IBERGRID

CSIC
Consejo Superior de Investigaciones Científicas

IFCA
Instituto de Física de Cantabria

Dr. Isabel Campos Plasencia
IFCA – CSIC
Director of the Spanish National Grid
& European Grid Initiative Executive Board

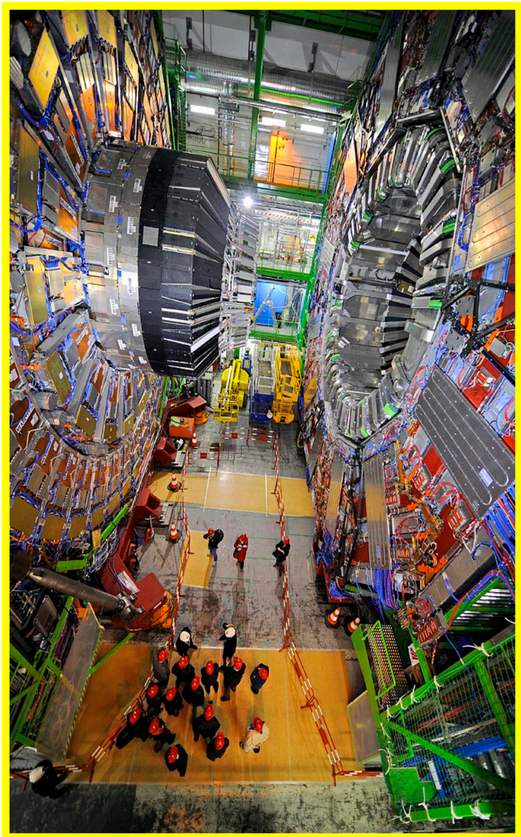
DataScience in Aviation Workshop,

Madrid, October 15th 2013

Overview

1. **Overview**
2. **The Challenge of the Large Hadron Collider: which data?**
3. **The Grid**
4. **Model of Data Distribution & tools @ LHC**
5. **Visualization**
6. **Conclusions**

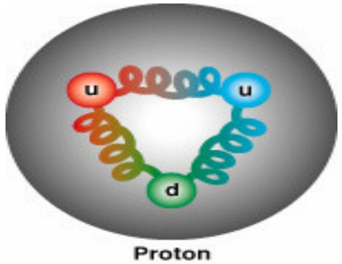
The LHC Challenge

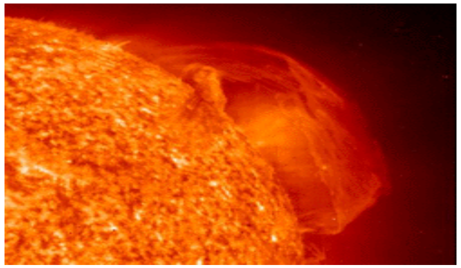


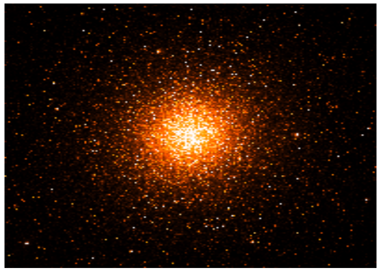
CSIC

High Energy Physics

- Has to do with the **fundamental components of matter**:
  - **Particles of (stable) matter**: quark ***u,*** quark ***d, electrons***
    - All the matter observed in the Universe is made with those 3 particles.





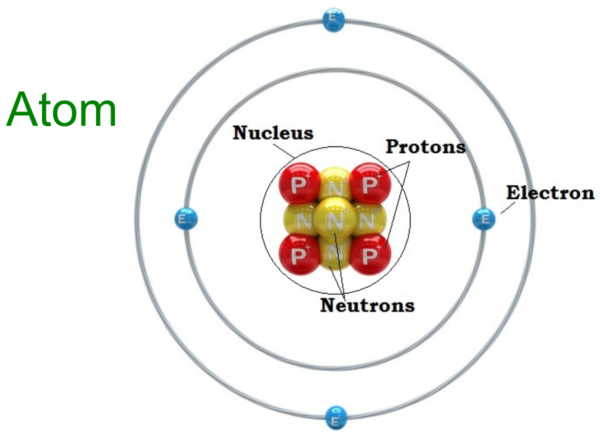


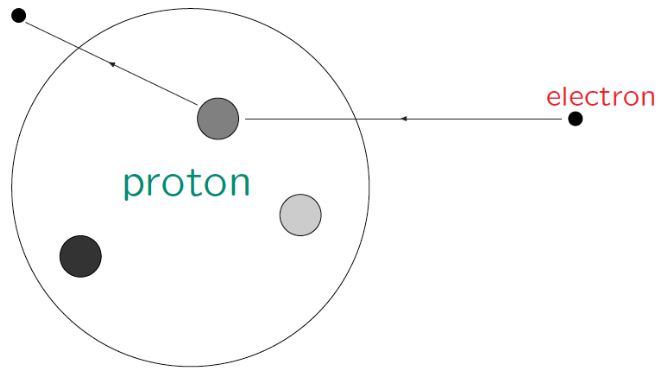
  - **Particles mediating the interactions of matter**
    - photons: mediate in electromagnetic interactions (light)
    - Gravitons (apples fall, earth circles around the Sun): never seen, but little doubts about its existence
    - Z/W bosons, responsible for the nuclear fusion: i.e. Sun
    - Gluons: keep atomic nucleus together (very useful)

CSIC

How do we find a new Law in Physics?

Looking for elementary particles





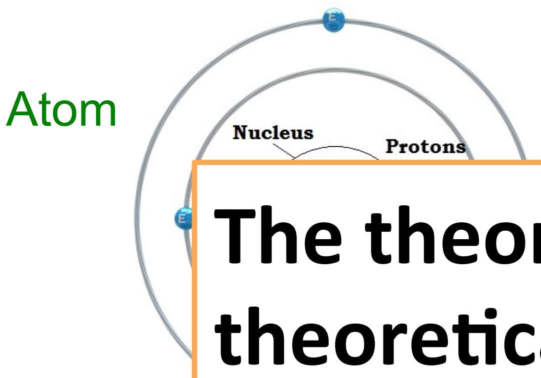
First, we guess it

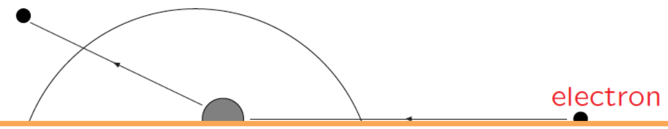
Second, we compute the mathematical consequences of our guess: How much does the electron deviate?

Third, we compare our result with Nature by performing an Experiment: How much does the electron really deviate

How do we find a new Law in Physics?

Looking for elementary particles





The theory is correct <u>if</u> our theoretical prediction agrees with experiment

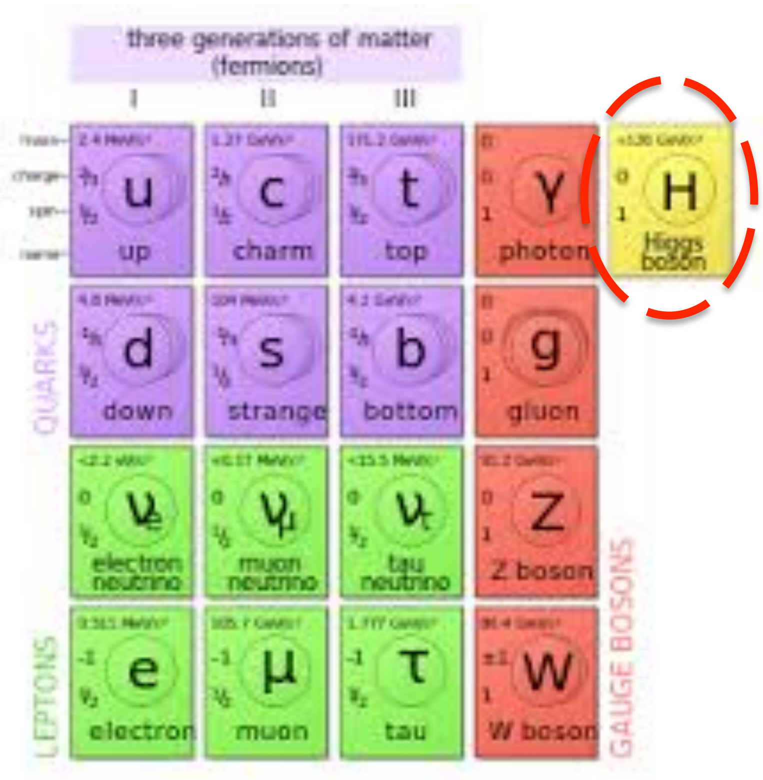
First, we guess it

mathematical consequences of our guess:
How much does the electron deviate?

Third, we compare our result with Nature by performing an Experiment:
How much does the electron really deviate

CSIC

Our current preferred theory (guess) for Nature is the ***Standard Model***

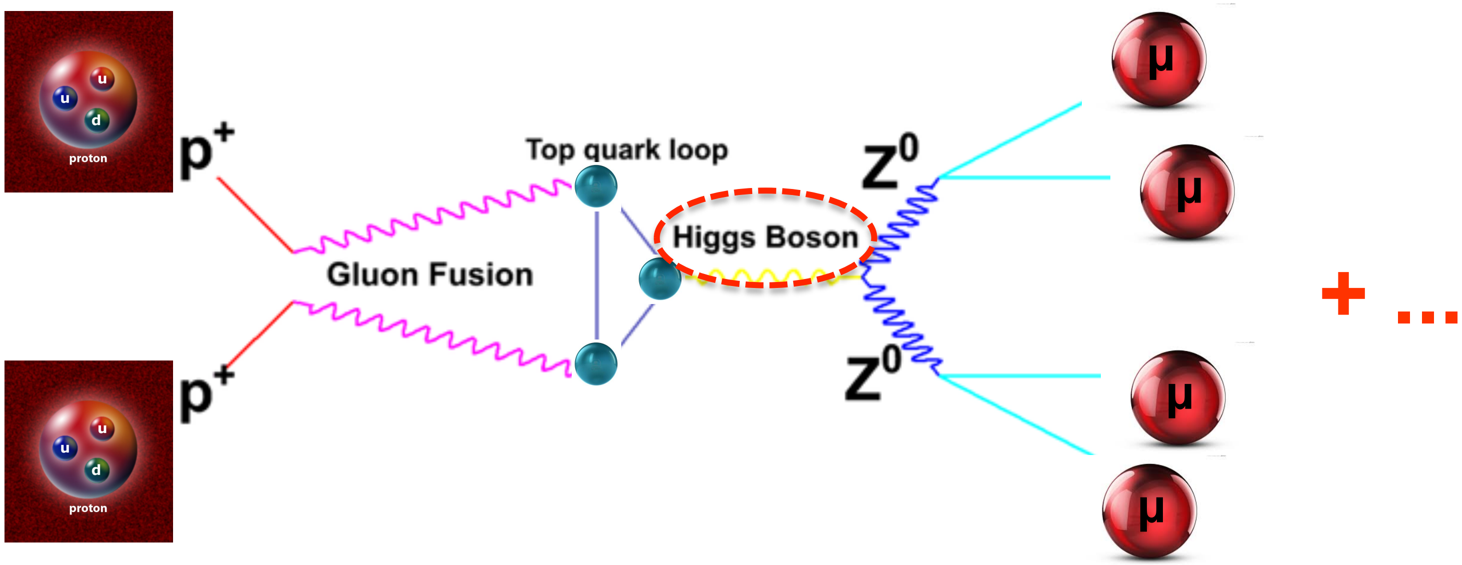


OPEN QUESTIONS

- How do elementary particles obtain the mass
- Is there a Higgs particle? Are there several Higgs particles?
- Do all forces of nature arise from a single fundamental interaction?
- Are there more than 3 dimensions of space?
- Can dark matter be produced in a laboratory?

- BTW: Gravity is missing here

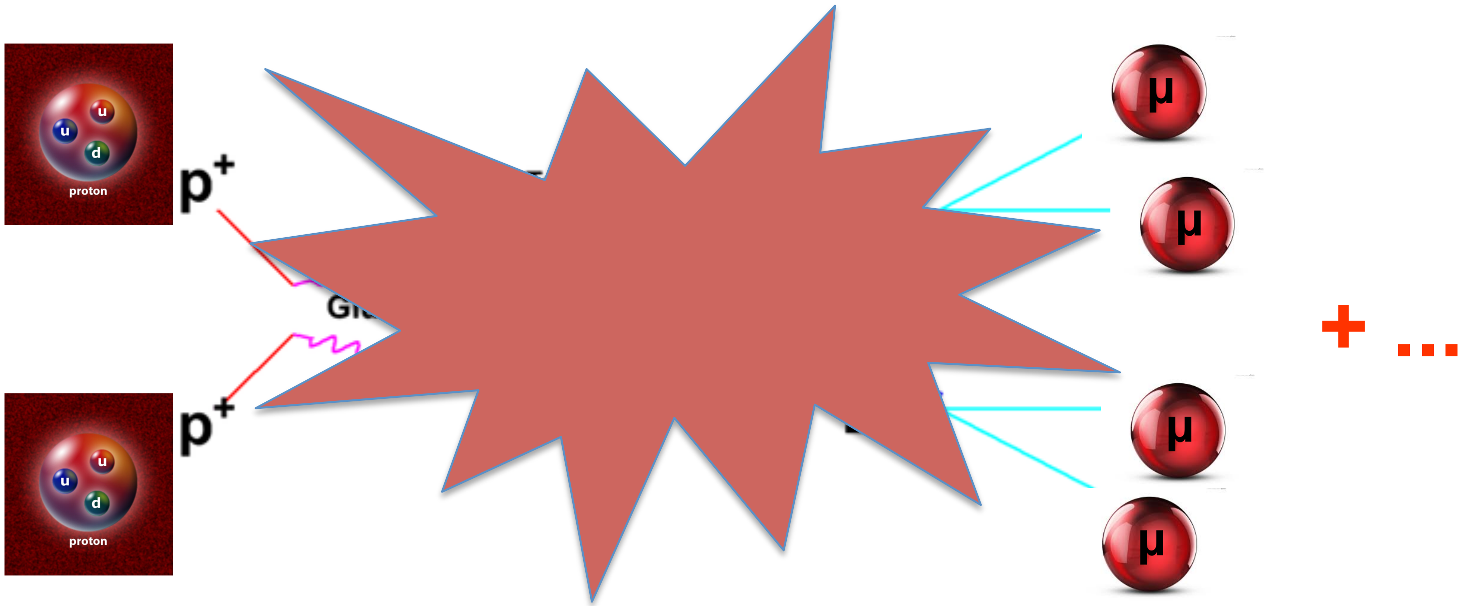
What does our calculation in the Standard Model tell us about **the behavior of the Higgs particle ?**



- During "an infinitesimal moment" the components of the proton are free
- According to our calculations in the Standard Model, the Higgs should appear "visible" for an infinitesimal amount of time
- A the detectors of the LHC we take a "photo" of that "instant in time"

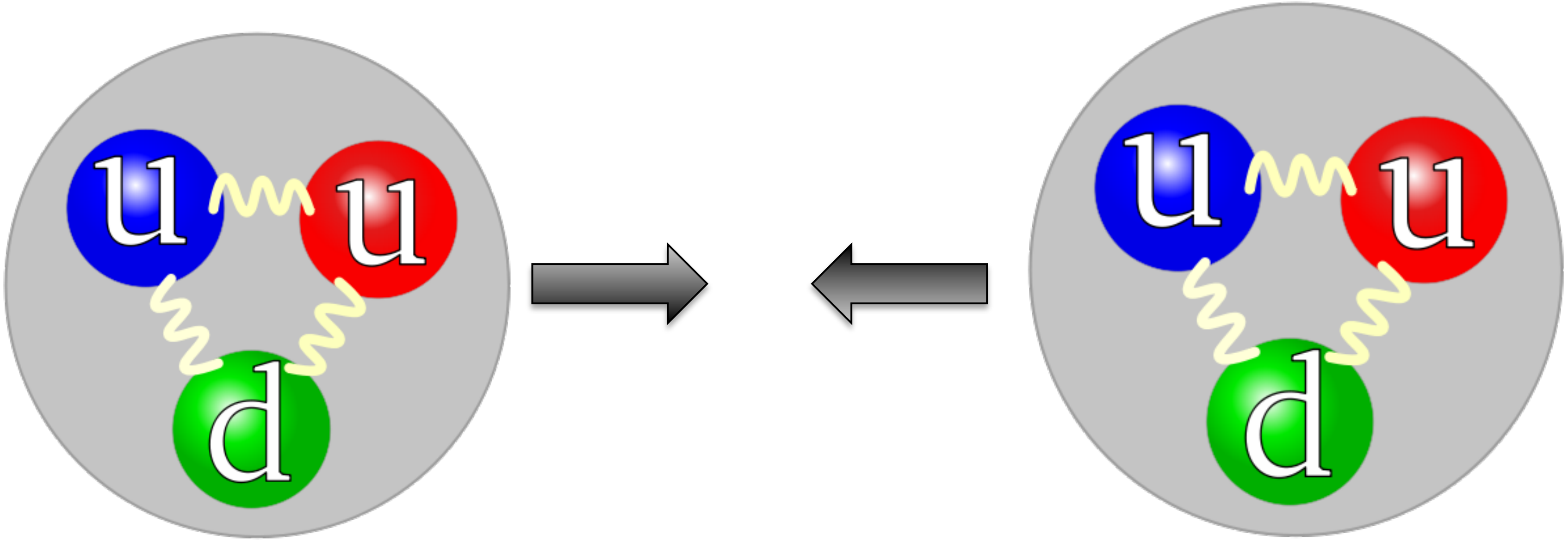
What does our calculation in the Standard Model tell us about

the behavior of the Higgs particle ?



- During “an infinitesimal moment” the components of the proton are free
- According to our calculations in the Standard Model, the Higgs should appear “visible” for an infinitesimal amount of time
- A the detectors of the LHC we take a “photo” of that “instant in time”

At the LHC we perform the experiment:
we smash two protons against each other

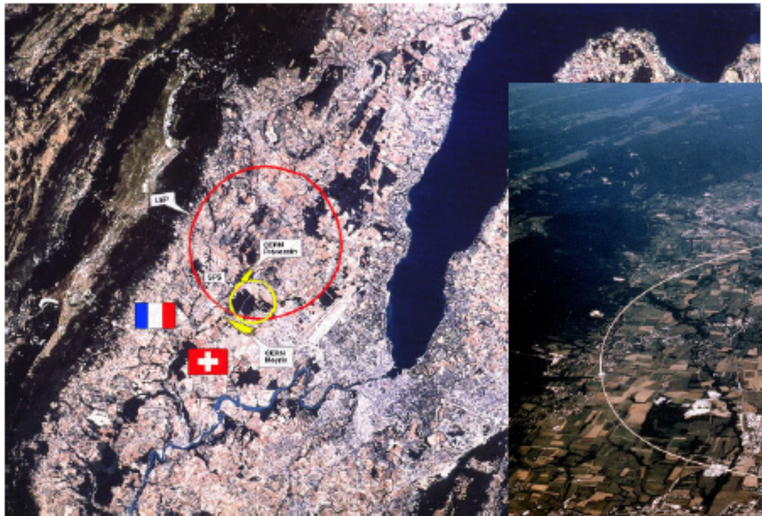


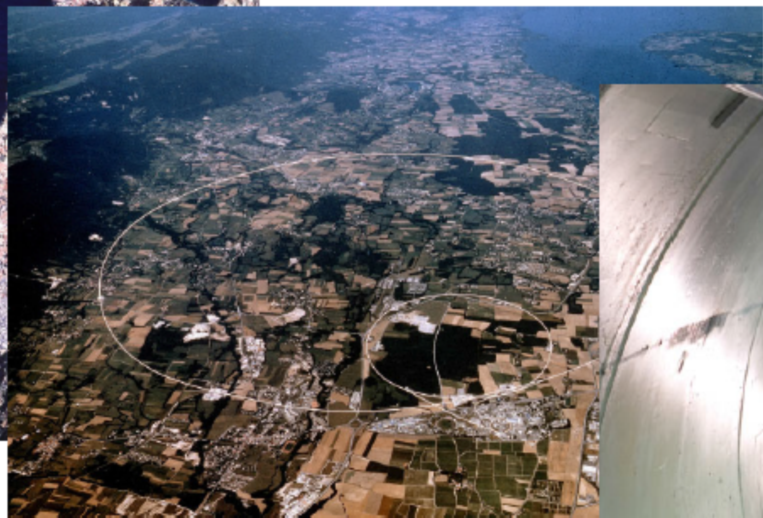
- During “an infinitesimal moment” proton breaks into pieces
- A the detectors of the LHC we take a “photo” of that “instant in time”
- And compare with our calculation

CSIC

We smash protons against each other as hard as technology can:

The fastest racetrack on the planet…



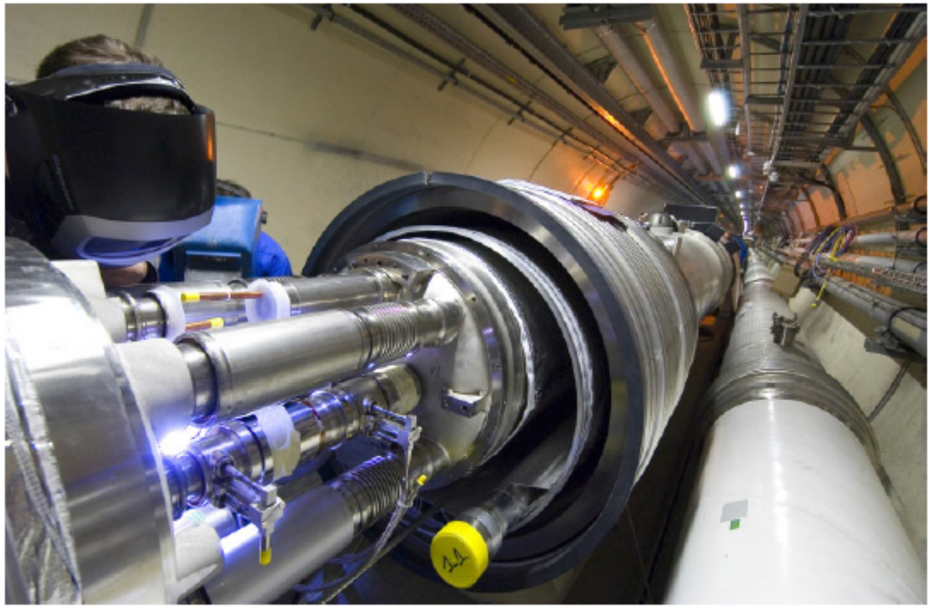


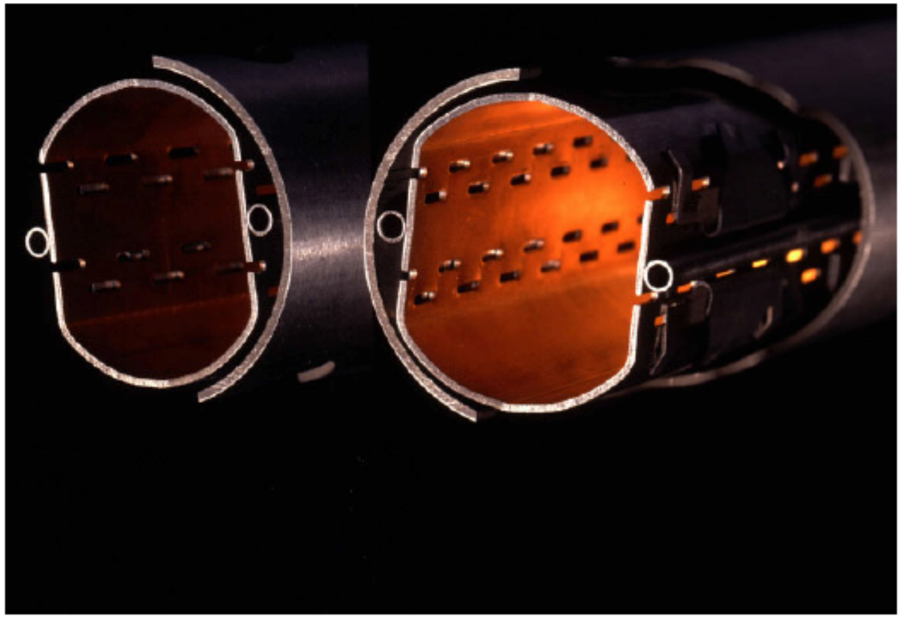
Trillions of protons will race around the 27km ring in opposite directions over 11,000 times a second, travelling at 99.999999991 per cent the speed of light.

CSIC

We smash protons against each other as hard as technology can:

The emptiest space in the solar system…





To accelerate protons to almost the speed of light requires a vacuum as empty as interplanetary space. There is 10 times more atmosphere on the moon than there will be in the LHC.

CSIC

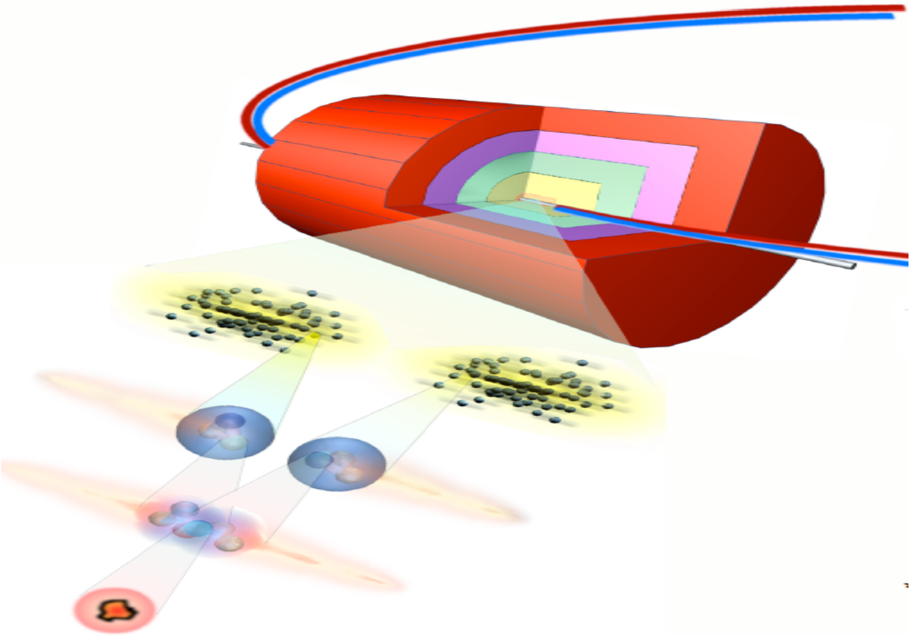
We smash protons against each other as hard as technology can:

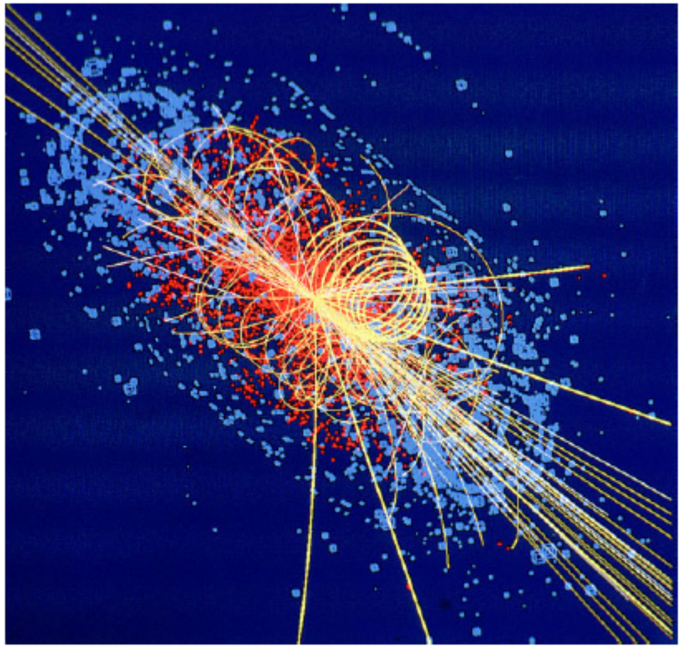
One of the coldest places in the universe…

With an operating temperature of about -271 degrees Celsius, just 1.9 degrees above absolute zero, the LHC is colder than outer space.

We smash protons against each other as hard as technology can:

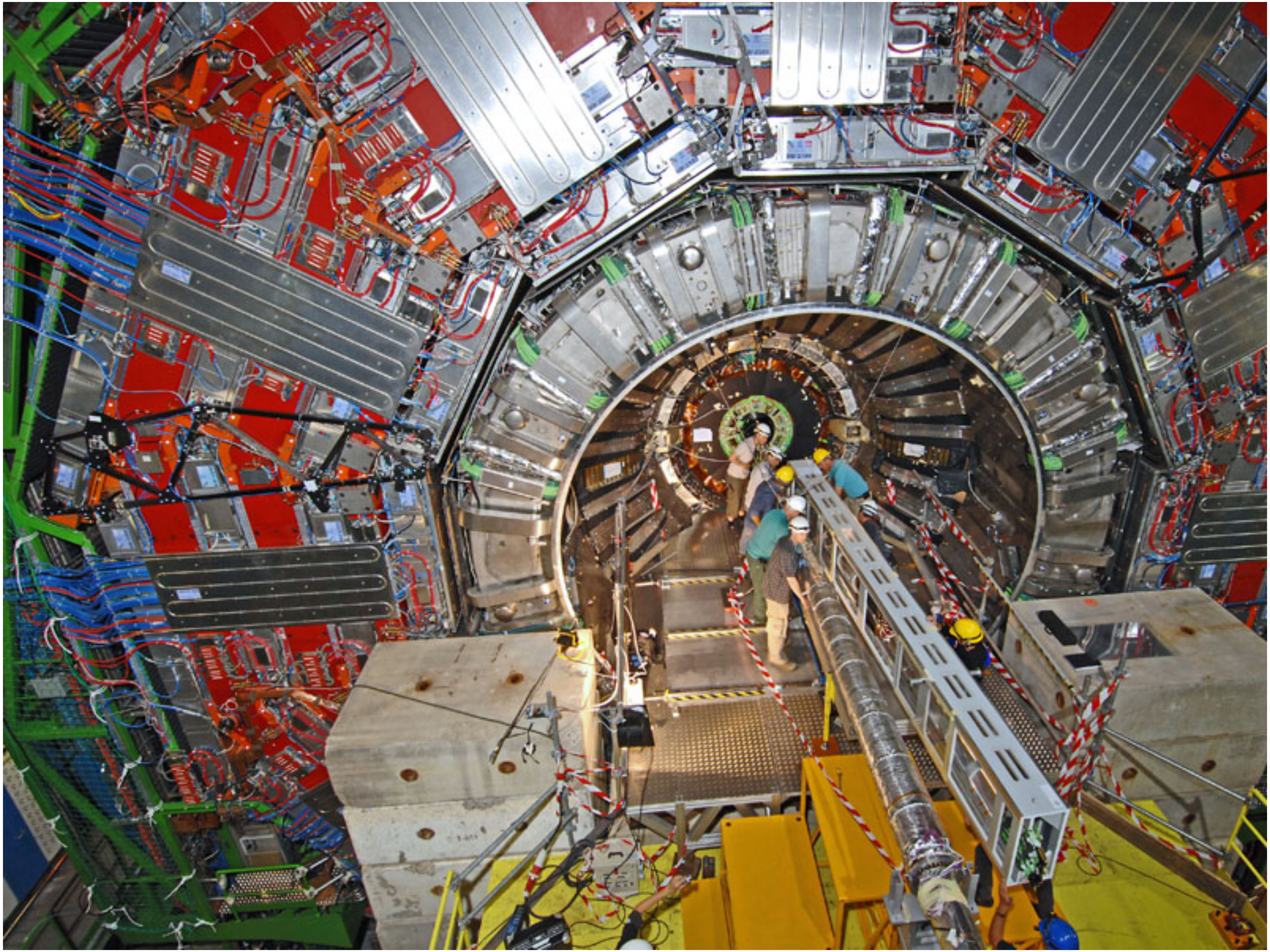
The hottest spots in the galaxy…





When two beams of protons collide, they will generate temperatures 1000 million times hotter than the heart of the sun, but in a minuscule space.

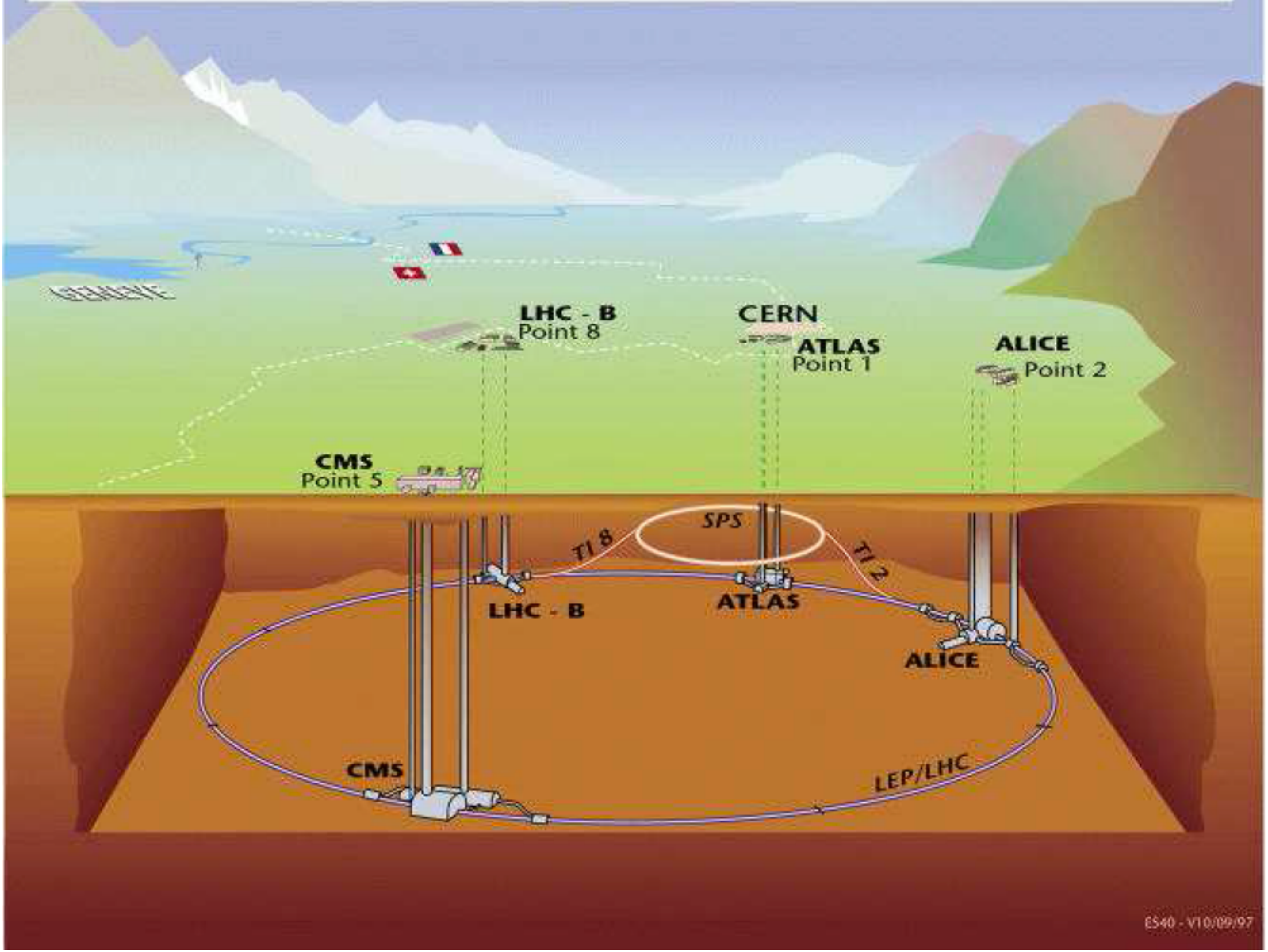
CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



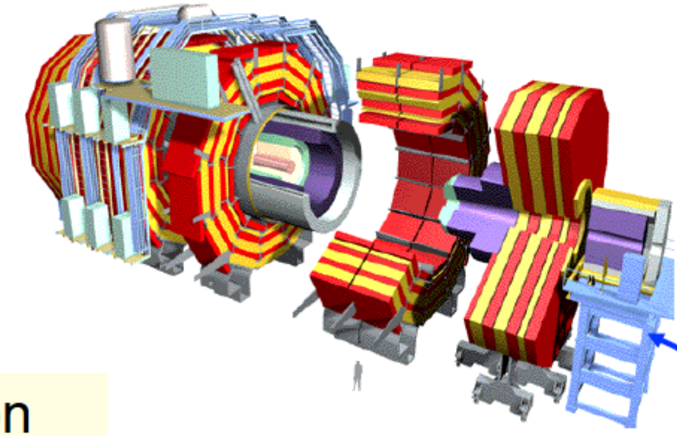
The image is for comparison
(It is not on the surface !)
• Diameter 25 m -- Length : 46 m
• Overall weight 7 000 tonnes

The biggest crane on earth needed
to be built at CERN to place the detectors
in the underground 175 m. deep

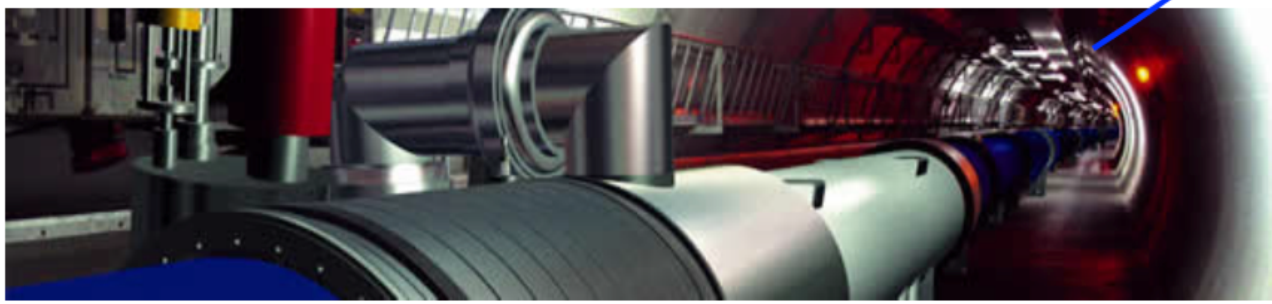
Overall view of the LHC experiments.



Accelerator & Detector

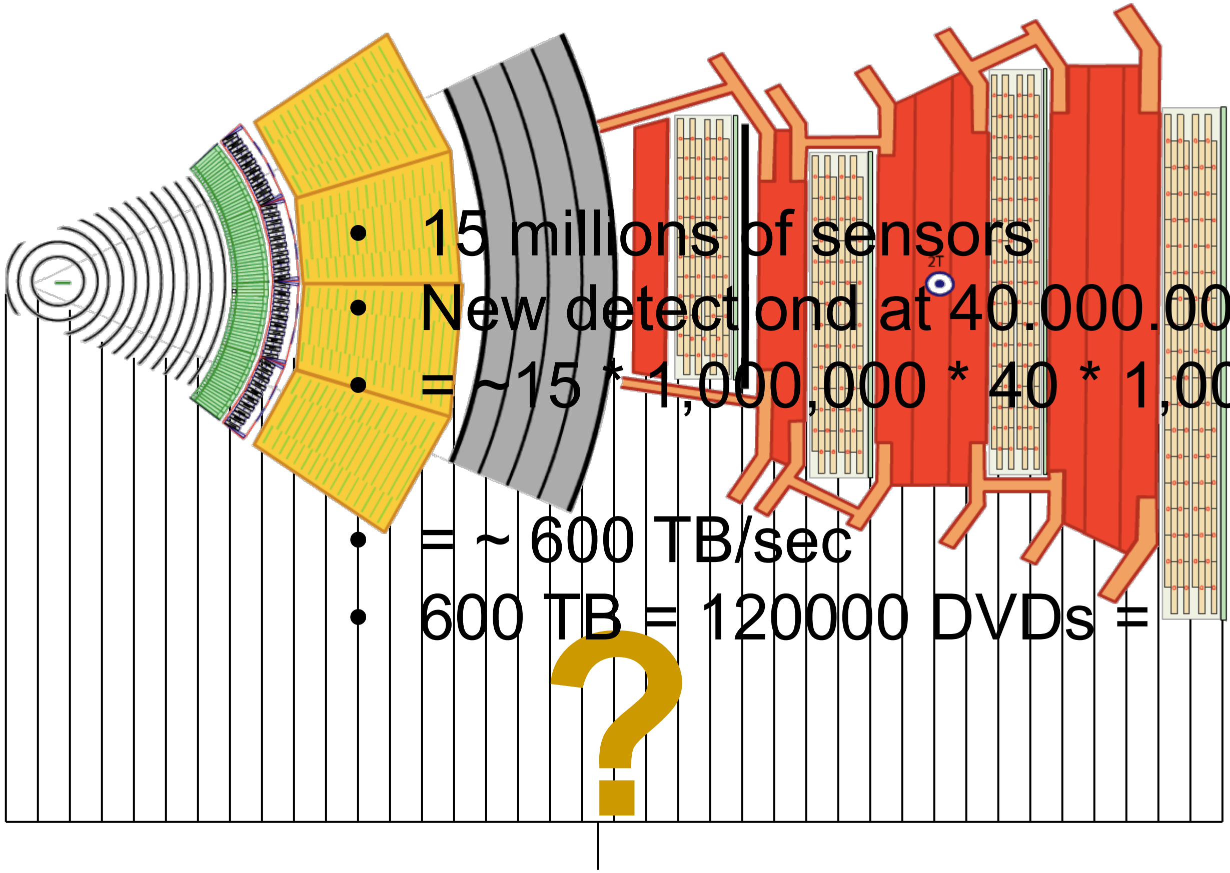


The accelerator generates 40 million particle collisions (events) every second at the centre of each of the four experiments' detectors



At the detector: looking for a needle in a haystack

- 15 millions of sensors
- New detectiond at 40.000.000 / second
- = ~15 * 1,000,000 * 40 * 1,000,000 bytes

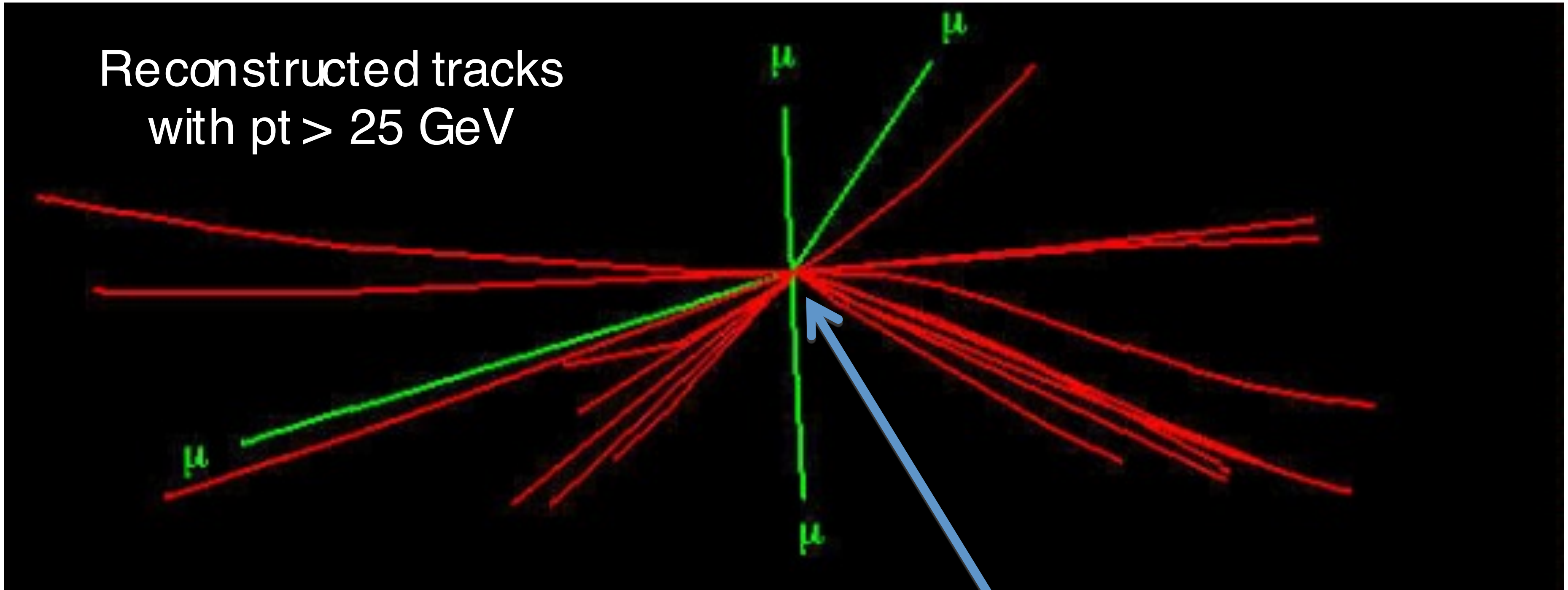


- = ~ 600 TB/sec
- 600 TB = 120000 DVDs =

A detector is like a 3-D camera taking a photo 40 millions times per second, without the possibility of data compression

We need to find ONE event (few Mbytes) inside 600 TB of information

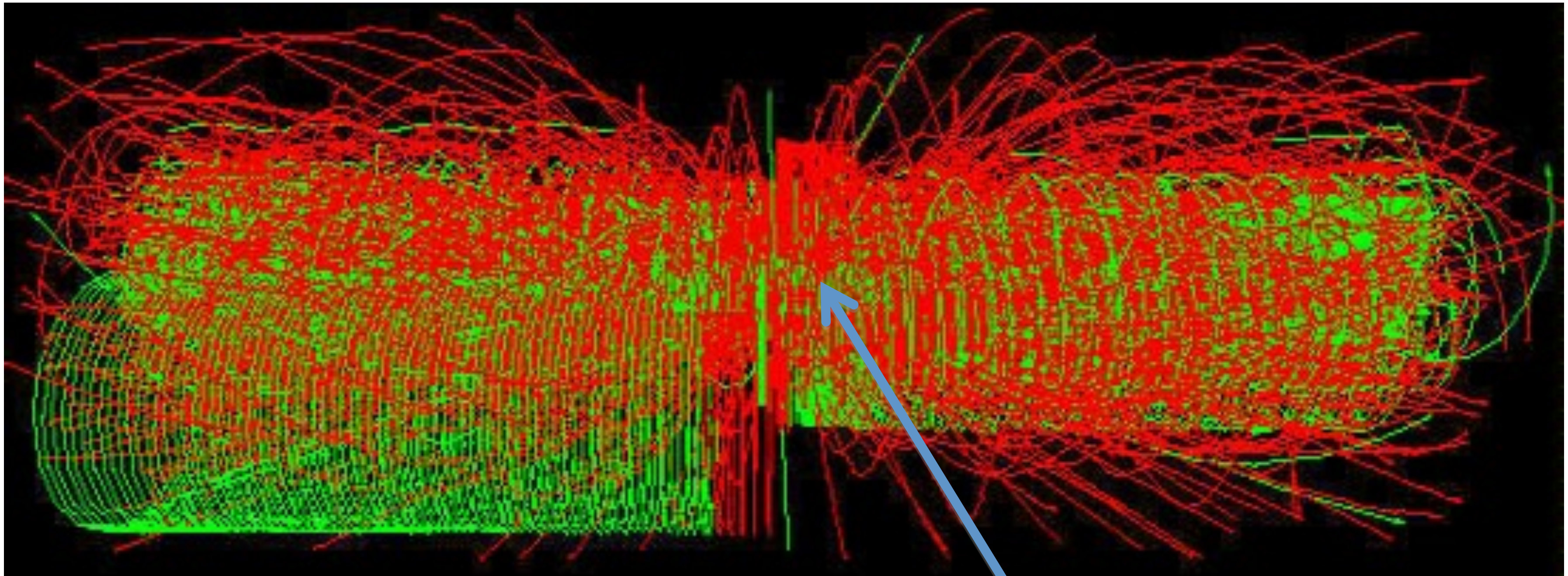
The CMS (Compact Muon Solenoid) detector is searching
For Higgs particles coming out from this interaction: 4-muons decay



The theoretical expectation is that one Higgs particle
Will be produced in every 40 Million proton-proton collisions.
This is about 1 Higgs particle per day.

CSIC

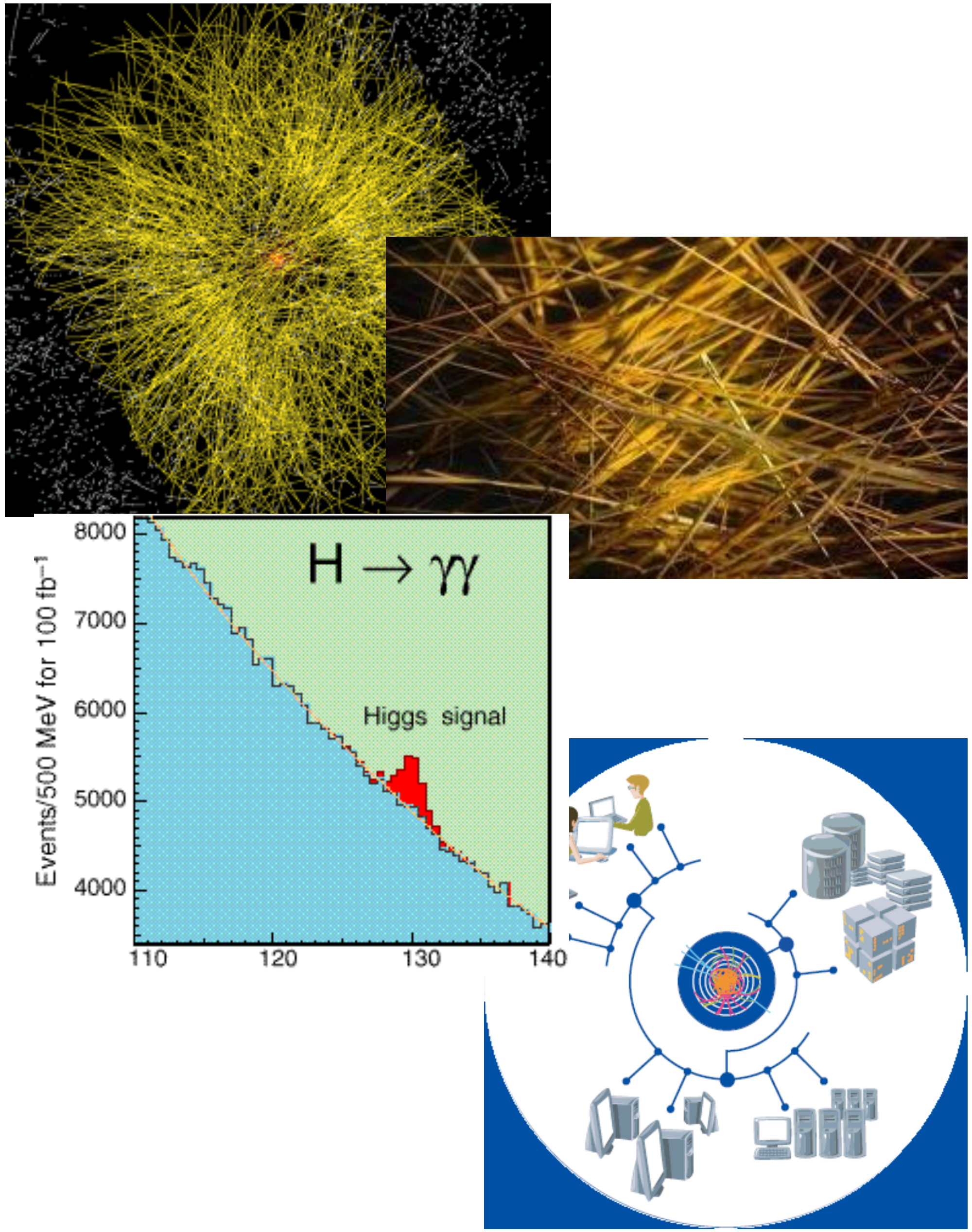
The CMS (Compact Muon Solenoid) detector is searching
For Higgs particles coming out from this interaction: 4-muons decay



This is what we find: more than 1000 particles
We get this 40.000.000 times per second!

At the detector: looking for a needle in a haystack

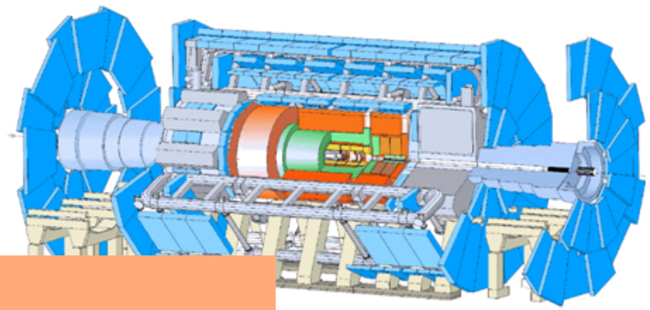
- Signal/Noise: 10^{-9}
- Data volume
  - High rate * large number of channels * 4 experiments
  - ➔ **15 PetaBytes of new data each year**
- Compute power
  - Event complexity * Nb. events * thousands users
  - ➔ **100 k of (today's) fastest CPUs**
  - ➔ **45 PB of disk storage**
- Worldwide analysis & funding
  - Computing funding locally in major regions & countries
  - Efficient analysis everywhere



At the detector:

Data detection and “reduction”

- Level 1 (PB/sec): Fast Detector Triggers (calorimeter, muon detectors)
- Level 2 (75GB/sec): Full detector info in (triggered) regions of interest
- Level 3 (5GB/sec): Event filter with calibration and alignment



After Reduction we are left (for each of the 4 detectors)

- RAW data: 400
  - Hits and pulse h
  - On the order of

  from each detector

With Data Taking 0.1 – 1 GBytes/seg

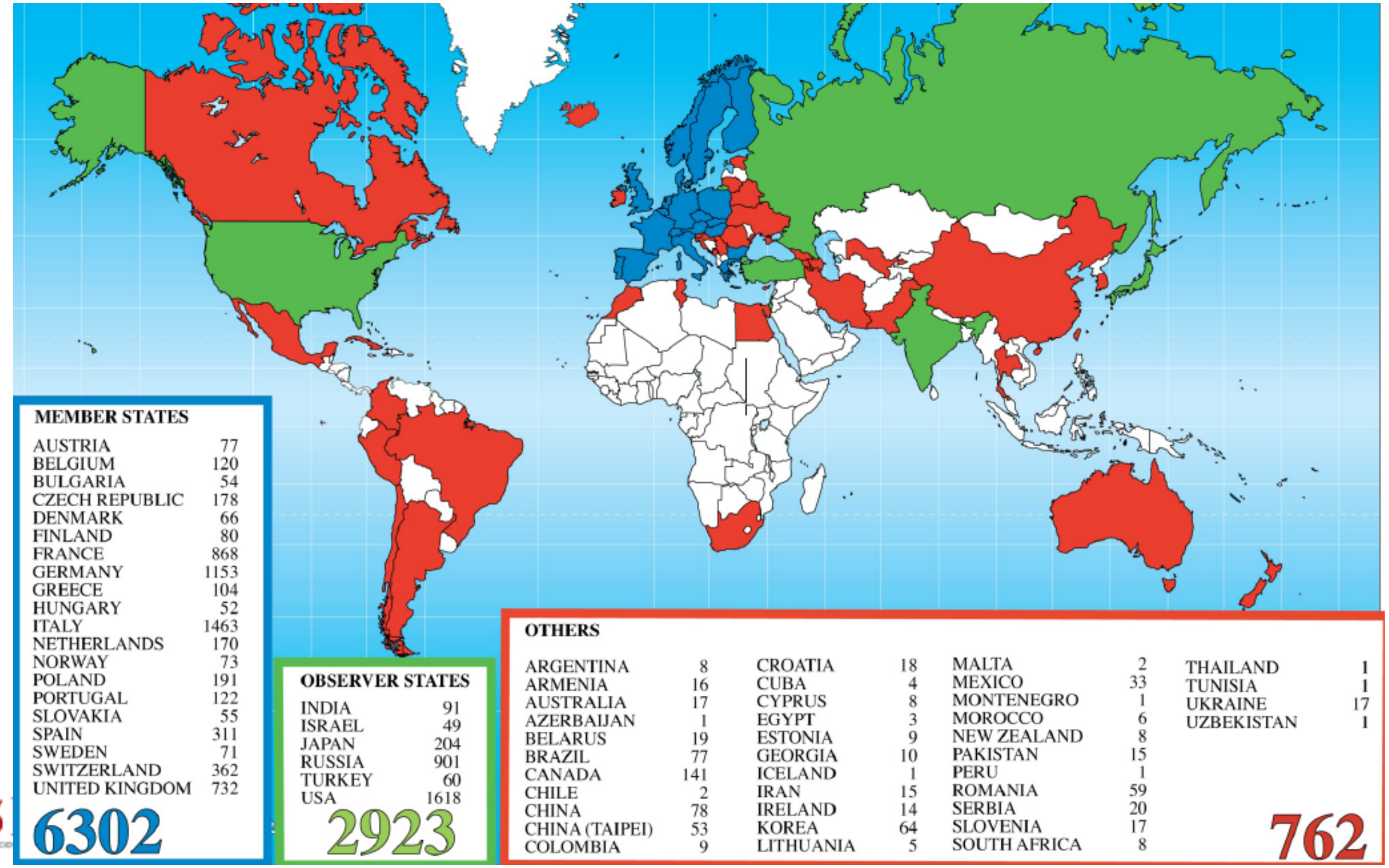
storage of about 15 PetaBytes/year

Computing power to analyse this data of about 100000 core

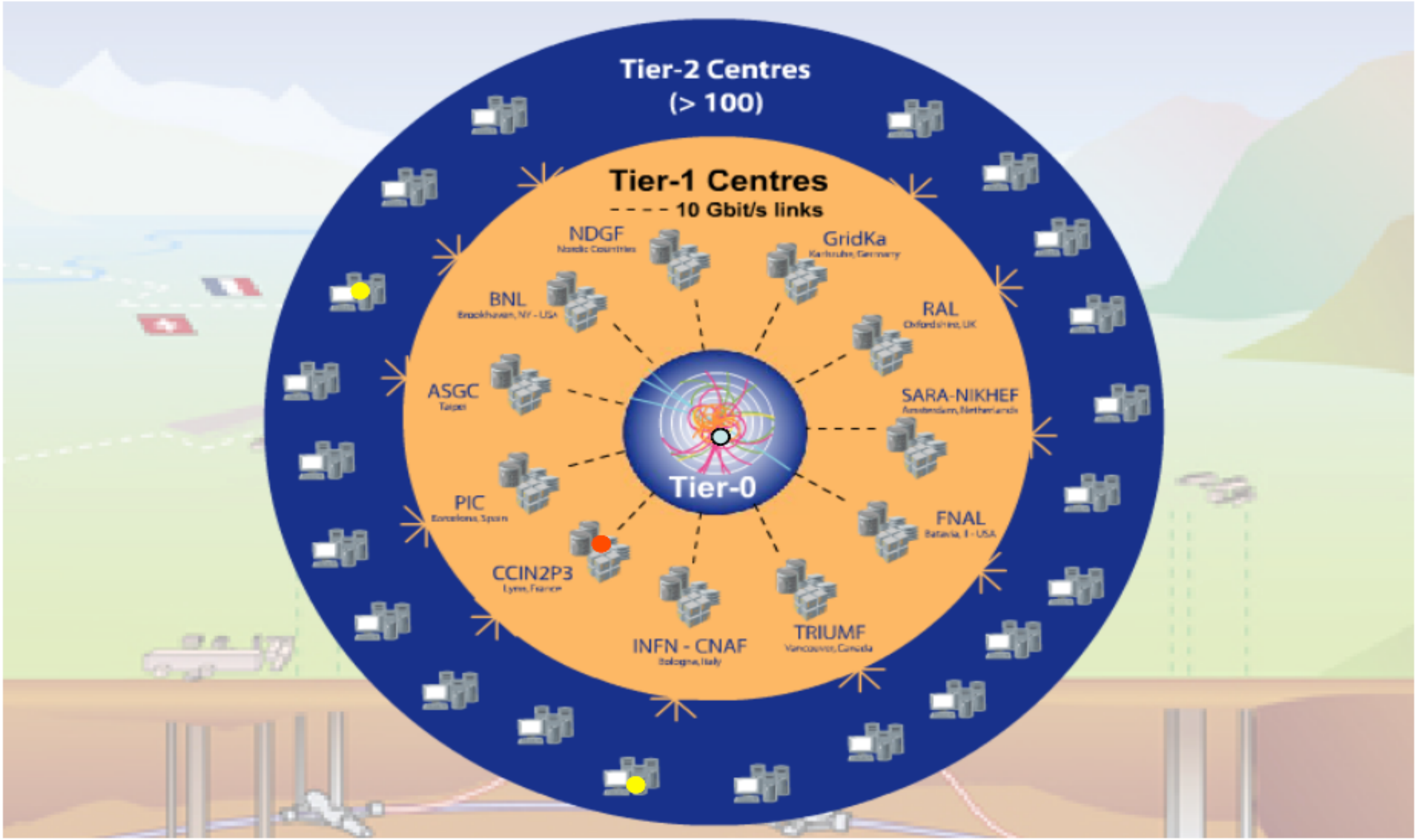
--> THE GRID

HEP: a very motivated user community

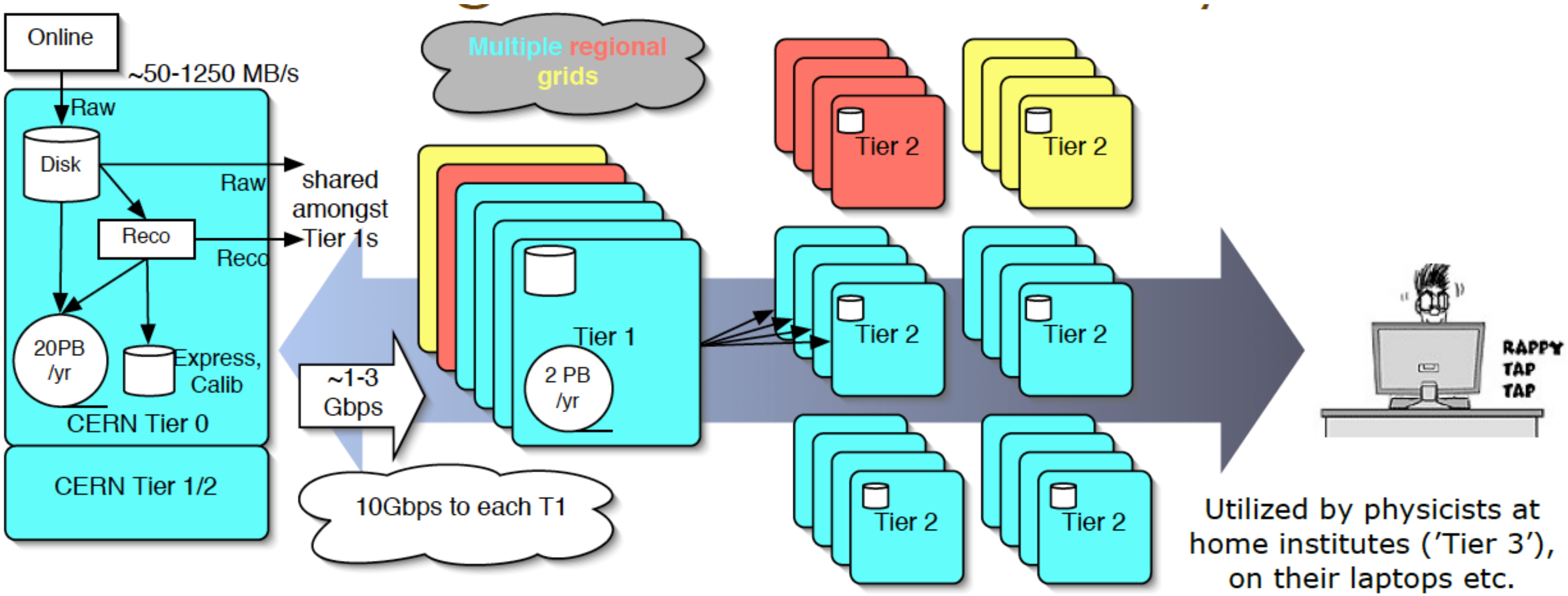
- **Many (hundreds) of research groups distributed through the world**
- **Long history in analyzing large amounts of experimental data: working like in a “Grid” but without the right tools for it**
- **Research & development teams very motivated due to the success of the** *World Wide Web* **(which was developed at CERN)**



Breaking up a massive data set into small pieces around the world



The Tier(s) data distribution



CERN

Buffering from online
Express, calibration reco
Prompt first pass reco (not HI)
Archival copy of RAW, ESD data
Second copy sent to T1s
Reprocessing during shutdowns
Analysis facility with AOD, ESD
Analysis: ~100 users/expt

Tier 1s: ~6-10 per experiment, 10 total

Data import, serving, export
Share of RAW+ESD
Archive for locally produced data
ESD fraction, AOD on disk
Re-reconstruction
(Primary reconstruction if needed)
Calibration
Simulation
Analysis skims, user analysis

Tier 2s: ~3-5 per Tier 1, ~40 overall

Skims from Tier 1
ESD, AOD subsets on disk
User analysis - regional and/or by physics topic
Simulation (most experiments)
~20 active users (widely variable)

PC today ~1 kSi2k

CSIC

For example for the CMS experiment

More than 50 CMS centers, in more then 20 countries

Flags taken from Wikipedia:
http://de.wikipedia.org/wiki/Liste_der_Nationalflaggen

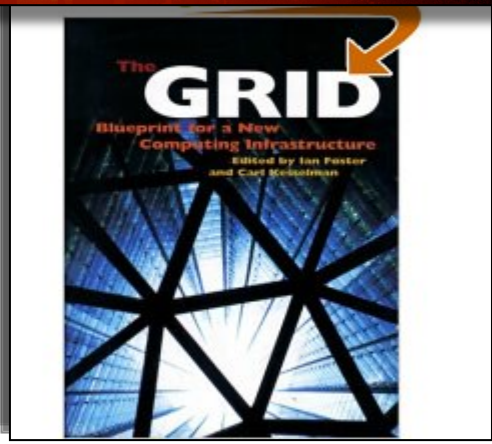
- Tier 0
  - prompt reconstruction
  - store RAW data and export to T1s
- Tier 1
  - re-recontruction
  - long term storage of RAW and MC data
- Tier 2
  - MC production
  - User analysis
- Tier 3
  - Mainly user analysis

Distributed Computing in Grids

- **"The Grid", Chapter 1, (Ian Foster & Carl Kesselman, 1998)**

Todays state of computing should be analogous to the status of development of electricity power to households at the beginning of last century. The true revolution was not electricity by itself, but the capability to distribute it over a network

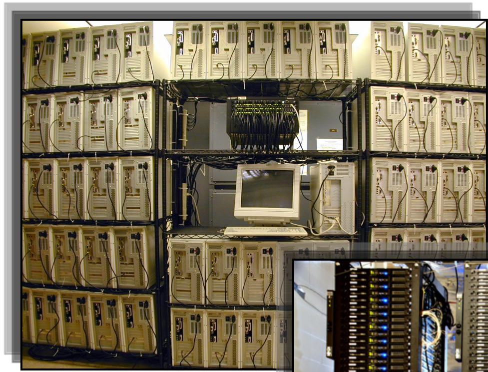
- **Access to computational resources (data storage and CPU should be as transparent as plugin a power switch**
  - The user does not need to know from where the electricity is generated
  - At the end of the service, you pay as you much as you have consumed

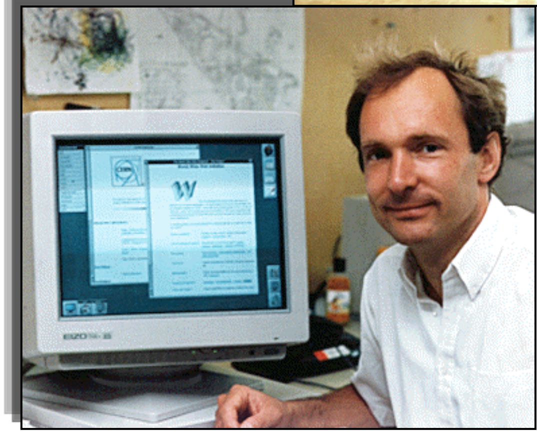


CSIC

Background & History

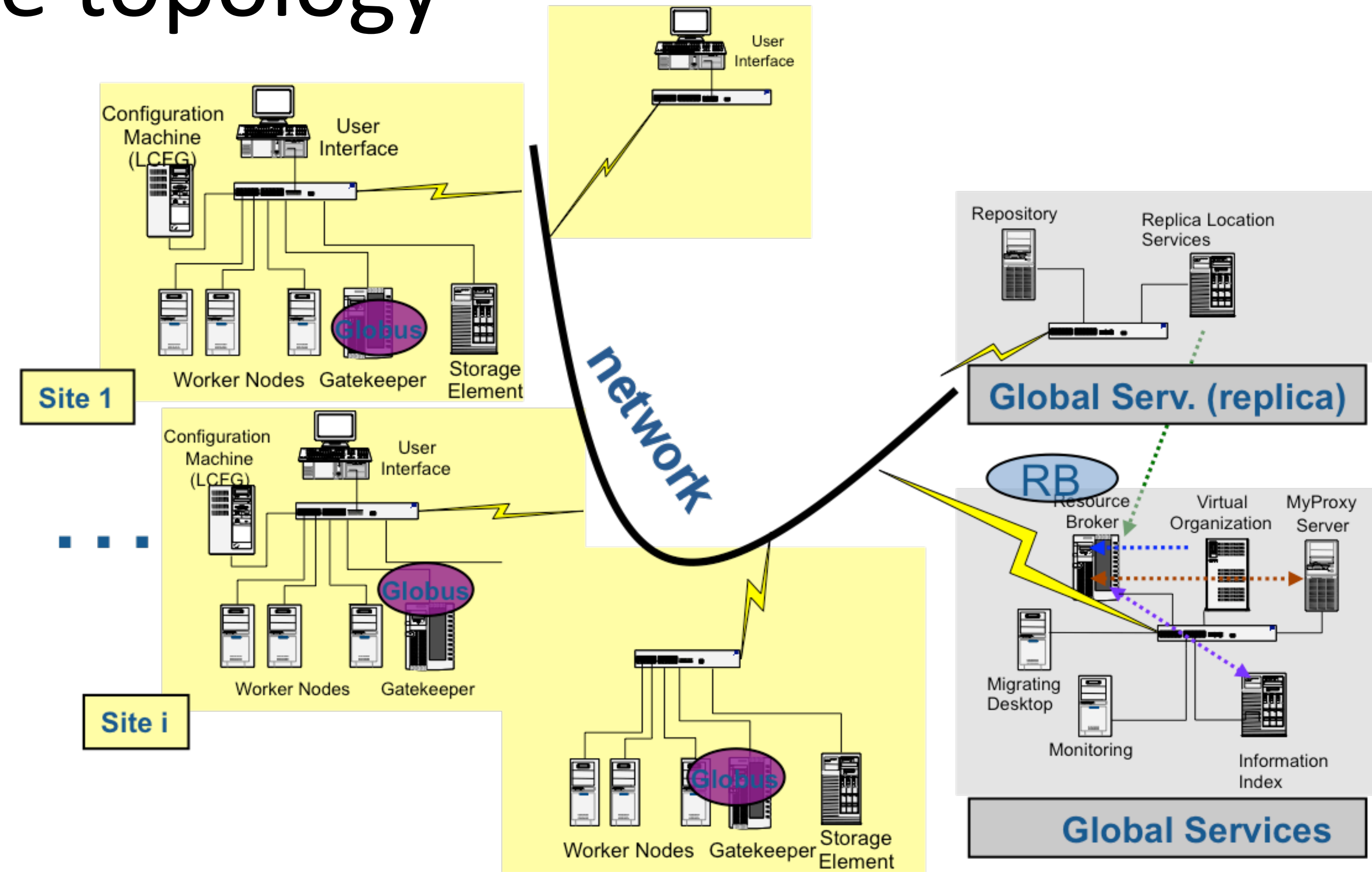
- **1998: the boom in commodity computing starts**
  - Hardware well prized due to massive production (Game industry!)
  - Communication networks becoming more reliable. Internet is starting to become the global source of information
  - Linux, and in general, GNU Software is setting down as de facto standard in the scientific world
    - Beowulf clusters multiply
- **The Globus project Globus Alliance:**
  - Development of software components to implement the Grid phylosophy (open source)
    - First release in 1998: Globus 1.0
    - Last release Feb. 2013: Globus 5.2.4
- Set of **software services and libraries**
  - **Remote management of distributed computing resources**
  - **Access to resources and distributed data**
- Environment for users to access remote resources and at the same time, resource owners keep full control on who access those resources





CSIC

The topology



N computing centers
(distributed in many countries)

Replicated global services (reliability)

CSIC

The Grid Infrastructure in numbers

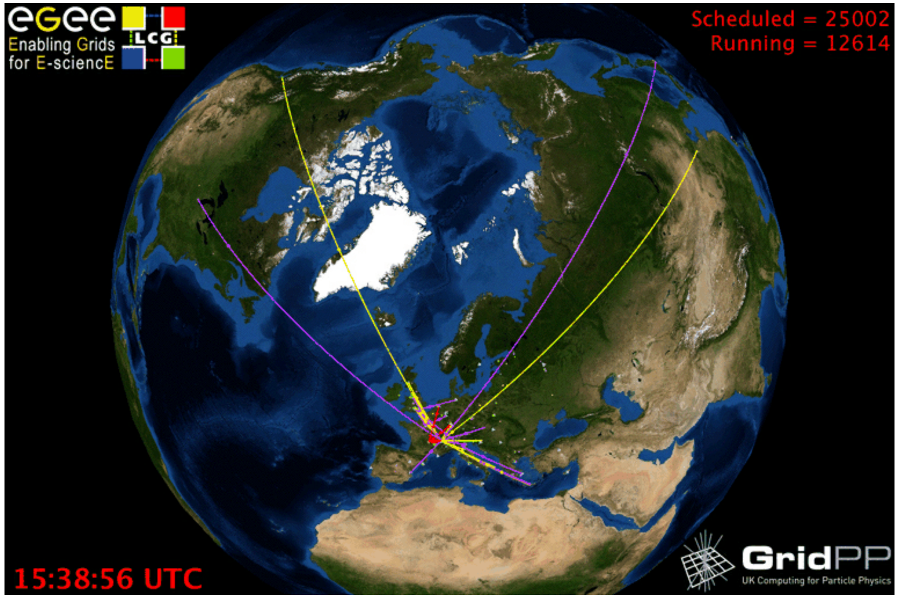
The largest computing Infrastructure in the world has been created to support the LHC Data analysis (2006-….

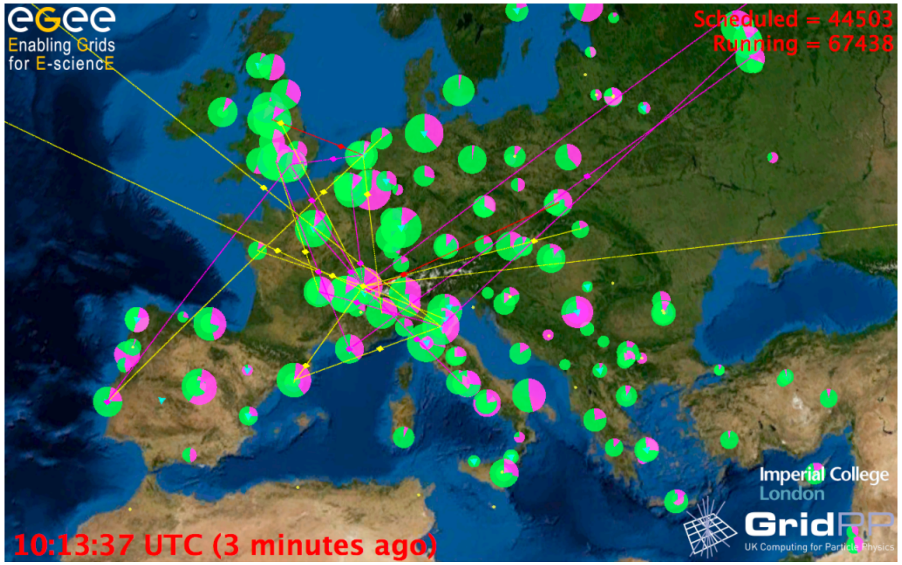
The Middleware used to “glue” the infrastructure is called **GLITE**:

It is based on Globus as base software

- Security system **gsi**,
- Data transfer using **gridftp**
- Job submission **globus-job-run**

- **GLITE contains enhancements to deal with the massive data and job distribution which imply the LHC data**





CSIC

RAW data -> ESD -> AOD -> Statistical Data

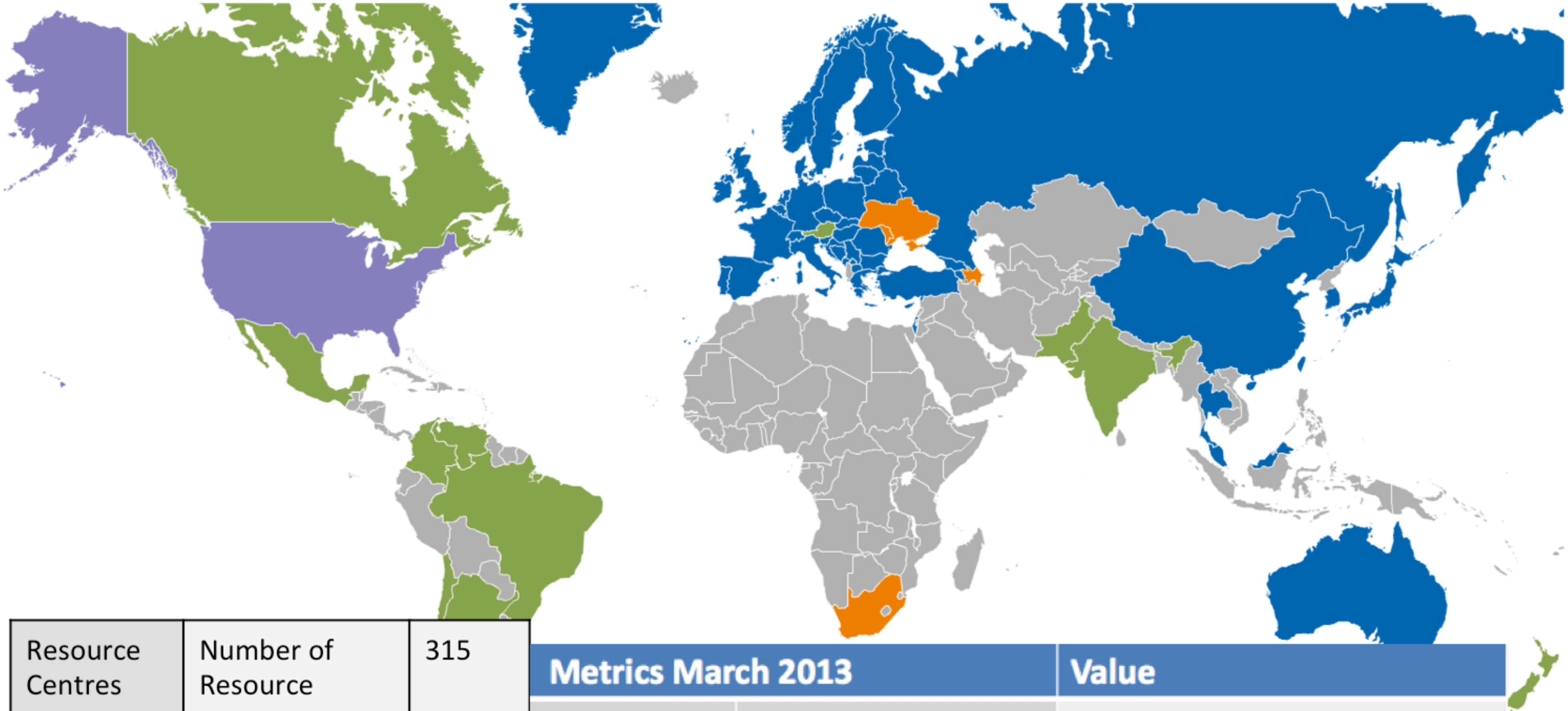
- **Raw data**
  - hits, pulse heights
- **Reconstructed data (ESD)**
  - tracks, clusters…
- **Analysis Objects (AOD)**
  - Physics Objects
  - Summarized
  - Organized by physics topic
- **Ntuples, histograms, statistical data**

Tier-0

Tier-1(s)

Tier-2(s)

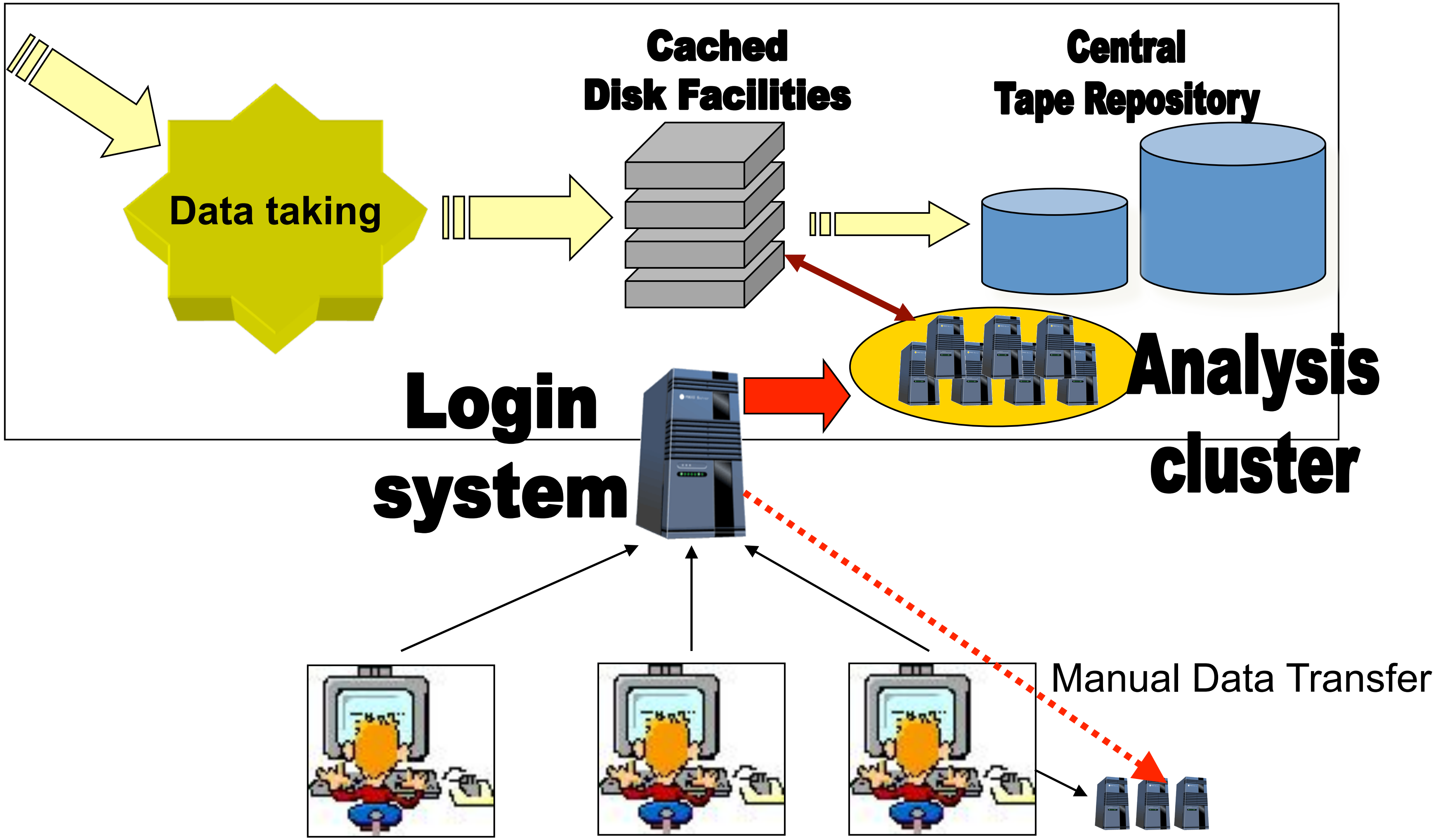
The Grid Infrastructure in numbers



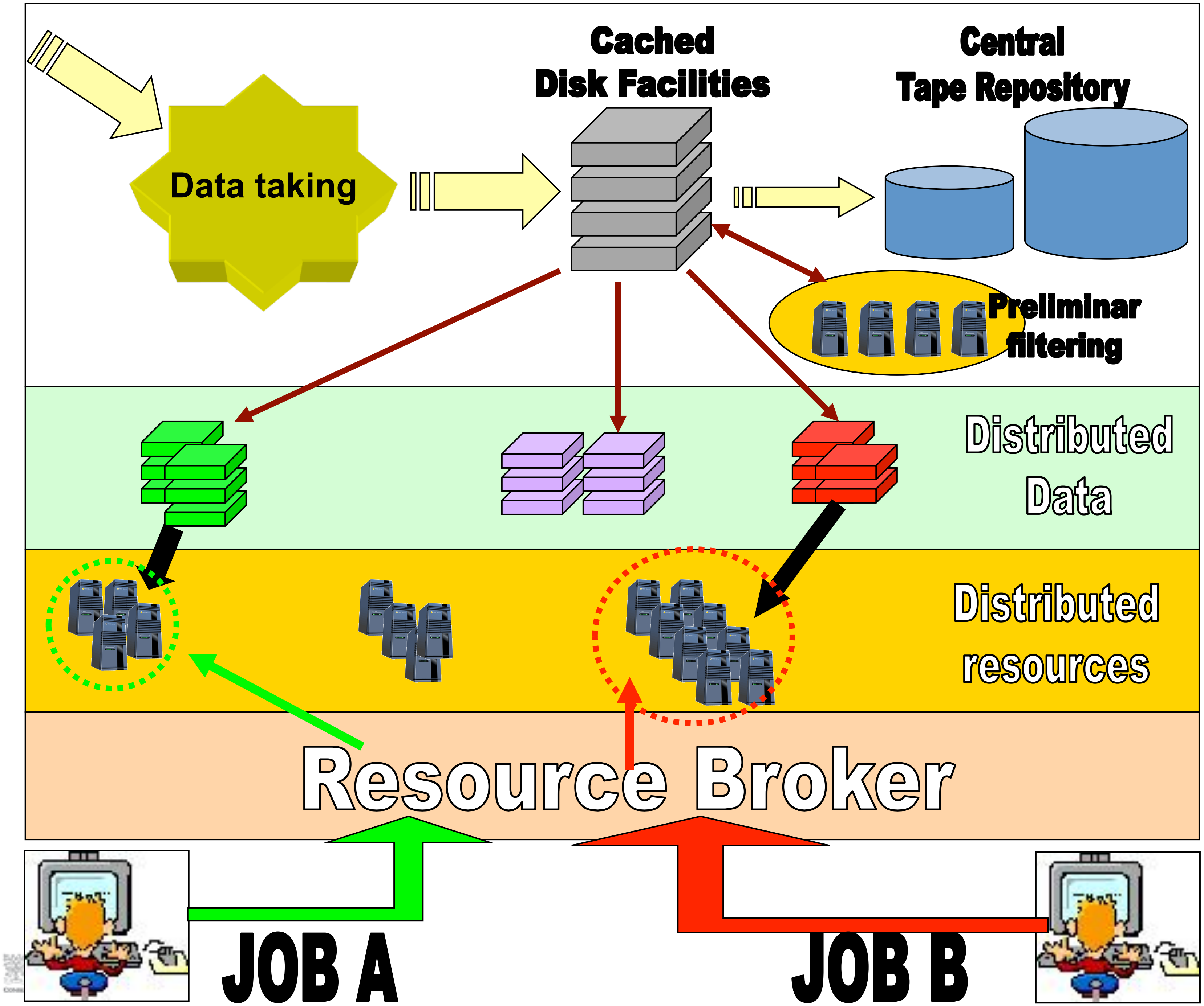
Resource Centres	Number of Resource Centers	315
Countries	Number of European Countries	42
	Including associated countries	54

Metrics March 2013		Value
Capacity	CPU cores (EGI and integrated resource providers) Disk/Tape (PB) Countries	372,612 (315 resource centres) 180/167 56
Jobs	Average Job/day (Million)	1.67 (2.25 including local computation)

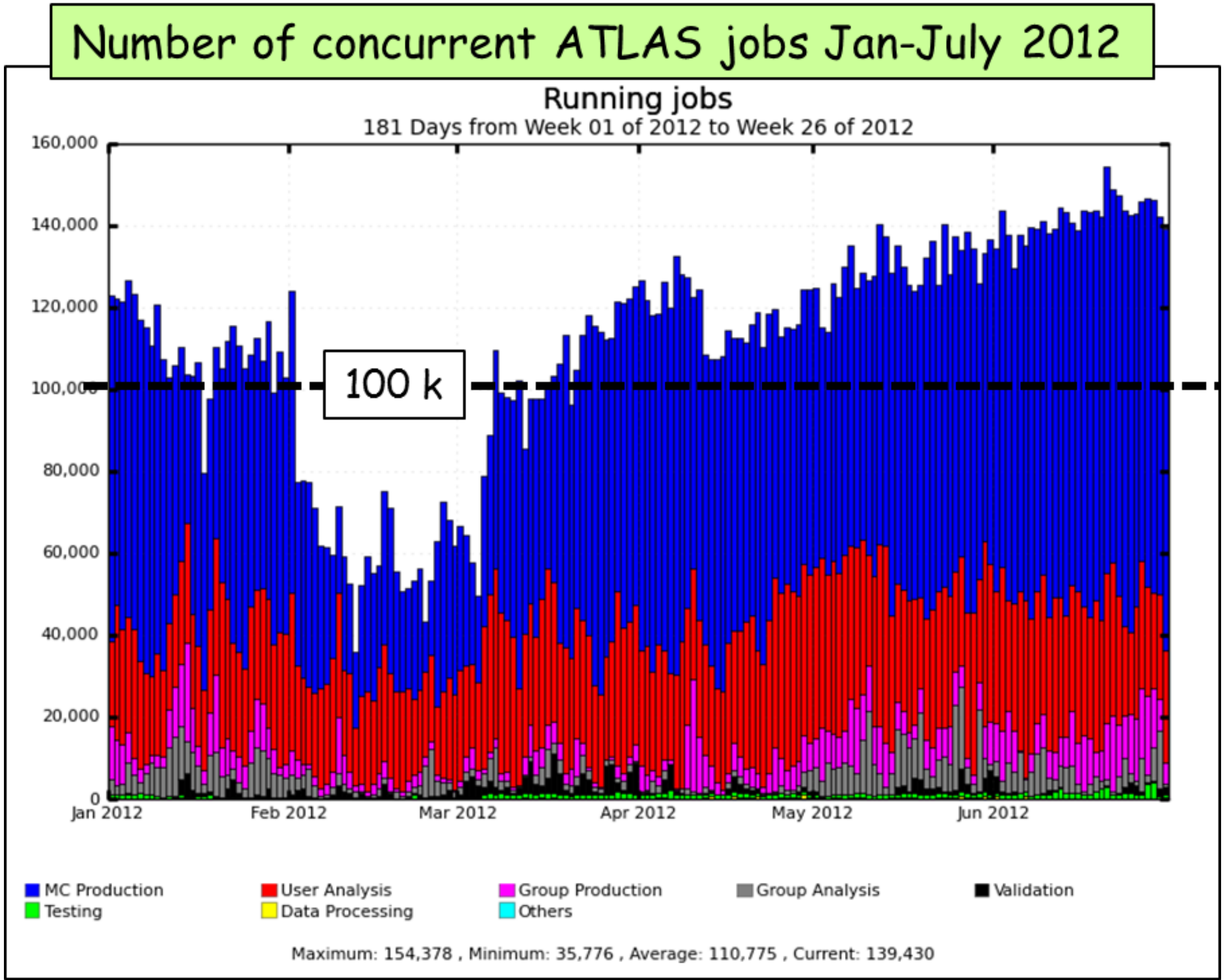
Without Grid



CSIC


Data taking
Cached
Disk Facilities
Central
Tape Repository
Preliminar
filtering
Distributed
Data
Distributed
resources
Resource Broker
JOB A
JOB B

Data are analized in Parallel



CSIC

Grid Middleware

Basic Elements of a Grid Site

- **Computing Element**
  – Standard Interface to access the local CPUs from the Grid
- **Storage Element**
  – Standard Interface to access the locally available storage system
- **Resource Broker**
  – Tool to analyze user job requests (input data sets, cpu time, data output requirements) and route there to sites according to data and cpu time availability
- **Data Placement is very important: jobs run where data is**
- **Experiments do not know a priori which dataset will be popular**
  – CMS has 8 orders of magnitude in access between most and least popular
  – **Dynamic data replication** create copies of popular datasets at multiple sites

CSIC

Grid Middleware – Data Distribution
File Transfer System (FTS)

- **FTS is the service responsible for distributing the majority of LHC data across the WLCG infrastructure (transferred 25PetaBytes in 2012)**
- Used by the Data Management teams
  - Rarely by single users
- It allows the participating sites to control the network resource usage
- Users interact with FTS by submitting transfer jobs, that simply say

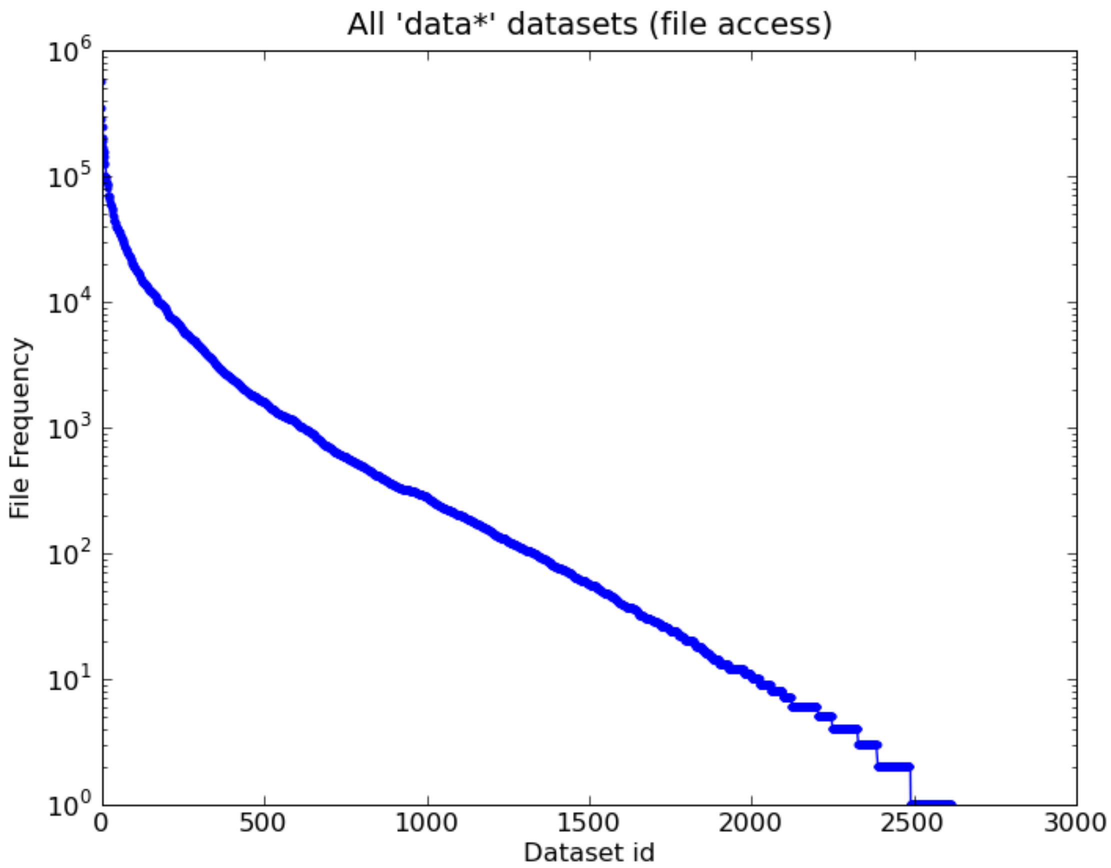
  "copy <source URL> to <destination URL>

  - FTS then queues, schedules and performs the transfer, retrying it if necessary
- Once data are distributed, **a file catalogue is needed to record which files are available where (at which Tier-1 or at which Tier-2 ?)**
- The **LCG File Catalogue (LFC)** was designed as a distributed catalogue (to avoid single points of failure)
  - All the **Tier-2 have a copy of the LFC** relative to their experiment

CSIC

We do not know a priori which data will be popular

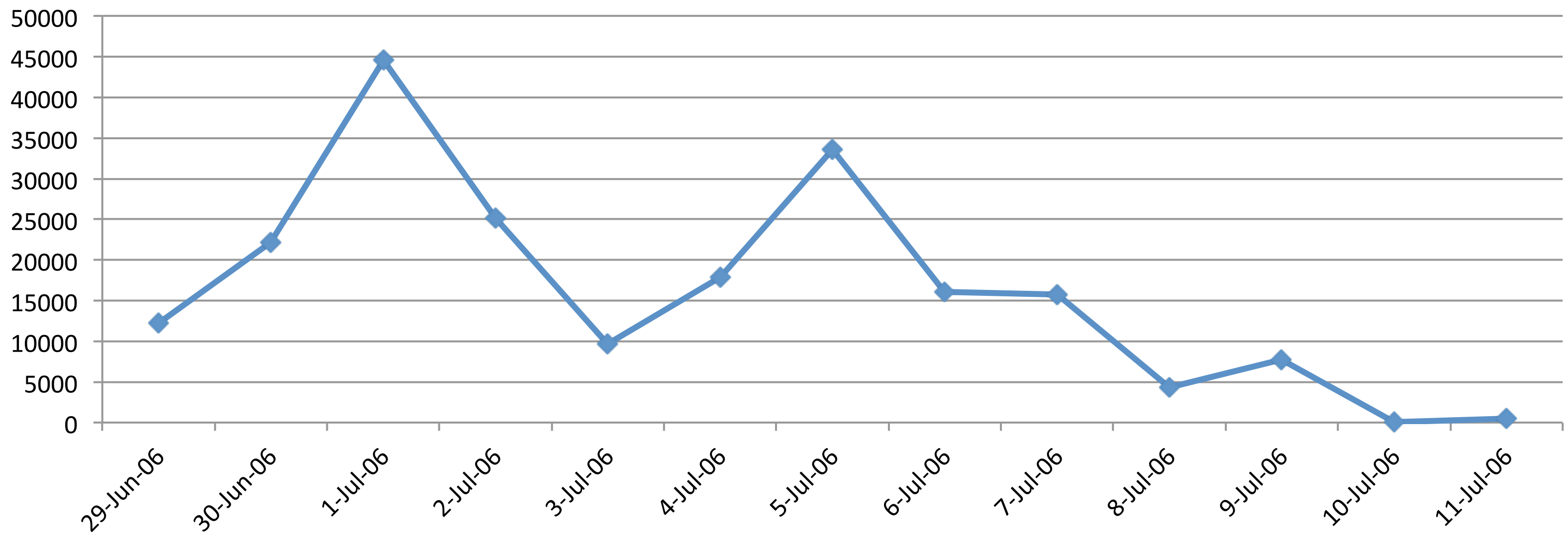
- A small fraction of the data we distribute is actually used
- Data* datasets
- Counts dataset access
- Only by official tools
- There are ~200k datasets



Datasets are popular for a very short time

- Data is popular for a very short time
- Dataset: data10_7TeV.00158116.physics_L1Calo.recon.ESD.f271
- Dataset Events: 99479
- Replicas: 6, Files: 6066, Users: 35, Dataset Size: 17.1 TB

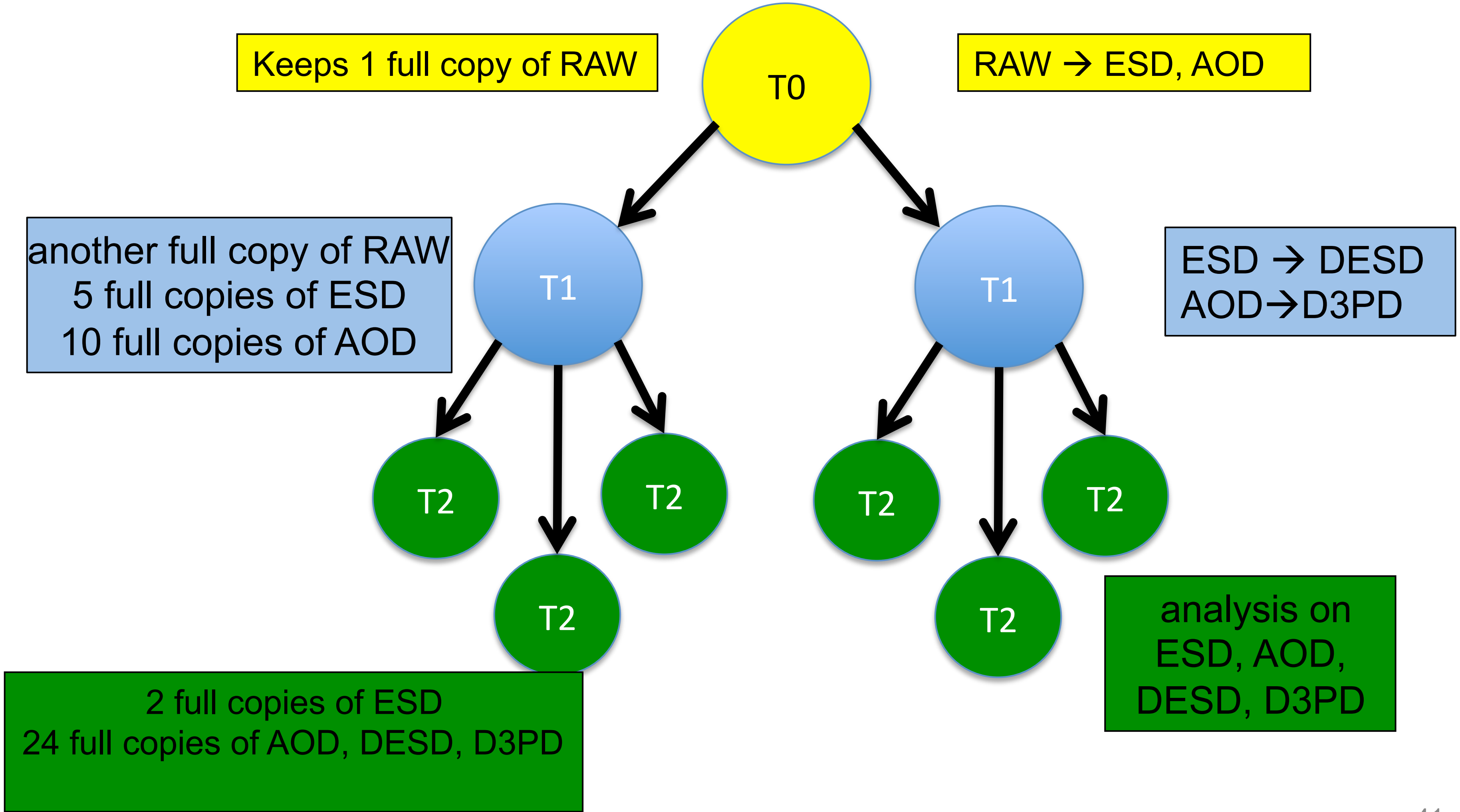
File Access



Note: Search was for the last 120 days, but only used for 13 days

Data placement model

Starting with 2 PB of RAW from the detector we end up with 14 PB of derived data for analysis.Very many copies in Tier-1's and Tier-2's to allow efficient analysis

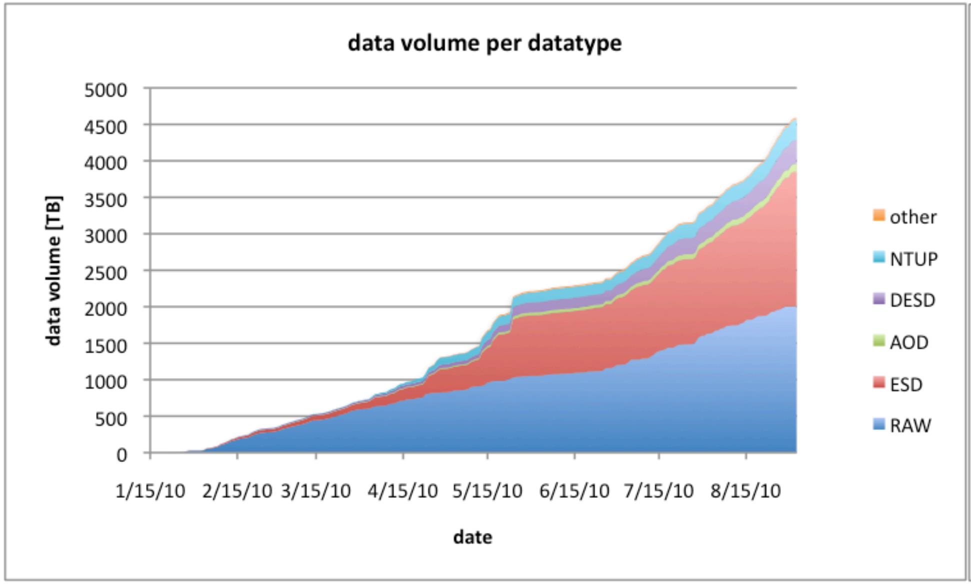


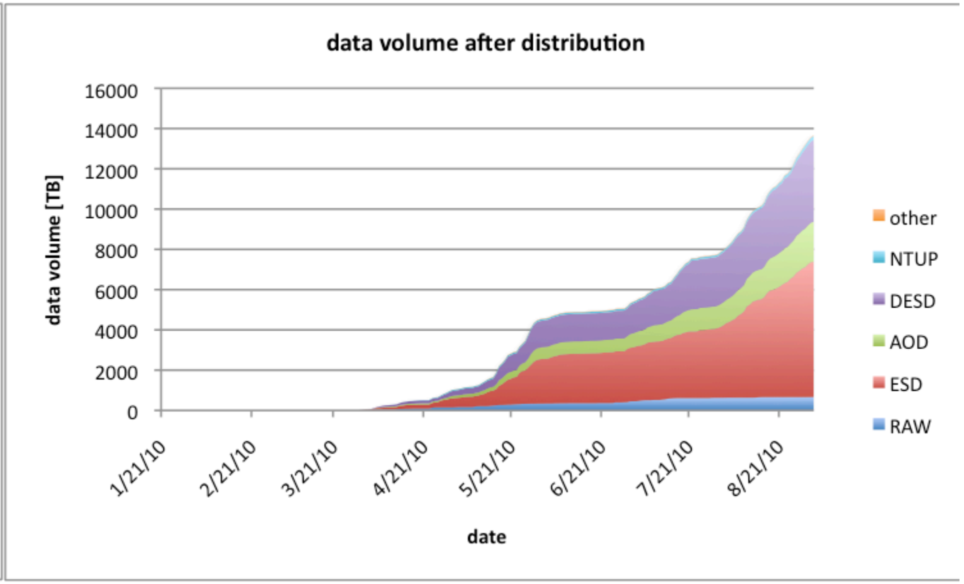
Influence of the Network

- Starting with 2 PB of RAW from the detector
- We end up with 14 PB of derived data for analysis (ignoring simulated data)
- Very many copies in Tier-1's and Tier-2's to allow efficient analysis

Caching data instead

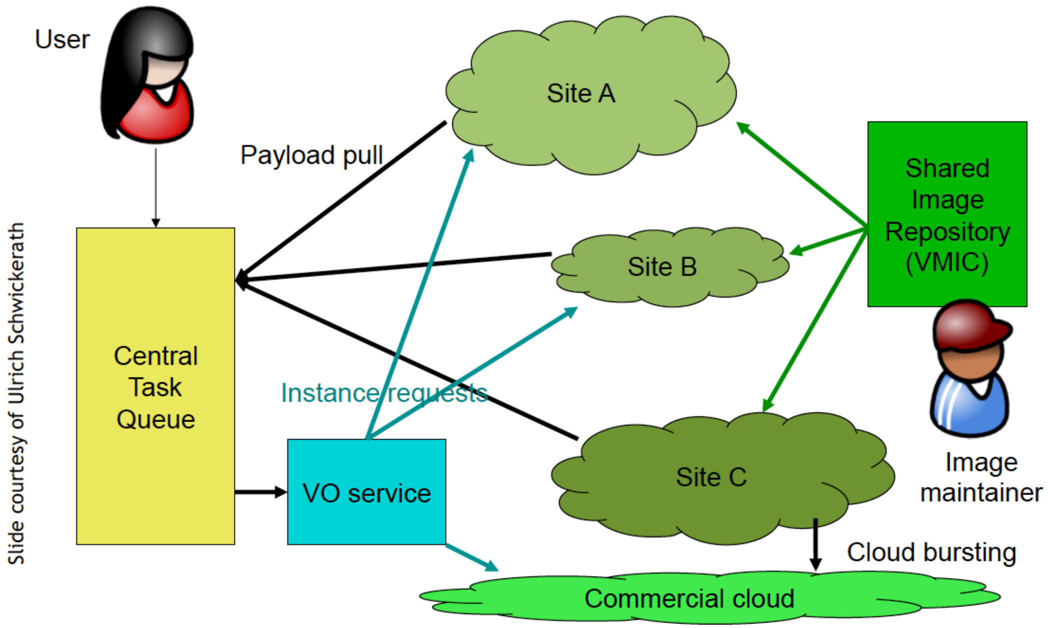
- With a well performing network we could do as well with fewer copies
- Download data needed for analysis → automatic selection of popular data
- Possibility to use Tier-0 and Tier-1's and Tier-2's as data source





Conclusions

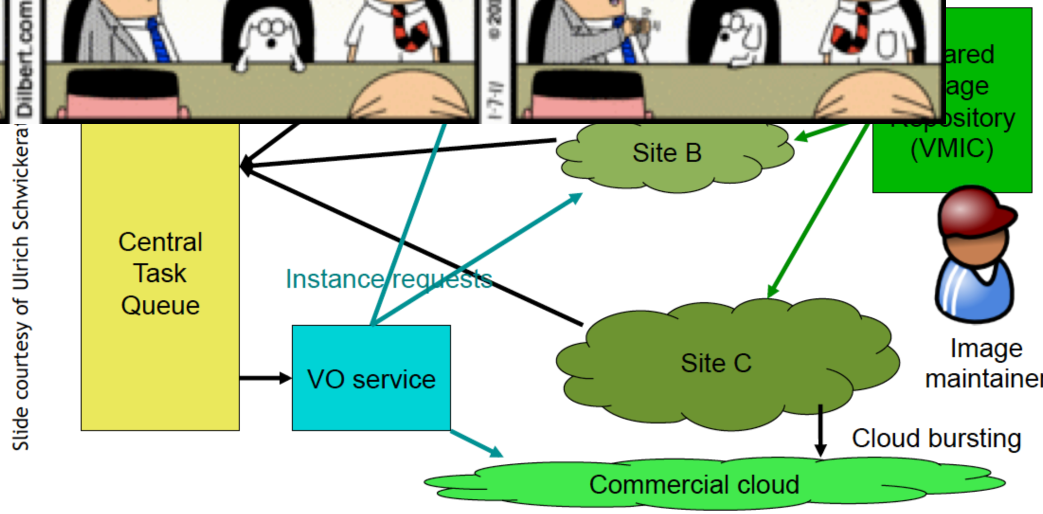
- Data analysis at Large Scientific Experiments is an **organizational (human) and technical challenge**
  - It has pushed the frontiers of physics and computing
- **Constant evolution** due to the improvements in
  - Deployment of Virtual Resources: the famous cloud
  - Network improvement: changes the model of computing-data movement



CSIC

Conclusions

- Data analysis at Large Scientific Experiments is an **organizational (human) and technical challenge**
  - It has pushed the frontiers of physics and computing
- **Constant evolution** due to the improvements in
  -
  -



CSIC