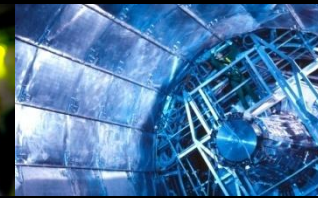


# World LHC Computing Grid

Юлия Андреева  
ЦЕРН

В презентации использованы слайды  
коллег из IT отдела ЦЕРНа

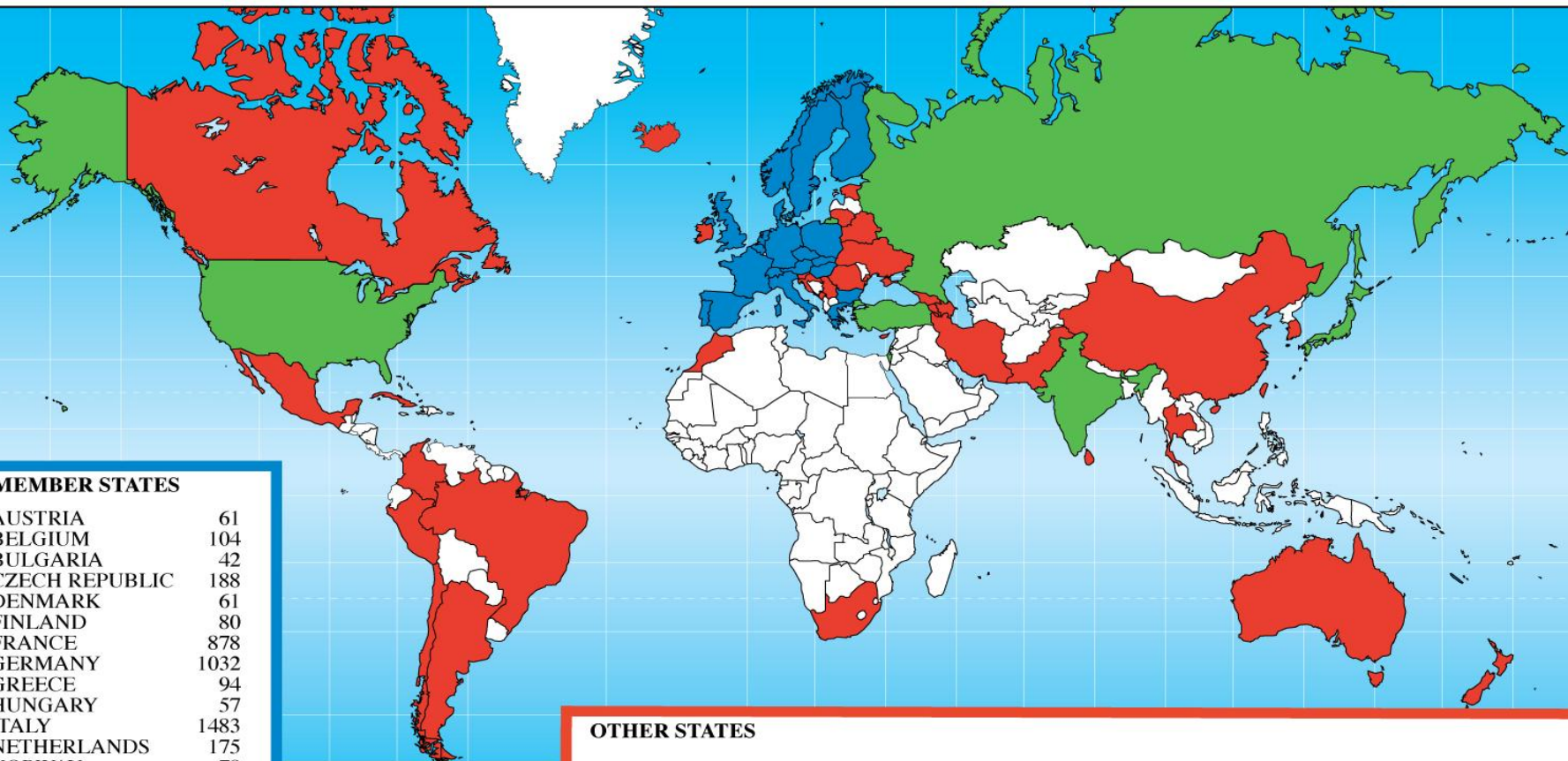


# 20 стран участниц ЦЕРН...



...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						

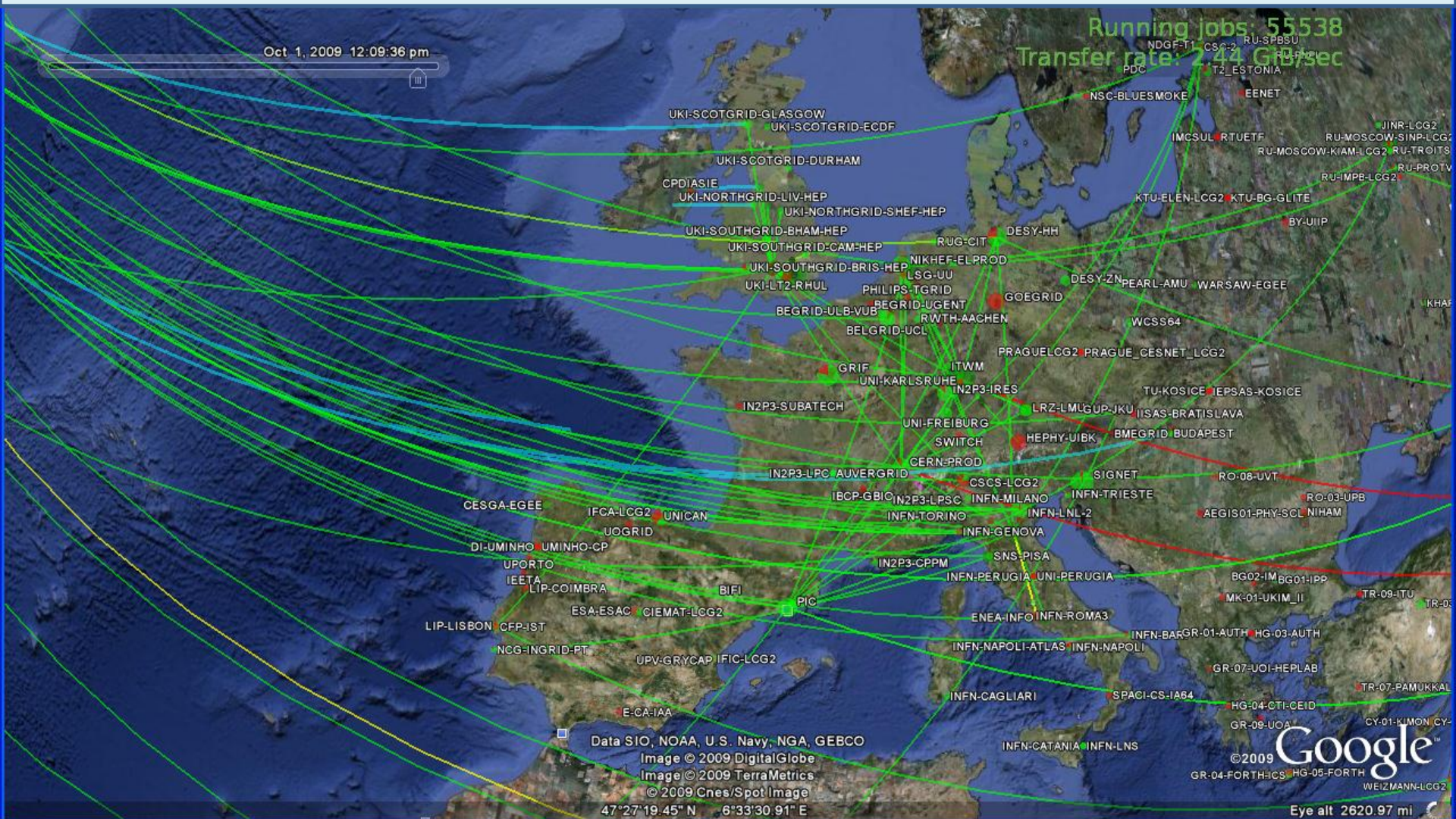
**696**



# 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

# The Worldwide LHC Computing Grid (WLCG)





# 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

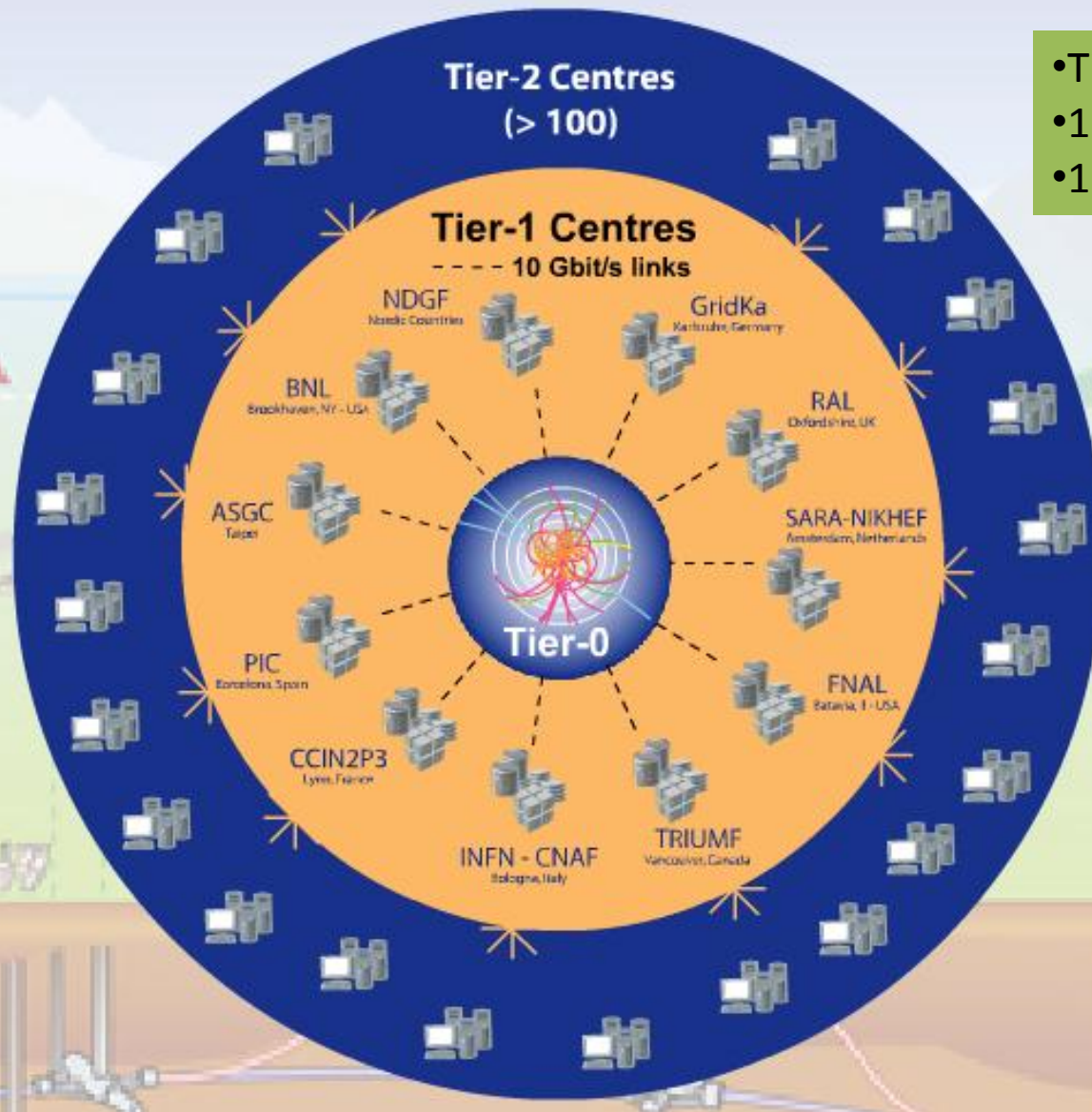
- Объем данных
  - Высокая скорость набора данных помножить на большой объем данных помножить на 4 эксперимента
  - 20 Petabytes новых данных в год
  - Все данные должны архивироваться на лентах
- Вычислительные мощности
  - Сложные события × большое количество событий × тысячи пользователей: ~ 100,000 процессоров
- Распределенные ресурсы и финансирование
- Неоднородное матобеспечение
- Федерация ГРИД систем (EGEE, OSG, NorduGrid)
- Большое количество пользователей
  - 5000 ученых
  - 500 институтов



- WLCG инфраструктура работает с 2005 года

# WLCG структура

- Tier-0 (CERN)
- 11 Tier-1
- 130 Tier-2





# WLCG сегодня и завтра

## • WLCG 2010

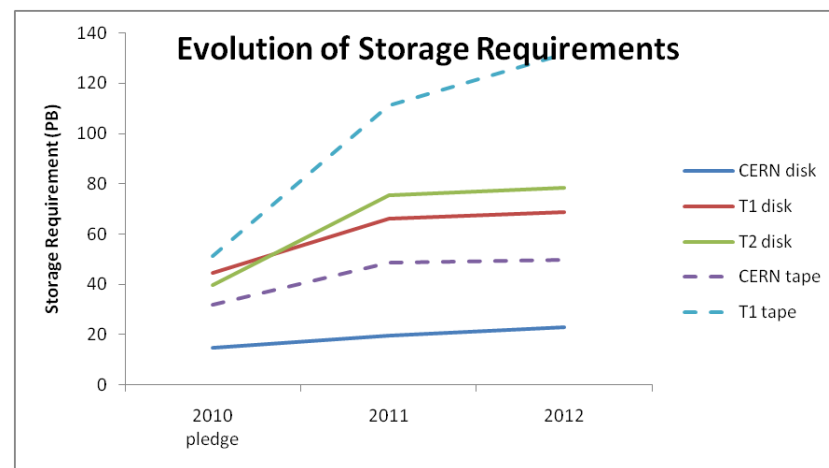
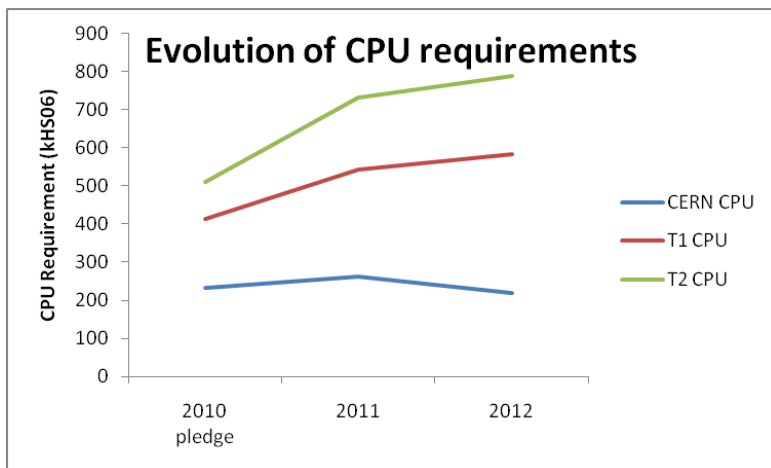
- CPU ~ 100,000 cores
- Диски ~ 100 PB
- Ленты ~ 100 PB
- Скорость передачи данных из ЦЕРНа ~ 5 GB/s

## • WLCG завтра (2012)

- + 30% of CPU
- + 50% дискового пространства
- + 200% пространства на лентах

Totals	2010	2011	2012
CERN CPU	233	263	220
CERN disk	15	20	23
CERN tape	32	49	50
T1 CPU	394	543	584
T1 disk	49	66	69
T1 tape	56	111	132
T2 CPU	563	730	787
T2 disk	47	75	78

CPU in kHEP-SPEC06  
Disk and tape in PB



# Задачи решаемые на WLCG

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

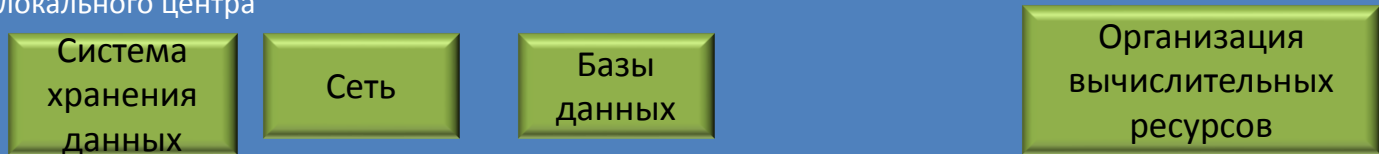
# Архитектура WLCG

Матобеспечение экспериментов

Матобеспечение ГРИД



Уровень локального центра





# Обработка пользовательской задачи

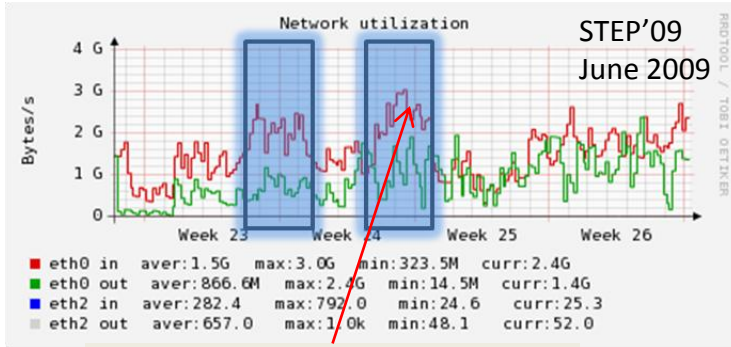


Пользователю нужно запустить задание, которое использует входные данные X и версию матобеспечения Y



# Передача данных

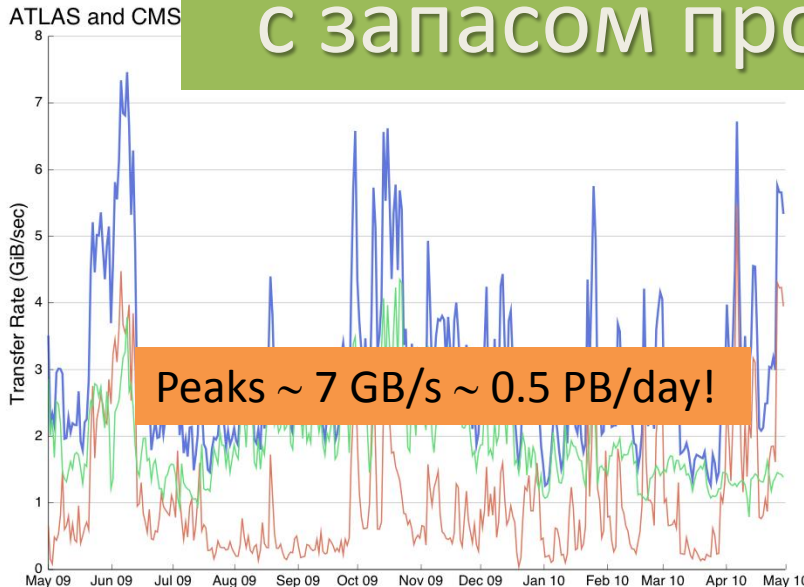
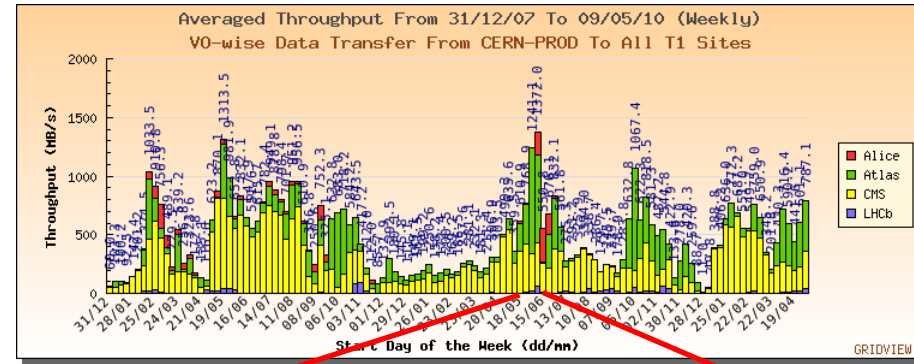
## Tier-0: writing to tape



> 3 GB/s not a problem

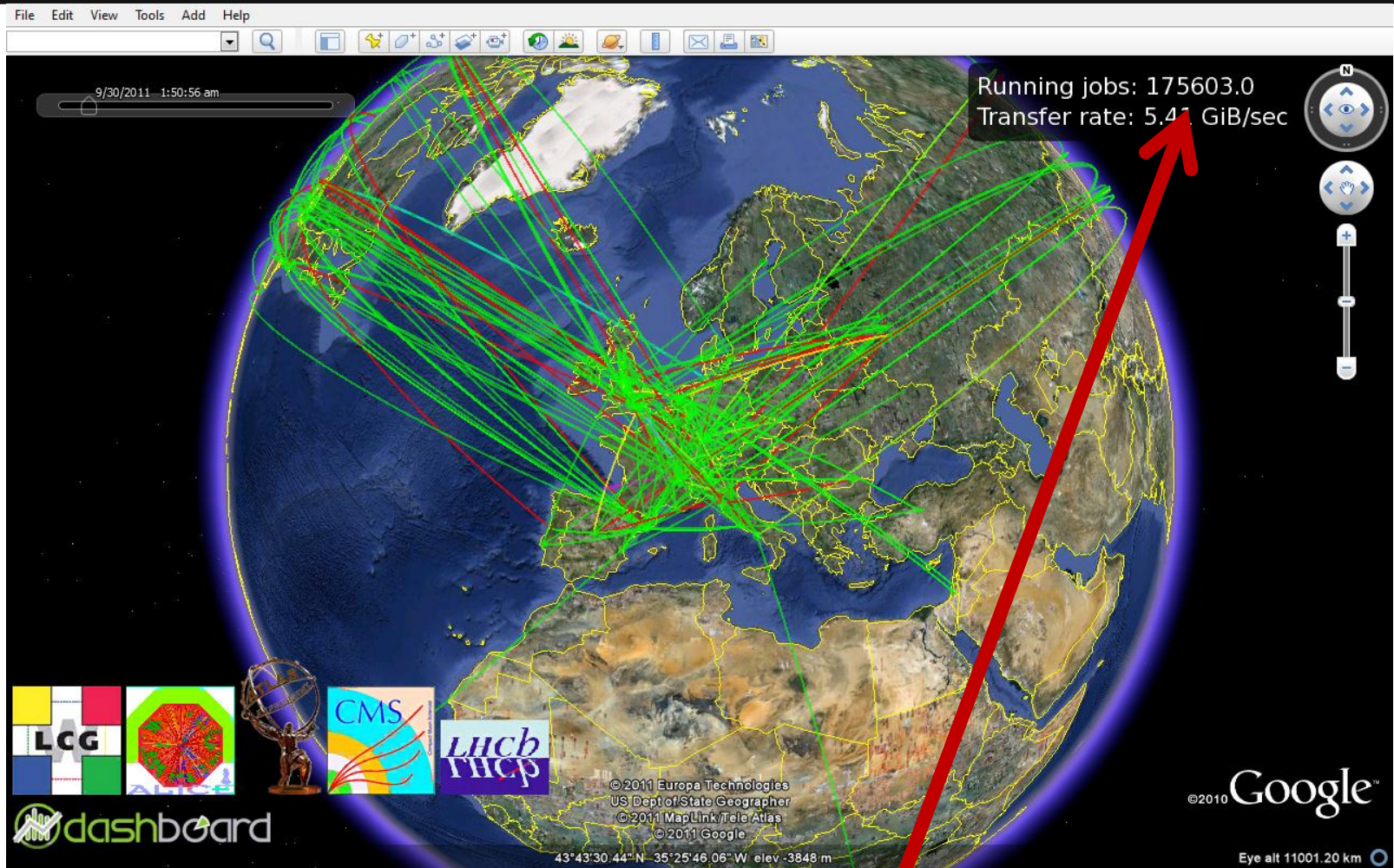
Tar  
Ag

Все поставленные задачи решены  
с запасом прочности в два раза



Export rates from Tier-0 up to ~ 3 GB/s  
Baseline is 1.5 GB/s

# Вычисления

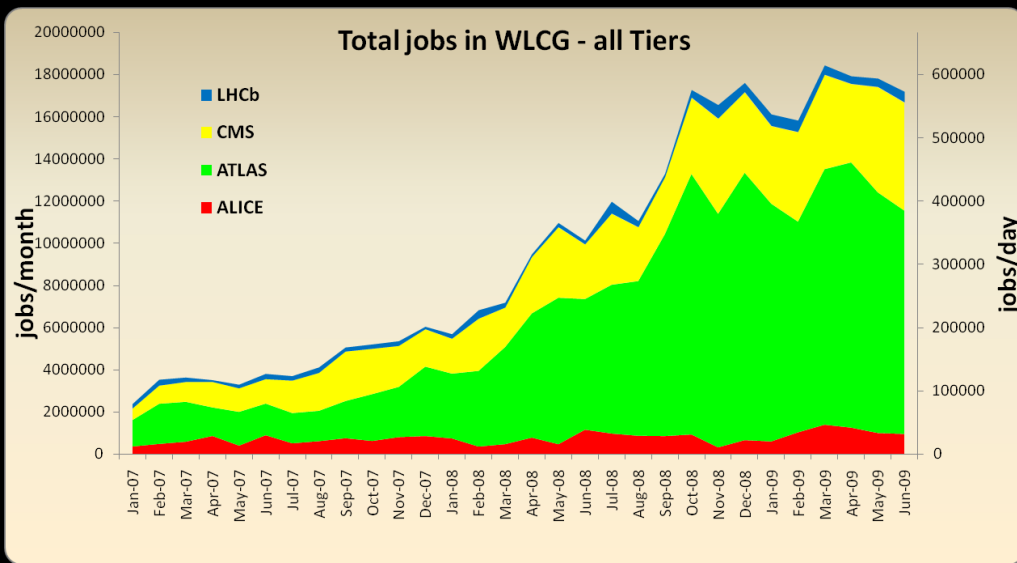


100-200 тысяч задач решается параллельно  
Около миллиона задач обрабатывается ежедневно

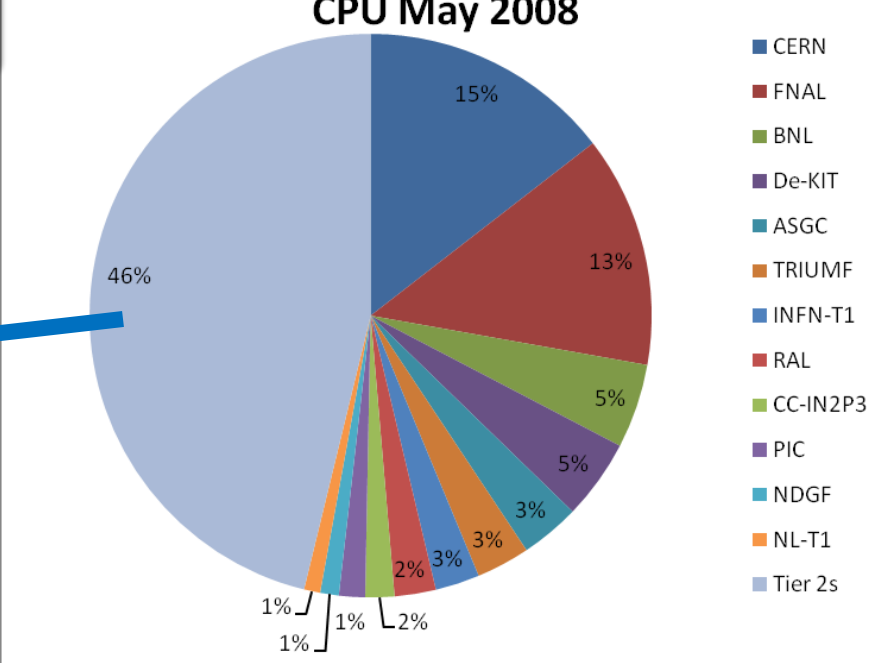


# Распределение задач по компьютерным центрам

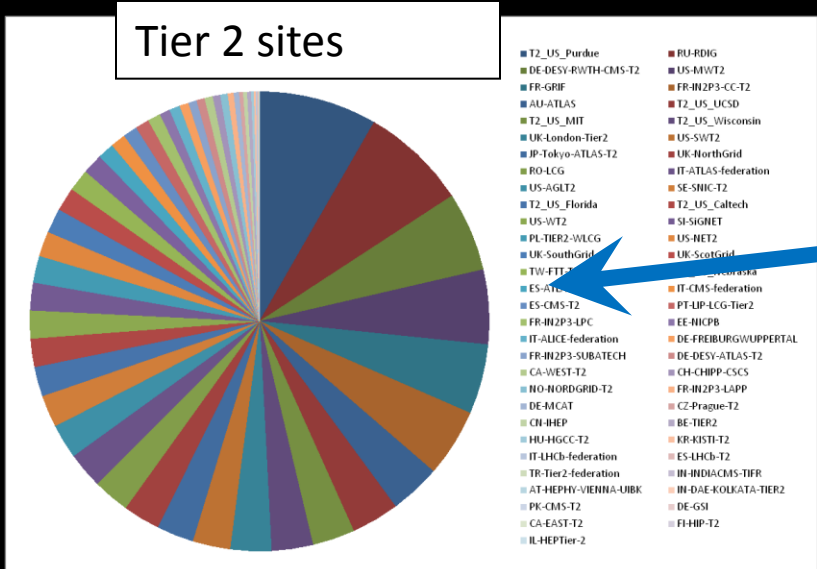
- Распределение задач по компьютерным центрам демонстрирует важность использования ГРИД системы
  - Вклад центров второго уровня ~ 50%;
  - >85% вычислений производится вне ЦЕРНа



Tier 0 + Tier 1 sites



Tier 2 sites



# 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 Киеве
- Являются центрами второго уровня

# 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





# 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



# 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

- 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

# 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