

Распределенные вычисления и большие данные

Кореньков Владимир Васильевич

директор Лаборатории информационных технологий ОИЯИ

зав. кафедрой «Распределенные информационно-
вычислительные системы» Университета Дубна

Руководитель лаборатории облачных вычислений и
аналитики больших данных РЭУ им. Г.В. Плеханова

Лекция в ЦЕРН
28 марта 2015 года

Наука с интенсивной обработкой данных – новая парадигма

Известный специалист MicroSoft в области хранения информации Джим Грэй предсказал, в 2005 г., что вступление научных исследований в эпоху пета- и экса-данных должно неизбежно потребовать развития новой **науки с интенсивной обработкой** и назвал это изменение **«четвертой парадигмой науки»**, в дополнение к трем предыдущим научным парадигмам — экспериментальной, теоретической и вычислительной

На торжестве по поводу получения Нобелевской премии за открытие бозона Хиггса директор ЦЕРНа Рольф Хойер прямо назвал грид-технологии одним из трех столпов успеха (наряду с ускорителем LHC и физическими установками).

Без организации грид-инфраструктуры на LHC было бы невозможно обрабатывать и хранить колоссальный объем данных, поступающих с коллайдера, а значит, - и совершать научные открытия.

Сегодня уже ни один крупный проект не осуществим без использования распределенной инфраструктуры для обработки данных.

ЦЕРН входит в эру Больших Данных и является одним из примеров (наряду с созданием в ЦЕРНе WWW-всемирной паутины), когда разработки в области физики частиц начинают влиять на исследования в других научных областях.



Концепция Грид

«Грид - это система, которая:

- координирует использование ресурсов при отсутствии централизованного управления этими ресурсами
- использует стандартные, открытые, универсальные протоколы и интерфейсы.
- обеспечивает высококачественное обслуживание»

(Ian Foster: "What is the grid? ", 2002 г.)

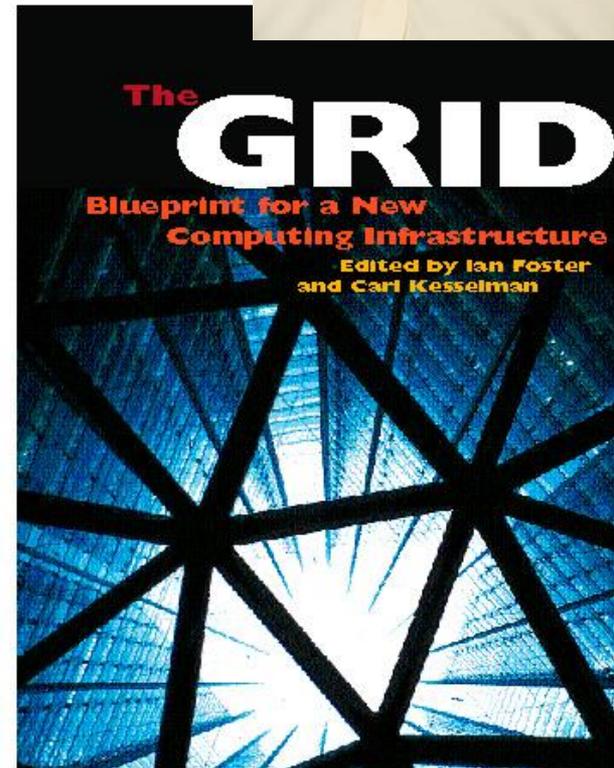


Модели грид:

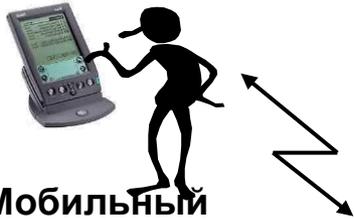
- ❖ **Distributed Computing**
- ❖ **High-Throughput Computing**
- ❖ **On-Demand Computing**
- ❖ **Data-Intensive Computing**
- ❖ **Collaborative Computing**

Междисциплинарный характер грид: развиваемые технологии применяются в физике высоких энергий, космофизике, микробиологии, экологии, метеорологии, различных инженерных и бизнес приложениях.

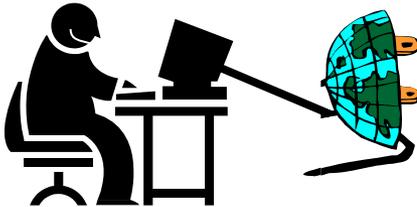
Виртуальные организации (VO)



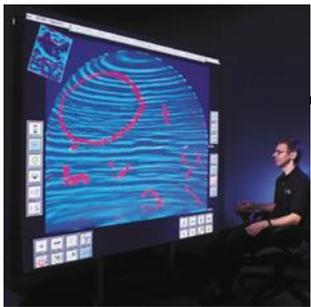
Грид - это средство для совместного использования вычислительных мощностей и хранилищ данных посредством интернета



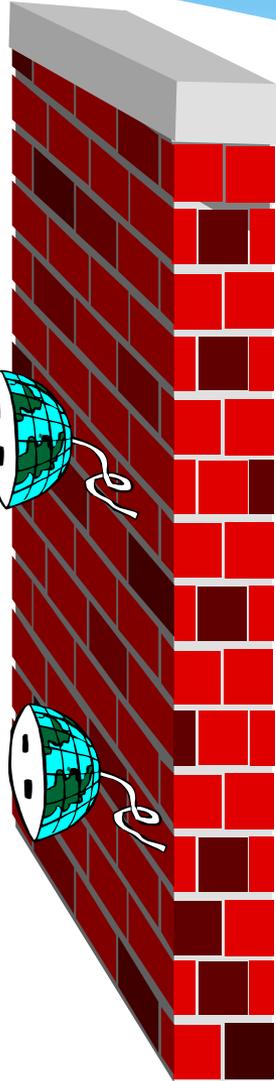
Мобильный доступ



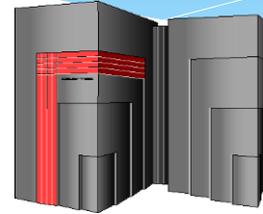
Рабочие станции



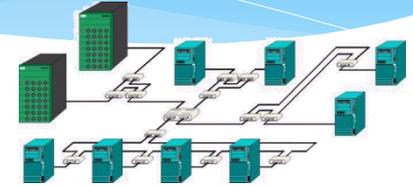
Визуализация



П
Р
О
М
Е
Ж
У
Т
О
Ч
Н
О
Е
П
Р
О
Г
Р
А
М
М
Н
О
Е



Суперкомпьютеры, ПК- кластеры



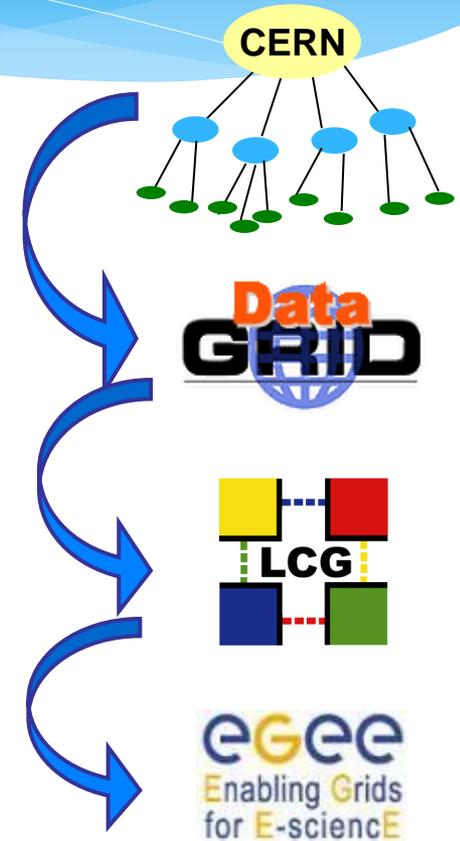
Массовая память, сенсоры, эксперименты



Интернет, сети

Some history

- * **1999 – Monarc Project**
 - * Early discussions on how to organise distributed computing for LHC
- * **2001-2003 - EU DataGrid project**
 - * middleware & testbed for an operational grid
- * **2002-2005 – LHC Computing Grid – LCG**
 - * deploying the results of DataGrid to provide a production facility for LHC experiments
- * **2004-2006 – EU EGEE project phase 1**
 - * starts from the LCG grid
 - * shared production infrastructure
 - * expanding to other communities and sciences
- * **2006-2008 – EU EGEE-II**
 - * Building on phase 1
 - * Expanding applications and communities ...
- * **2008-2010 – EU EGEE-III**
- * **2010-2012 - EGI-InSPIRE**



Status Jan 2011 (yearly increase)

- 13800 users: +38%
- 288000 LCPUs (cores): +18.5%
- 117PB disk: +192.5%
- 91.5PB tape: +50 %
- 28 million jobs/month: +86.7%
- 340 sites: +7.25%
- 56 countries: +7.7%
- 217 VOs: +24%
- 30 active VOs: constant

Scheduled = 21539

Running = 25374

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences

...

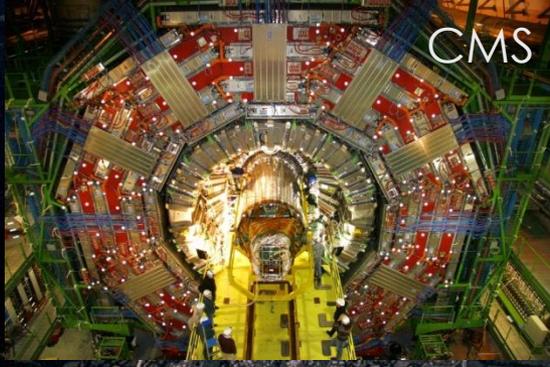
21:13:50 UTC

Cloudscape III - EGI Use Case



GridPP₃

UK Computing for Particle Physics



Exploration of a new energy frontier
Proton-proton and Heavy Ion collisions
at E_{CM} up to 14 TeV



Параметры детектора АТЛАС

Энергия центра масс 14 TeV

Частота столкновений пучков
40 MHz

Светимость :

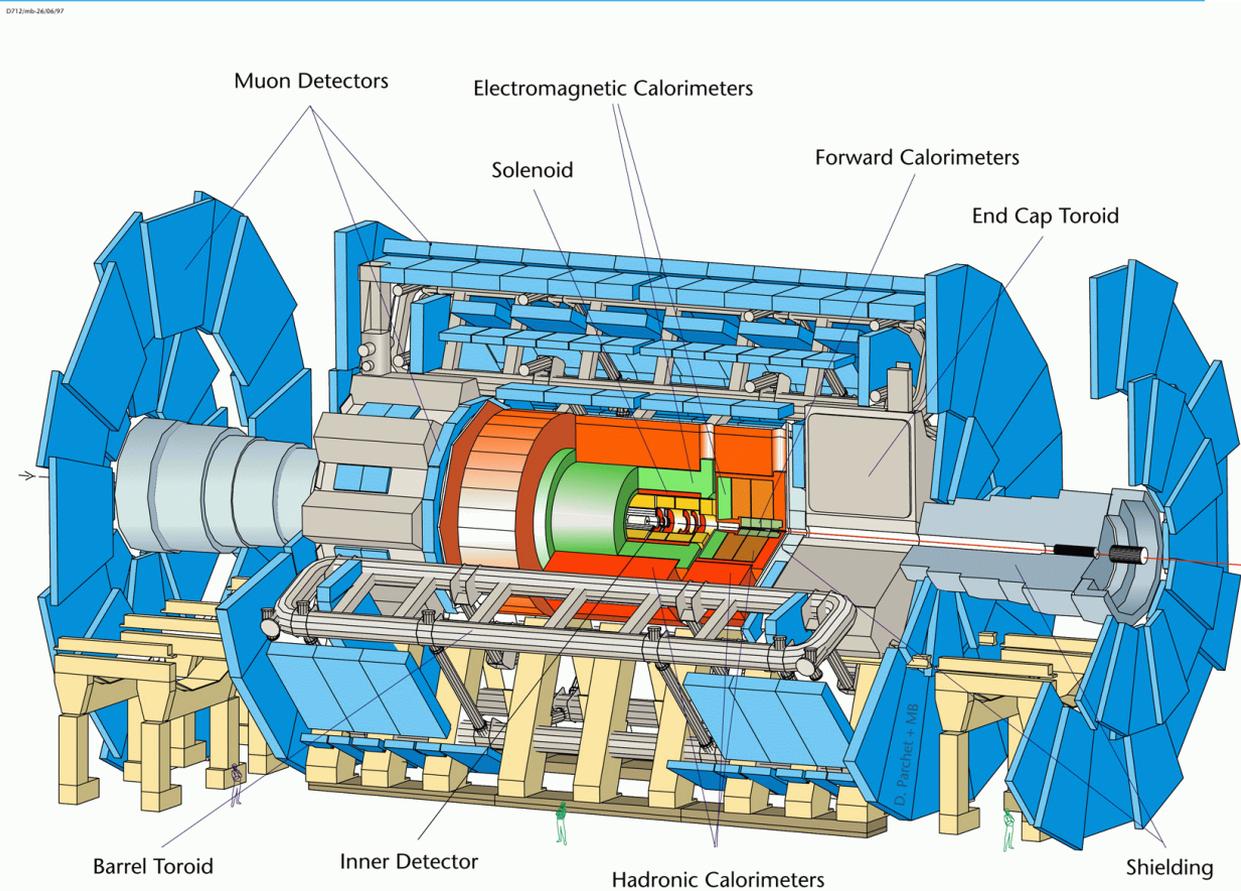
- начальная: $10^{31} \text{ см}^{-2}\text{с}^{-1}$
- низкая: $2 \cdot 10^{33} \text{ см}^{-2}\text{с}^{-1}$
- целевая: $10^{34} \text{ см}^{-2}\text{с}^{-1}$

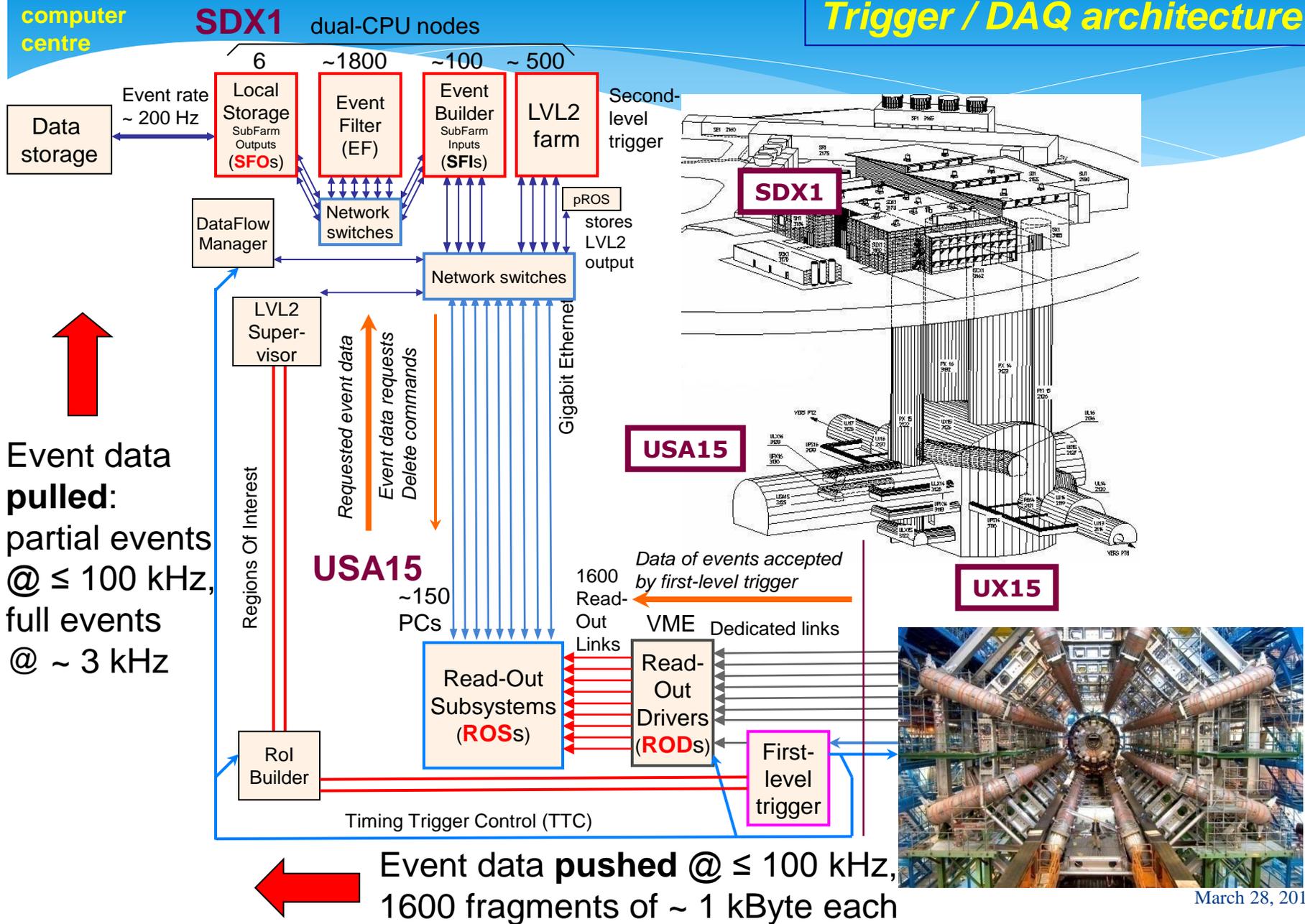
Вес 7000 тонн,

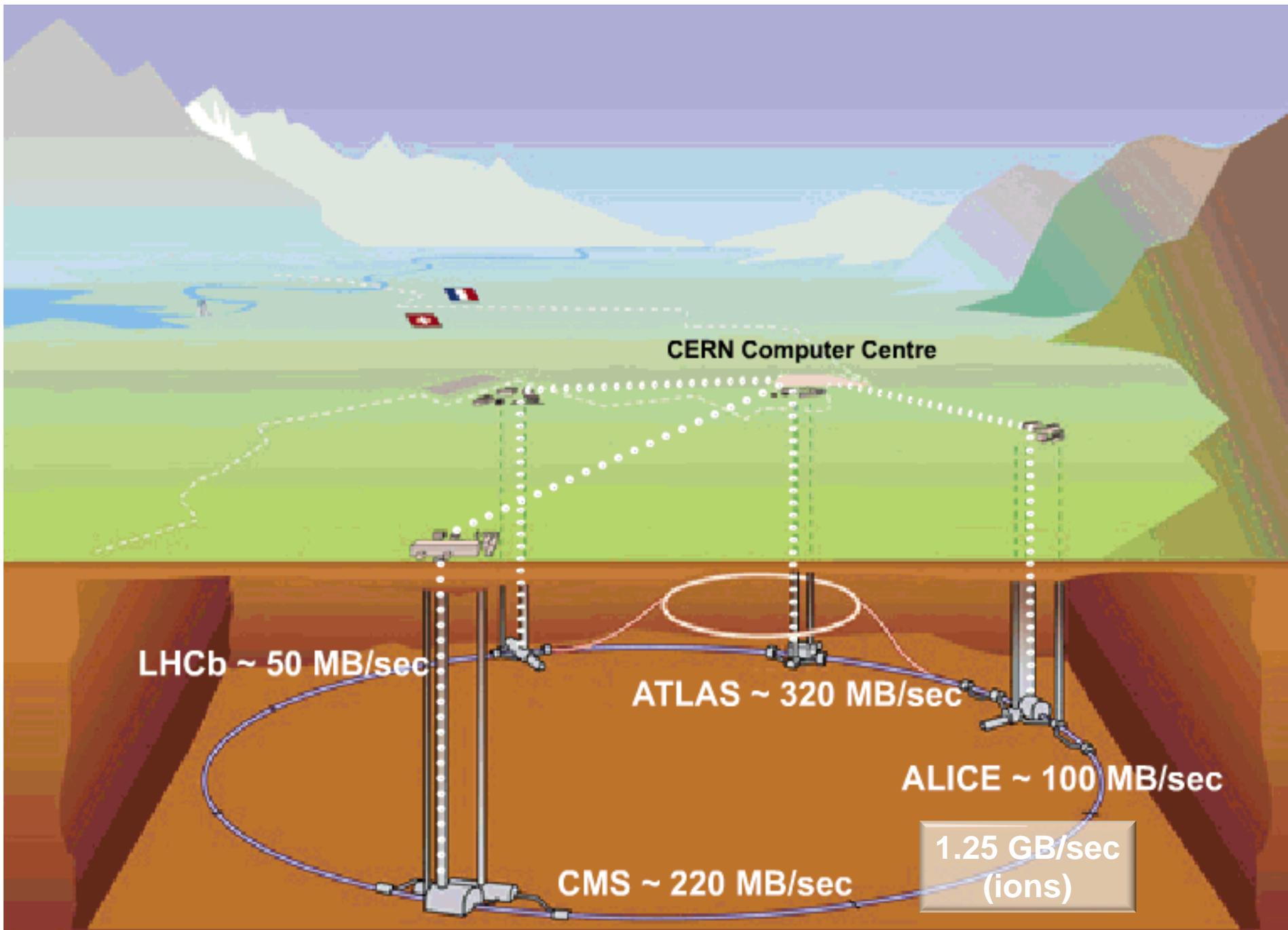
Диаметр 24м,

Длина 46м

Количество регистрирующих
каналов 140 000 000







CERN Computer Centre

LHCb ~ 50 MB/sec

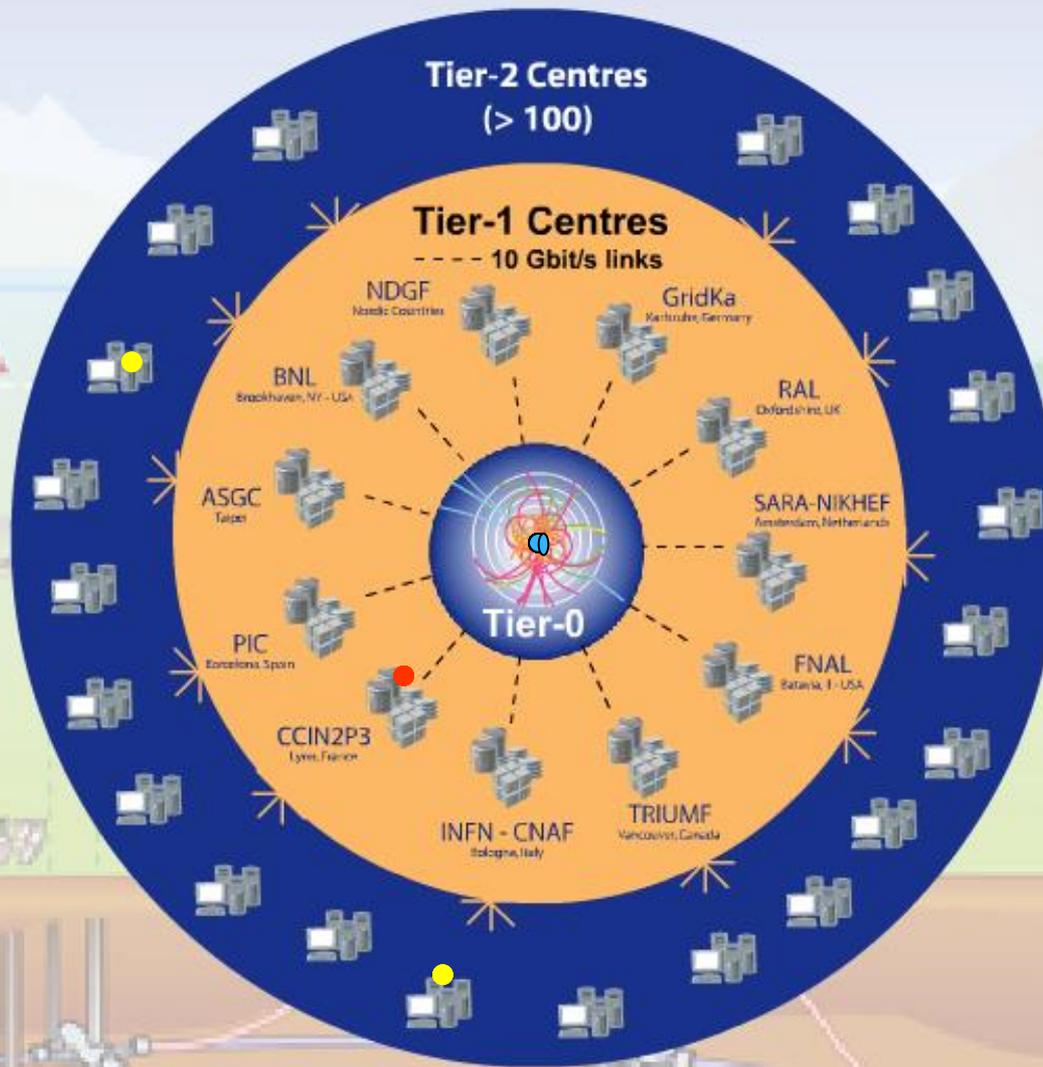
ATLAS ~ 320 MB/sec

ALICE ~ 100 MB/sec

CMS ~ 220 MB/sec

1.25 GB/sec
(ions)

Модель компьютеринга на LHC



Tier-0 (CERN):

- Прием данных
- Начальная реконструкция данных
- Распределение данных

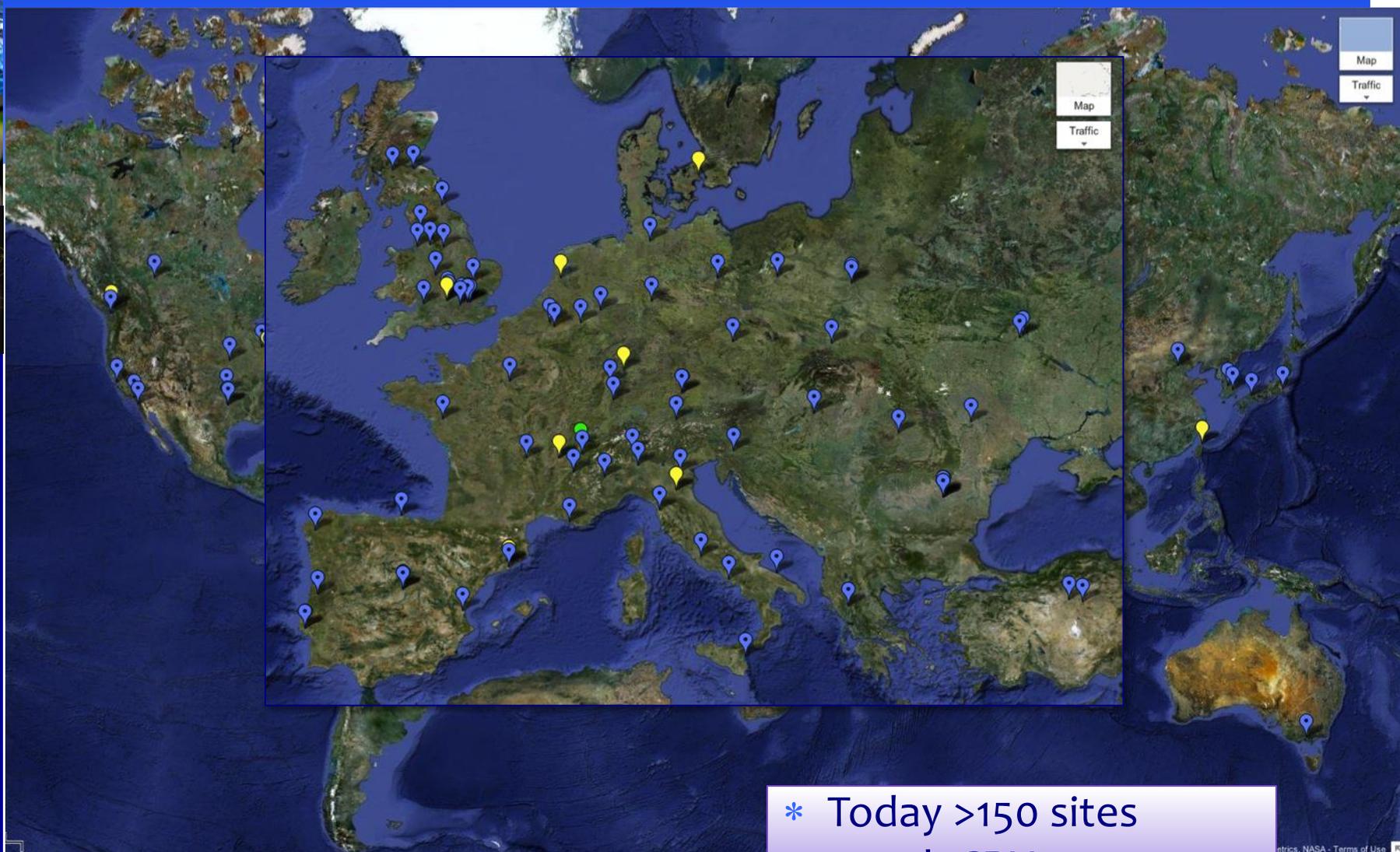
Tier-1 (11 centres):

- Постоянное хранение данных
- Реконструкция и обработка
- Анализ

Tier-2 (>200 centres):

- Моделирование
- Физический анализ

WLCG Grid Sites



- * Today >150 sites
- * >300k CPU cores
- * >250 PB disk



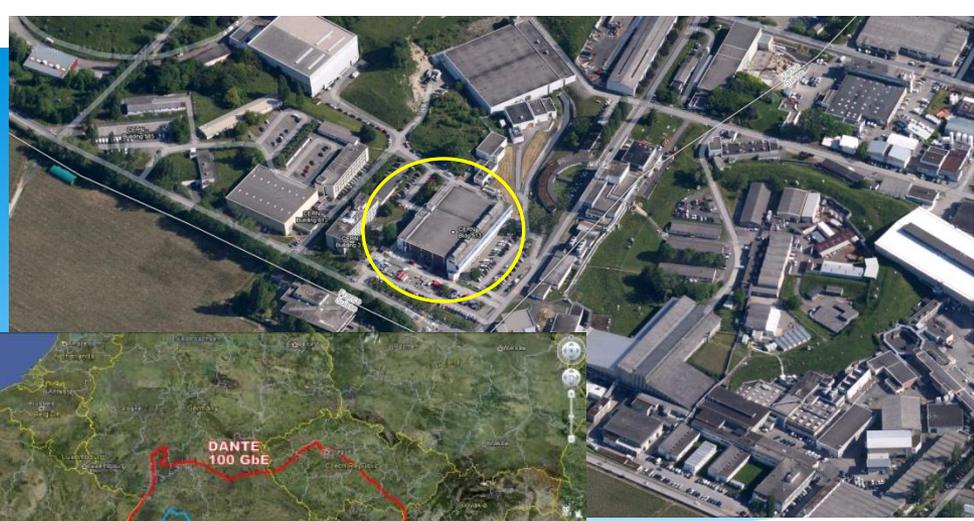
CERN Computer Centre

CERN Computer Centre:

- Built in the 70s on the CERN site (Geneva)
- ~3000 m² (3 main machine rooms)
- 3.5 MW for equipment

New extension:

- Located at Wigner (Budapest)
- ~1000 m²
- 2.7 MW for equipment
- Connected to the Geneva CC with 2x100Gb links



Грид-инфраструктура LHC – ЦЕРН (Tier-0)



CERN “Big Data”



Service	Size	Files
AFS	290 TB	2.3 B
CASTOR	89.0 PB	325 M
EOS	20.1 PB	160 M

Physics Data on CASTOR/EOS

- LHC experiments produce ~10GB/s
25PB/year

User Data on AFS & DFS

- Home directories for 30k users
- Physics analysis development
- Project spaces for applications

Service Data on AFS/NFS/Filers

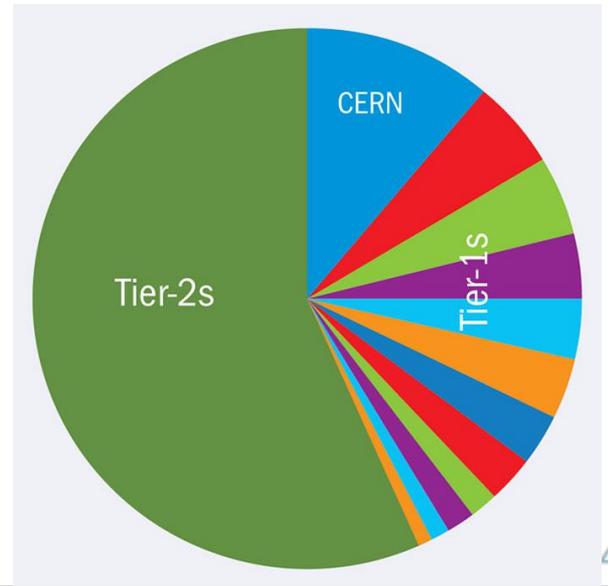
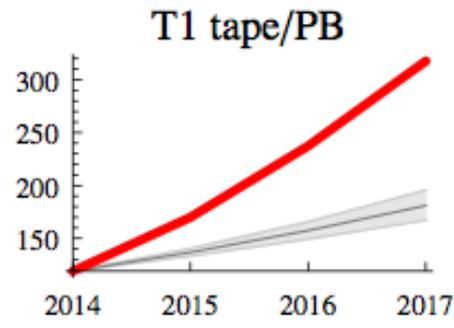
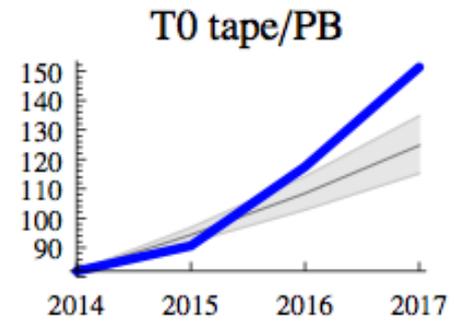
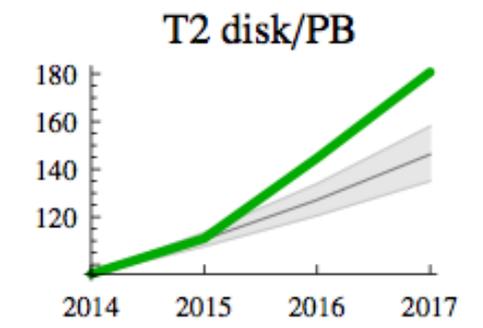
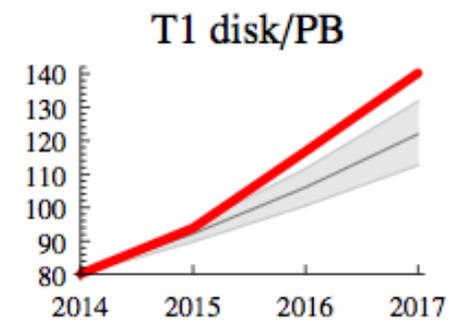
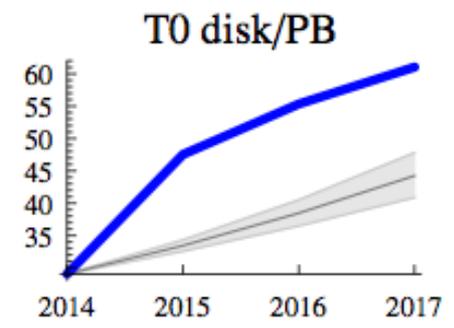
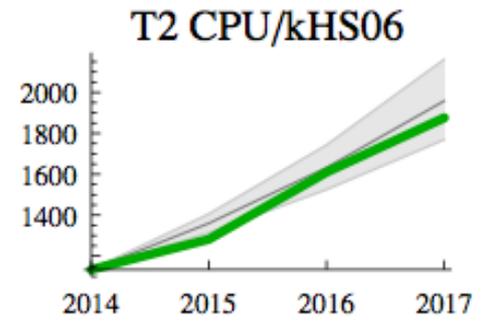
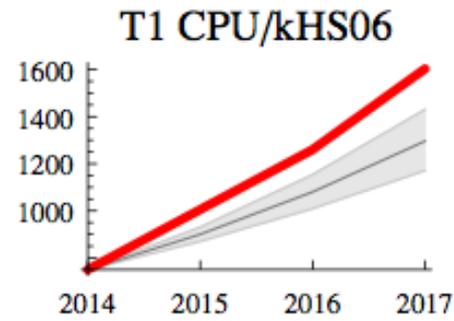
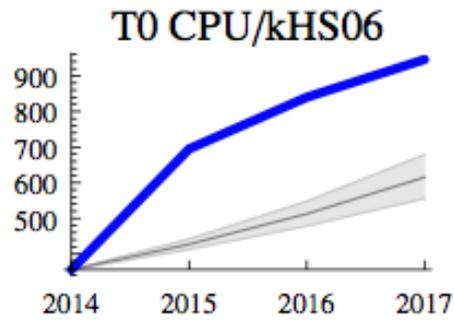
- Databases, admin applications

Tape archival with CASTOR/TSM

- Physics data (RAW and derived/additional)
- Desktop/Server backups



Combined (sum of experiments)



Starting from 2014 pledges

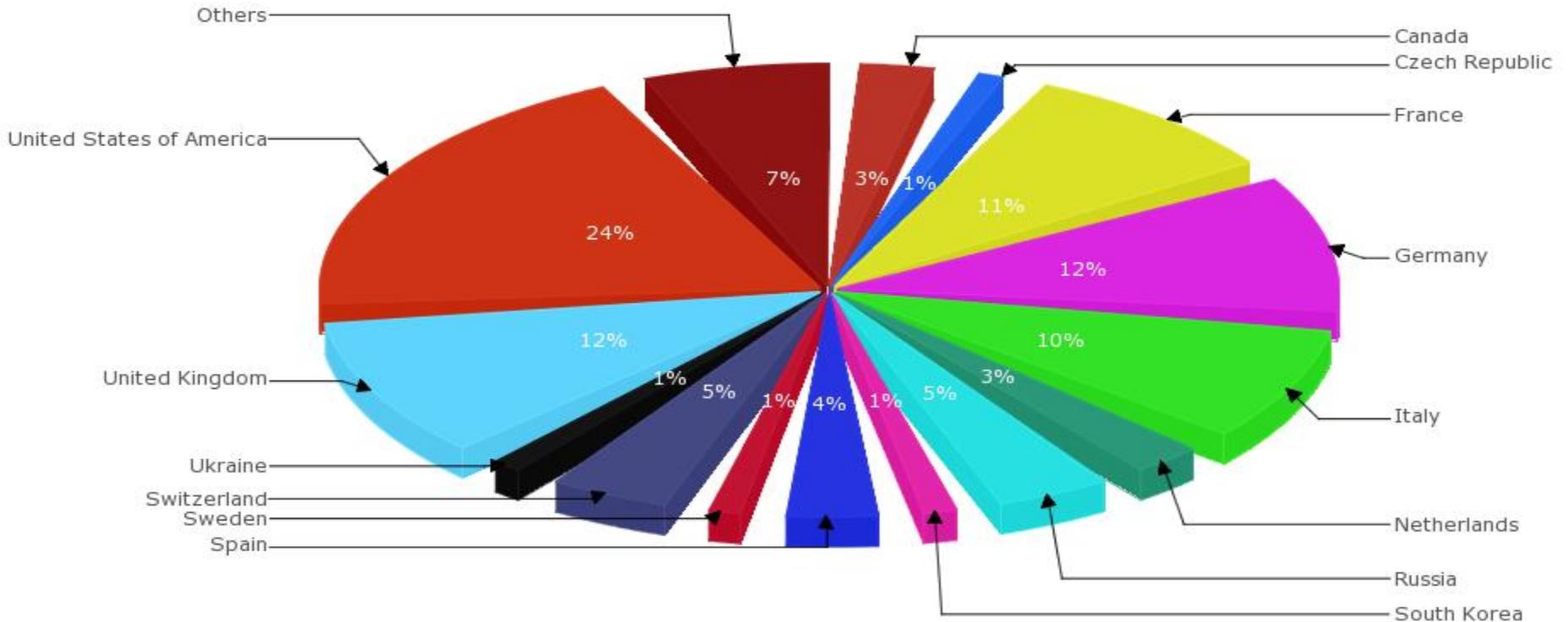


Country Normalized CPU time 2014

LCG View: / normcpu / 2014: 11-2015: 3 / COUNTRY-VO / lhc (x) / GRBAR-LIN / 1

20

COUNTRY Normalised CPU time (kSI2K) per COUNTRY



All Country - 23,107,789,664
Job 724,967,359

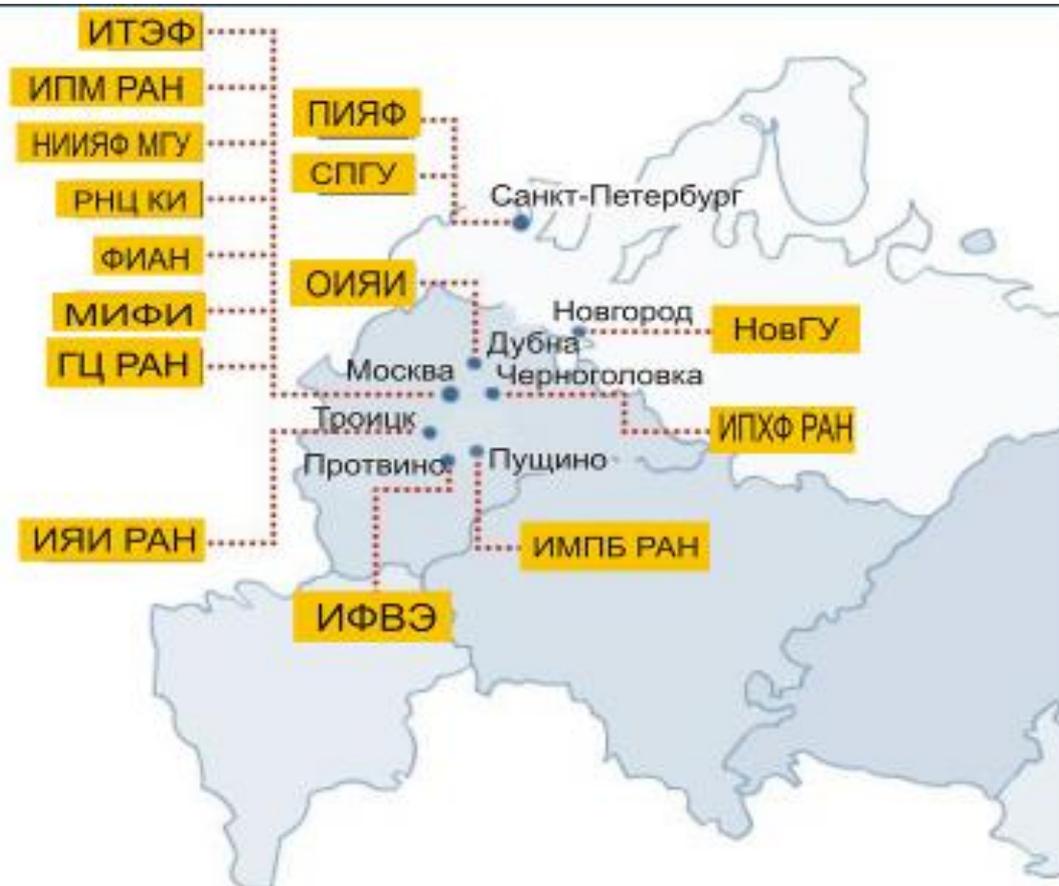
Russia- 1,021,388,412
34,061,750

Russian Data Intensive Grid infrastructure (RDIG)

The Russian consortium RDIG (Russian Data Intensive Grid), was set up in September 2003 as a national federation in the EGEE project. Now the RDIG infrastructure comprises **10 Resource Centers** with **> 12000 CPU** and **> 10000 TB** of disc storage.

RDIG Resource Centres:

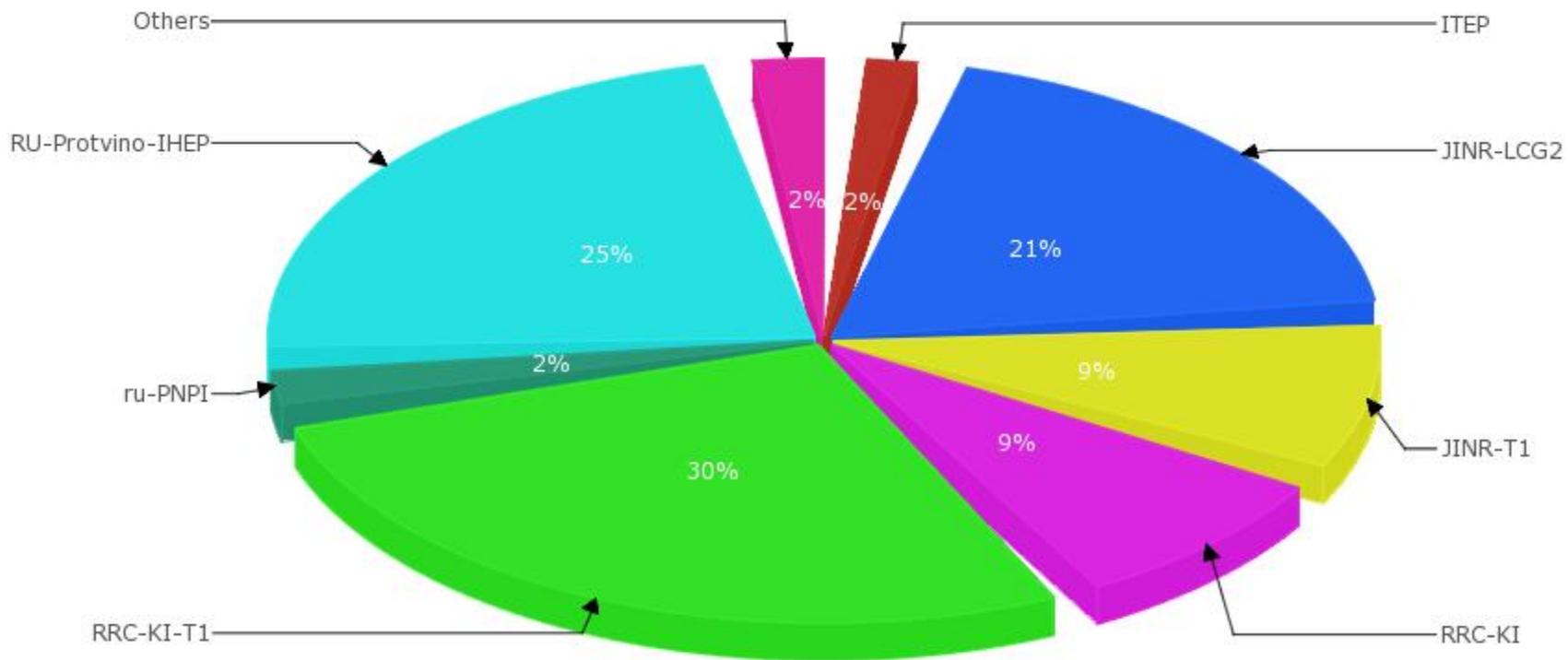
- ITEP
- JINR-LCG2 (Dubna)
- RRC-KI
- RU-Moscow-KIAM
- RU-Phys-SPbSU
- RU-Protvino-IHEP
- RU-SPbSU
- Ru-Troitsk-INR
- ru-IMPB-LCG2
- ru-Moscow-FIAN
- ru-Moscow-MEPHI
- ru-PNPI-LCG2 (Gatchina)
- ru-Moscow-SINP
- **Kharkov-KIPT (UA)**
- **BY-NCPHEP (Minsk)**
- **UA-KNU**
- **UA-BITP**





RDIG Normalized CPU time (2014)

Russia Normalised CPU time (kSI2K) per SITE



Worldwide LHC Computing Grid Project (WLCG)

Основной задачей проекта WLCG является создание глобальной инфраструктуры региональных центров для обработки, хранения и анализа данных физических экспериментов LHC.

Грид-технологии являются основой построения этой инфраструктуры. *Протокол между ЦЕРН, Россией и ОИЯИ об участии в проекте LCG был подписан в 2003 году.*

МоУ об участии в проекте WLCG был подписан в 2007 году.

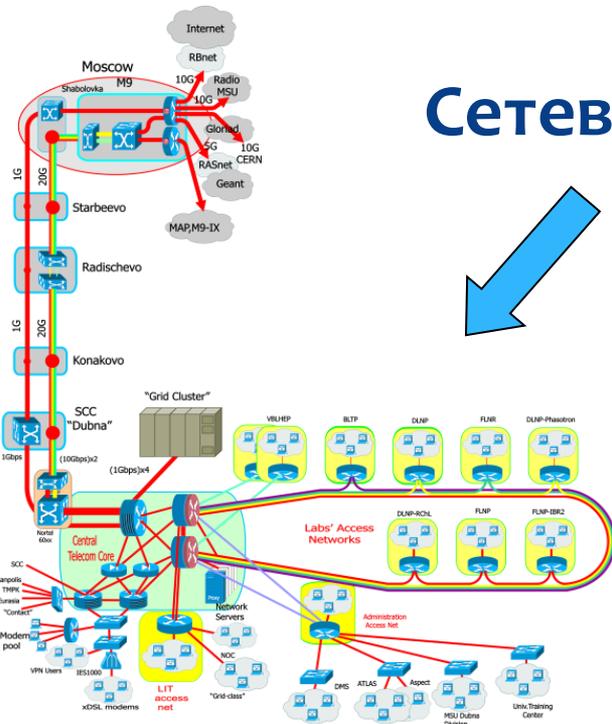
Задачи российских центров и ОИЯИ в проекте WLCG в 2014 году:

- Создание комплекса тестов для gLite
- Внедрение сервисов WLCG для экспериментов
- Развитие систем мониторинга WLCG
- Система глобального мониторинга Tier3 центров
- Развитие пакетов моделирования для экспериментов
- Разработка архитектуры Tier1 центра в России

Основные направления деятельности ЛИТ

- Развитие ИТ-инфраструктуры
- Математическая, алгоритмическая и программная поддержка исследований, проводимых в ОИЯИ и странах-участницах
- Корпоративная информационная система ОИЯИ

Три основных уровня Сетевой, ресурсный, прикладной



Всего сотрудников – 324 из 18 стран мира, 30 – докторов наук, 65-кандидаты наук,
В 2014 году опубликовано: **270** статей в реферированных журналах, **2** монографии,
112 статей в трудах конференций

JINR Central Information and Computing Complex (CICC)



CICC comprises 4780 Cores
Disk storage capacity 4200 TB

Availability and Reliability = 99%

RDIG Normalized CPU time (HEPSPEC06) per site (2010-2015)

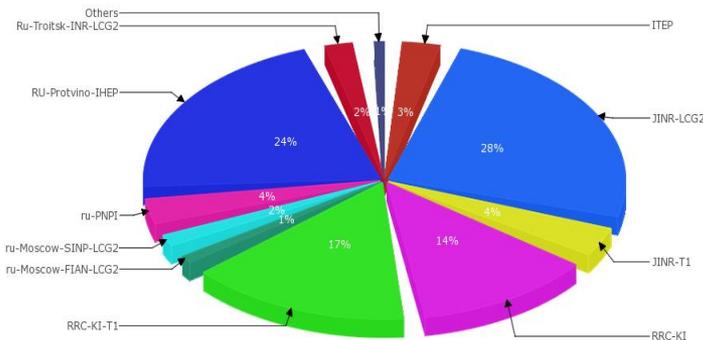
30 million Jobs and 800 million normalized CPU time were executed in 2010-2015

Foreseen computing resources to be allocated for JINR CICC

CE9SA*EGI View: / normcpu / 2010: 1-2015: 3 / SITE-VO / lhc (x) / GRBAR-LIN / 1

Russia Normalised CPU time (kS12K) per SITE

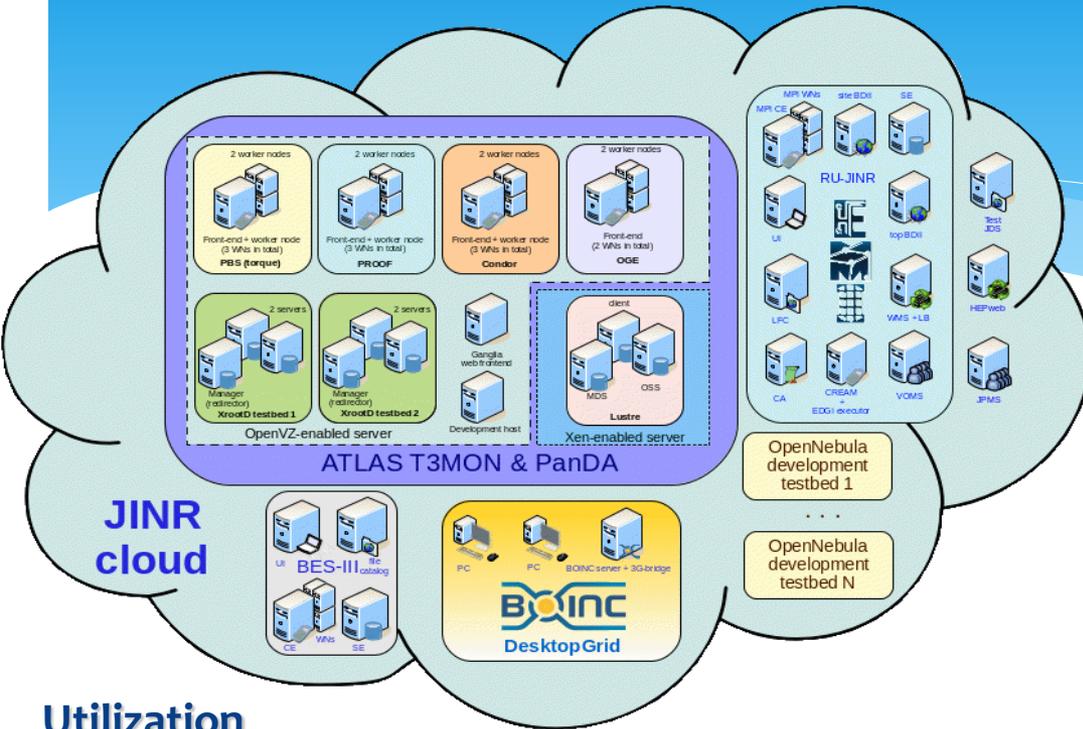
2015-03-18 18:37



JINR covers 33% of the RDIG share to the LHC

	2015	2016
CPU (HEPSPEC06)	28 000	40 000
Disk storage (TB)	4 000	8 000
Tape Library (TB)	5 000	10 000

JINR cloud infrastructure



Purpose:

- increase the efficiency of hardware and proprietary software utilization,
- improve IT-services management.

Implementation:

- Cloud platform: OpenNebula
- Two types of VM managers: OpenVZ (linux only) and KVM (any OS)
- Two types of VMs: highly reliable (run on RAID1 disks) and normal
- user access: web GUI and command line interface
- VM access: rsa/dsa-key or password
- Authentication: JINR central user database (LDAP+Kerberos)
- HTTPS protocol (more secure information exchange between clients and service)

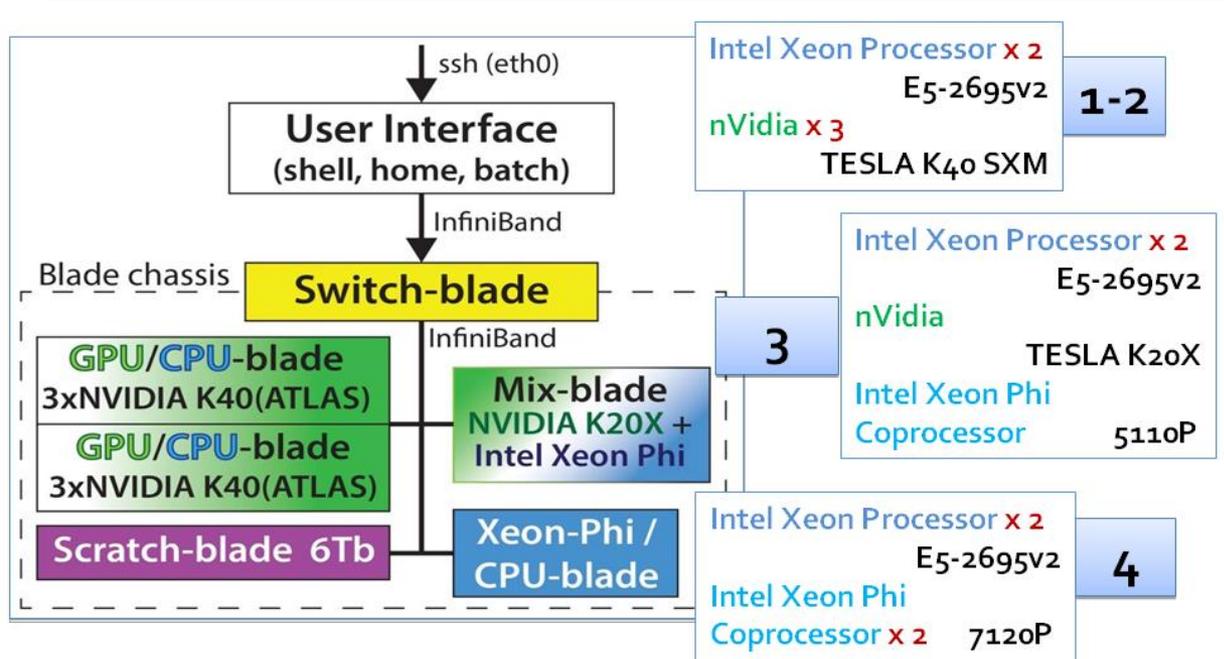
Utilization

- EMI testbed
- ATLAS T3MON + PanDA
- BES-III (DIRAC)
- DesktopGrid testbed
- HEPWEB
- Test JINR document server (JDS)
- JINR project management service (JPMS)
- OpenNebula testbeds for development and tests

Service URL:
<http://cloud.jinr.ru>
(Authentication needed)

The screenshot shows the login page for the JINR cloud service. It features the JINR logo at the top, which includes the text 'JOINT INSTITUTE FOR NUCLEAR RESEARCH' and 'JINR'. Below the logo, there are two input fields: 'Username' and 'Password'. At the bottom left, there is a checkbox labeled 'Keep me logged in'. At the bottom right, there is a 'Login' button.

The heterogeneous computing cluster **HybriLIT**



Resources

CPU 96 cores	RAM 512 Gb
GPU 19968 cuda cores	EOS storage 14 Tb
PHI 182 cores	Ethernet
	InfiniBand 40 Gb/s

Performance

Max. single-precision	46,914 Tflops
Max. double-precision	17,979 Tflops
Power consumption:	10 KW

Installed software

Scientific Linux 6.5.
CUDA Toolkit 5.5, CUDA Toolkit 6.0
OpenMPI 1.6.5, 1.8.1
OpenMP GCC, ICC
Intel Cluster Studio 2013
JDK-1.7.0, JDK-1.8.0

Used compilers

nvcc
mpic++, mpicc, mpicxx, mpif77, mpif90,
mpifort, icc, ifort
mpiicc, mpiifort

Учебные курсы на **HybriLIT**

hYBRI

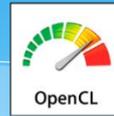


Parallel programming technologies on hybrid architectures



NVIDIA

MPI OpenMP



7 – 17 July, 2014

Participants
From Mongolia,
Romania,
Russia



27 August, 2014

Participants from CIS and Russian institutes and companies



MPAMCS 2014

International Conference for Young Scientists
«MODERN PROBLEMS OF APPLIED
MATHEMATICS & COMPUTER SCIENCE»

August 25 - 29 2014, Dubna, Russia

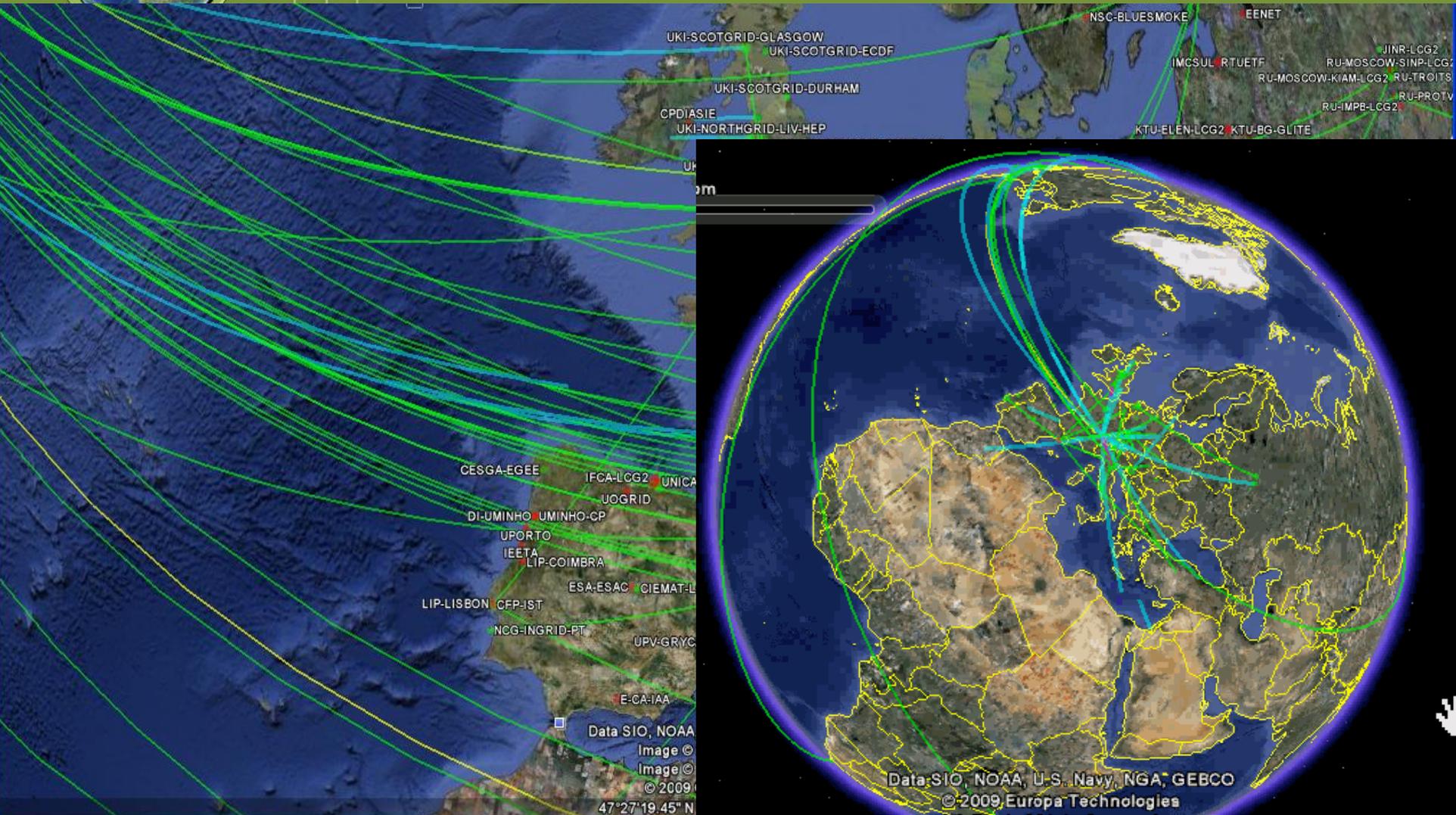
1 and 5 September, 2014

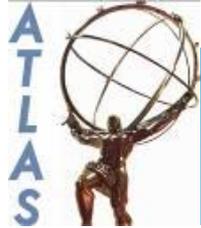
Participants from India, Germany, Japan, Ireland, Austria,
Ukraine, Russia



Полное число участников учебных курсов – студентов, аспирантов и молодых ученых составило порядка **60** из Германии, Индии, Монголии, Украины, России, Румынии и др.

WLCG Google Earth Dashboard





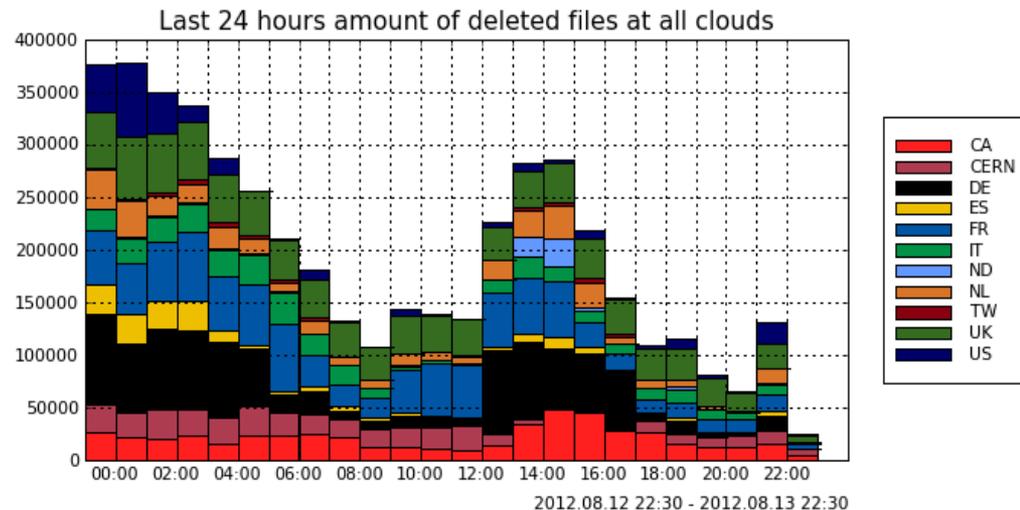
Архитектура системы распределенного хранения данных эксперимента ATLAS

Разработка новой архитектуры сервиса удаления данных для обеспечения целостности распределенного хранения информации эксперимента ATLAS.

- * обслуживает запросы на удаление,
- * организует балансировку нагрузки,
- * обеспечивает масштабируемость и отказоустойчивость системы,
- * корректную обработку исключений, возникающих в процессе работы,
- * стратегию повтора операций в случае возникновения отказов.

Разработаны:

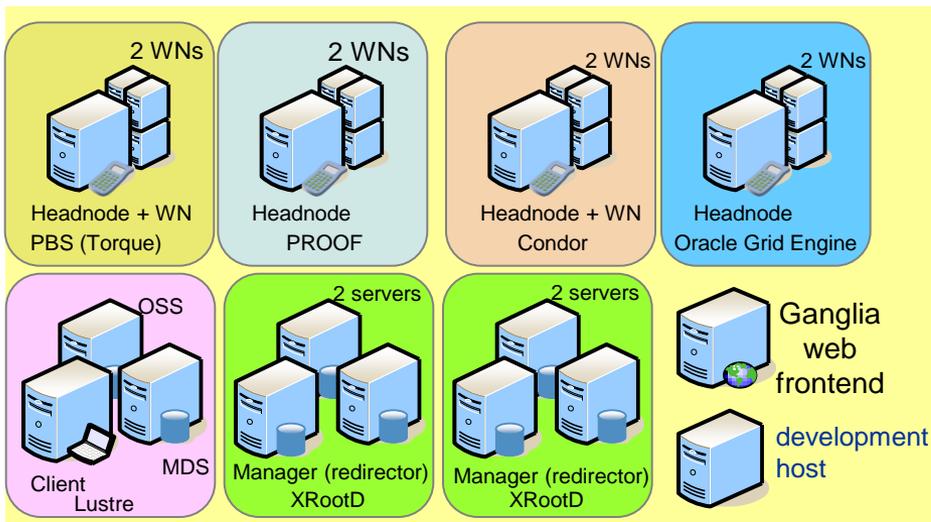
- * новый интерфейс между компонентами сервиса
- * создана новая схема базы данных,
- * перестроено ядра сервиса,
- * созданы интерфейсы с системами массового хранения,
- * модернизирована система мониторинга работы сервиса.



Созданный сервис обеспечивает целостность хранения информации в географически распределенной среде. Данные эксперимента ATLAS распределены более, чем на 130 грид-сайтах с общим объемом дискового пространства более 120 петабайт, в котором хранятся сотни миллионов файлов. Недельный объем удаляемых данных составляет 2 Пб (20 000 000 файлов).

Система мониторинга Tier3-центров для анализа данных экспериментов БАК

Для исследования центров уровня Tier3 в ОИЯИ была разработана архитектура тестовой инфраструктуры, которая позволяет создавать прототипы различных конфигураций центров уровня Tier3. С применением технологий виртуализации такая инфраструктура была реализована на базе виртуальных кластеров, что позволило разработать для каждого варианта конфигурации документацию, настроить или разработать средства локального мониторинга, выработать полные рекомендации по системе сбора информации для глобального мониторинга центров уровня Tier3.



Реализация системы мониторинга Tier3 – центров (не-грид) имеет огромное значение для координации работ в рамках виртуальной организации, так как обеспечивается глобальный взгляд на вклад Tier3 сайтов в вычислительный процесс.



Глобальная система мониторинга передачи данных в инфраструктуре WLCG

Суть проекта состоит в создании универсальной системы мониторинга, способной собирать информацию :

- о каждой передаче данных (Петабайта в день),
- независимо от метода осуществления (несколько протоколов и сервисов передачи файлов, FTS, xROOTd),
- уровень ресурсного центра (Tier-0, Tier-1, Tier-2, Tier-3)
- принадлежности данных определенной виртуальной организации;
- передавать с высокой степенью надежности собранную информацию в центральное хранилище (ORACLE, HADOOP);
- обрабатывать собранные данные для предоставления различным потребителям и программные интерфейсы для получения данных.



Система позволяет полностью удовлетворить потребности в информации различных типов пользователей и администраторов инфраструктуры WLCG.

LIT JINR - China collaboration

LIT team is a key developer of the BES-III distributed computing system

A prototype of BES-III Grid has been built (9 sites including IHEP CAS and JINR). Main developments have been done at IHEP and JINR. The Grid is based on DIRAC interware.

Monitoring

- BES-III grid monitoring system is operational since February 2014.
- Implementation of the new monitoring system based on DIRAC RSS service are in progress

Job management

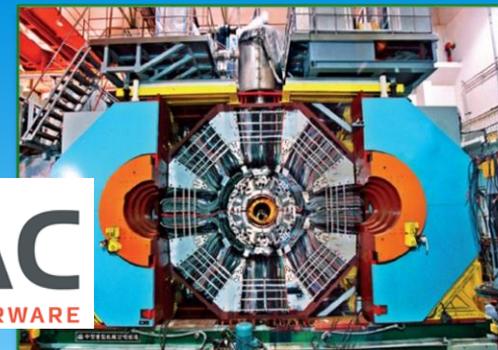
- Advising on the CE's installation and management
- BES-III jobs can be submitted on JINR cloud service now

Data management

- Installation package for Storage Element was adopted for BES-III Grid
- Solution on dCache-Lustre integration was provided for main data storage in IHEP
- Research on the alternative DB and data management service optimization is in progress

Infrastructure

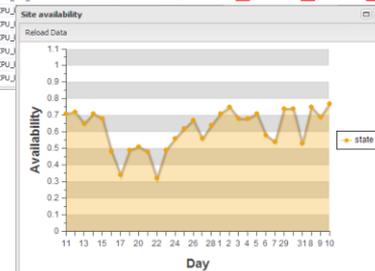
- Creation of the back-up DIRAC services for BES-III grid at JINR is in progress



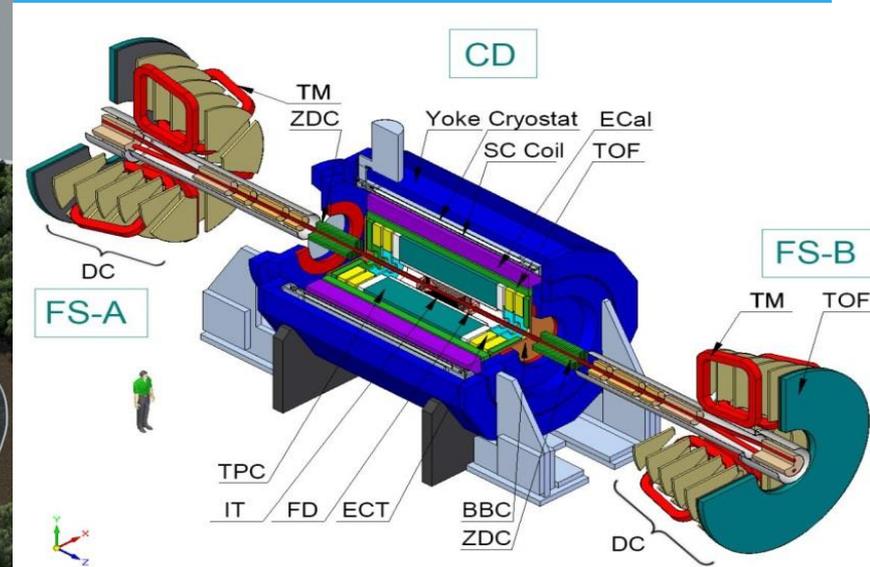
Source	Destination	Latency(sec)
IHEPD-USER	IHEPD-USER	2.678
IHEPD-USER	JINR-USER	16.316
IHEPD-USER	USTC-USER	15.932
IHEPD-USER	WHU-USER	6.728
JINR-USER	IHEPD-USER	14.322
JINR-USER	JINR-USER	14.24
JINR-USER	USTC-USER	14.827
JINR-USER	WHU-USER	8.516
USTC-USER	IHEPD-USER	3.677
USTC-USER	JINR-USER	17.855
USTC-USER	USTC-USER	2.746
USTC-USER	WHU-USER	624.375
WHU-USER	IHEPD-USER	5.727
WHU-USER	JINR-USER	20.227
WHU-USER	USTC-USER	9.199
WHU-USER	WHU-USER	3.092

Site	Service	Test	Result	Description	24h Reliability	48h Reliability	Week Reliability
BES.LUCAS.cn	WMS	WMS_jcmd_test	Fail	Failed after 30 ...	■	■	■
BES.IHEP-PBS.cn	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.JINR.ru	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.LMHL.us	WMS	WMS_jcmd_test	Fail	Failed after 30 ...	■	■	■
BES.LMHL.us	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.LSTC.cn	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.WHU.cn	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.IHEP-Torim	WMS	WMS_jcmd_test	Success	Remote call	■	■	■
BES.IHEP-IM.cn	WMS	WMS_jcmd_test	Fail	Failed after 30 ...	■	■	■
BES.LUCAS.cn	WMS	BOSS_work_test	Fail	Failed after 90 ...	■	■	■
BES.IHEP-PBS.cn	WMS	BOSS_work_test	Success	Success	■	■	■
BES.JINR.ru	WMS	BOSS_work_test	Success	Success	■	■	■
BES.JINR.ru	WMS	BOSS_work_test	Fail	Failed after 90 ...	■	■	■
BES.LMHL.us	WMS	BOSS_work_test	Success	Success	■	■	■
BES.LSTC.cn	WMS	BOSS_work_test	Fail	boss.exe not fo...	■	■	■
BES.WHU.cn	WMS	BOSS_work_test	Success	Success	■	■	■
BES.IHEP-Torim	WMS	BOSS_work_test	Success	Success	■	■	■
BES.IHEP-IM.cn	WMS	BOSS_work_test	Fail	Failed after 90 ...	■	■	■
BES.LUCAS.cn	WMS	CPU_jcmd_test	Fail	Failed after 30 ...	■	■	■
BES.IHEP-PBS.cn	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.JINR.ru	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.LMHL.us	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.LSTC.cn	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.WHU.cn	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.IHEP-Torim	WMS	CPU_jcmd_test	Success	Success	■	■	■
BES.IHEP-IM.cn	WMS	CPU_jcmd_test	Success	Success	■	■	■

Site	Host	24h	24h	24h	48h	48h	48h	Week	Week	Week	Week	F
BES.IHEP-PBS.cn	gridb002.ihep.ac.cn	2	2	2	2	2	2	2	2	2	2	■
BES.LMHL.us	twins-a04.spa.umn.edu	1	1	1	1	1	1	1	1	1	1	■
BES.JINR.ru	vm362.jinr.ru				1	1	1	1	1	1	1	■
BES.IHEP-CLOU...	diraccloudint1403249980							7	7	7	7	■
BES.IHEP-CLOU...	diraccloudint1403250760											■
BES.LMHL.us	twins-b14.spa.umn.edu	1	1	1	1	1	1	1	1	1	1	■
BES.LMHL.us	twins-a24.spa.umn.edu				1	1	1	1	1	1	1	■
BES.IHEP-CLOU...	diraccloudint1403250400											■
BES.JINR.ru	vm000.jinr.ru				1	1	1	1	1	1	1	■
BES.JINR.ru	vm400.jinr.ru				1	1	1	1	1	1	1	■
BES.JINR.ru	vm323.jinr.ru	1	1	1	1	1	1	1	1	1	1	■
BES.IHEP-CLOU...	diraccloudint1403490272	1	1	1	1	1	1	14	14	14	14	■
BES.WHU.cn	cu33	6	6	6	6	6	6	6	6	6	6	■
BES.LMHL.us	twins-b03.spa.umn.edu				1	1	1	1	1	1	1	■
BES.IHEP-CLOU...	diraccloudint1403254687							5	5	5	5	■
BES.IHEP-CLOU...	diraccloudint1403495687	2	2	2	2	2	2	7	7	7	7	■
BES.JINR.ru	vm324.jinr.ru				1	1	1	1	1	1	1	■



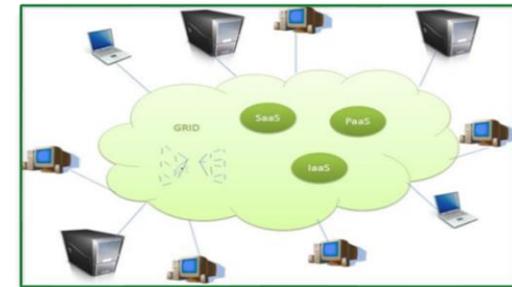
Ускорительный комплекс НИКА



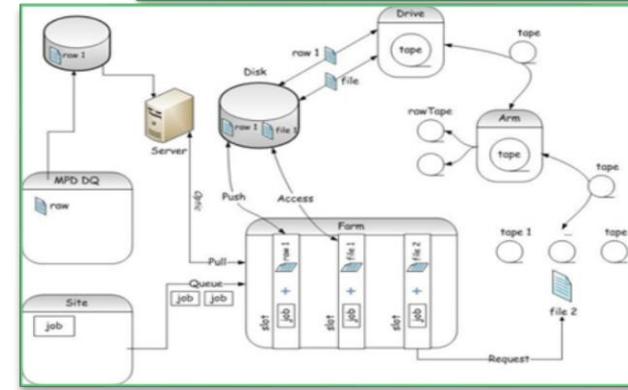
Для проекта НИКА поток данных имеет следующие параметры:

- высокая скорость набора событий (до 6 КГц),
- в центральном столкновении Au-Au при энергиях НИКА образуется до 1000 заряженных частиц,
- Прогнозируемое количество событий 19 миллиардов;
- общий объем исходных данных может быть оценен в 30 PB ежегодно, или 8.4 PB после обработки.

Моделирование распределенной компьютерной инфраструктуры



Создана модель для изучения процессов:
✓Tape robot,
✓Disk array,
✓CPU Cluster.



Проект создания центра уровня Tier1 в России

С 2011 года идут подготовительные работы по созданию центра уровня Tier1 в России для обработки, хранения и анализа данных с Большого адронного коллайдера (БАК) на базе НИЦ КИ и ОИЯИ. В настоящее время в рамках ФЦП Министерства образования и науки РФ финансируется проект создания прототипа этого комплекса.

28 сентября 2012 года на заседании Наблюдательного Совета проекта WLCG (глобальная GRID-инфраструктура для БАК) был принят план создания Tier1 в России.

В этом плане 3 основных этапа:

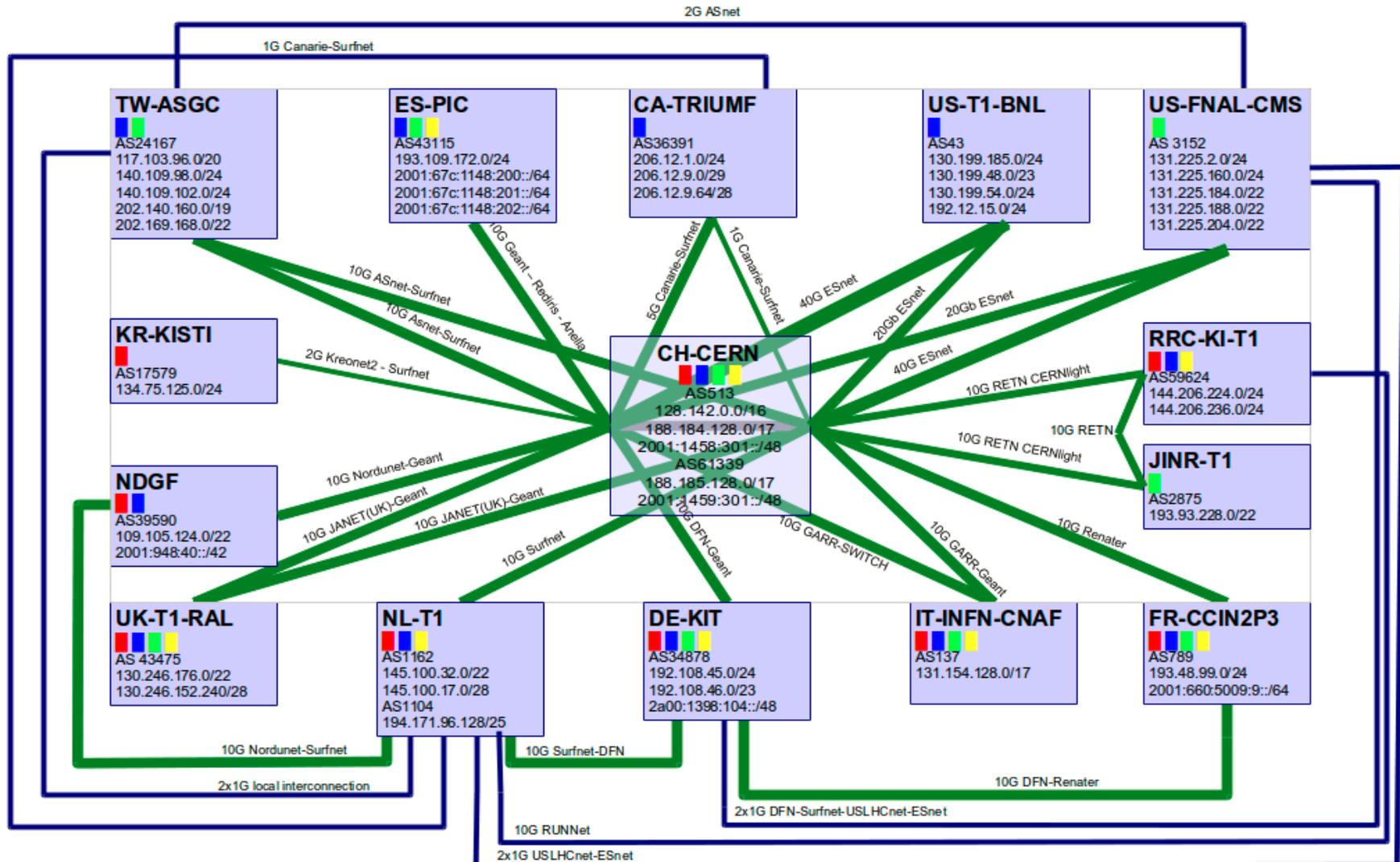
Первый этап (декабрь 2012 года-закончен) - создание прототипа Tier1 в НИЦ КИ и ОИЯИ.

Второй этап (ноябрь 2013 года) – установка оборудования для базового Tier1 центра, его тестирование и доведение до необходимых функциональных характеристик.

Третий этап (март 2015 года) - дооснащение этого комплекса и ввод в эксплуатацию полномасштабного Tier1 центра в России.

В дальнейшем потребуется планомерное повышение вычислительной мощности и систем хранения данных в соответствии с требованиями экспериментов.

LHCOPN

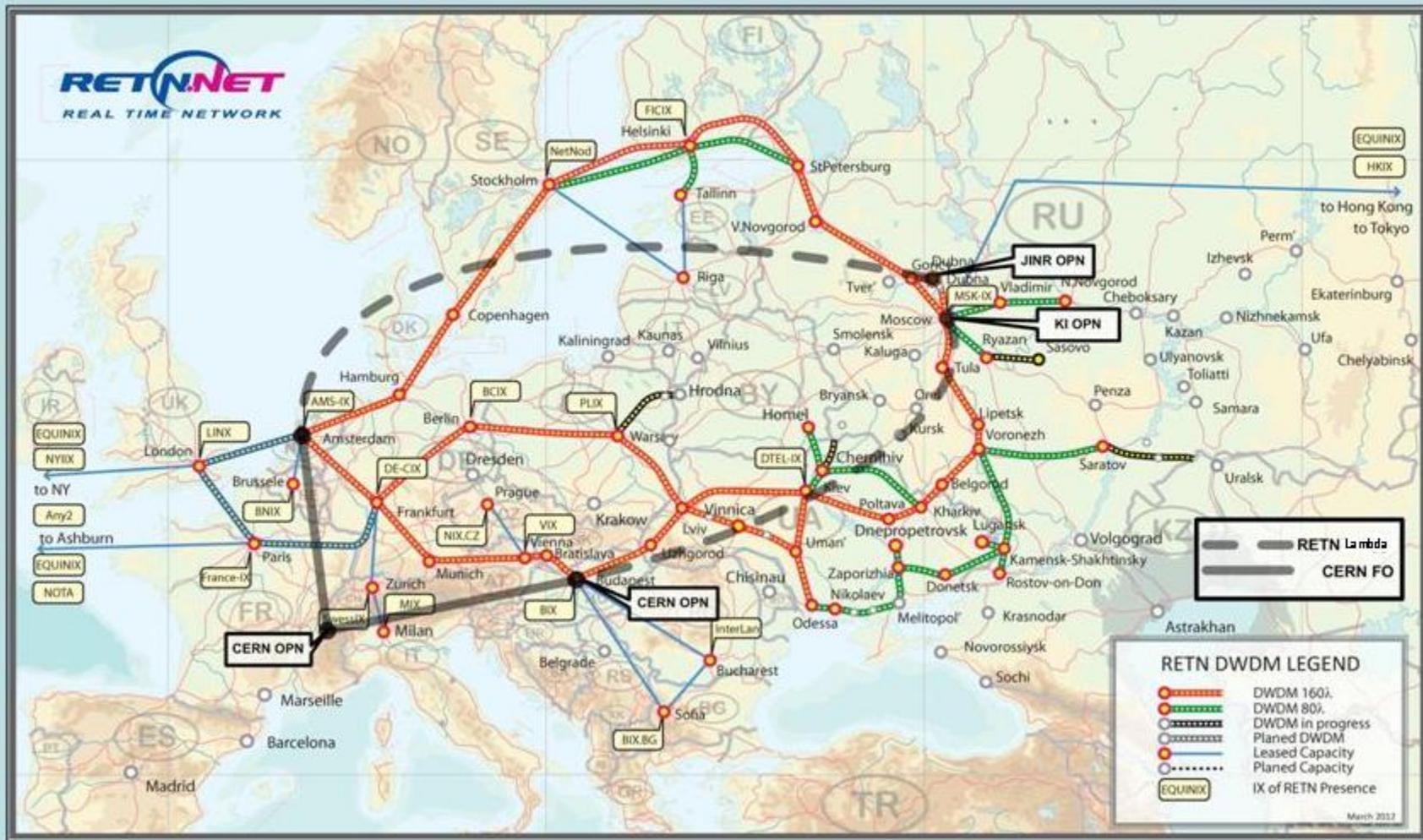


— T0-T1 and T1-T1 traffic
— T1-T1 traffic only
- - - Not deployed yet
— (thick) >=10Gbps
— (thin) <10Gbps

■ = Alice ■ = Atlas
■ = CMS ■ = LHCb

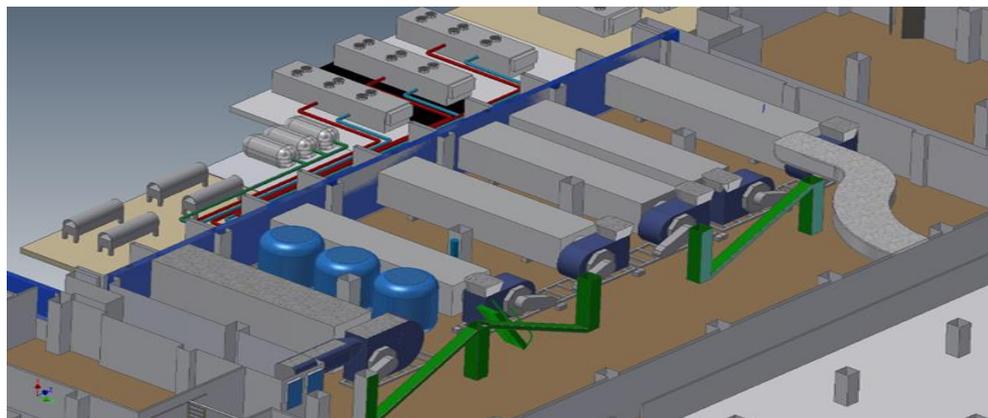
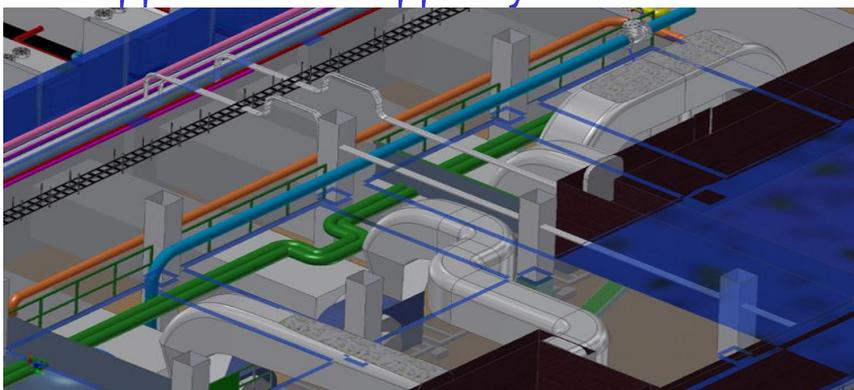
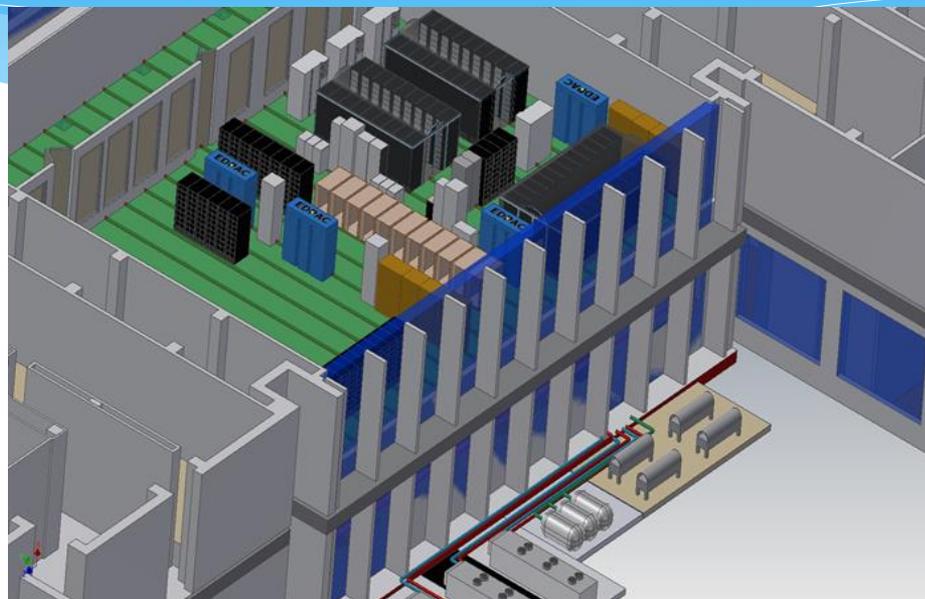
p2p prefix: 192.16.166.0/24 - 2001:1458:302::/48
 edoardo.martelli@cern.ch 20141212

JINR Tier1 Connectivity Scheme

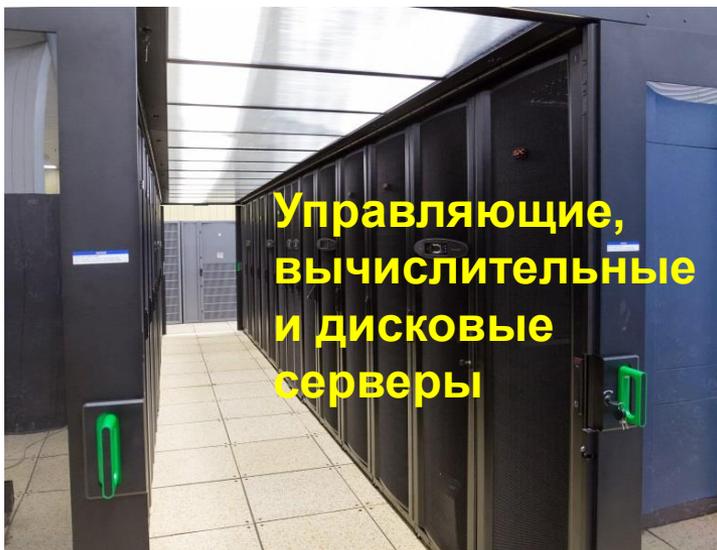


Создание CMS Tier1 в ОИЯИ

- Инженерная инфраструктура (система бесперебойного электропитания, климат-контроля);
- Высокоскоростная надежная сетевая инфраструктура с выделенным резервируемым каналом в ЦЕРН (LHCOPN);
- Вычислительная система и система хранения на базе дисковых массивов и ленточных библиотек большой емкости;
- Надежность и доступность 100%



Компоненты центра Tier1 в ЛИТ ОИЯИ



2015 Март

2400 ядер (~ 30 кНС06)

5 PB ленты (IBM TS3500)

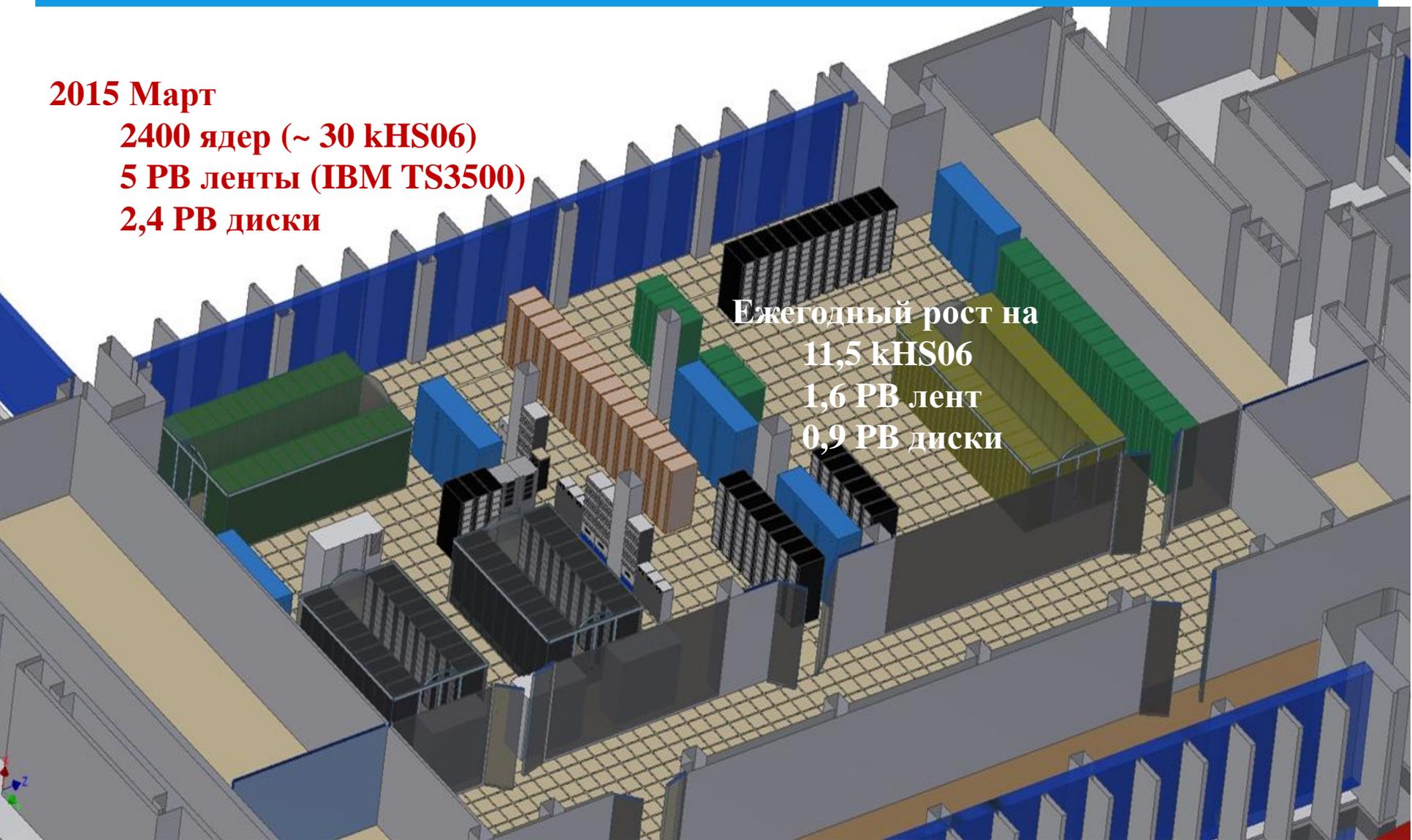
2,4 PB диски

Ежегодный рост на

11,5 кНС06

1,6 PB лент

0,9 PB диски





Russia:
NRC KI



US-BNL



Amsterdam/NIKHEF-SARA



Taipei/ASGC



Bologna/CNAF



Ca-
TRIUMF



JINR



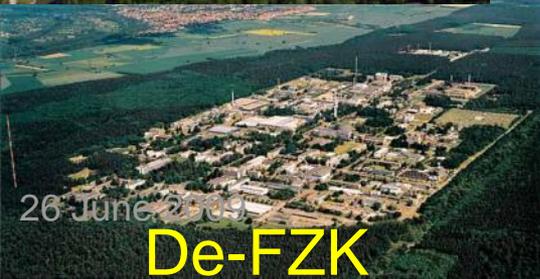
CERN



NIDGE



US-FNAL



De-FZK



Barcelona/PIC



Lyon/CCIN2P3



UK-RAL

Проекты в области распределенных вычислений

- **Worldwide LHC Computing Grid (WLCG)**
- **EGI-InSPIRE**
- **RDIG Development**
- **Project BNL, ANL, UTA “Next Generation Workload Management and Analysis System for BigData”**
- **Tier1 Center in Russia (NRC KI, LIT JINR)**
- **6 Projects at CERN**
- **CERN-RFBR project “Global data transfer monitoring system for WLCG infrastructure”**
- **BMBF grant “Development of the grid-infrastructure and tools to provide joint investigations performed with participation of JINR and German research centers”**
- **“Development of grid segment for the LHC experiments” was supported in frames of JINR-South Africa cooperation agreement;**
- **Development of grid segment at Cairo University and its integration to the JINR GridEdu infrastructure**
- **JINR - FZU AS Czech Republic Project “The grid for the physics experiments”**
- **NASU-RFBR project “Development and implementation of cloud computing technologies on grid-sites of Tier-2 level at LIT JINR and Bogolyubov Institute for Theoretical Physics for data processing from ALICE experiment”**
- **JINR-Romania cooperation Hulubei-Meshcheryakov programme**
- **JINR-Moldova cooperation (MD-GRID, RENAM)**
- **JINR-Mongolia cooperation (Mongol-Grid)**
- **JINR-China cooperation (BES-III)**

Значение Больших данных

Ключевым фактором успеха для многих отраслей экономики становится возможность эффективно обрабатывать огромные массивы и потоки информации. Поиск оптимальных способов обработки больших объемов данных приобретает важнейшее значение для развития перспективных направлений науки и индустрии, таких как биоинформатика, энергетика нового поколения, экономическое моделирование, социология и др. Сегодня уже очевидно, что Большие Данные — одно из ключевых направлений компьютерной науки, открывающее новые перспективы для исследований в сфере представления, анализа и извлечения полезных знаний из больших объемов данных различной природы.

Что такое «Большие данные»

Простое определение: Большие Данные те, что слишком велики и сложны, чтобы их можно было эффективно запомнить, передать и проанализировать **стандартными средствами доступных баз данных и иных имеющихся систем хранения, передачи и обработки.**

Кроме объема, следует учитывать и другие их характеристики. Еще в 2001 году Мета Групп ввела в качестве определяющих характеристик для больших данных так называемые **«три V»:**

- * **объём** (*volume*, в смысле величины физического объёма),
- * **скорость** (*velocity* в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов),
- * **многообразие** (*variety*, в смысле возможности одновременной обработки различных типов структурированных и неструктурированных данных)

Однако, когда **общий поток данных растет экспоненциально, удваиваясь каждый год**, за счет революционных технологических изменений, к 2014 году даже эту "3V" модель предлагают расширить, добавляя новые и новые «V»,

- **Validity** (обоснованность, применимость),
- **Veracity** (достоверность),
- **Value** (ценность, полезность),
- **Visibility** (обозримость, способность к визуализации) и т. д.

Какая польза от Больших данных?



Большие должны быть доступны для поисковых систем, проанализированы в центрах бизнеса, производства, медицины, науки и просто индивидуумов, которые их могут затребовать.

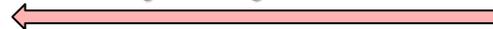
Что это может дать? – Упростить документооборот, но главное - прогноз!

Пример с вирусом H1N1: Google в 2009 г. точно предсказал распространение гриппа в разных штатах США, проанализировав 3 млрд запросов на лекарства от гриппа на их корреляцию с распространением гриппа во времени и пространстве.

Большие данные помогут решению насущных глобальных проблем, таких как борьба с изменением климата, искоренение болезней, а также содействие эффективному управлению и экономическому развитию. Однако для этого придется пересмотреть традиционные представления об управлении, принятии решений, человеческих ресурсах и образовании.

Как это делается?

читай эту книгу!



<http://www.livelib.ru/book/1000755419>

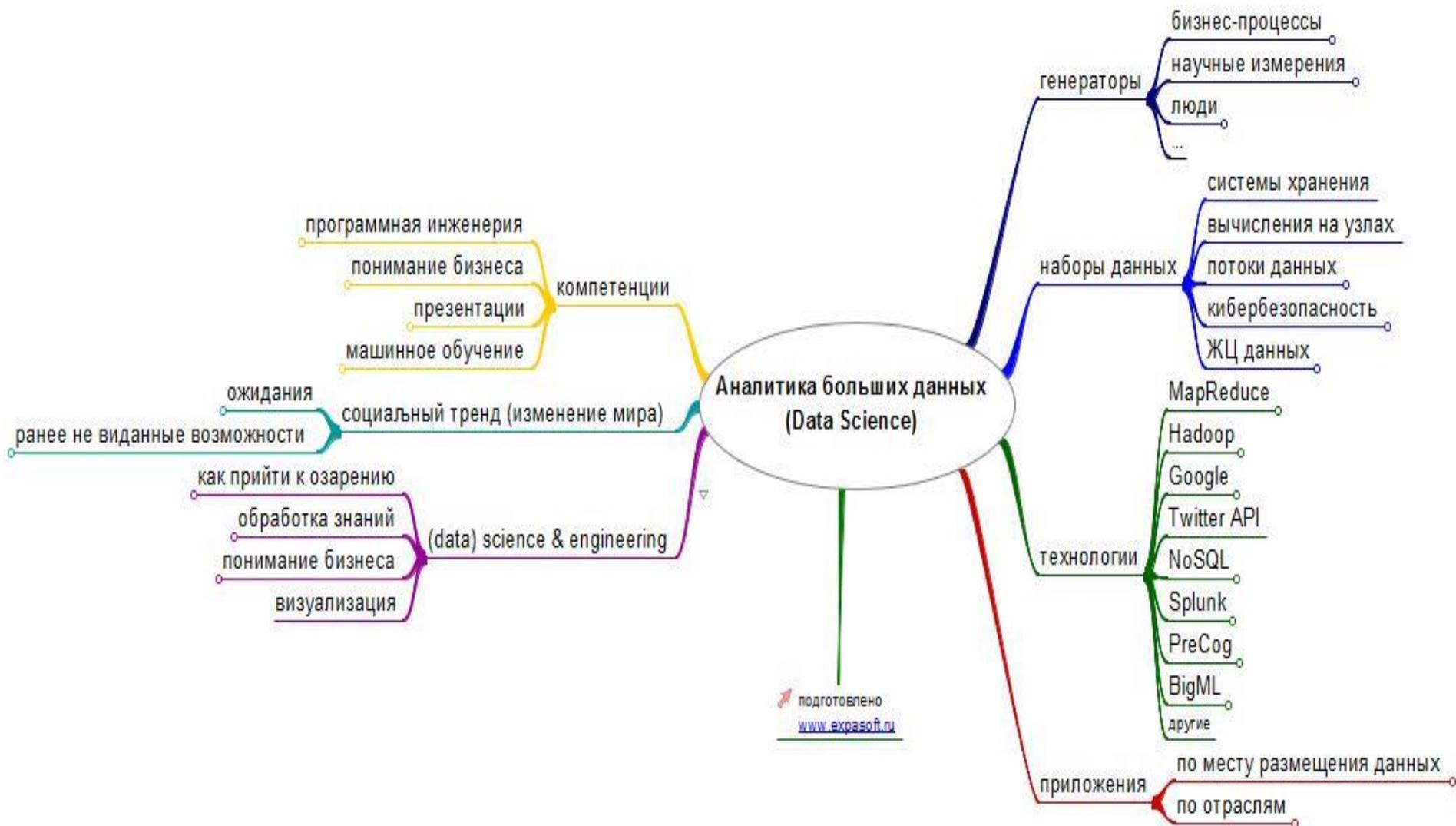
- Используется не отобранная малая часть данных (выборка), а ВСЁ доступное множество.
- Ради этого приходится отказываться от деталей (пренебрегать точностью) и исследовать не причинную связь (почему происходит явление), а корреляционные связи, объясняющие, что именно происходит (полезно, опасно).

Большие данные – горячая тема, потому что технологии сделали возможным анализ ВСЕХ доступных данных

Эффективно с точки зрения затрат управлять и анализировать **ВСЕ** доступные данные, в их первоначальном виде – структурированные, неструктурированные, потоковые



Первый взгляд на большие данные



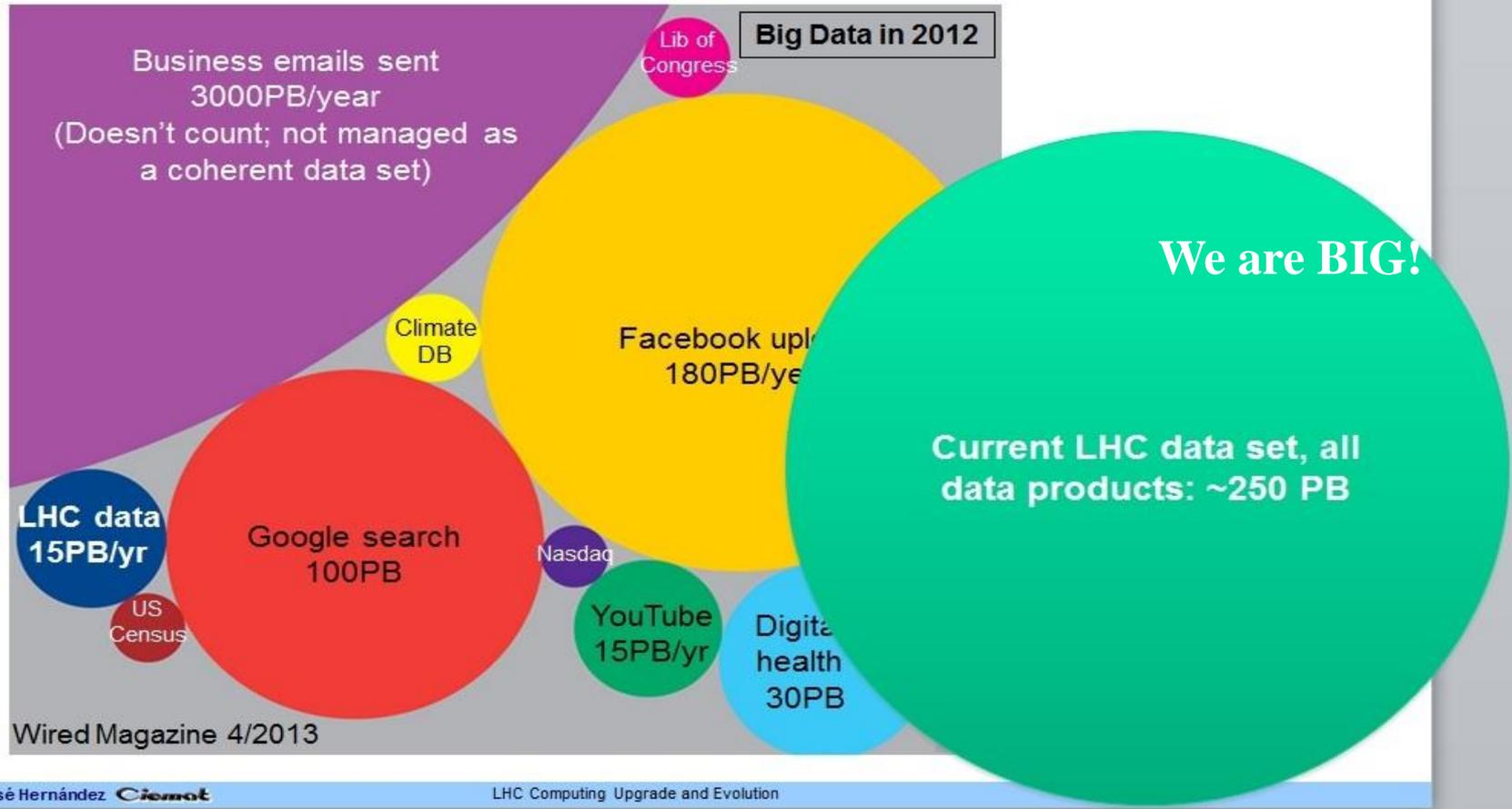
Что может аналитика Больших данных?

Современные технологии делают возможным обработку и анализ огромного количества данных, в некоторых случаях — всех данных, касающиеся того или иного явления, и позволяют достичь качественно нового уровня обработки и анализа данных, принципиально невозможного без применения данных технологий.

Во всем мире ведутся исследования в области аналитики Больших данных — развиваются методы анализа и предсказательные модели (математическая статистика, анализ временных рядов, кластерный анализ, корреляционный анализ, регрессионный анализ, нейронные сети, генетические алгоритмы, нечеткие алгоритмы, распознавание образов и др.). С развитием интереса к Большим данным у бизнеса, государства, науки все больше усилий предпринимается во всем мире по разработке новых высокоэффективных методов.

Вступление в эру Big Data

Where is LHC in Big Data Terms?



Сравнительная диаграмма обрабатываемых данных наглядно показывает, что исследования в ЦЕРНе идут в **условиях Больших Данных**.

После модернизации и запуска ЛНС в 2015 году поток данных возрастет в 2.5 раза, что потребует увеличения ресурсов и оптимизации их использования.



PanDa в эксперименте ATLAS



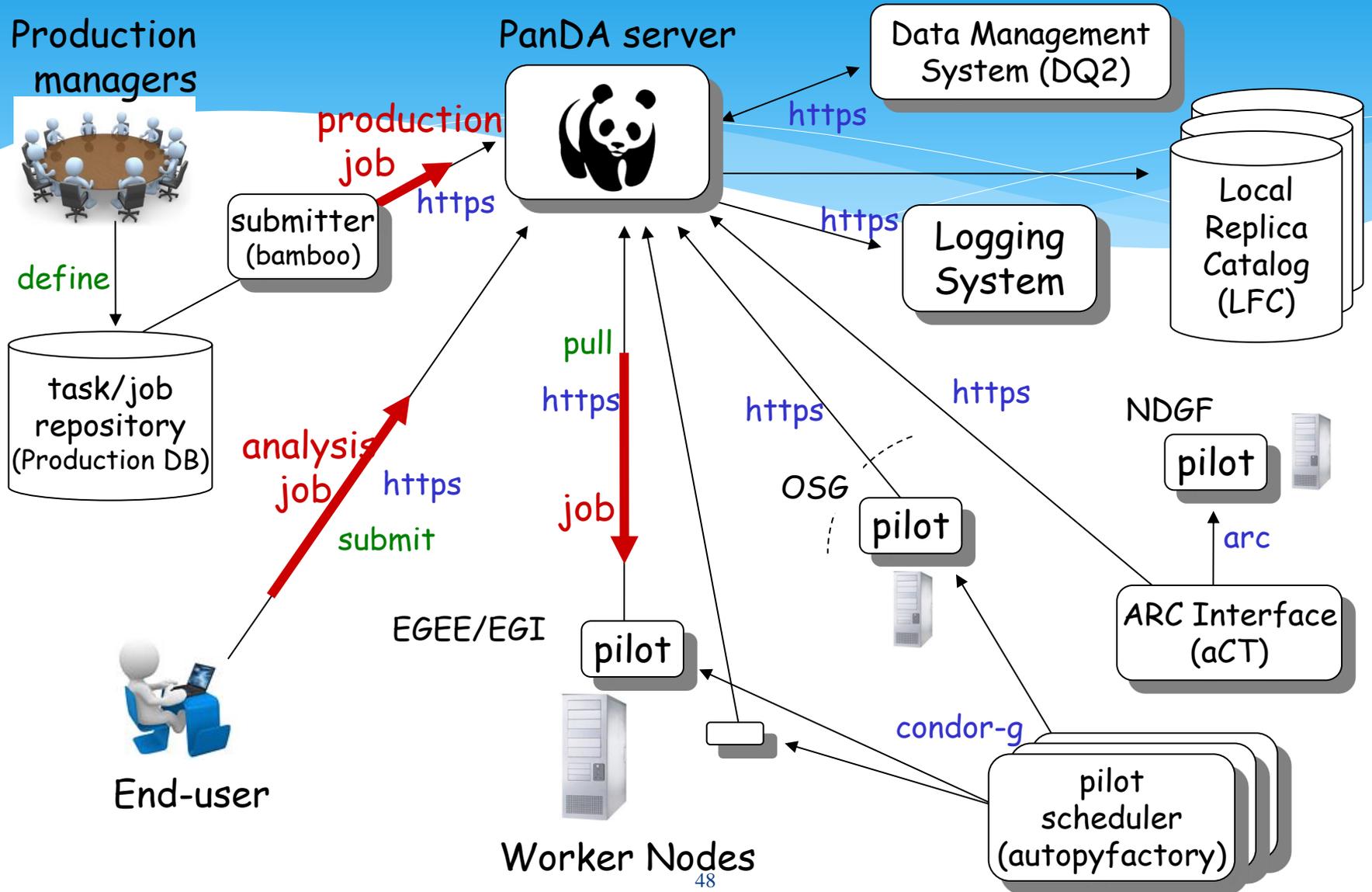
В эксперименте ATLAS на Большом адронном коллайдере разработана платформа для управления вычислительными ресурсами PanDA Workload Management System (WMS), которая обладает следующими возможностями:

- Проект PanDA начался в 2005 году группами BNL и UTA - **Production and Data Analysis system**.
- Автоматизированная и гибкая система управления заданиями, которая может оптимально сделать распределенные ресурсы доступными пользователю.
- С помощью PanDA, физики видят единый вычислительный ресурс, который предназначен для обработки данных эксперимента, даже если дата-центры разбросаны по всему миру
- PanDa изолирует физиков от аппаратного обеспечения, системного и промежуточного программного обеспечения и других технологических сложностей, связанных с конфигурированием сети и оборудования.
- Вычислительные задачи автоматически отслеживаются и выполняются. Могут выполняться групповые задачи физиков

В настоящее время PanDa контролирует:

- **сотни дата - центров в 50 странах мира**
- **сотни тысяч вычислительных узлов**
- **сотни миллионов заданий в год**
- **тысячи пользователей**

PanDA Workload Management



Интеграция технологий распределенных вычислений для управления большими данными

PanDa, как платформа, обеспечивающая прозрачность процесса хранения, обработки и управления данными для приложений с большими потоками данных и массивными вычислениями.

Развитие PanDA в направлении интеграции различных систем распределенных и параллельных вычислений (грид, cloud, кластеры, ЦОД, суперкомпьютеры) с целью создания универсальной платформы для крупных проектов управления большими данными в науке, государственном управлении, медицине, высокотехнологической промышленности, бизнесе.

За большими данными следит ПАНДА

Топик
А. Климентов, сотрудник Бружневской национальной лаборатории,
А. Ваняшин, сотрудник Аргонской национальной лаборатории,
Ф. Жоренко, ОИЯИ, г. Дубна
Иллюстрация Владимир Камышев



Броский и звучный термин «Big Data» прочно занял свое место в лексиконе не только ученых и компьютерных профессионалов, но и политиков разного ранга.

Суперкомпьютеры №15, 2013, стр.56

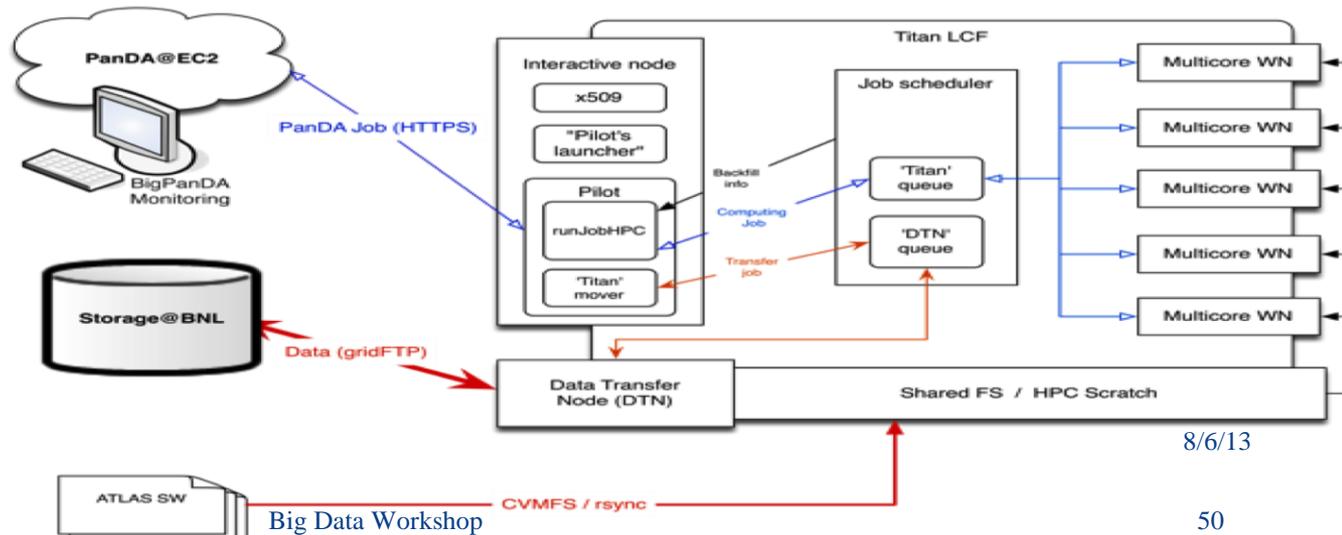
Пример использования системы PanDA (суперкомпьютер Titan в Oak Ridge)



Titan System (Cray XK7)			
Peak Performance	27.1 PF 18,688 compute nodes	24.5 PF GPU	2.6 PF CPU
System memory	710 TB total memory		
Interconnect	Gemini High Speed Interconnect	3D Torus	
Storage	Lustre Filesystem	32 PB	
Archive	High-Performance Storage System (HPSS)	29 PB	
I/O Nodes	512 Service and I/O nodes		

Конфигурация суперкомпьютера Titan и схема управления заданиями с помощью платформы PanDA

12 OLCF | 20



8/6/13

Ресурсы, доступные с помощью платформы PanDA



Many Others



OLCF



Titan System (Cray XK7)			
Peak Performance	27.1 PF 18,688 compute nodes	24.5 PF GPU	2.6 PF CPU
System memory	710 TB total memory		
Interconnect	Gemini High Speed Interconnect	3D Torus	
Storage	Lustre Filesystem	32 PB	
Archive	High-Performance Storage System (HPSS)	29 PB	
I/O Nodes	512 Service and I/O nodes		



Google Cloud Platform



Проекты с суперкомпьютерными центрами НИЦ КИ, ННГУ, Острова (Чехия)
Центры Больших данных в НИЦ КИ, ЛИТ ОИЯИ, РЭУ им. Плеханова

LIT traditional conferences



**Distributed Computing and Grid-technologies
in Science and Education**



Mathematics. Computing. Education

LIT schools



JINR / CERN

GRID AND ADVANCED INFORMATION SYSTEMS

SCHOOL ON NUCLEAR ELECTRONICS &
COMPUTING BASED ON

[XXIV INTERNATIONAL SYMPOSIUM ON NUCLEAR
ELECTRONICS & COMPUTING](#)



MPAMCS 2012

International Conference-School for Young Scientists
"Modern Problems of Applied Mathematics &
Computer Science"

August 22 - 27, 2012, Dubna, Russia

INFORMATION



**Mathematical Modeling and
Computational Physics**



NEC'2013



XXIV International Symposium on Nuclear Electronics & Computing



**DIGITAL LIBRARIES:
ADVANCED METHODS AND TECHNOLOGIES
DIGITAL COLLECTIONS**

Planned Schools In 2015

AIS-GRID-2015

Oktober

*In LIT holds regular
tutorial courses and
traineeship of young
scientists and students
from the JINR Member
States*

NEC'2015

XXV International Symposium
on Nuclear Electronics & Computing



- * On 28 September – 02 October, 2015, Montenegro (Budva), will host the regular JINR XXV Symposium on Nuclear Electronics and Computing - NEC'2015 and students' schools on advanced information technologies
- * <https://indico-new.jinr.ru/conferenceDisplay.py?confId=60>
- * **Chairpersons**
Vladimir Korenkov, JINR
Ian Bird, CERN

The International Conference Mathematical Modeling and Computational Physics, 2015



Stará Lesná, High Tatra Mountains, Slovakia, July 13 — July 17, 2015

<http://web.tuke.sk/mmcp/mmcp2015/index.php>

[Laboratory of Information Technologies](#), Joint Institute for Nuclear Research, Dubna, Russia, the [Institute of Experimental Physics](#), Slovak Academy of Sciences, Košice, the [Technical University of Košice](#), the [Pavol Jozef Šafárik University](#), Košice, the [Slovak Physical Society](#), Košice, Slovakia, [IFIN-HH, Bucharest](#), Romania

Conference Chairmen:

V.V. Korenkov (JINR Dubna) and G. Semanišin (UPJŠ Košice)