.

Computing infrastructure, data acquisition and data processing in HEP from experiments

Vladimir Korenkov
(JINR, University “Dubna”)

CERN, 04.11.09

Laboratory of Information Technologies

Specific tasks:

1. Provision of JINR and its Member States with high-speed telecommunication data links
2. Creation of a high-speed, reliable and protected local area network (LAN) of JINR
3. Creation and maintenance of the distributed high-performance computing infrastructure and mass storage resources
4. Provision of information, algorithmic and software support of the JINR research-and-production activity
5. Elaboration of the JINR Grid-segment and its inclusion in European and global Grid-structures
6. Methods and tools for modeling physical processes and experimental data analysis
7. Methods and numerical algorithms for modeling magnetic systems and charged particle beam transport
8. Software and computer complexes for experimental data processing
9. Numerical algorithms and software for the simulation of complex physical systems
10. Methods, algorithms and software of computer algebra
11. New generation computing tools

LIT JINR: two important projects completed

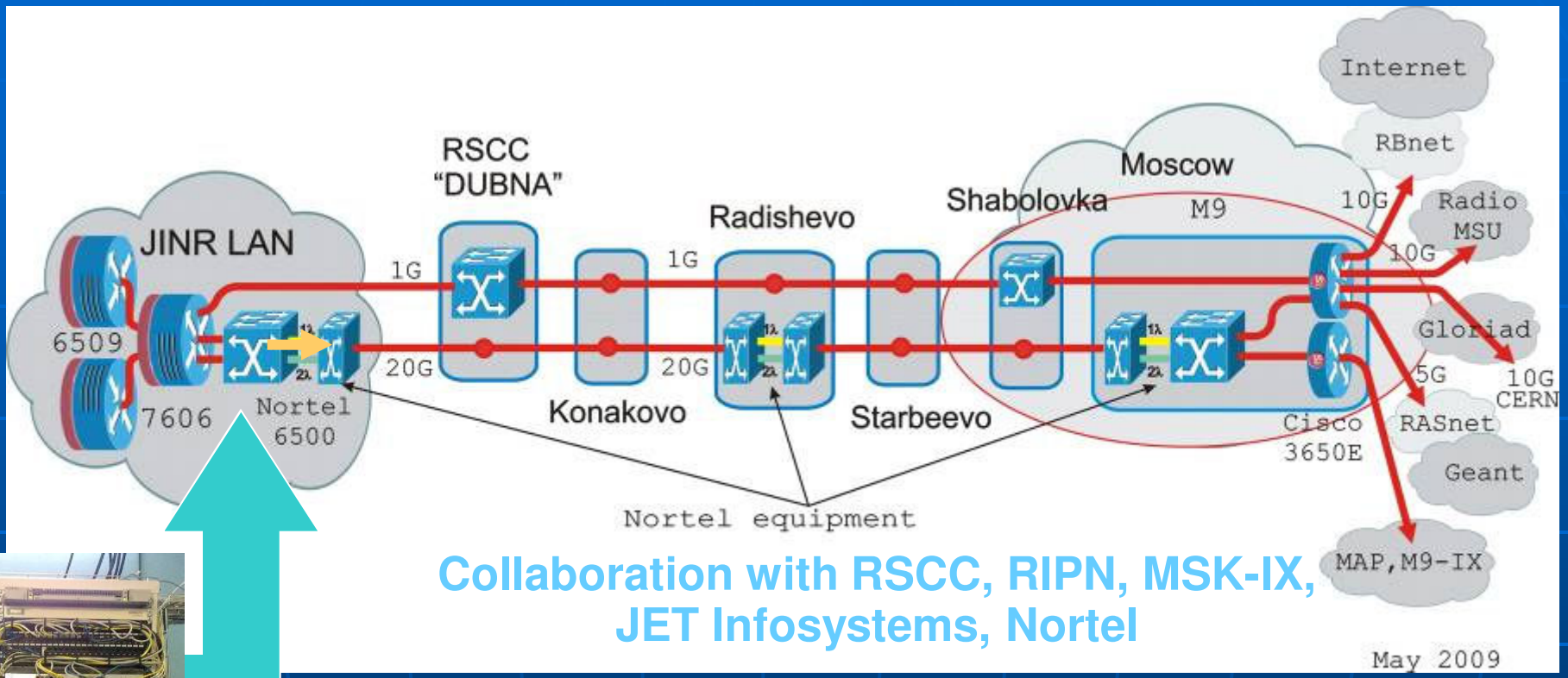
Two large-scale very important projects have been completed.

A grand presentation was held simultuneously with the meeting of the Programme Advisory Committee for Particle Physics on June 10.

All organizations involved in the projects realisation together with the representatives of Ministry of Communications, Federal Agency for Science and Innovation attended the Presentation.

Our community colleagues and prospective users - residents of the special economical zone - were also invited.

JINR - Moscow telecommunication channel



Collaboration with RSCC, RIPN, MSK-IX, JET Infosystems, Nortel

Russian Satellite Communications Company
Federal State Unitary Enterprise

RIPN Russian Institute for Public Networks

MSK IX Московский Internet Exchange

JINR Central Information and Computing Complex

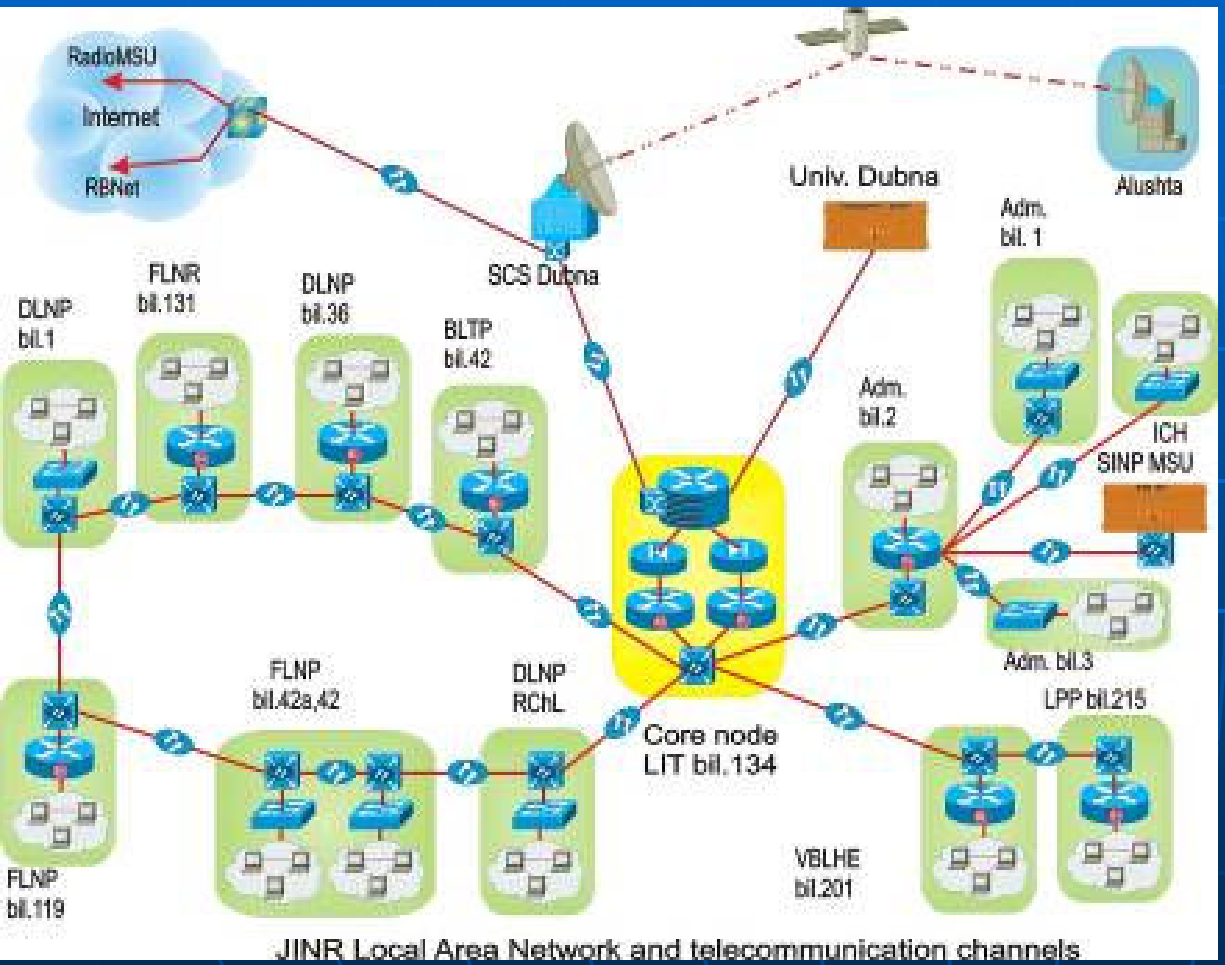
In 2008, total CICC performance was 1400 kSI2K, the disk storage capacity 100 TB

At present, the CICC performance equals 2300 kSI2K and the disk storage capacity 500 TB

PLAN

	2008	2009	2010
CPU (kSI2K)	1250	1750	2500
Disk (TB)	400	800	1200
Tapes (TB)	-	100	200

Локальная сеть ОИЯИ



JINR Local Area Network and telecommunication channels

- Comprises 7488 computers and nodes, Users - 4669
Remote VPN users – more 2000;
- High-speed transport (1Gbps);
-

Планы

Поэтапная модернизация локальной сети ОИЯИ на основе технологии 10 GEthernet

Развитие системы управления информационными потоками и механизмов информационной безопасности

Industry Journey

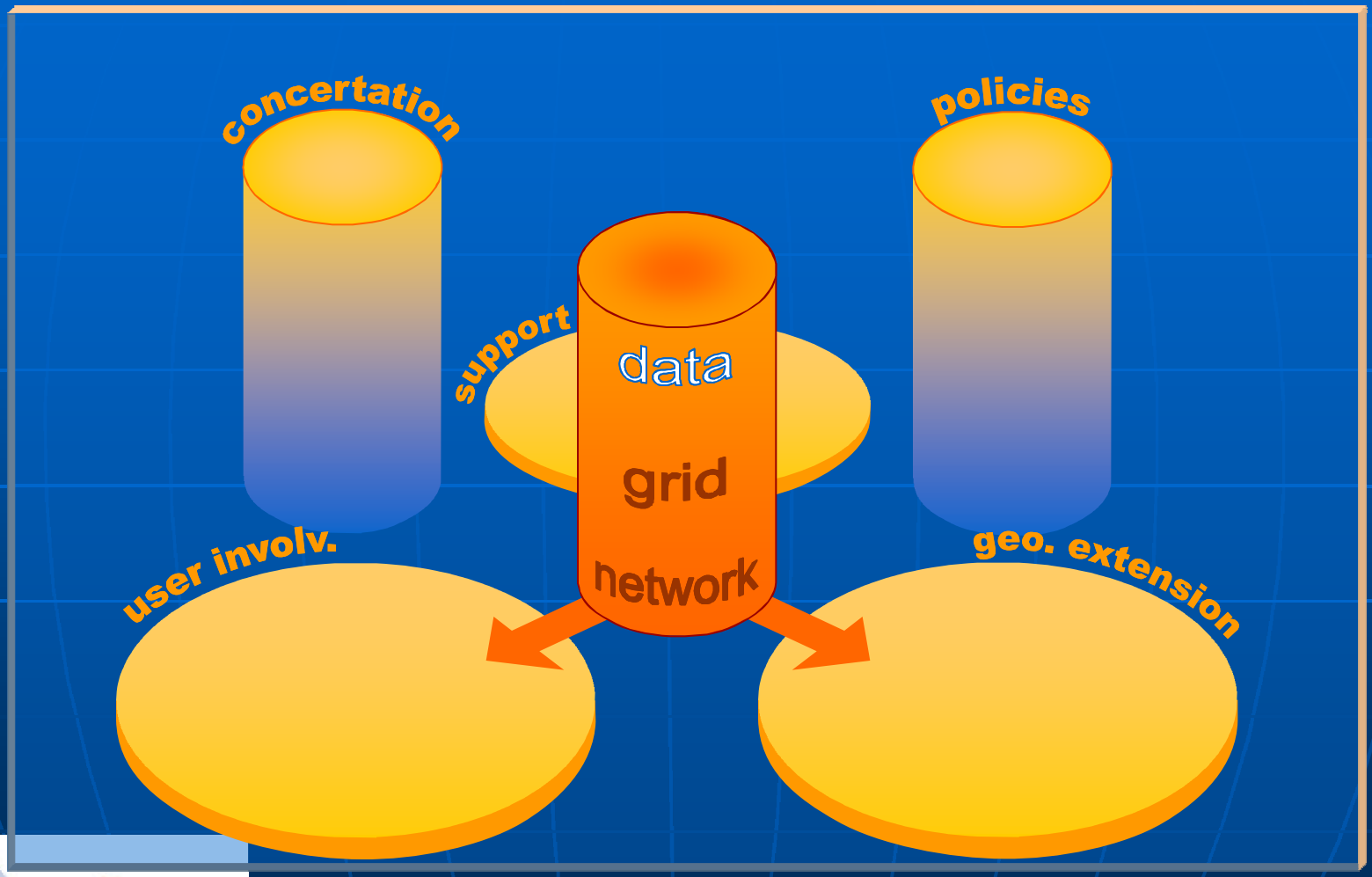
Old World

- Static
- Solo
- Physical
- Manual
- Application

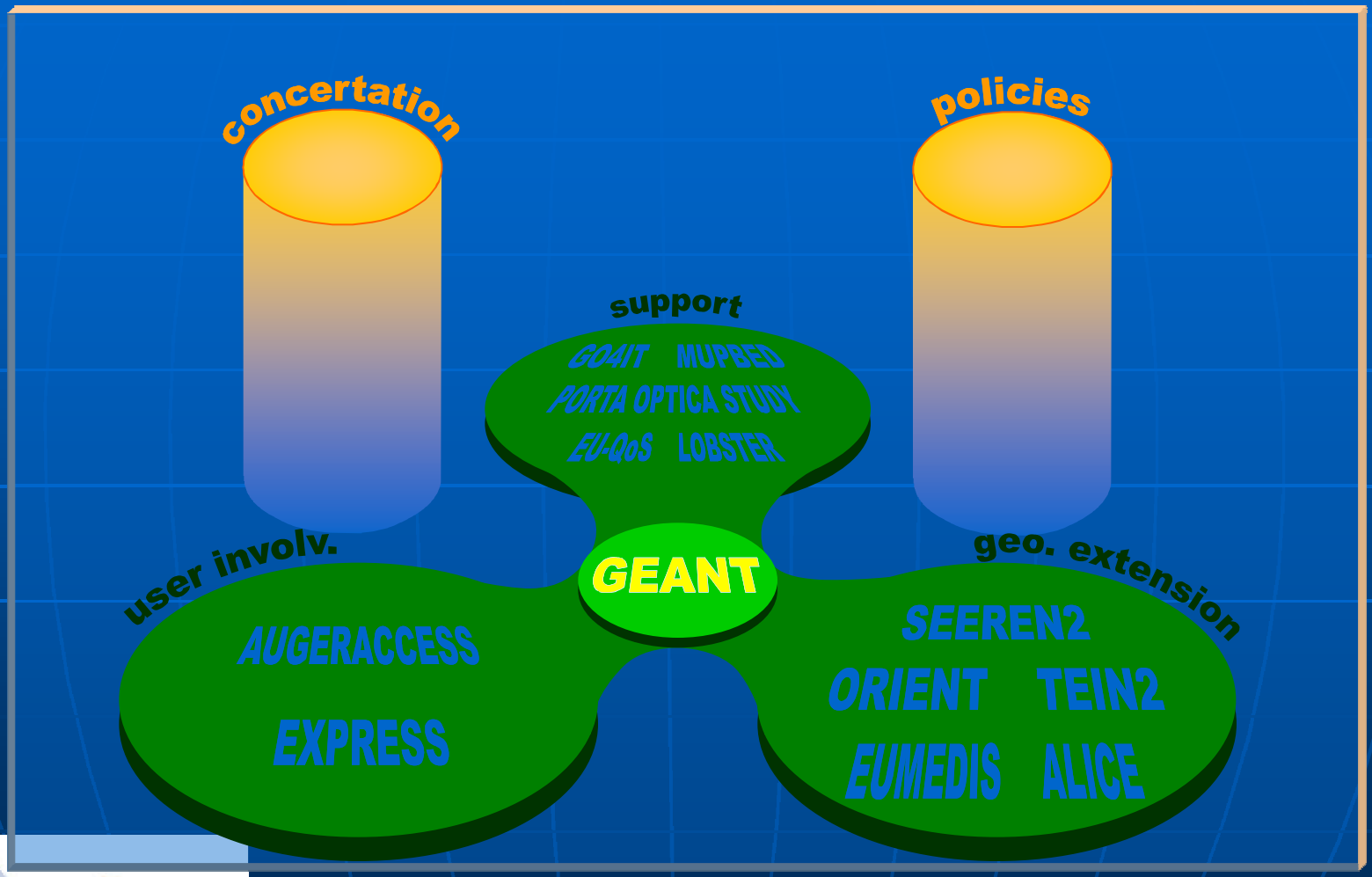
New World

- Dynamic
- Shared
- Virtual
- Automated
- Service

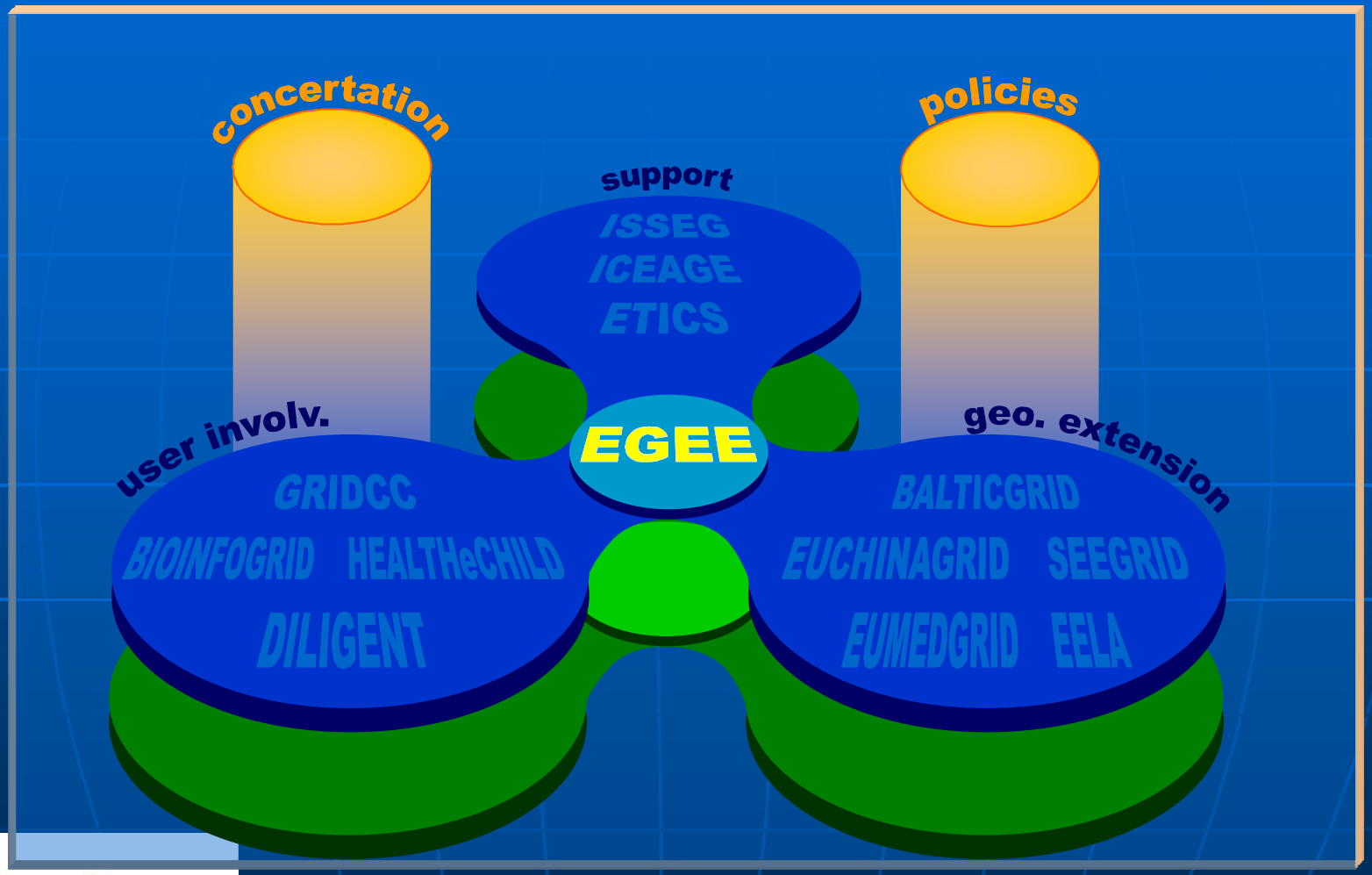
e-Infrastructures: reference model



e-Infrastructures: network layer

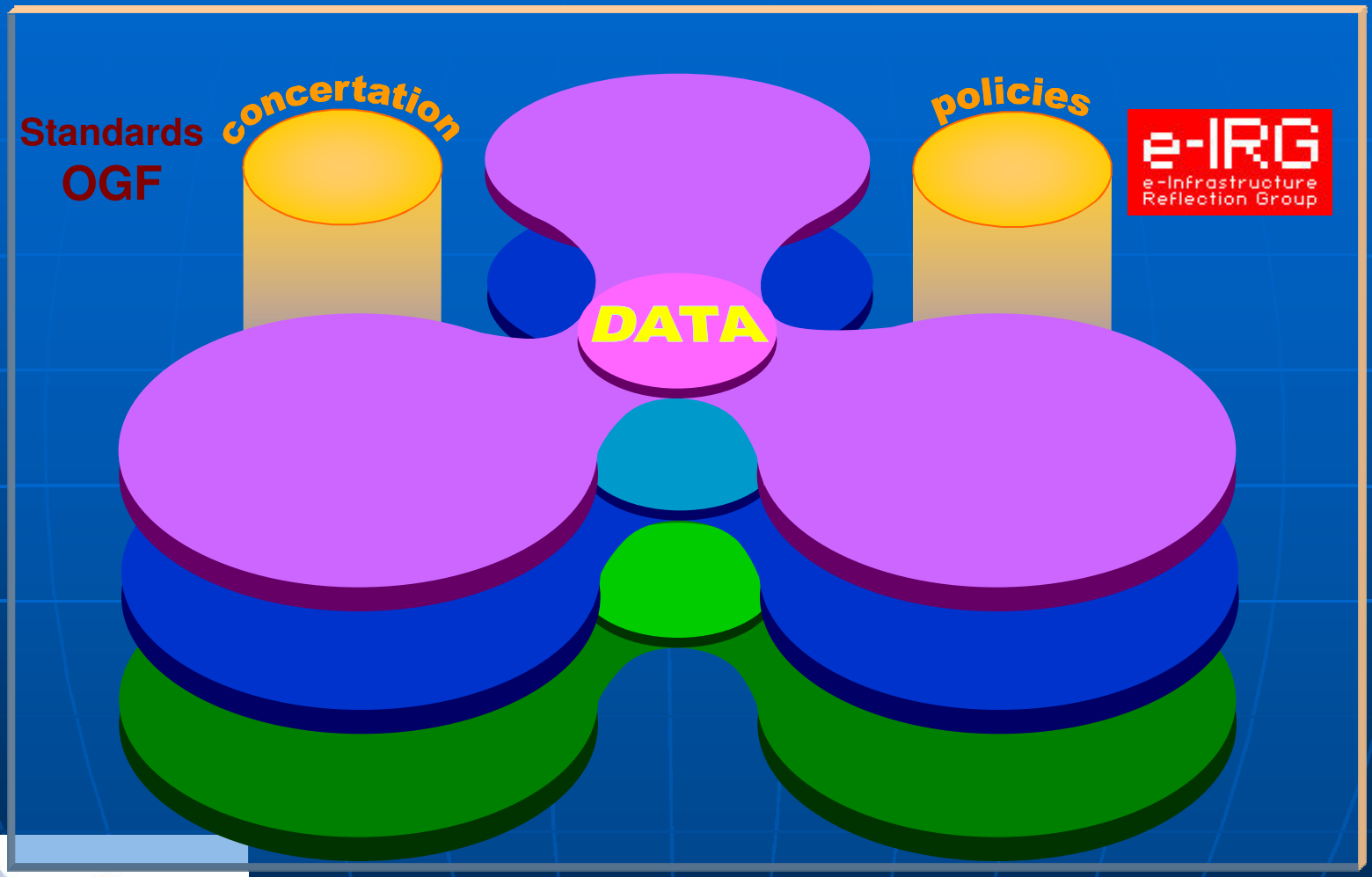


e-Infrastructures: grid layer



SEVENTH FRAMEWORK PROGRAMME

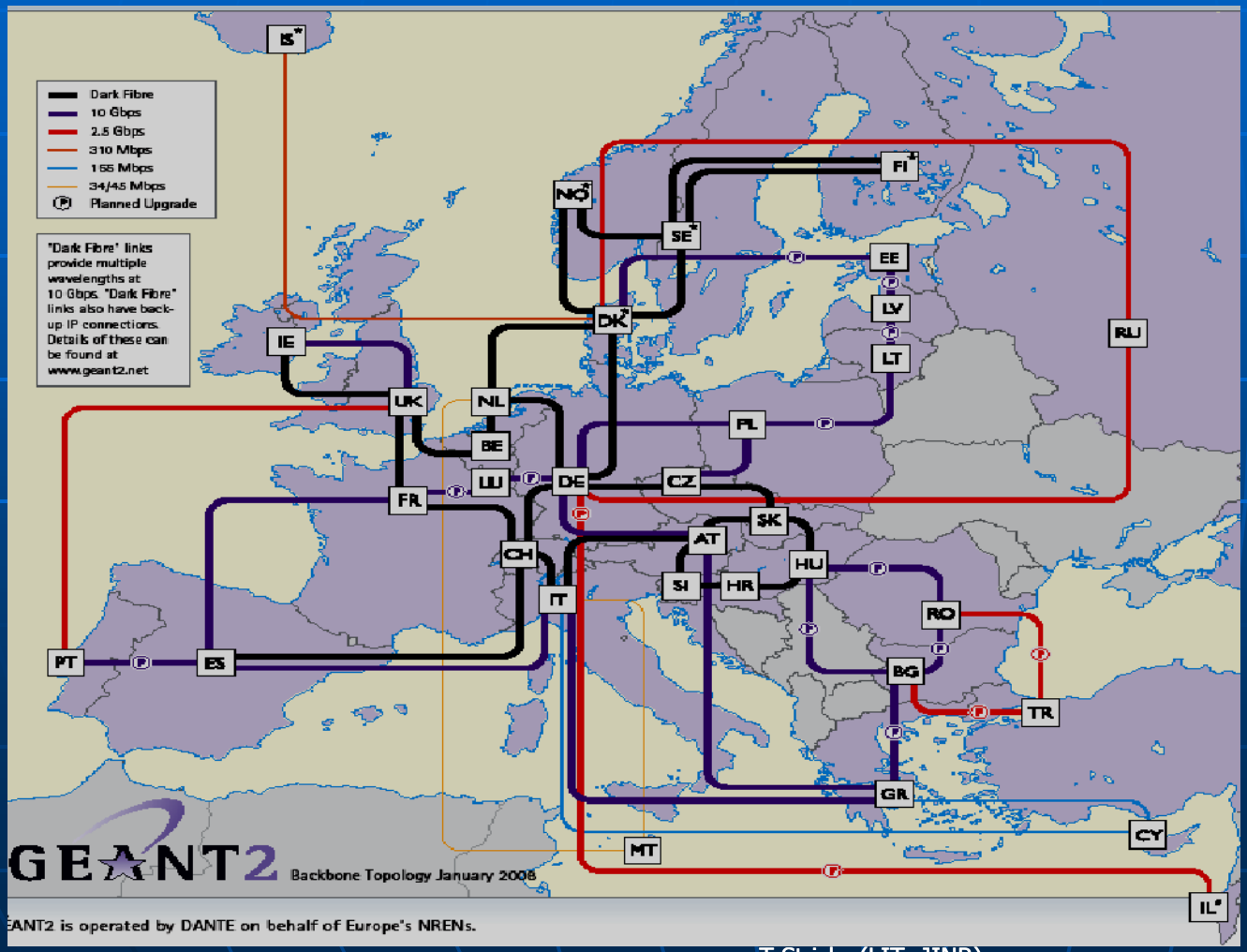
e-Infrastructures: data layer



SEVENTH FRAMEWORK
PROGRAMME

GÉANT2 Pan-European Backbone

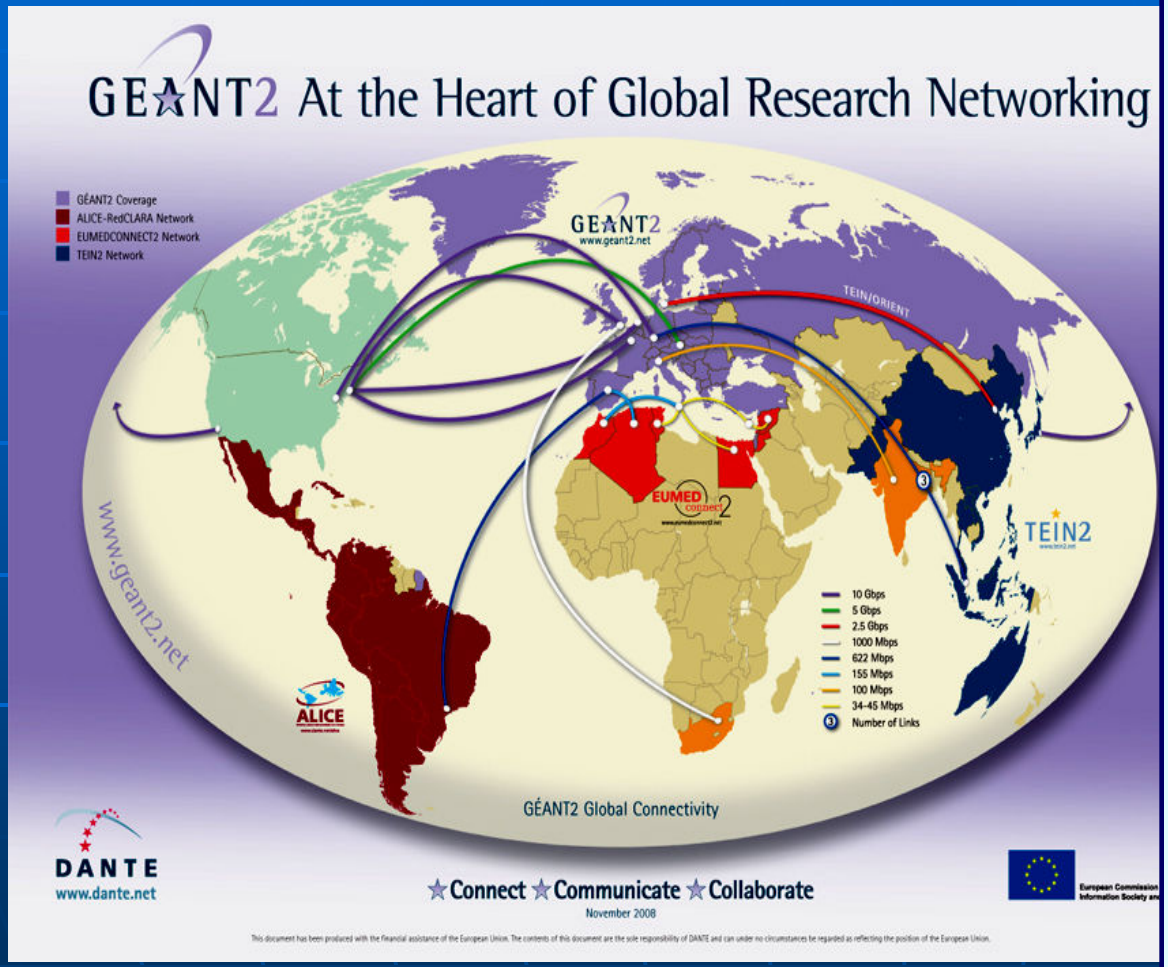
34 NRENs, ~30M Users; 50k km Leased Lines
12k km Dark Fiber; Point to Point Services
GN3 Next Gen. Network Projected Start Q2 2009



Dark Fiber Core Among 19 Countries:

- Austria
- Belgium
- Croatia
- Czech Republic
- Denmark
- Finland
- France
- Germany
- Hungary
- Ireland
- Italy
- Netherlands
- Norway
- Slovakia
- Slovenia
- Spain
- Sweden
- Switzerland
- United Kingdom

GEANT2 International Connectivity



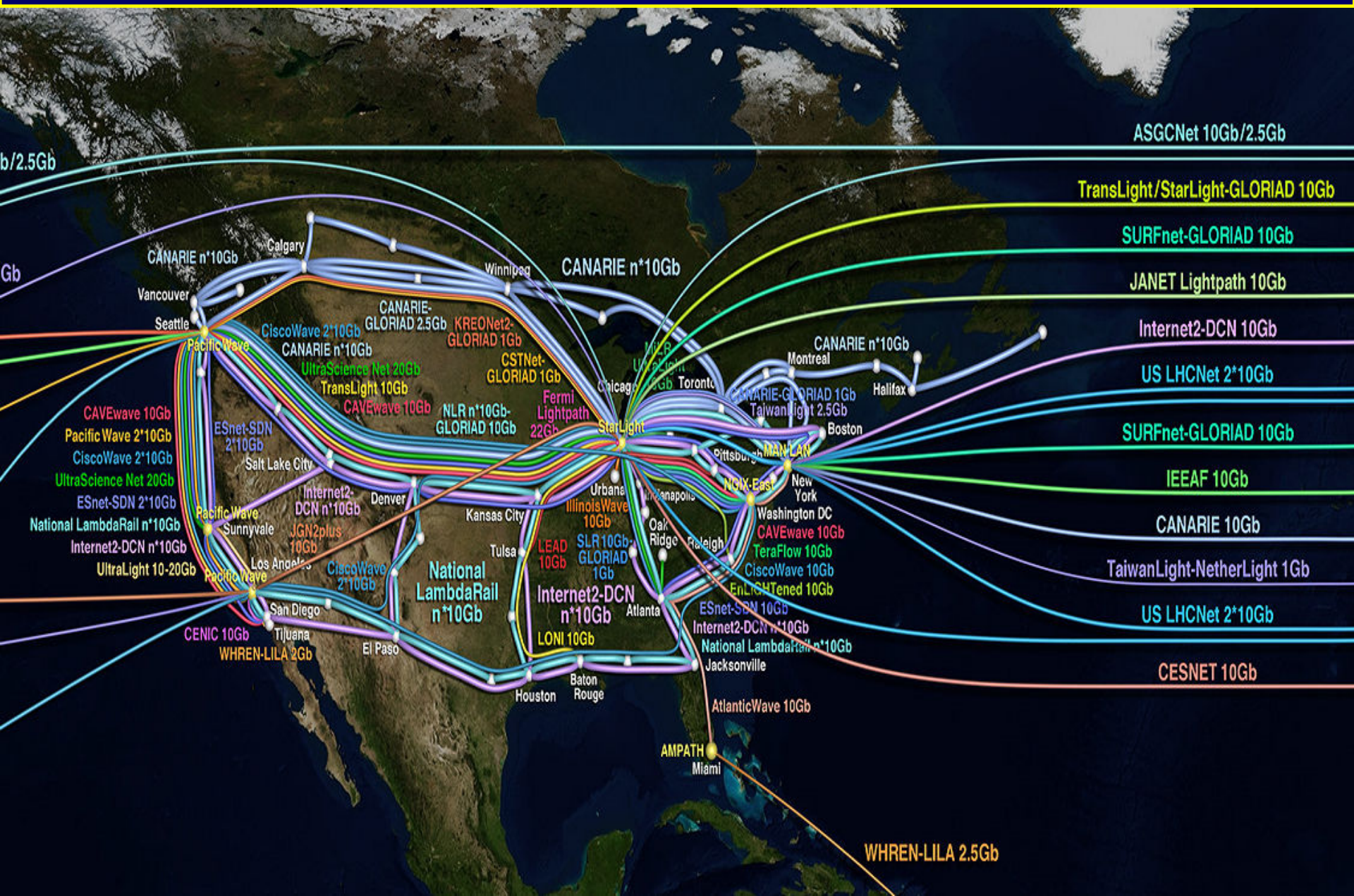
New Projects in 2008-9

- ALICE2 in and to Latin America
- CAREN for Central Asia
- FEAST: Feasibility Study for Sub-Saharan Africa

Inter-Regional Connections

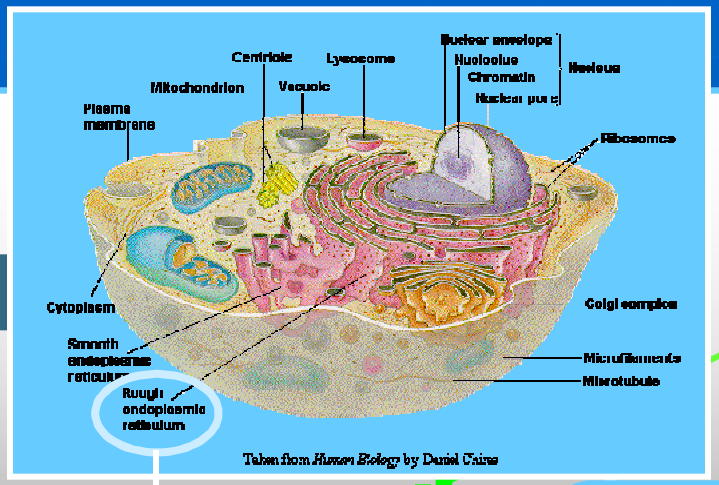
- GEANT2/ESNet over US LHCNet for US Tier2/ EU Tier1 Connections
- 4 10G Links to US R&E Nets
- 1 GbE Connection to the Ubuntunet Alliance (Africa); also TENET in South Africa
- ORIENT: CERNET + CSTNET (China) to Europe at 2.5G
- TEIN2: 9 Asian Countries + Australia to Europe; TEIN3 Approved 2008-2011
- EUMEDConnect: 11 Mediterranean + N. Africa
- ALICE: 12 Latin American countries to GEANT2 at 622 Mbps
- SuperSINET (Japan) at New York: 10G
- India: 100 Mbps

A Global Partnership of R&E Networks and Advanced R&D Projects Supporting the LHC Program



Real World Problems Taking Us

BEYOND PETASCALE



1 ZFlops
100 EFlops
10 EFlops
1 EFlops
100 PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops

1993 1999 2005 2011 2017 2023 2029

SUM
Of Top500

#1

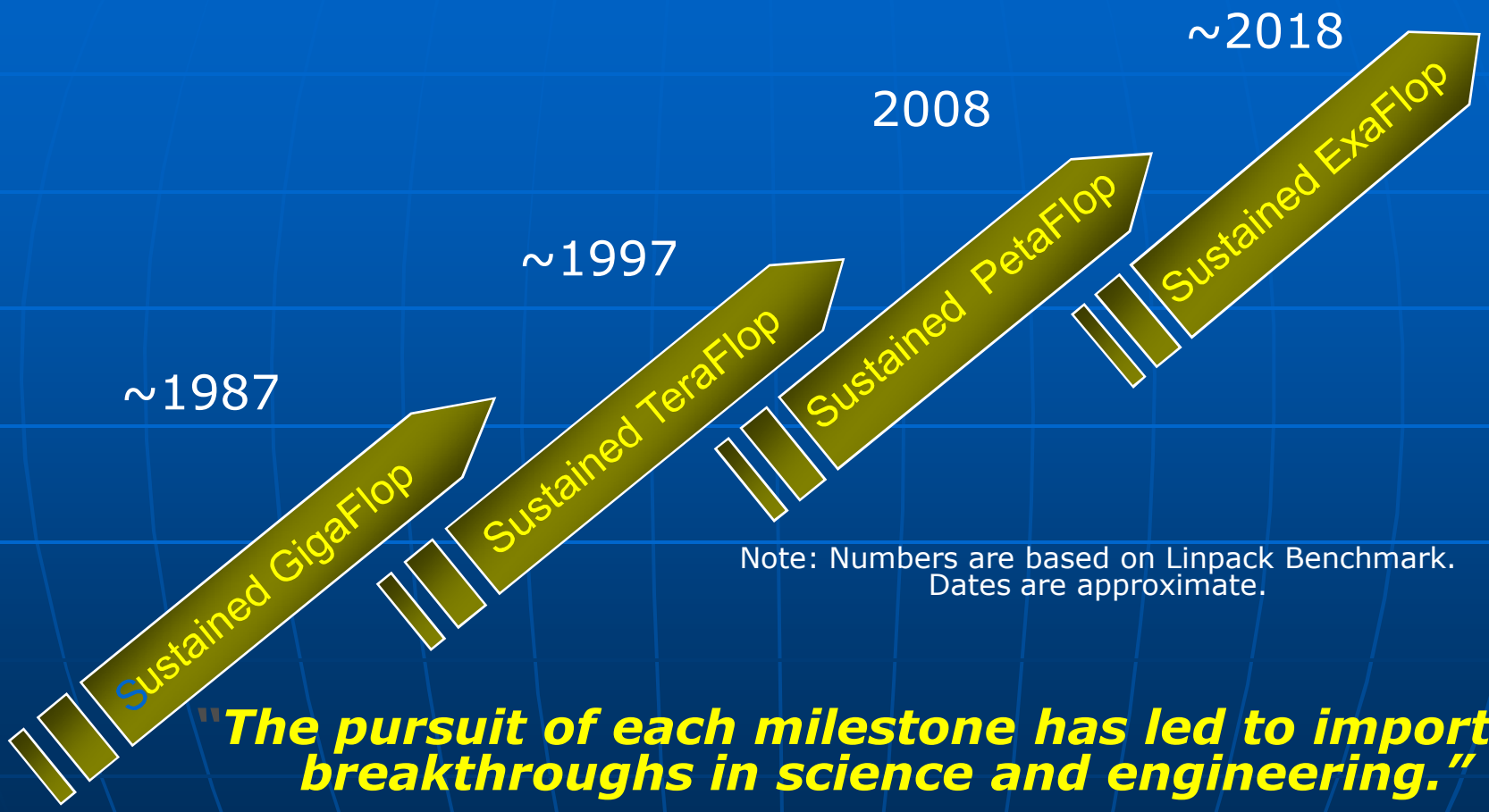
What we can just model today with <100TF

Example Real World Challenges:

- Full modeling of an aircraft in all conditions
- Green airplanes
- Genetically tailored medicine
- Understand the origin of the universe
- Synthetic fuels everywhere
- Accurate extreme weather prediction

Reach Exascale by 2018

From GigFlops to ExaFlops



Note: Numbers are based on Linpack Benchmark.
Dates are approximate.

"The pursuit of each milestone has led to important breakthroughs in science and engineering."

Source: IDC "In Pursuit of Petascale Computing: Initiatives Around the World," 2007

A look at CERN's Computing Growth

CERN Tape Library

Tape Space (PetaByte)

30,000 Cores @ 2800 SI2K per core

Computing

21,500 Cores @ 1400 SI2K per core

Disk Space (PetaByte)

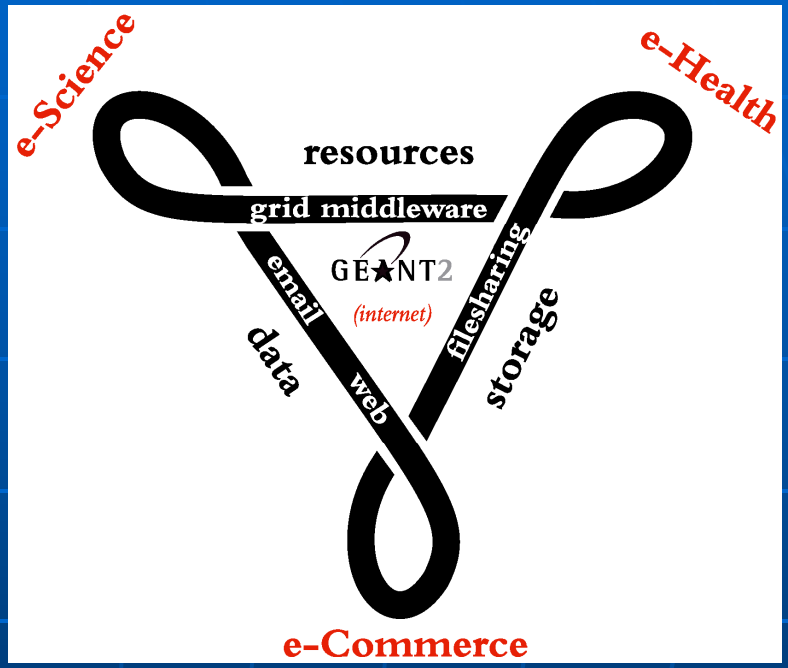
Source: CERN, Jarp Sverre

120
100
80
60
40
20
0

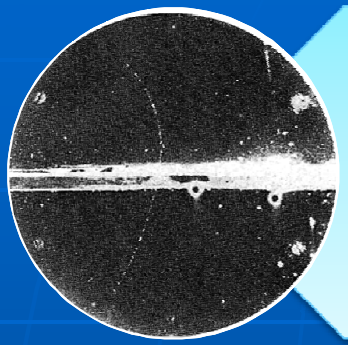
2007 2008 2009 2010 2011 2012 2013 2014 2015 2016

We do *e-Science*

- "e" like in *e*-mail:
  - digital, distributed
- Science that is:
  - Computationally intensive
  - Operates on massive digital data sets
  - Carried out in a distributed network environment
- High-**Throughput** vs High-Performance Computing:
  - HTC: distributed (serial tasks), free cycles, cheap
  - HPC: compact (parallel tasks), booked years ahead, expensive
- High-Energy Physics is a textbook example of e-science

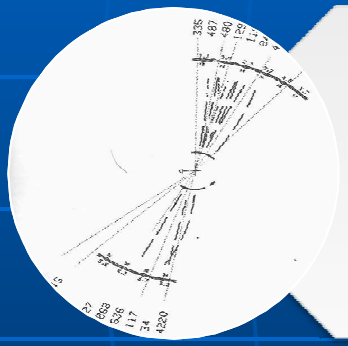


New instruments, more data, more scientists, more computers



A discovery in 1930-ies

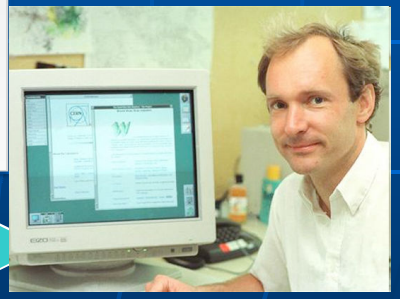
- ~2 scientists in 1country
- pen-and-paper

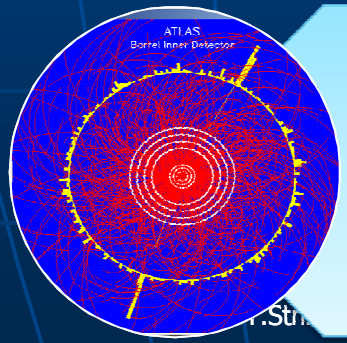


A discovery in 1970-ies

- ~200 scientists in ~10 countries
- mainframes

1989 - WWW born in CERN





A discovery tomorrow

- ~2000 scientists in ~100 countries
- **grids**

Modern HEP data processing: workflow of very different tasks

Monte Carlo

- Event generation (*Pythia*)
- Detector simulation (*Geant*)
- Hit digitization
- Reconstruction
- Analysis data preparation
- Analysis, results (*ROOT*)

Slide adapted from Ch.Collins-Tooth and J.R.Catmore

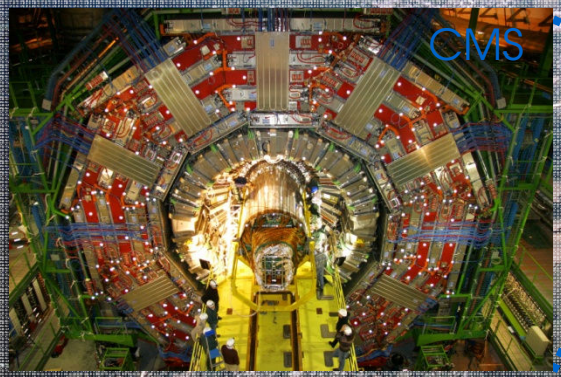
Software for HEP experiments

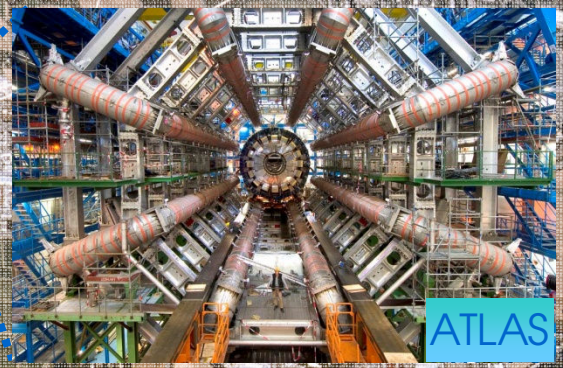
Massive pieces of software	• Written by very many different authors in different languages • Dozens of external components
Frequent releases	• Occupy as much as ~10 GB of disk space each release • Releases can come as often as once a month
Difficult to set up outside the lab	• Software is often tuned to the specifics of one lab (e.g. CERN SLC4)
Unfeasible to be maintained locally by small university teams	• Plan A: do everything centrally at the lab • Plan B: use Grid to connect large external computing "plants" managed by teams of experts

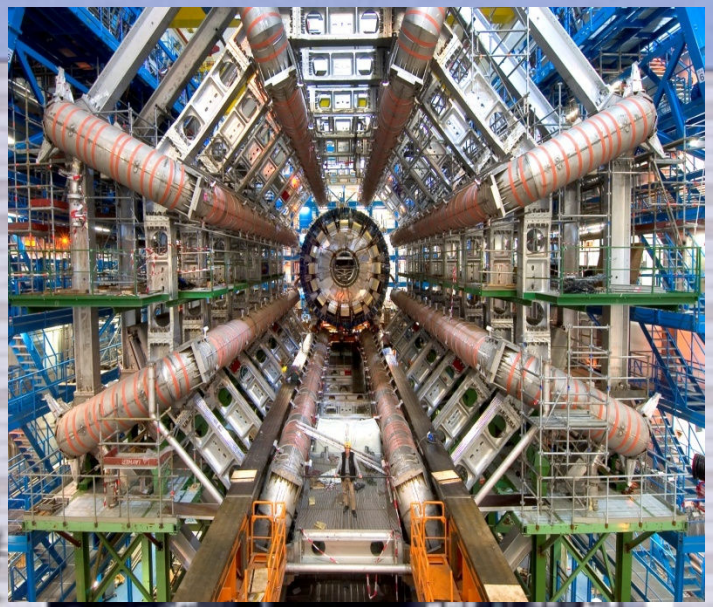
CERN -
- European Organization for Nuclear Research
Research/
Discovery
Technology
Training
Collaborating

Large Hadron Collider

Start-up of the Large Hadron Collider (LHC), one of the largest and truly global scientific projects ever, is the most exciting turning point in particle physics.







ATLAS superimposed to
the CERN 5-floors building 40

Параметры детектора АТЛАС

Энергия центра масс 14 TeV

Частота столкновений пучков 40 MHz

Светимость :

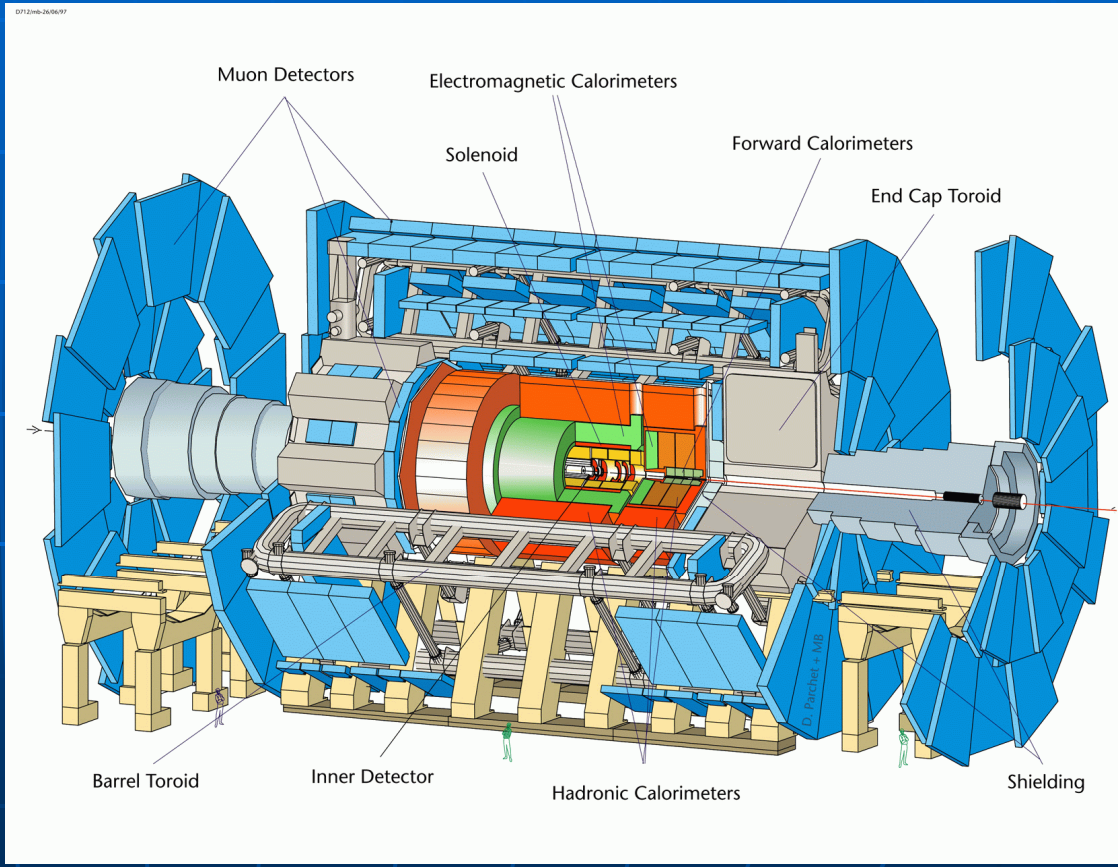
- начальная: 10^{31} $см^{-2}с^{-1}$
- низкая: $2*10^{33}$ $см^{-2}с^{-1}$
- целевая: 10^{34} $см^{-2}с^{-1}$

Вес 7000 тонн,

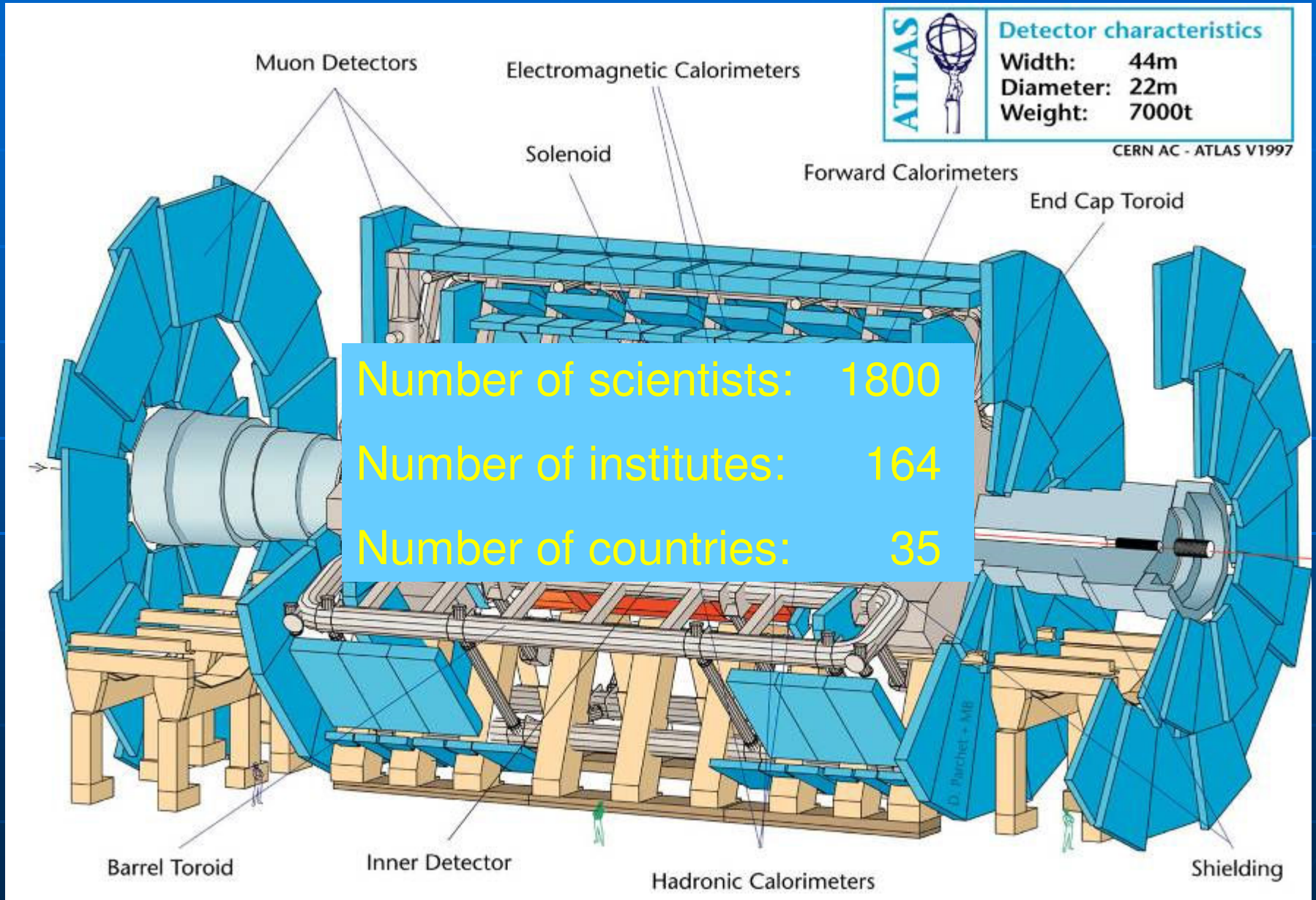
Диаметер 22м,

Длинна 46м

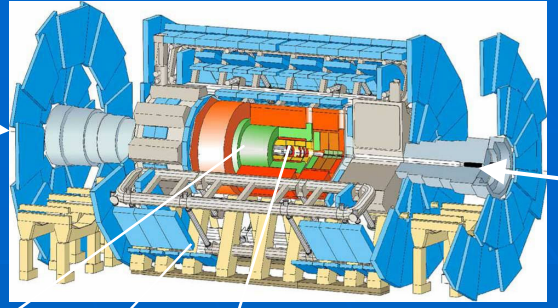
Количество регистрирующих каналов 140 000 000



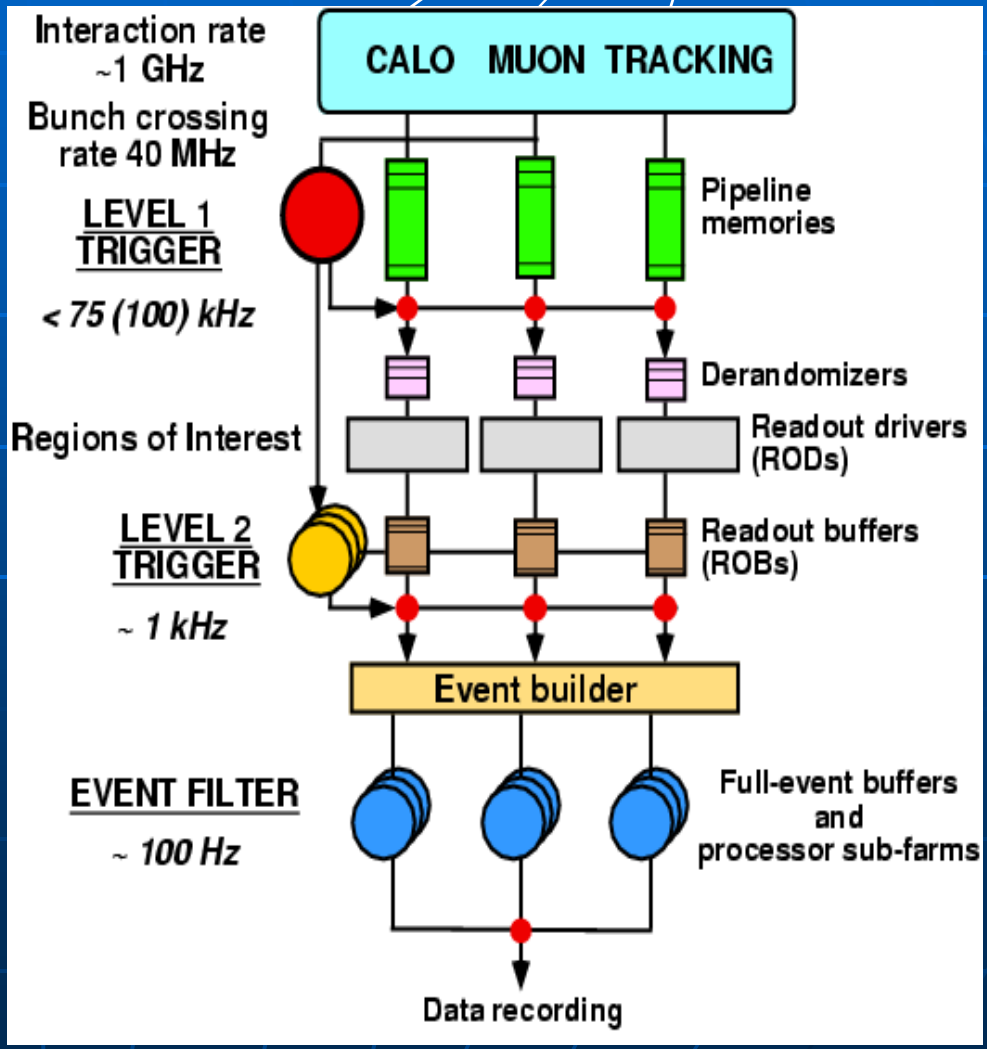
ATLAS


Muon Detectors
Electromagnetic Calorimeters
Solenoid
Forward Calorimeters
End Cap Toroid
Barrel Toroid
Inner Detector
Hadronic Calorimeters
Shielding
ATLAS
Detector characteristics
Width: 44m
Diameter: 22m
Weight: 7000t
CERN AC - ATLAS V1997
Number of scientists: 1800
Number of institutes: 164
Number of countries: 35

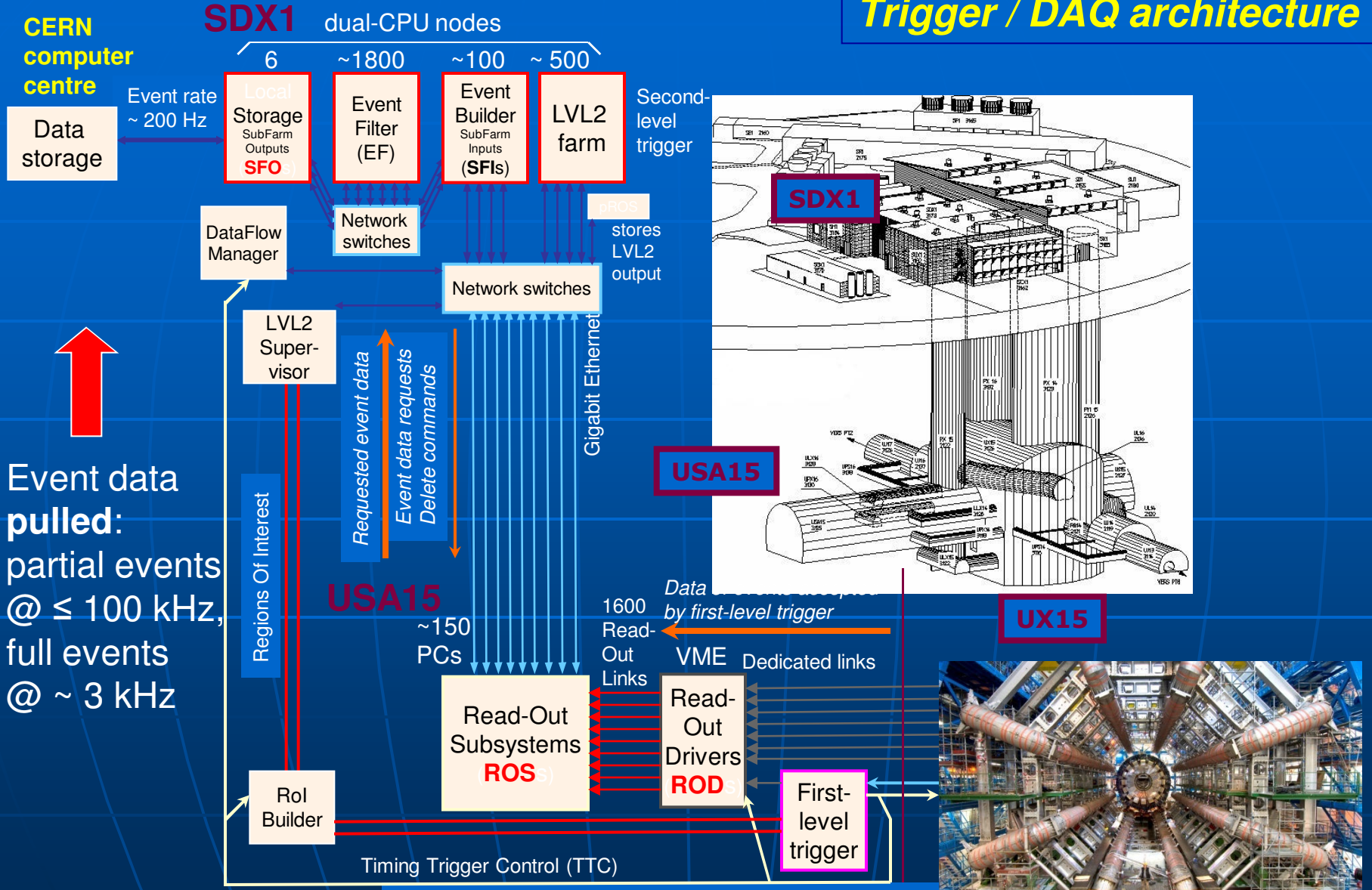
Система Селекции и Сбора Данных



- Триггер 1-го уровня:
  - Сигналы с Калориметров и Мюонной системы
  - Частота событий на выходе ~75 kHz
- Триггер 2-го уровня :
  - 500 компьютеров
  - Селекция на основе **"Region of Interest" (RoI)**
  - Специализированные быстрые алгоритмы
  - Время принятия решения 20 миллисекунд/событие
- Event Builder (EB):
  - 100 компьютеров
- Триггер 3-го уровня (Event Filter (EF)):
  - 2000 компьютеров
  - Стандартные оффлфйн алгоритмы
  - Время принятия решения 2 секунды/событие



Trigger / DAQ architecture



HLT Farms

Final size for max L1 rate ***(TDR)***

~ ***500 PCs for L2 + ~ 1800 PCs for EF***

(multi-core technology)

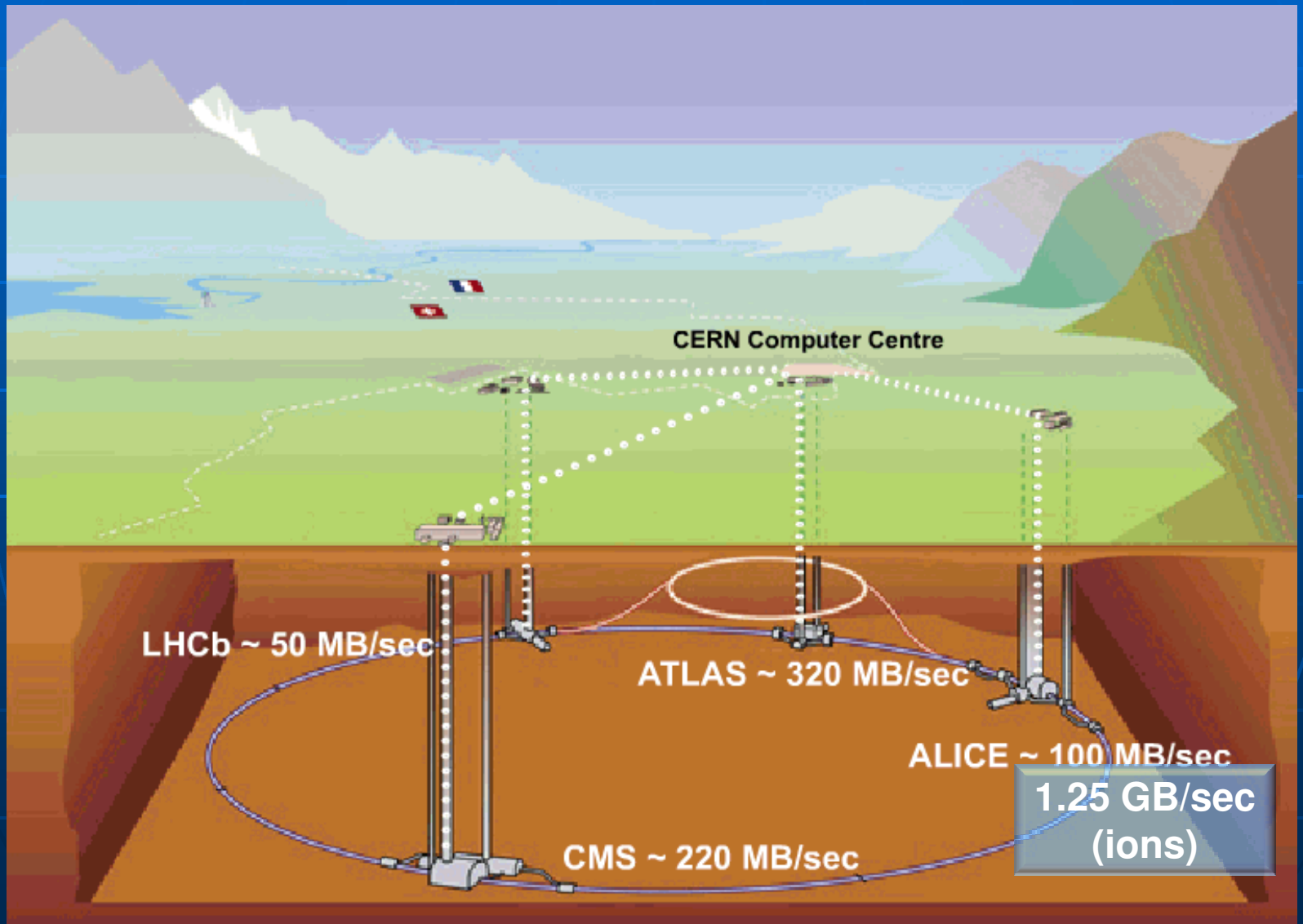
For 2008 : **850 PCs installed**
total of 27 XPU racks = 35% of final system

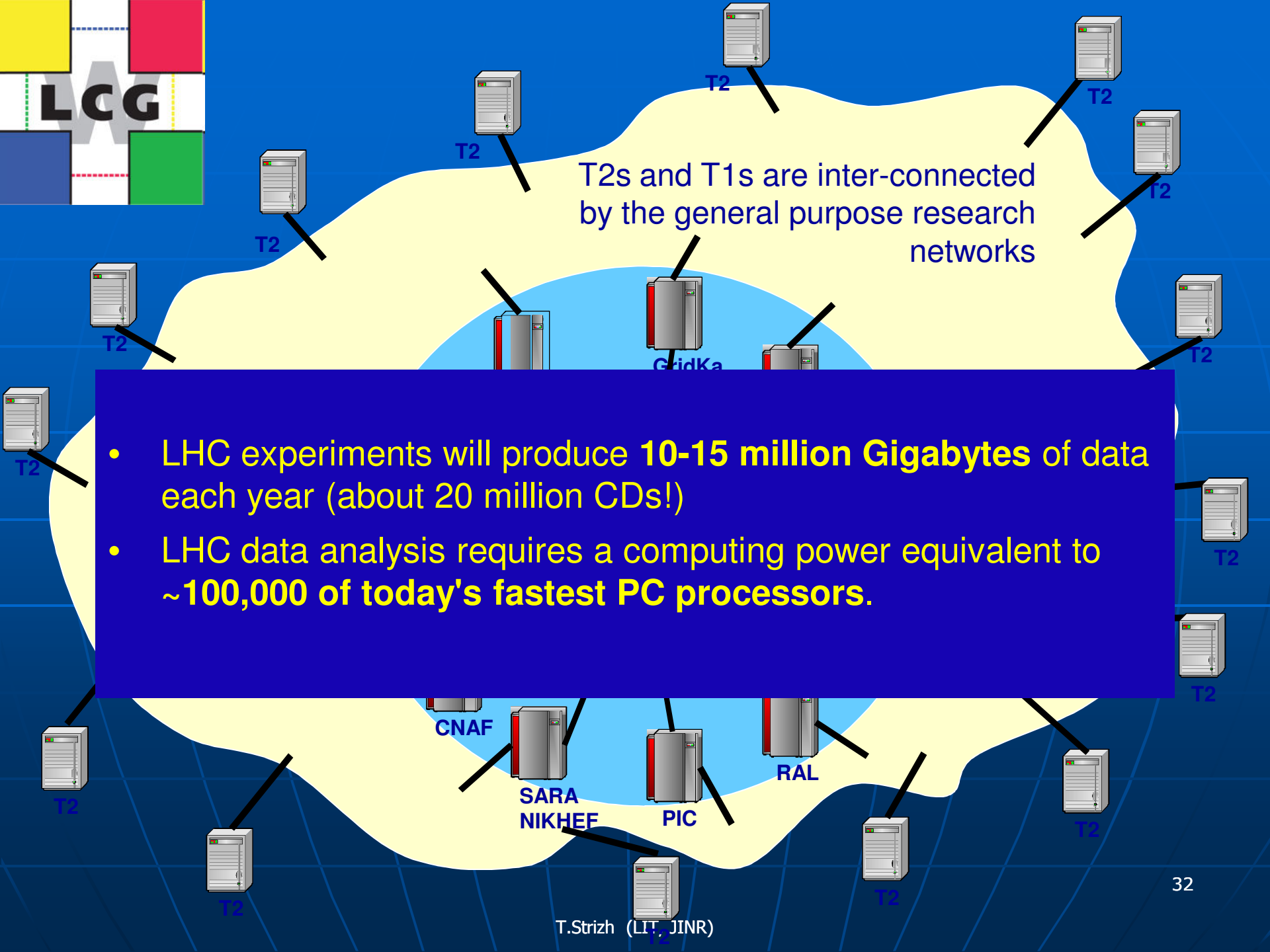
(1 rack = 31 PCs)
(XPU = can be connected to L2 or EF)

- ***x 8 cores***
- ***CPU: 2 x Intel Harpertown quad-core 2.5 GHz***
- ***RAM: 2 GB / core, i.e. 16 GB***

Final system : **total of 17 L2 + 62 EF racks**
of which 28 (of 79) racks as XPU

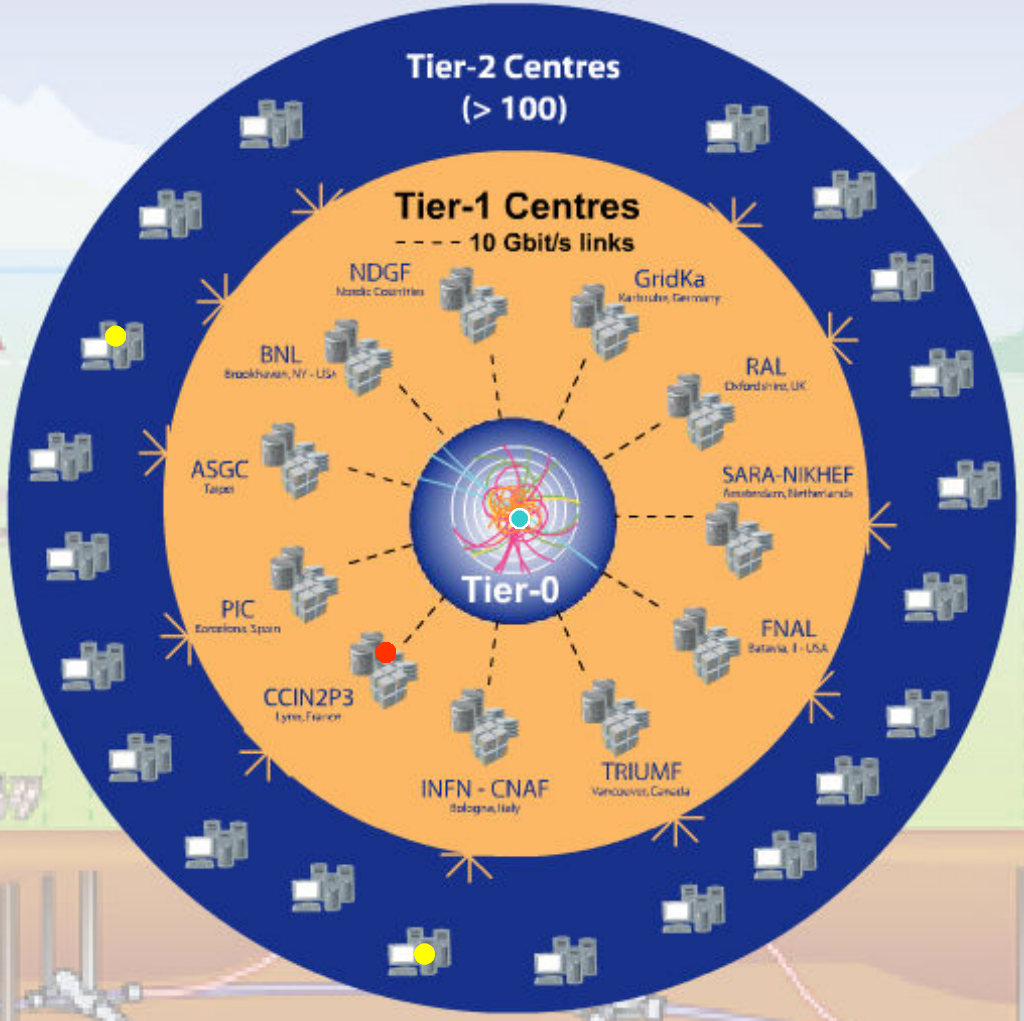
Потоки данных от физических установок БАК до вычислительного центра ЦЕРН (Tier 0 at CERN)





- LHC experiments will produce **10-15 million Gigabytes** of data each year (about 20 million CDs!)
- LHC data analysis requires a computing power equivalent to **~100,000 of today's fastest PC processors**.

Взаимодействие уровней Tier 0 – Tier 1 – Tier 2

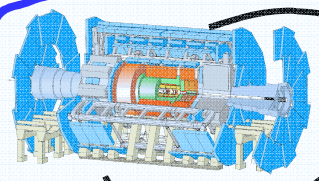


Tier-0 (CERN):
- **Прием данных**
- **Начальная реконструкция данных**
- **Распределение данных**

Tier-1 (11 centres):
- **Постоянное хранение данных**
- **Реконструкция и обработка**
- **Анализ**

Tier-2 (>200 centres):
- **Моделирование**
- **Физический анализ**

Data Handling and Computation for Physics Analysis



Tier-0

detector

event filter (selection & reconstruction)

reconstruction

raw data

event summary data

Tier-1

processed data

event reprocessing

batch physics analysis

analysis

event simulation

simulation

analysis objects (extracted by physics topic)

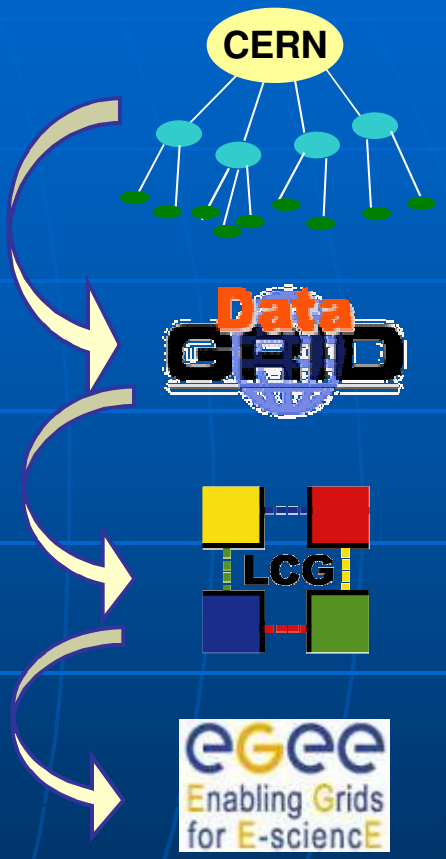
interactive physics analysis

Tier-2

les.robertson@cern.ch

Some history

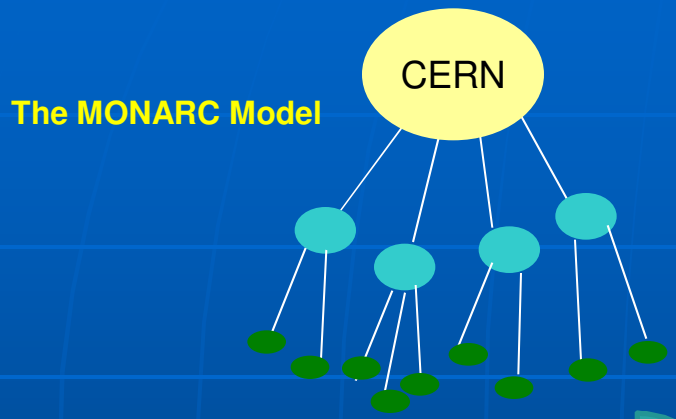
- 1999 – Monarc Project
  - Early discussions on how to organise distributed computing for LHC
- 2001-2003 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 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

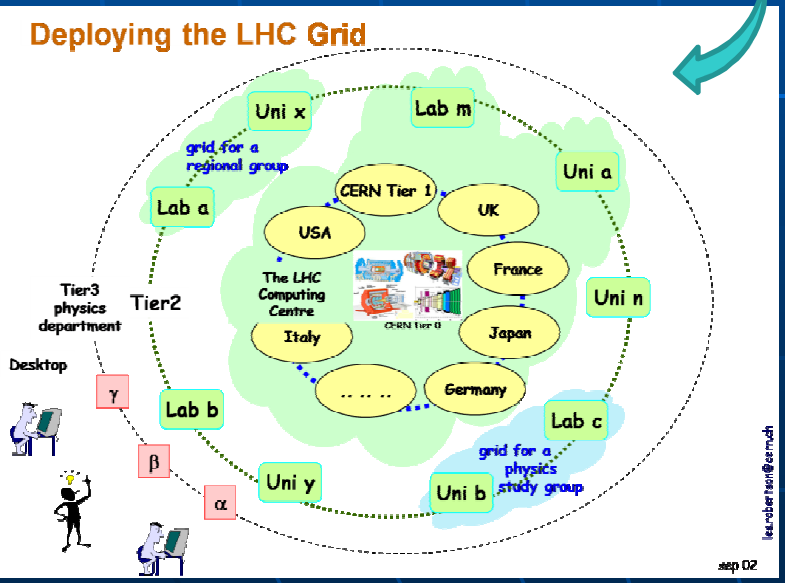


MONARC project

- 1998 – MONARC project
  - a distributed model
    - Integrate existing centres, department clusters, recognising that funding is easier if the equipment is installed at home
    - Devolution of control– local physics groups have more influence over how local resources are used, how the service evolves
  - a multi-Tier model
    - Enormous data volumes → looked after by a few (expensive) computing centres
    - Network costs favour regional data access
    - Simple model that HEP can develop and get into production ready for data in **2005**

Then came the Grid





- 1999 – Grid
  - More flexible → easier to use, adapt to the reality of data analysis
  - But – more complex to build and to manage
  - However – the basics were already there!
  - Prospect of a general science grid (memories of research networks)
  - With expectation of non-HEP funding
    - during development
    - and for long term operation
- Consensus on this approach emerged during CHEP 2000 in Padova

2000s - the Decade of the Grid

- E-Science in fashion
- Many grid/science/physics projects funded in Europe and US
  - Stimulated international collaboration – open to all LHC sites
  - WLCG could operate on top of these multi-science infrastructure grids – EGEE, OSG, ..

  → de-facto standards
  - Significant non-HEP funding was made available to LHC groups and centres – supporting operation, tools and middleware, application development and adaptation to grids
- But many other sciences and also some industries have ported applications to the HEP style of Grid

EGEE & OSG

WLCG depends on two major science grid infrastructures

EGEE - Enabling Grids for E-Science

OSG - US Open Science Grid

WLCG activities

Applications Area
Common projects
Libraries and tools,
data management

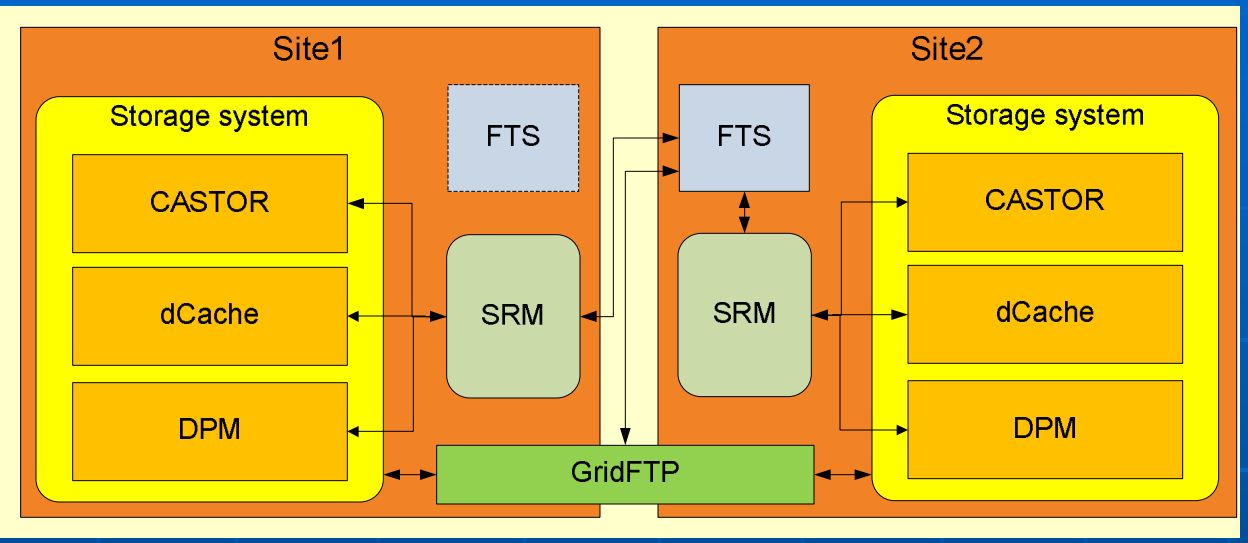
Distributed Analysis
Joint project on distributed analysis with the LHC experiments

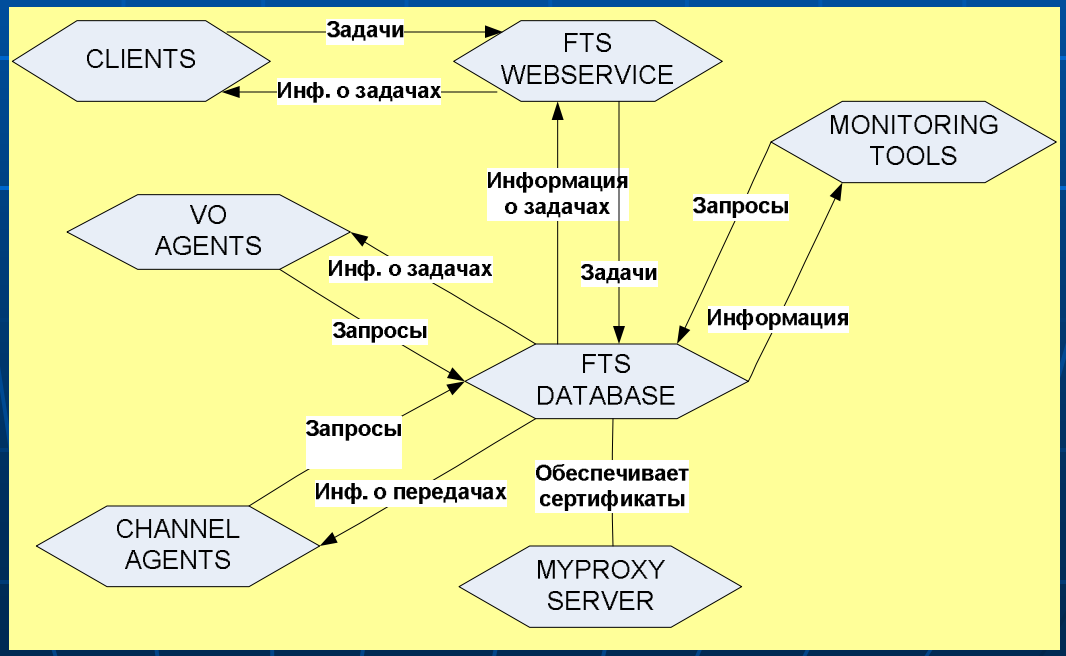
Middleware Area
Provision of grid middleware – acquisition, development, integration, testing, support

Grid Deployment Area
Establishing and managing the Grid Service - Middleware certification, security, operations. Service Challenges

CERN Fabric Area
Cluster management
Data handling
Cluster technology
Networking (WAN+local)
Computing service at CERN

File Transfere System Monitoring and Testing





Integration with Google Earth

Experiment specific monitoring systems provide input data
Dashboard agents publish this information in the KML format
Strong contribution to the development of Sergey Mitsyn (JINR)
Application will be shown during the LHC demo at the EGEE conference in Barcelona

RDMS CMS Data Bases

RDMS CMS Data Management system

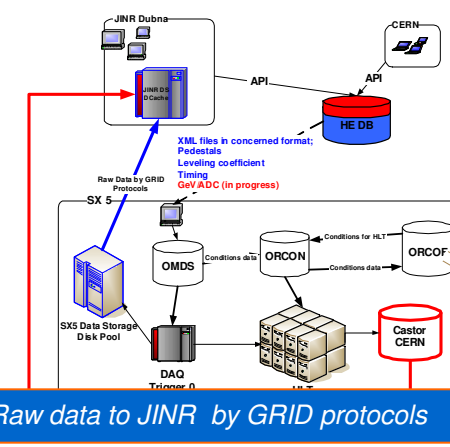
Data Bases

- **Equipment Data Bases**
  ME 1/1
  HE
- **Condition Data Base**
  Calibration data base HE

Services

- API to databases
- WEB Interfaces
- **Data storage and data management**
- Export data and integration with other databases

HE Calibration and Raw Data Flow



OMDS - **O**nline **M**aster **D**ata **S**torage
ORCON - **O**ffline **R**econstruction **C**onditions DB **ON**line subset
ORCOF - **O**ffline **R**econstruction **C**onditions DB **OF**fline subset

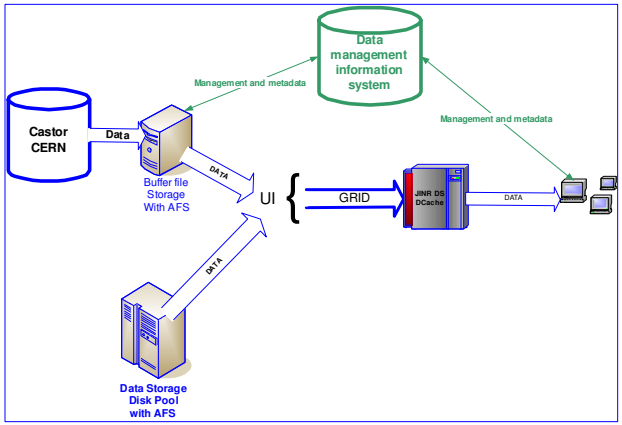
Raw data to JINR by GRID protocols

~1TB of raw data already transferred

HE Calibration DB Status

- System is online
- Full calibration cycle support
- Integrated into CMS computing environment
- ~30000 records
- ~500Mb
- ~600Gb raw data transferred to JINR

CERN - JINR Data management for MTCC: common schema of realization



Удалённый мониторинг

- Коллаборация АТЛАС:
  - 164 институтов из 35 стран
- Для эффективного участия в коллаборации пользователи должны иметь возможность:
  - Отслеживать общий статус эксперимента
  - Отслеживать состояние подсистем, разработанных в их институтах
- Эти задачи решаются с помощью системы удалённого мониторинга

Remote ATLAS Control Room in Dubna

MOTIVATION

- Monitoring of the detector at any time
- Participation of the subsystem experts from Dubna in the shifts and data quality checks remotely
- Training the shifters before they come to CERN

ACR at CERN

The goal - to have reduced copy at JINR