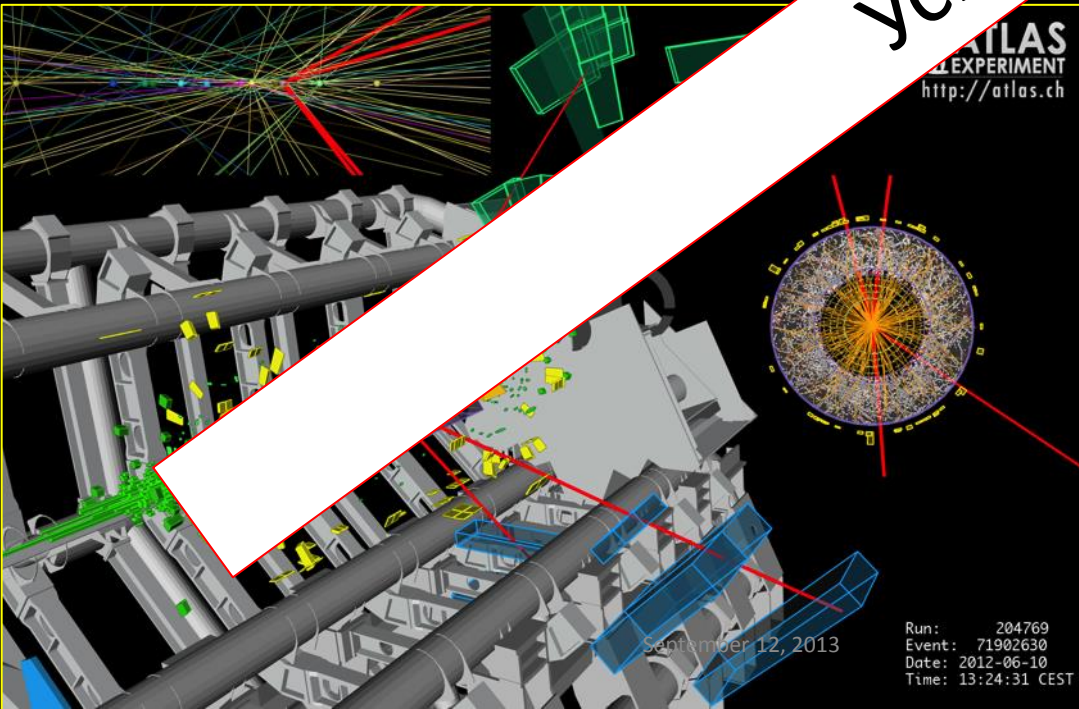
Results today only possible due to  
extraordinary performance of  
accelerators – experiments – Grid computing

Observation of a new particle consistent with  
a Higgs Boson (but which one...?)

Historic Milestone but only the beginning

Global Implications for the future

Успех!



## Higgs boson-like particle discovery claimed at LHC

COMMENTS (1665)

By Paul Rincon

Science editor, BBC News website, Geneva



The moment when Cern director Rolf Heuer confirmed the Higgs results

Cern scientists reporting from the Large Hadron Collider (LHC) have claimed the discovery of a new particle consistent with the Higgs boson.

Relat

Q&A:

# Заключение

- ГРИД стал реальностью
- WLCG инфраструктура – в настоящий момент самая большая существующая ГРИД инфраструктура. Она была создана для вычислений и хранения данных экспериментов БАК (Большого Адронного коллайдера). Эта инфраструктура объединяет более 150 компьютерных центров в разных уголках планеты и используется учеными 500 институтов
- Опыт использования инфраструктуры после запуска ускорителя доказал правильность технического решения имплементации компьютерной системы БАК в виде распределенной ГРИД системы
- В настоящее время рассматриваются новые технические решения, как например интеграция ГРИДа и облачных вычислений