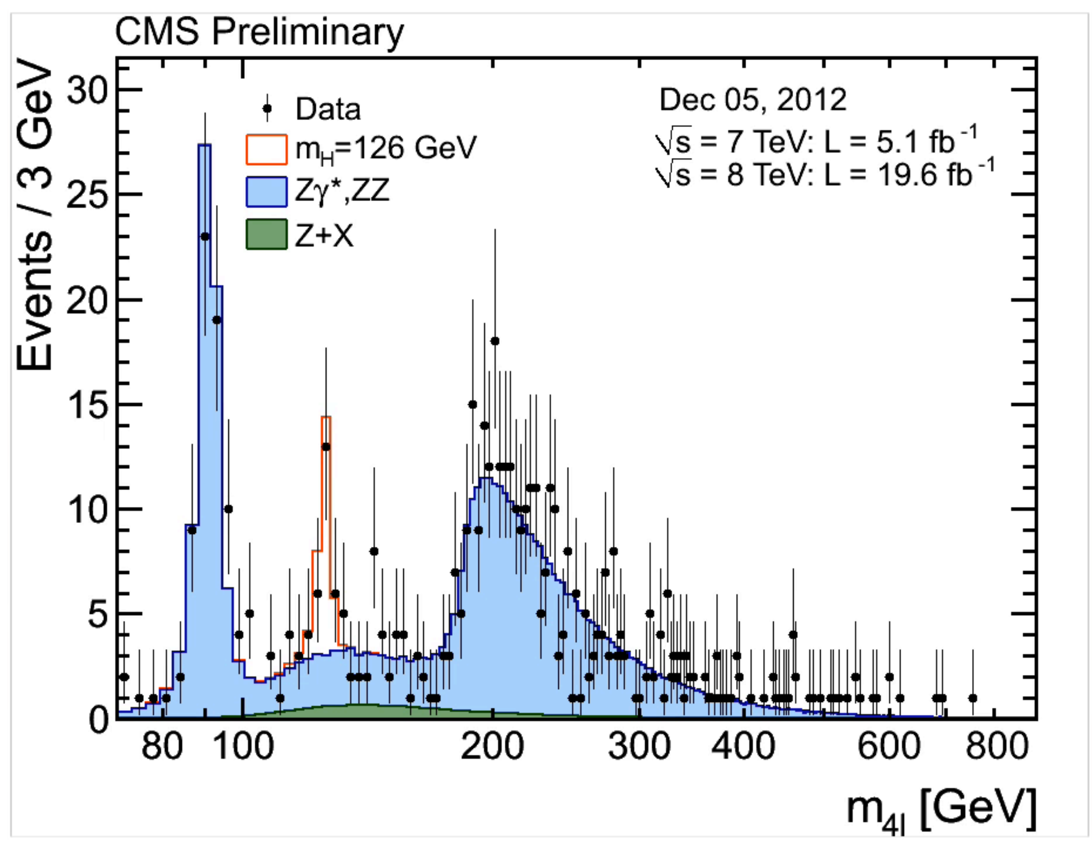
Computing: From Data to Physics



Click for animation: https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsHIG/HZZ4l_date_animated.gif

23. January 2014

Oliver Gutsche (Fermilab)

Computing and the Higgs

The Nobel Prize in Physics 2013
François Englert, Peter Higgs

The Nobel Prize in Physics 2013

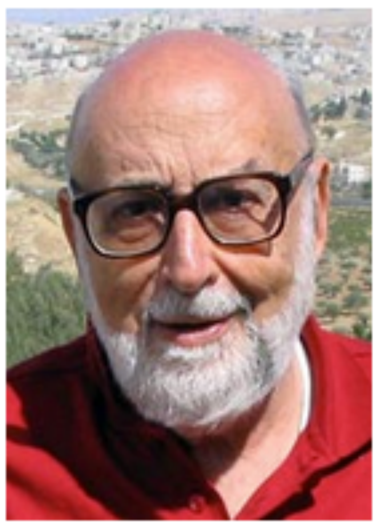


Photo: Pnicolet via Wikimedia Commons
François Englert

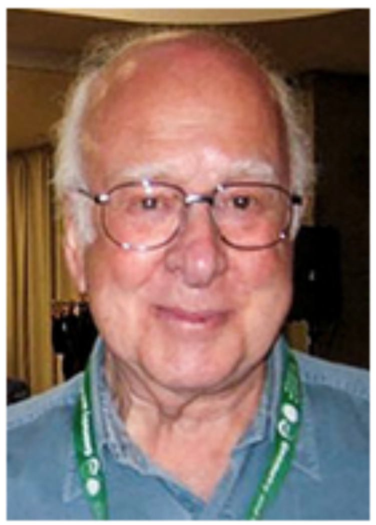
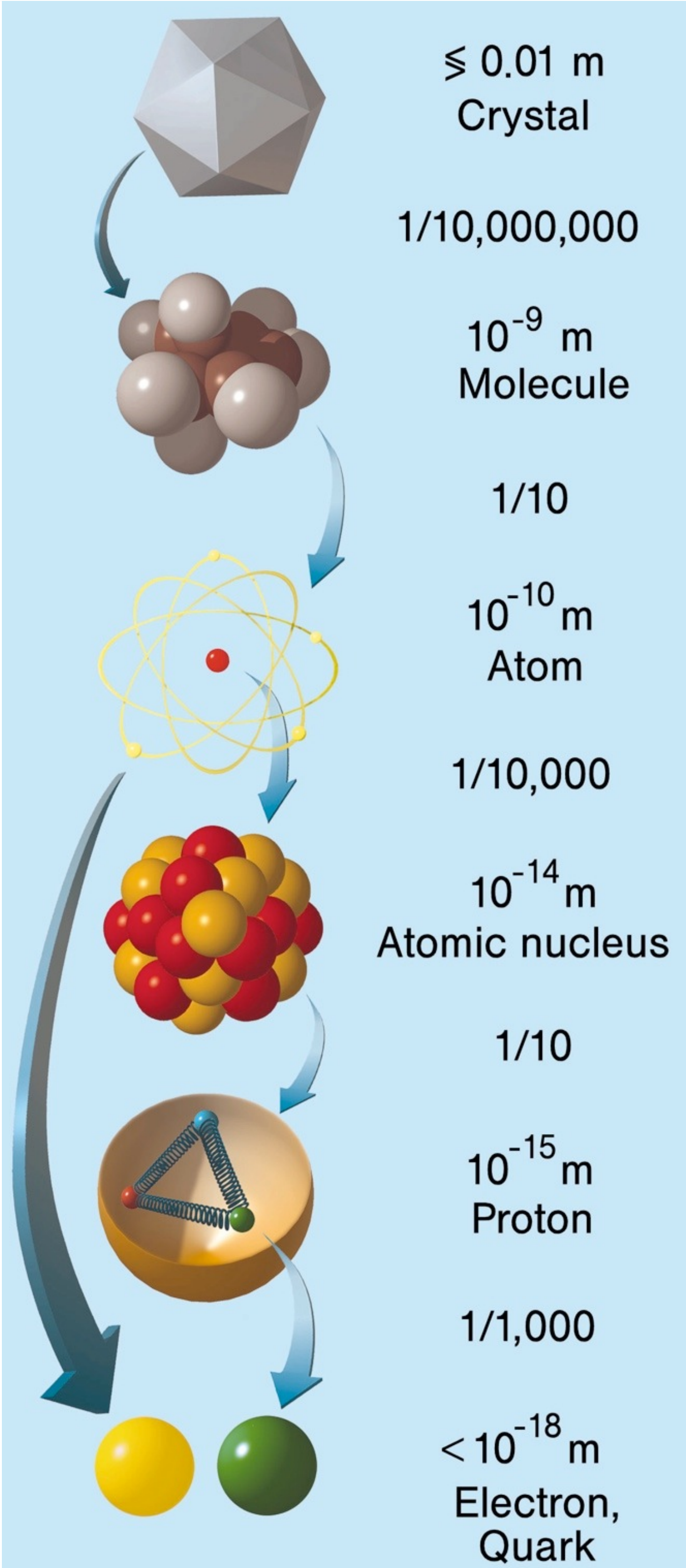
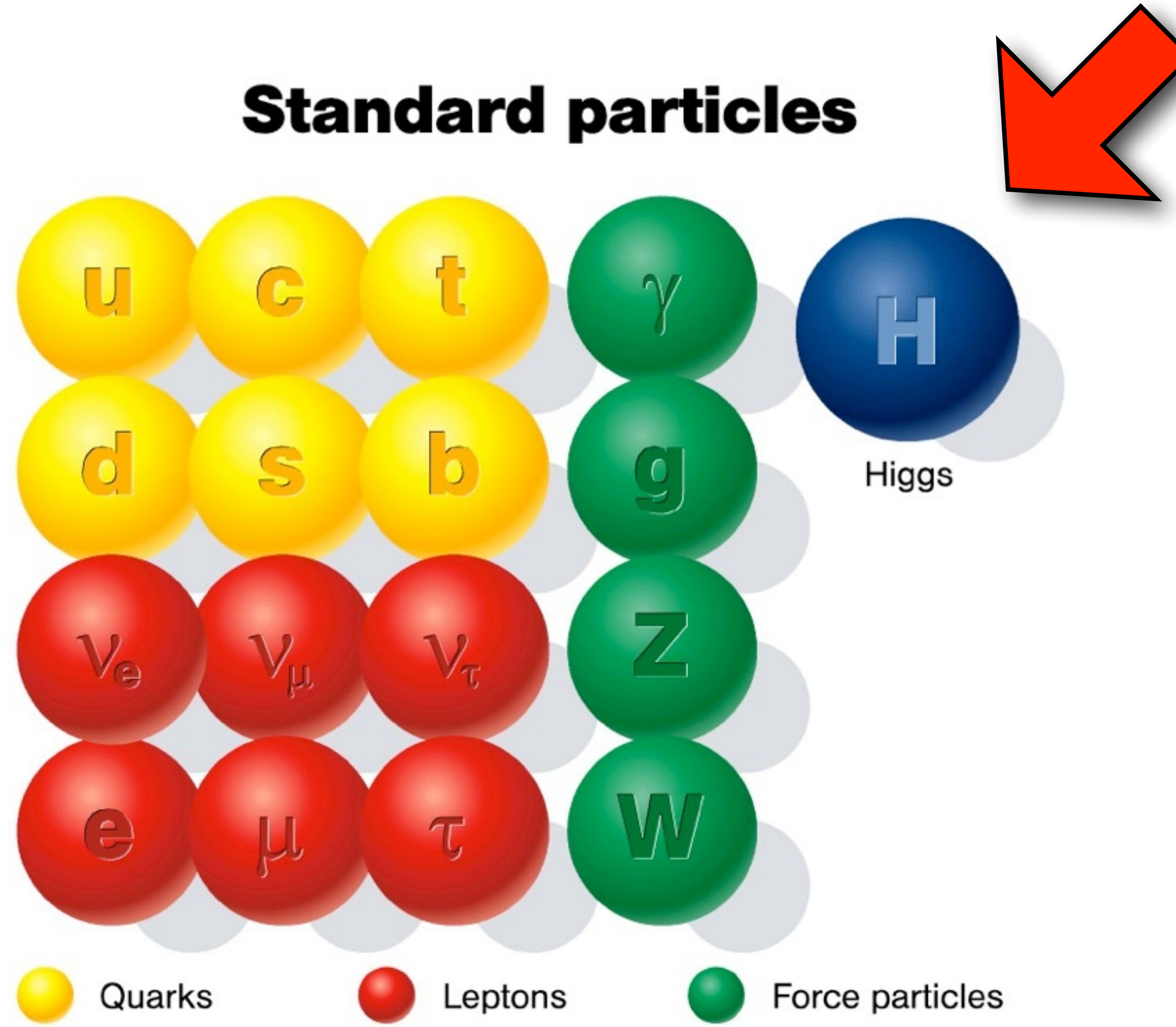


Photo: G-M Greuel via Wikimedia Commons
Peter W. Higgs

The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*

- The Higgs was discovered by the 2 experiments Atlas and CMS at the Large Hadron Collider (LHC) in Geneva, Switzerland
- Computing played a very important role in the discovery of the Higgs
- Today's question:
  - From Data to Physics, how did computing enable the Higgs discovery?

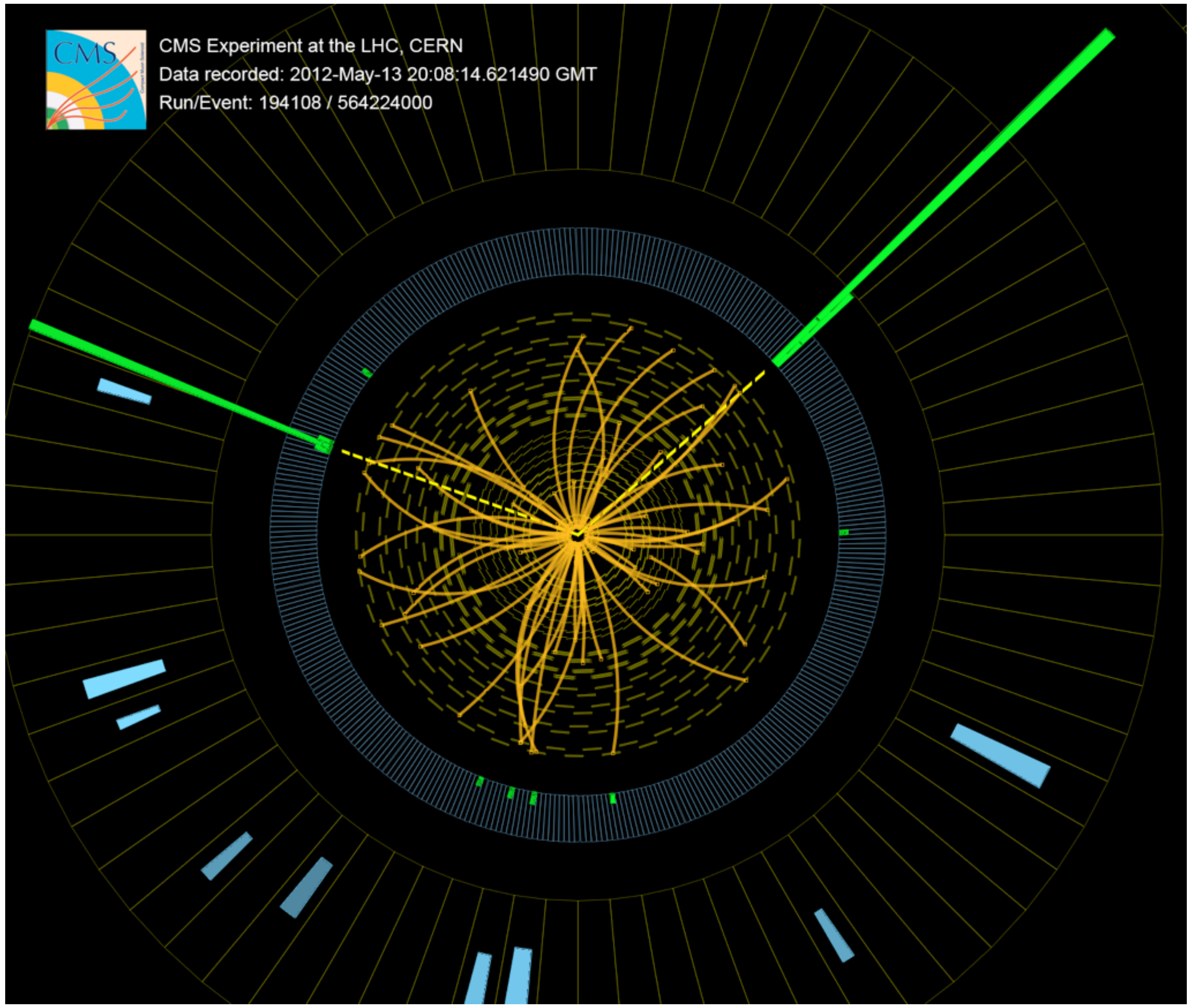

≲ 0.01 m
Crystal
1/10,000,000
10^{-9} m
Molecule
1/10
10^{-10} m
Atom
1/10,000
10^{-14} m
Atomic nucleus
1/10
10^{-15} m
Proton
1/1,000
< 10^{-18} m
Electron,
Quark


Standard particles
u
c
t
γ
H
d
s
b
g
Higgs
ν_e
ν_μ
ν_τ
Z
e
μ
τ
W
Quarks
Leptons
Force particles

- The Higgs is both a particle and a field.
- The Higgs field permeates the whole universe.
- The Higgs interacts with all particles, the stronger the interaction, the heavier the particle

How can we detect the Higgs?

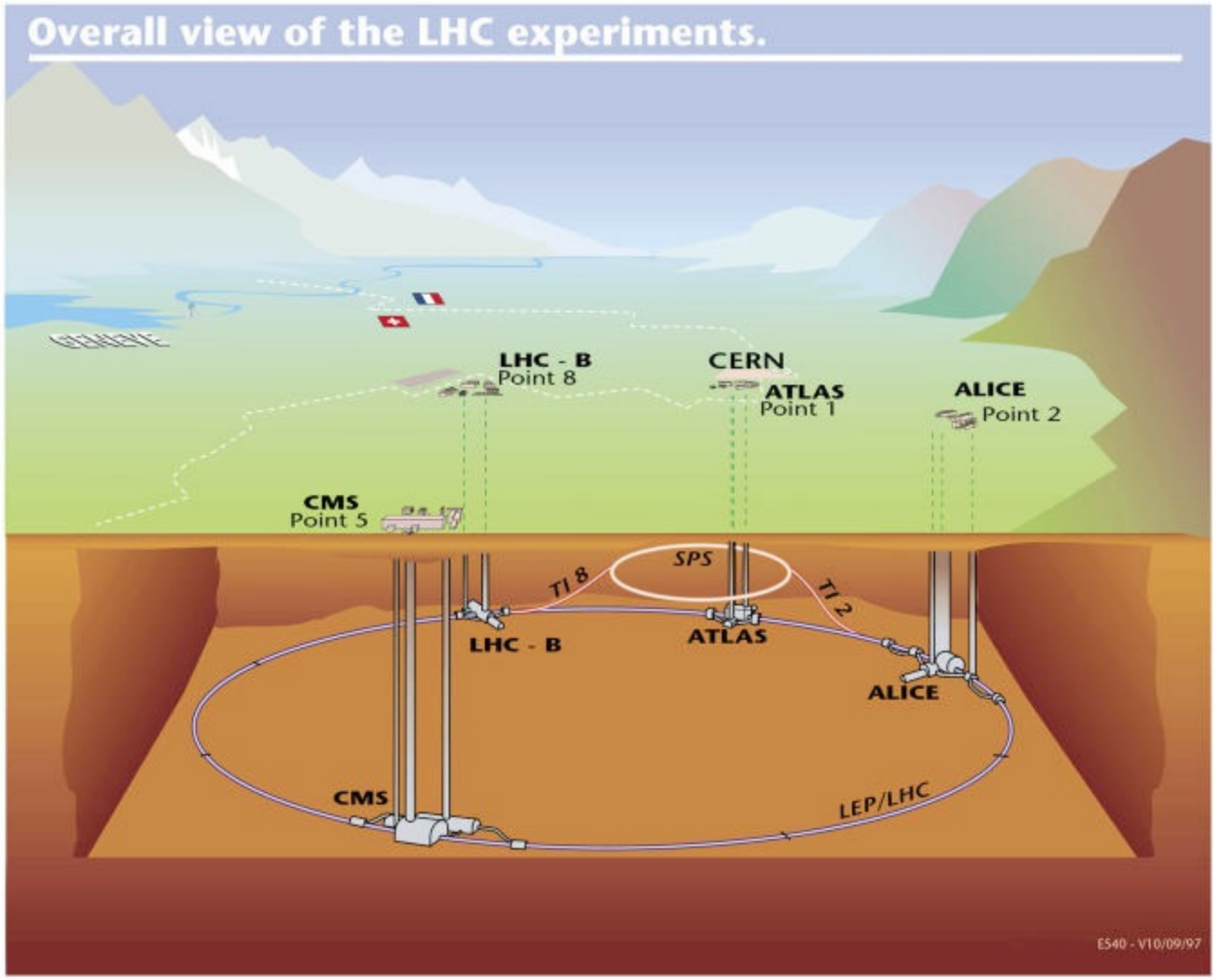
- We literally have to knock a Higgs particle out of space
  - We need a lot of energy concentrated in a very small volume of space.
- After being knocked out of space (produced), the Higgs vanishes again very quickly (decays into other particles)
- We cannot detect the Higgs itself, we only can detect particles that came from the Higgs particle



- Therefore we need
  - Particle accelerators for the energy
  - Particle detectors for the detection

Large Hadron Collider (LHC)

How to produce enough energy to knock a Higgs particle out of space → LHC

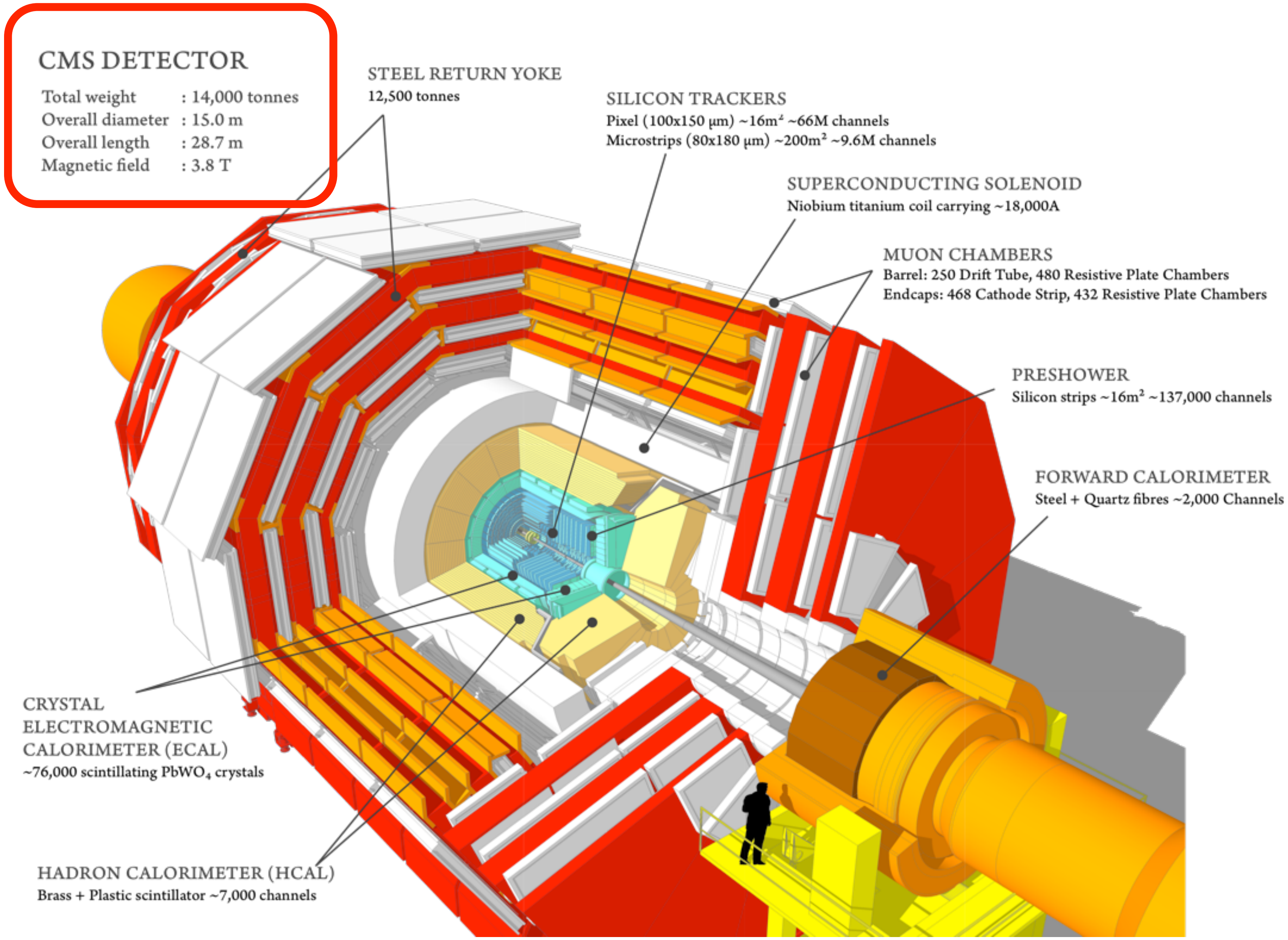


- Circumference: almost 17 Miles
- 2 proton beams circulating at 99.9999991% of the speed of light
- A particle beam consists of bunches of protons (100 Billion protons per bunch)
- Beams cross and are brought to collision at 4 points
  - 20 Million collisions per second per crossing point
- Energy stored in one LHC beam is equivalent to a 40t truck crashing into a concrete wall at 90 Mph

LHC guide: http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf

Compact Muon Solenoid (CMS)

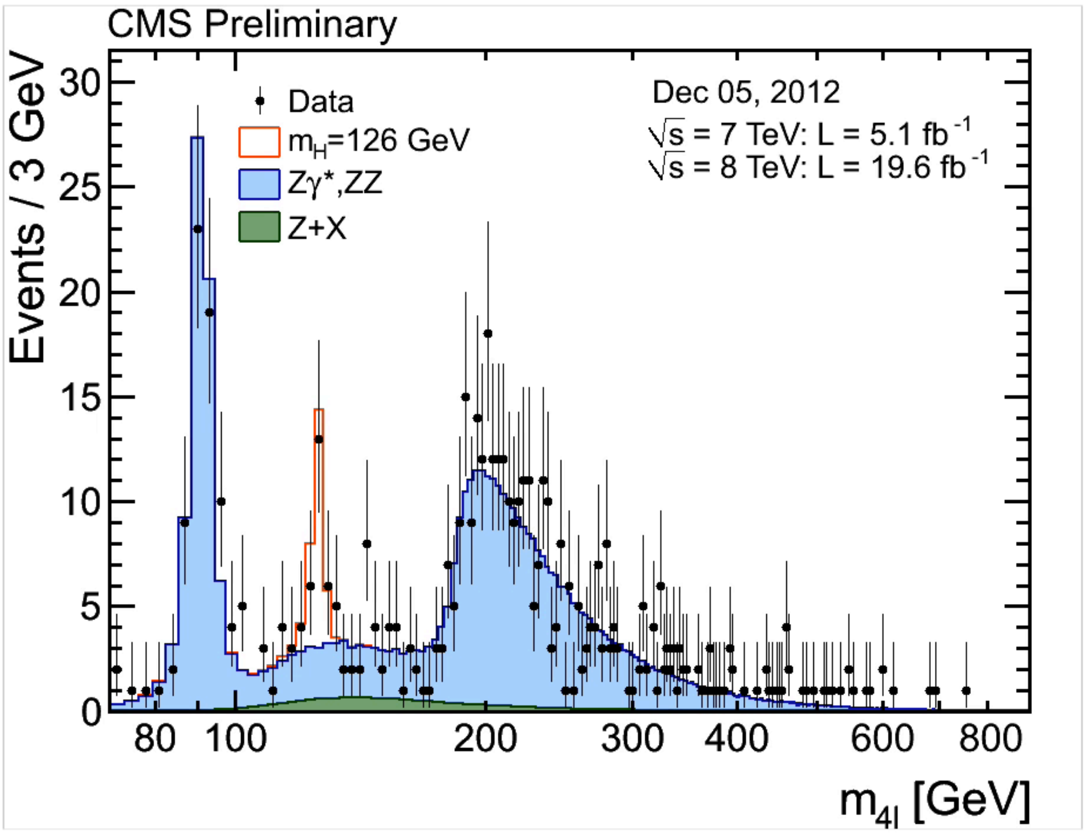
How do we measure the particles coming from the Higgs → CMS detector



- Detector built around collision point
- Records flight path and energy of all particles produced in a collision
- 100 Million individual measurements (channels)
- All measurements of a collision together are called: **event**

This is how we discovered the Higgs

Click for animation: https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsHIG/HZZ4l_date_animated.gif

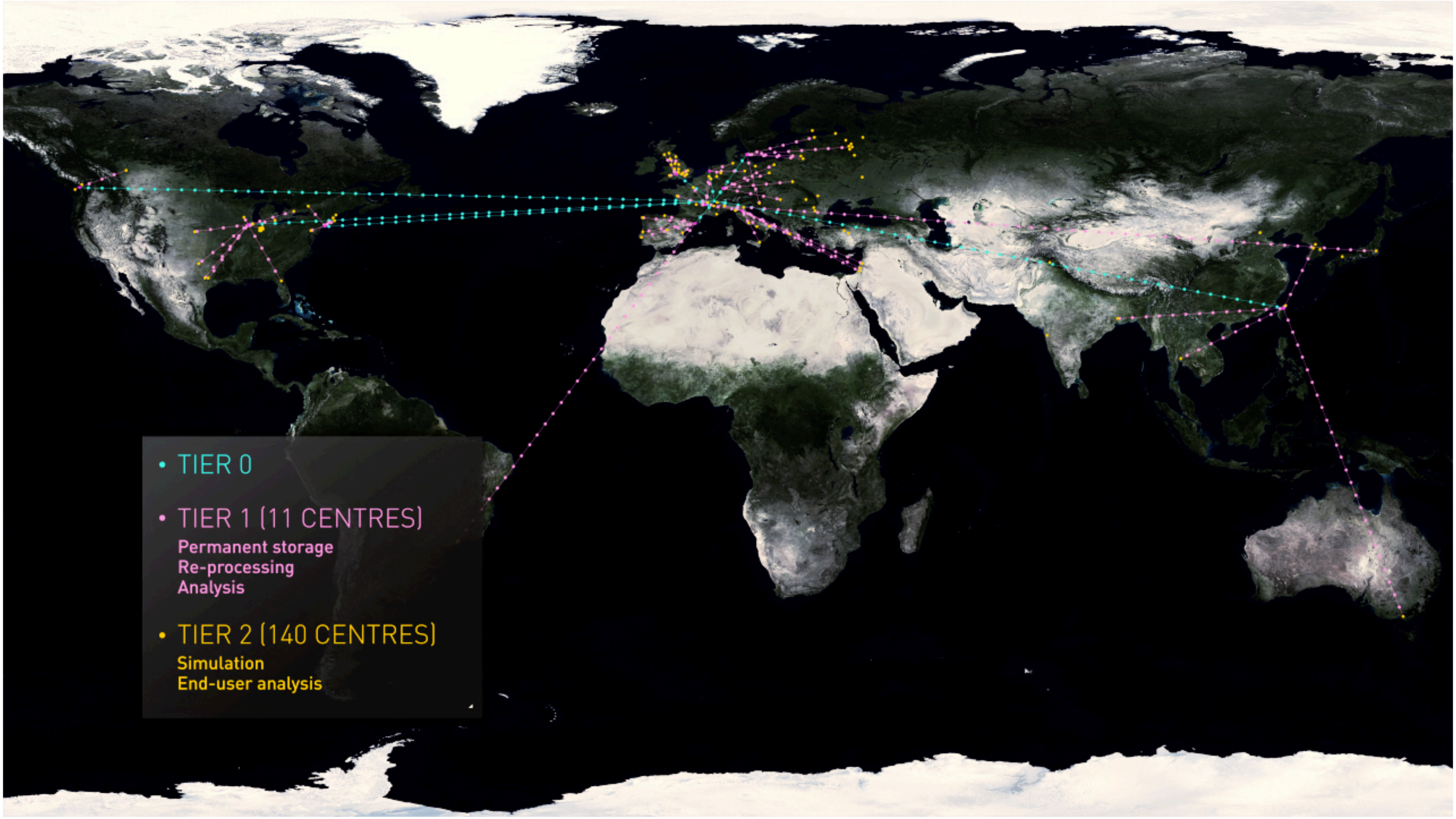


- We needed to
  - Record many Billions of events
  - Identify the Higgs events from all the events we collect with the detector
    - "Finding a needle in a haystack"

During 2010/2011 and 2012, we collected more and more events and finally collected enough to discovery the Higgs particle

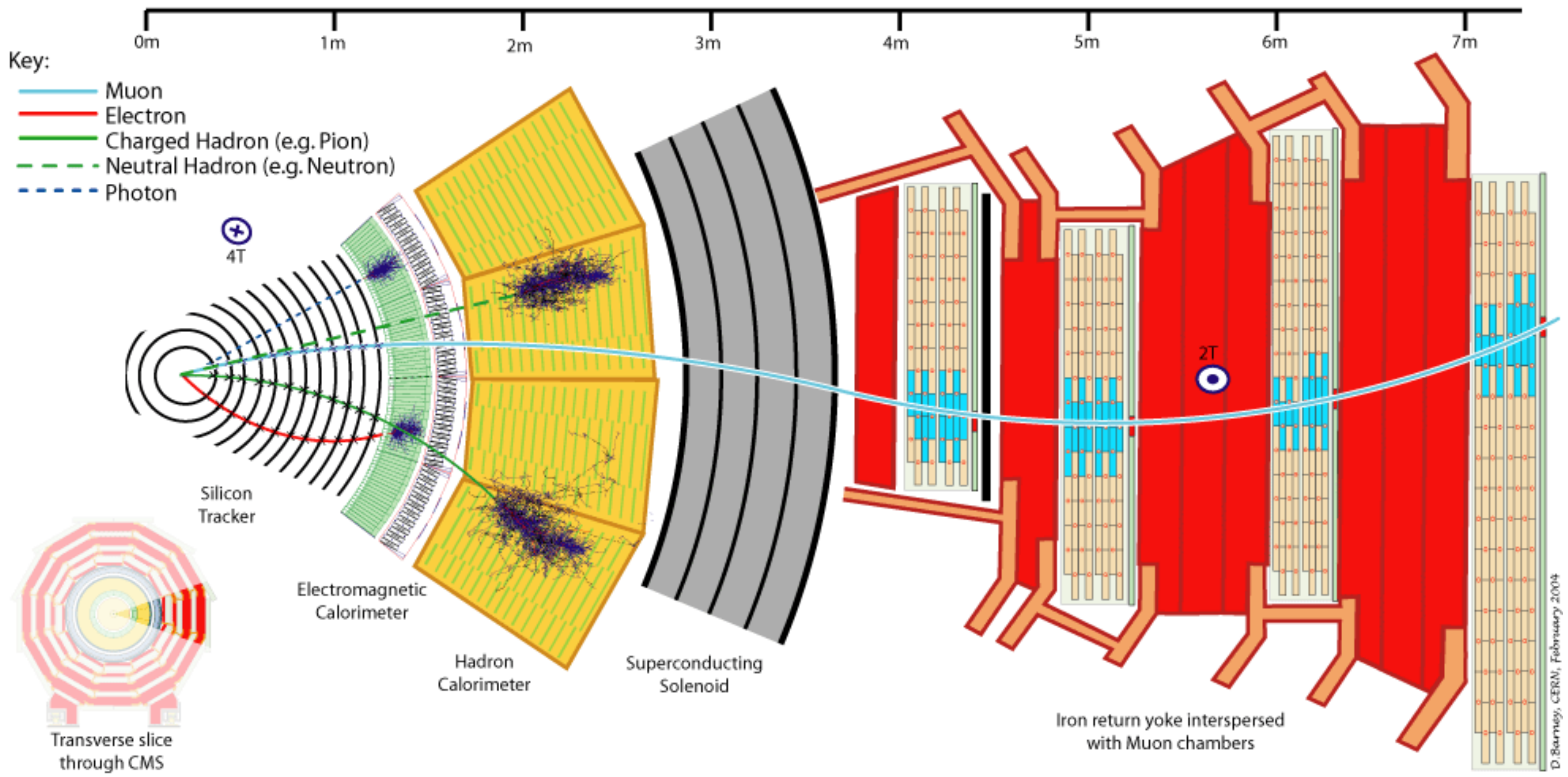
- For all these steps, we needed a lot of computing to make this possible!

Movie: Computing for LHC/CMS



Introduction movie to LHC/CMS computing: http://cds.cern.ch/record/1541893?ln=en

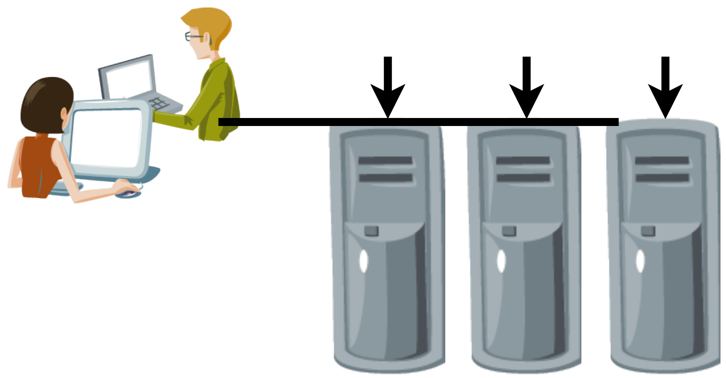
Reconstructing a collision



- The detector records digital signals from its detector components for each event
  - Goal is to measure flight path and energy of all particles produced in a collision
- Detector signals have to be reconstructed to give them meaning, for example:
  - The tracking detector consists of a 3D mesh of small detection points in space, a particle produces a signal in the points it passes through
  - Software uses these points to reconstruct the path of a particle through the detector
- A lot of computing power is needed to reconstruct Billions of events
  - Reconstruction of a single event takes 10-30 seconds
  - **The big advantage: every event can be reconstructed individually → this enables us to process events in parallel**

From Single Computer to the GRID

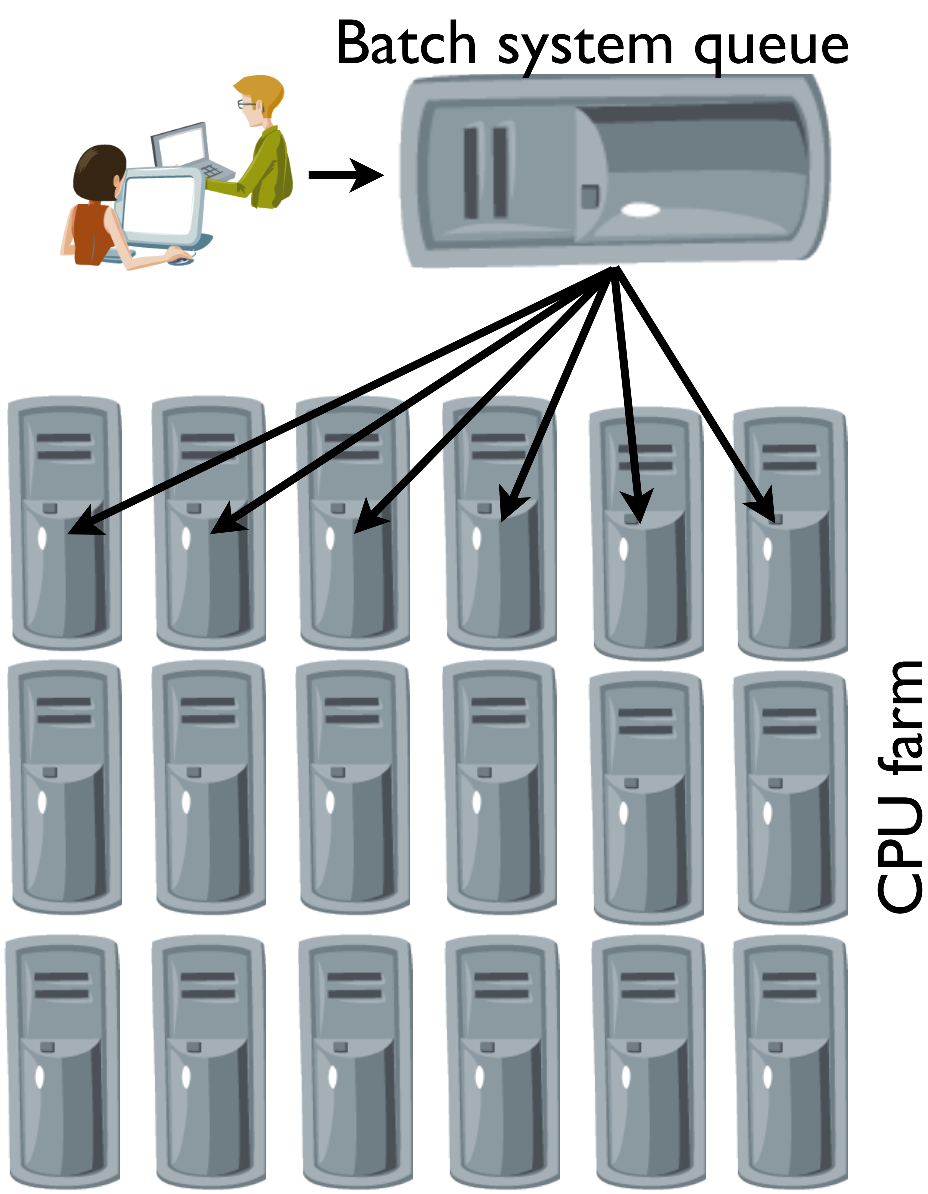
1. Start with one PC to reconstruct events
    - Modern PCs have a CPU (Central Processing Unit, the main chip in the PC) with multiple cores: in particle physics, we currently can process one event in parallel per core
  - Caveat: you need to manually start the reconstruction program for each core → **reconstructing Billions of events takes a long time and is work intensive**



2. Buy 2 more PCs to be faster
  - You have to start programs on each of the PCs manually
  - Even more work intensive

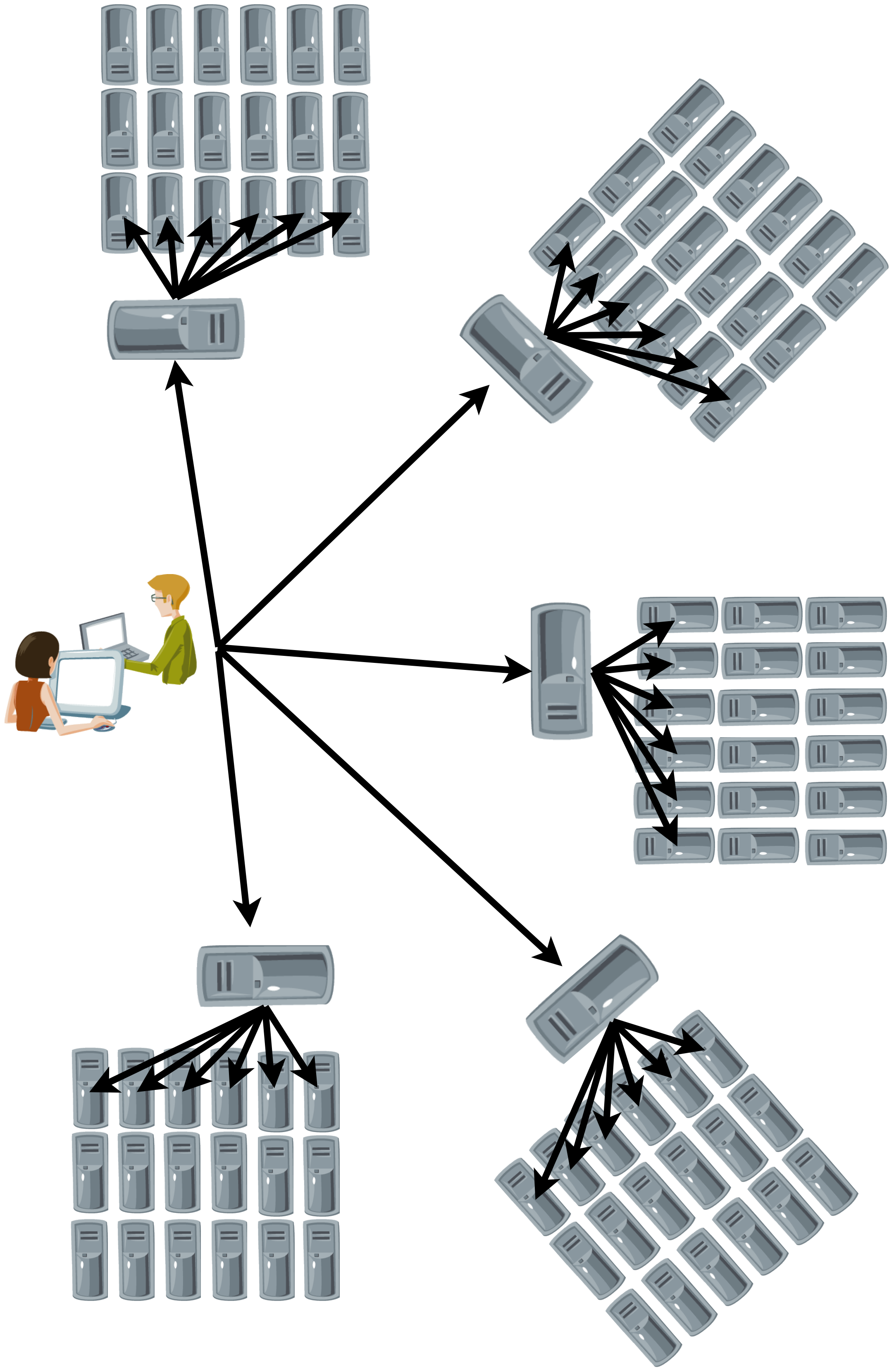
- CMS needed more than **100,000 cores** in parallel to find the Higgs

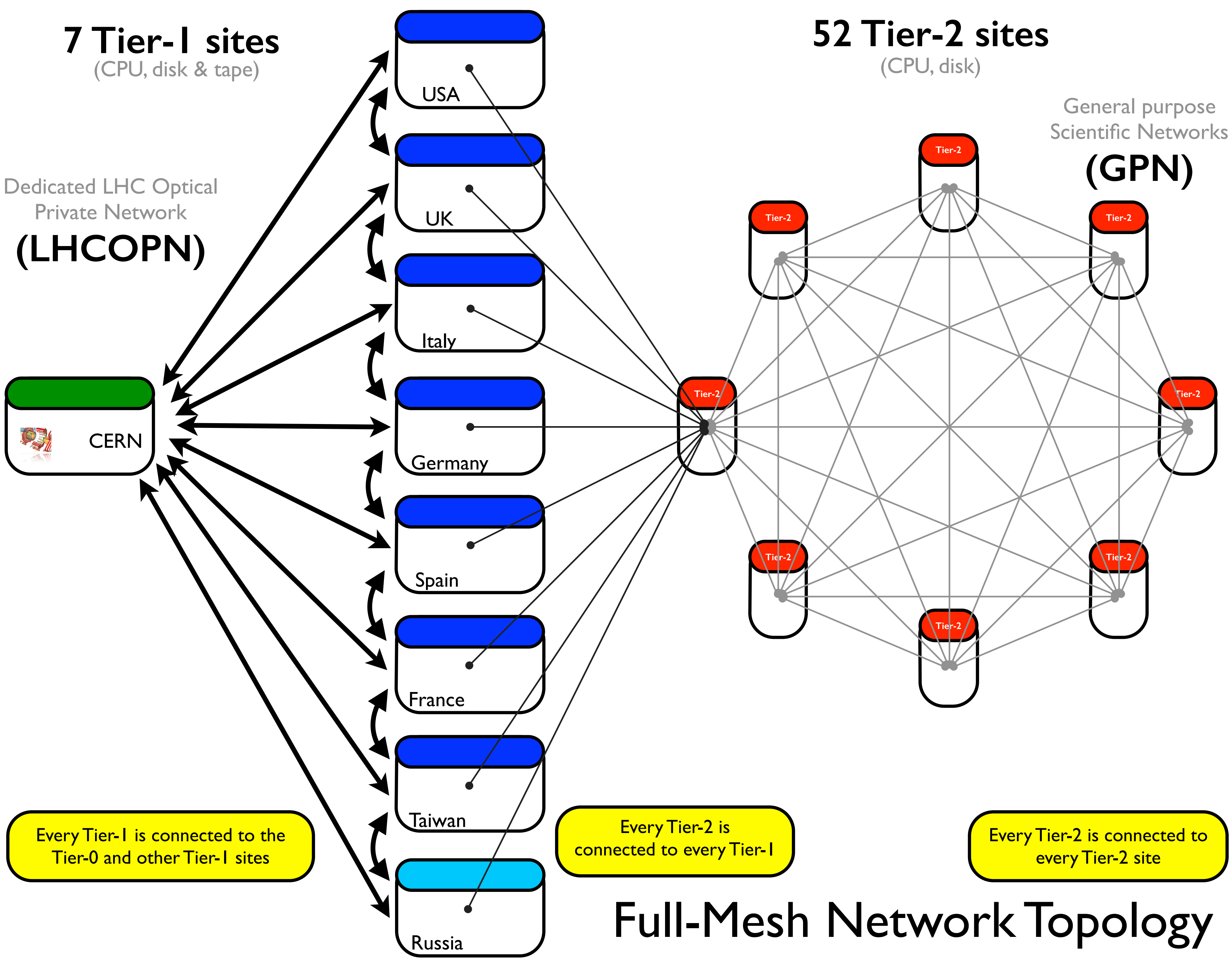
3. We use CPU farms and **Batch systems**:
  - Individual programs are put in a queue
  - The batch system has access to a farm of (many thousands of) cores
  - The batch system takes the first "job" in the queue and executes it on a free core → fills the farm with jobs



From Single Computer to the GRID

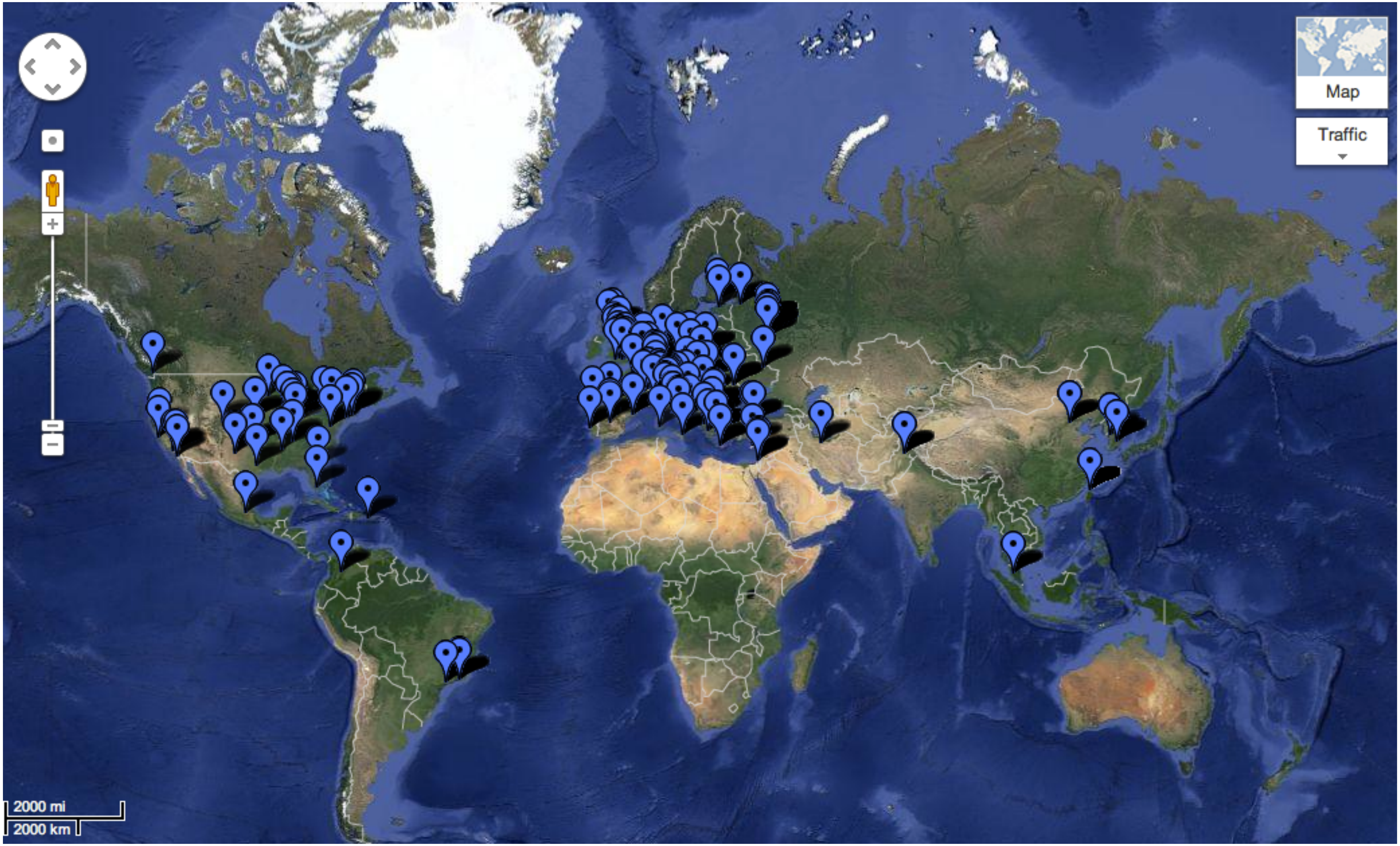
4. The GRID, an interconnected network of batch farms
   - Why not a single huge batch farm for CMS:
     - Running 100,000 cores in one installation is very difficult
       - You need a lot of cooling and large amounts of electricity
     - CMS is an international collaboration, funding agencies (like the Department of Energy in the US) prefer to spend research money in their home countries
   - The GRID enables CMS to have access to enough PCs despite being distributed over the world
   - The GRID software or middleware lets the individual computing centers or farm look like one big farm or center
     - Important: you need strong networks between the centers





Full-Mesh Network Topology

Distributed over the world



Data sizes

WHAT IS A PETABYTE?

TO UNDERSTAND A PETABYTE WE MUST FIRST UNDERSTAND A GIGABYTE.

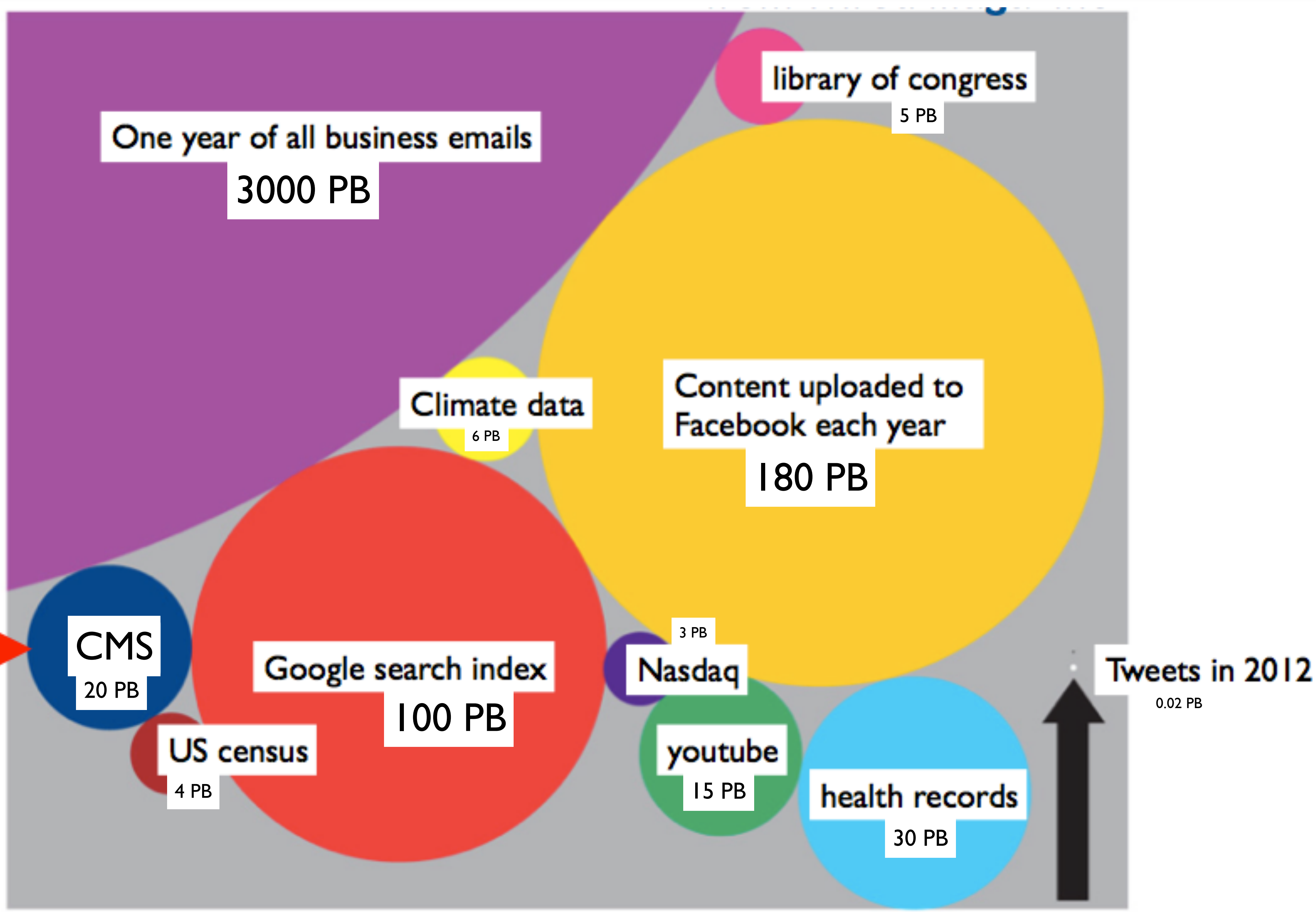
- **1 GIGABYTE** — 7 MINUTES OF HD-TV VIDEO
- **2 GIGABYTES** — 20 YARDS OF BOOKS ON A SHELF
- **4.7 GIGABYTES** — SIZE OF A STANDARD DVD-R

THERE ARE A MILLION GIGABYTES IN A PETABYTE

A PETABYTE IS A LOT OF DATA

- **1 PETABYTE** — 20 MILLION FOUR-DRAWER FILING CABINETS FILLED WITH TEXT
- **1 PETABYTE** — 13.3 YEARS OF HD-TV VIDEO
- **1.5 PETABYTES** — SIZE OF THE 10 BILLION PHOTOS ON FACEBOOK
- **20 PETABYTES** — THE AMOUNT OF DATA PROCESSED BY GOOGLE PER DAY
- **20 PETABYTES** — TOTAL HARD DRIVE SPACE MANUFACTURED IN 1995
- **50 PETABYTES** — THE ENTIRE WRITTEN WORKS OF MANKIND, FROM THE BEGINNING OF RECORDED HISTORY, IN ALL LANGUAGES

CMS is not the biggest fish in the pond



CMS produces per year **20 PB**

- Adapted from Wired: http://www.wired.com/magazine/2013/04/bigdata/

Where do we store all this data?

- Tape robots
  - A large shelf for tape cartridges (each 5 TB = 1000 GB)
  - A robot arm that can pick up a cartridge and insert it into a tape drive for reading or writing
  - Very cheap to store large amounts of data
  - Very slow to access
  - CMS has 50 PB of tape available to store it’s data

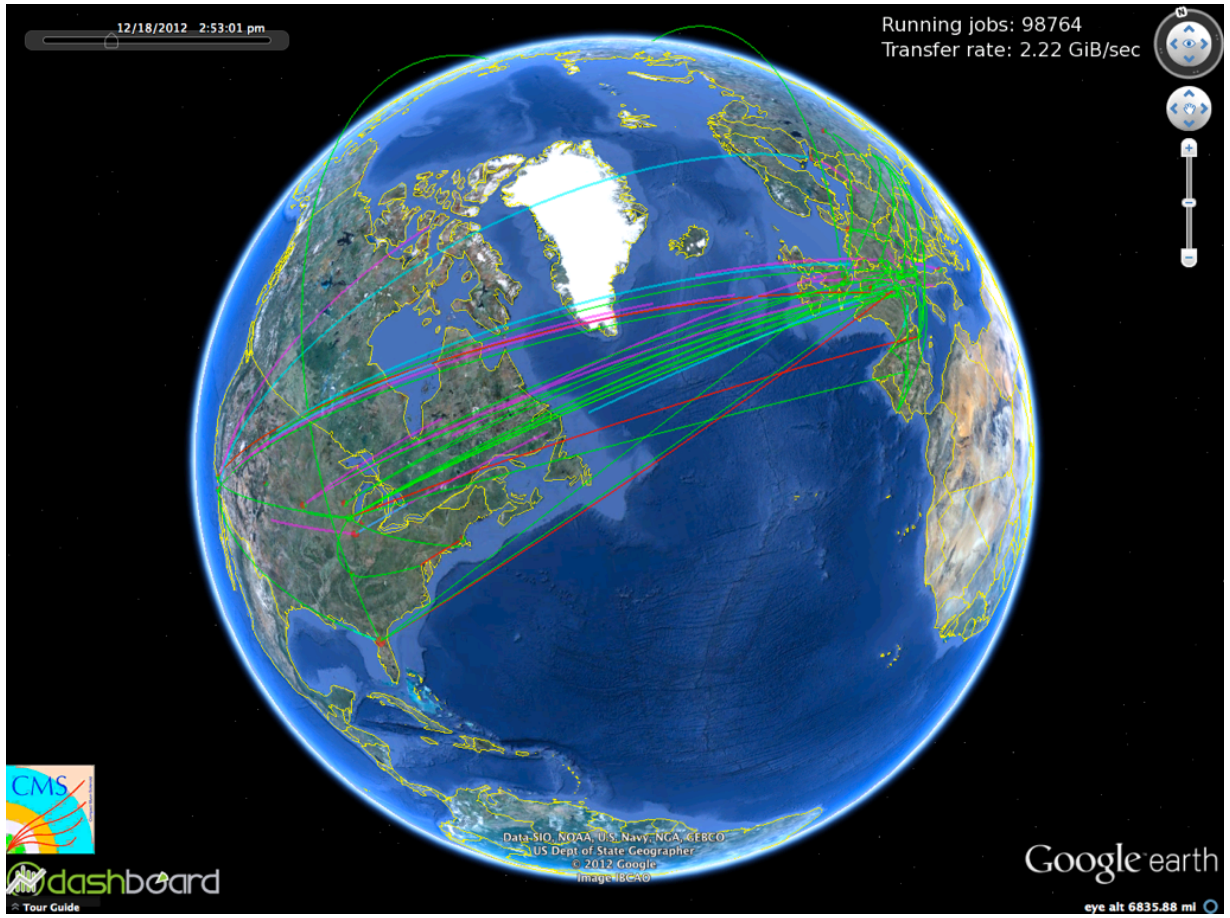
How do we access all this data

1 unit holds up to 72 hard drives

- **Large disk arrays**
  - Many thousands of normal hard drives are used to store data temporarily for fast access (data is cached)
    - Software systems make them appear as one big hard drive
    - Data has to be copied from tape to disk first before starting the reconstruction
    - Output produced at a center has to be copied to tape to keep it longterm
  - At each computing center, all the cores in the farm can access the disk through the local network
- At all sites together, CMS has over 80 PB of disk available

Fermilab

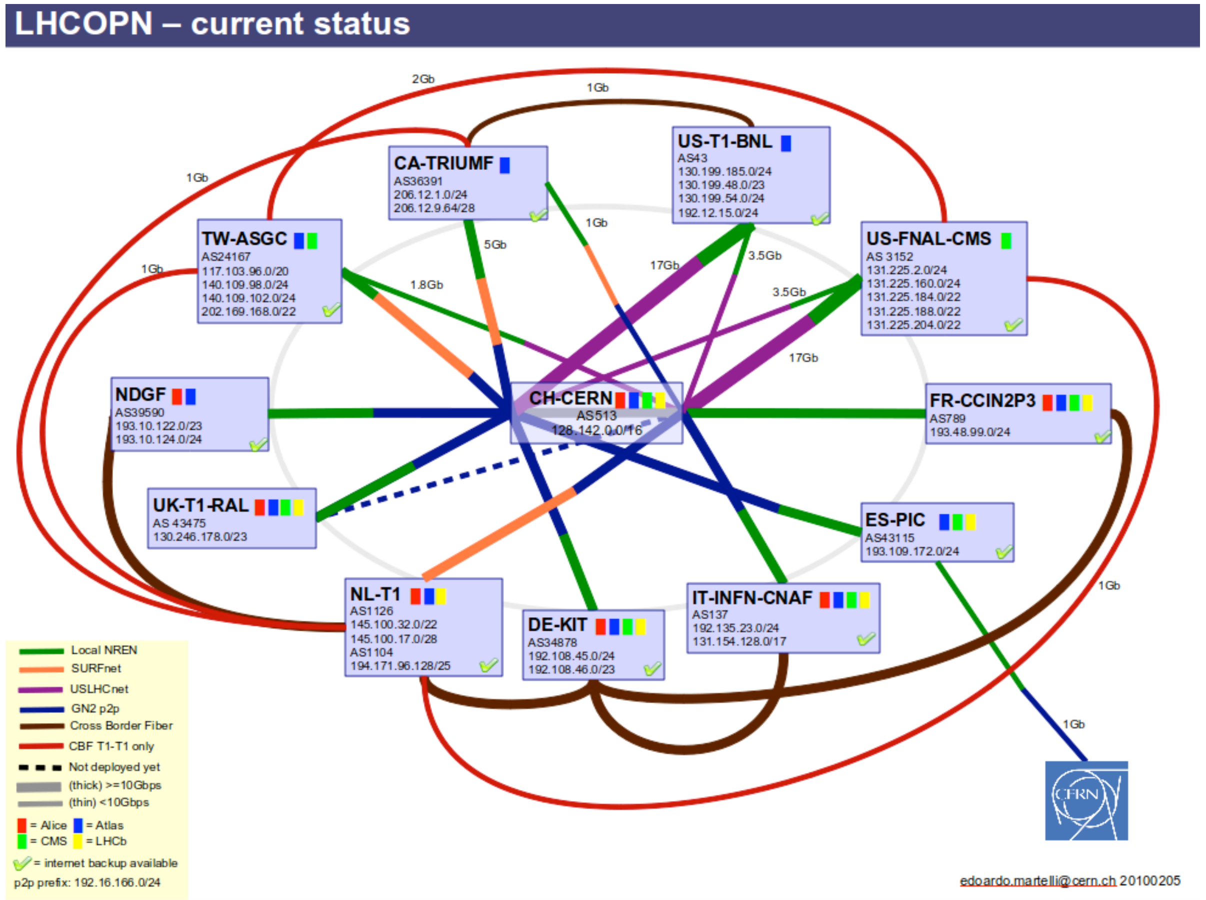
- Largest Tier-1 site for CMS
  - 40% of the computing resources on the Tier-1 level
    - FNAL: ~11,000 cores
  - Archives 40% of CMS data and simulation files on tape
    - FNAL: 22 PB
  - Large disk cache for efficient access to files
    - FNAL: 12 PB
  - Strong network connection to all of more than 60 CMS sites worldwide
    - FNAL: 80 Gbps → 10 GB per second
- Operated by Fermilab CMS facility team



The USCMS computing facilities at Fermilab are recognized world-wide as one of the best places for CMS central workflows and to perform LHC analysis

- We have to distribute our 20 PB of data across all CMS centers
  - Allow for access of data at the centers

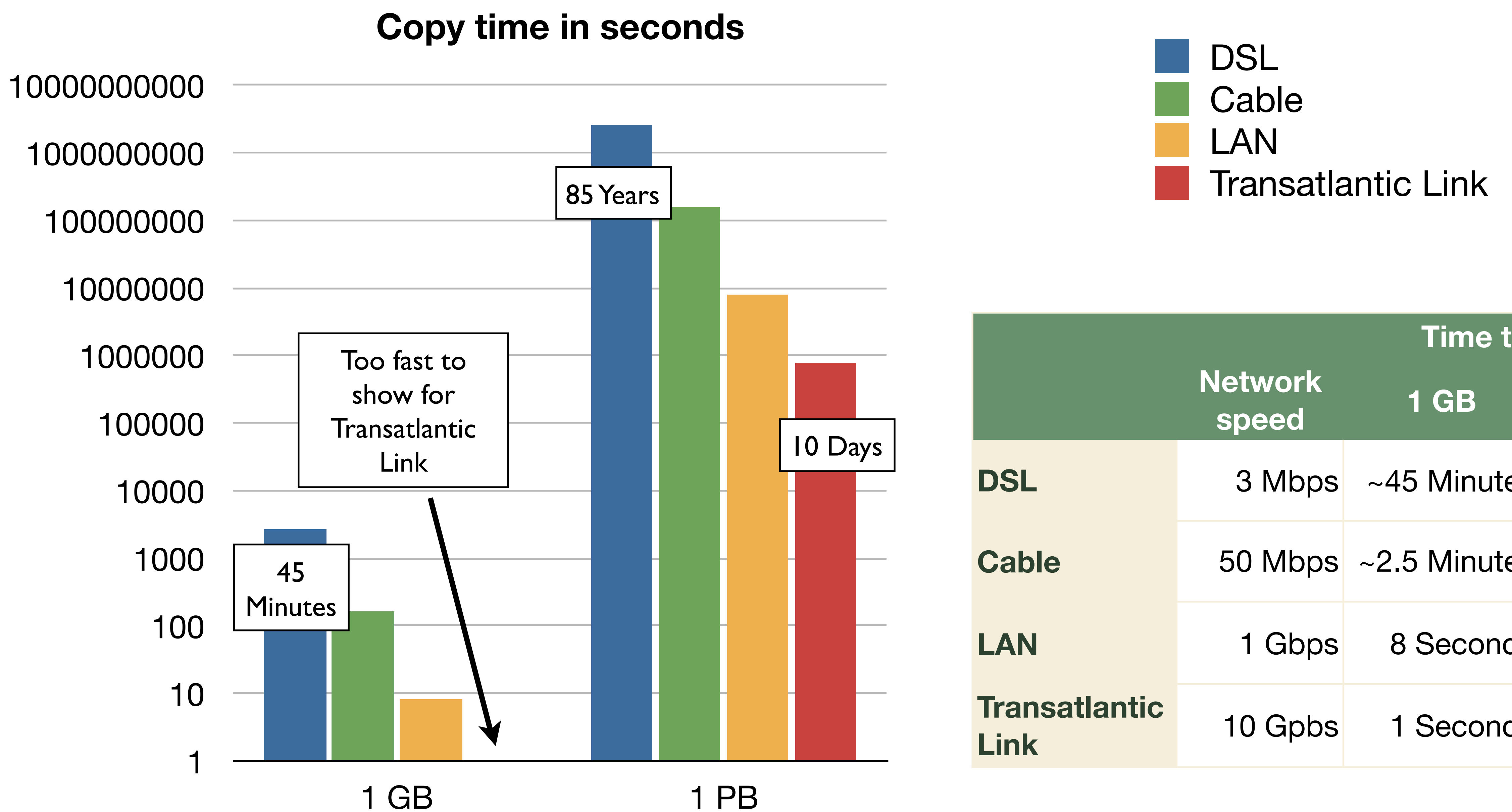
- We built dedicated networks between the Tier-0 at CERN and the Tier-1 sites to safely store all recorded collisions
  - Name: Large Hadron Collider Optical Private Network (LHCOPN)
    - We rented fibers from commercial vendors that also handle general internet traffic, including fibers across the Atlantic



- Network connectivity to all the remaining sites is provided by national science General Purpose Networks (GPN) in the participating countries
  - In the US: Internet2 and ESNet

Network speeds

- Comparison of network speeds and how long it takes to copy data
  - Byte: 8 Bit
  - Mbps: Megabit per Second, 1 Million Bits per Second
  - Gbps: Gigabit per Second, 1 Thousand Million Bits per Second



	Network speed	Time to transfer 1 GB	Time to transfer 1 PB
DSL	3 Mbps	~45 Minutes	85 Years
Cable	50 Mbps	~2.5 Minutes	5 Years
LAN	1 Gbps	8 Seconds	93 Days
Transatlantic Link	10 Gpbs	1 Seconds	10 Days

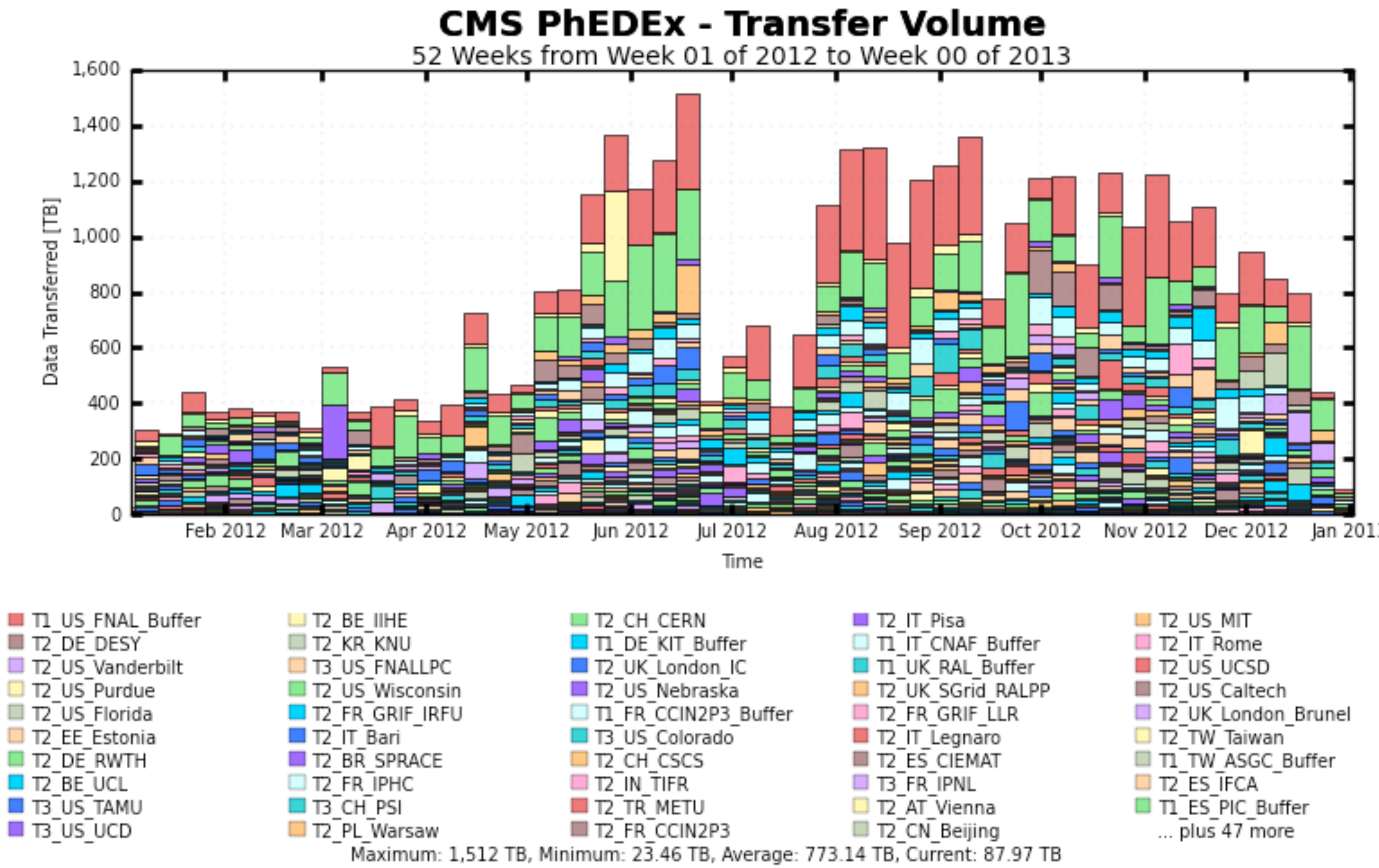
The data distribution problem

- To set LHC data movement into perspective: **Netflix**
  - Netflix delivers streaming video content to about 20M subscribers
  - Routinely quoted as the single largest user of bandwidth in the US
- CMS has a smaller number of clients and less distribution because of the total data volume
- Netflix has much less data, can duplicate all movies in different parts of the country many many times

	Netflix	CMS
Clients	20 Million	100,000
Total Data	12 TB	20 PB

It is easier to distribute a small amount of data to many clients.
Large volume data distribution is hard!

CMS data transfers in 2012



- In 2012, CMS transferred on average over ¾ PB per week
  - Peak: 1½ PB per week

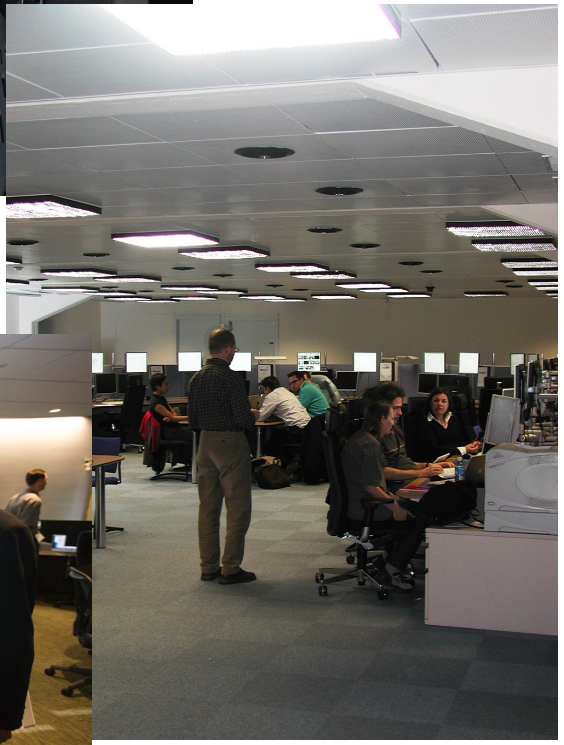
Lets put everything together

- To discover the Higgs, we needed
  - 100,000 cores of compute power
  - 80 PB of disk storage
  - 50 PB of tape storage
  - Over 60 individual computing centers distributed all over the world
  - Strong networks connecting the centers capable of transferring 2 PB per week and more
  - A lot of technology to make this all work together seamlessly and easy for all 2000 physicists of CMS

CERN

FNAL

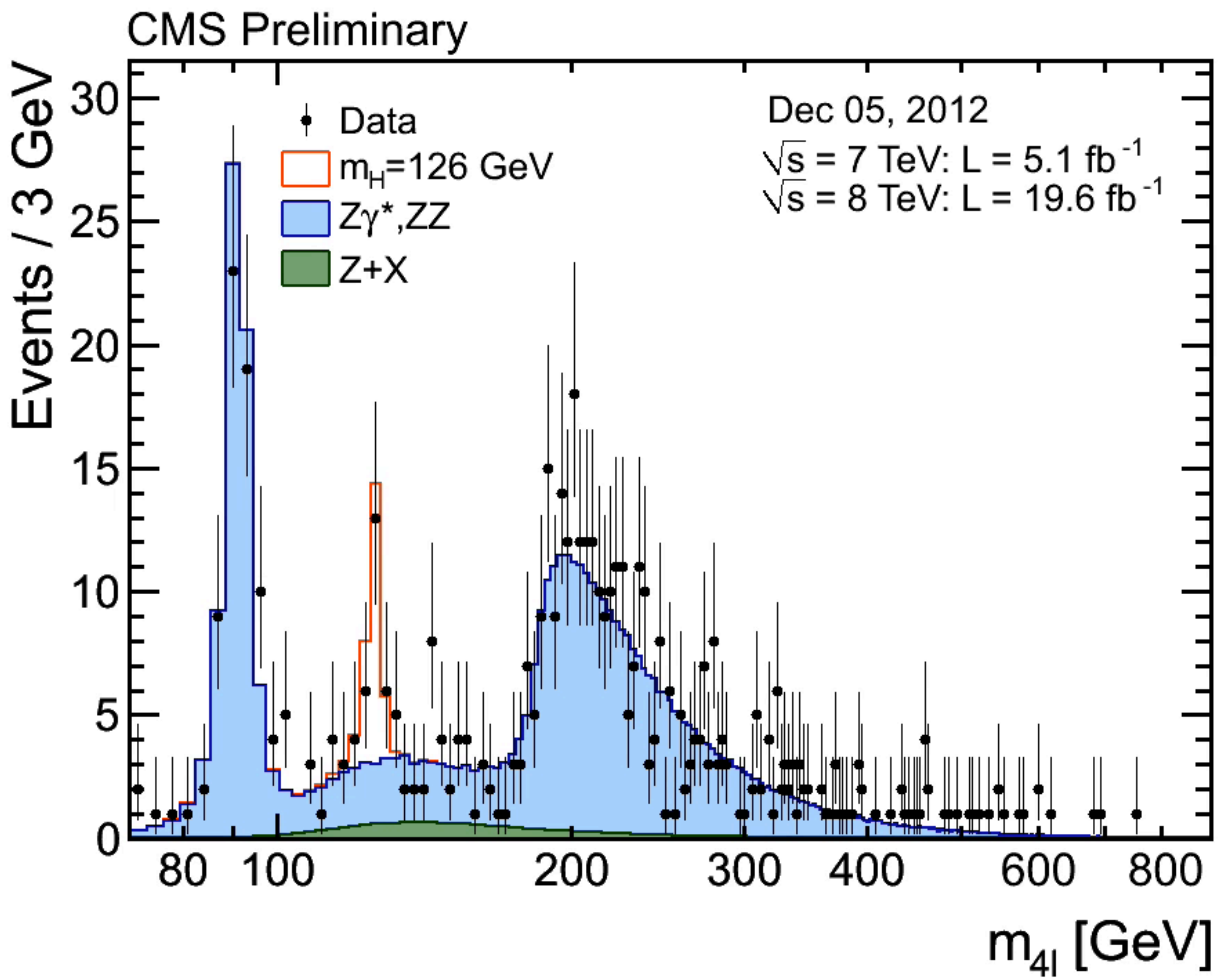
CERN



FNAL

The result

- All this was needed to find the "needle in the haystack" and discover the Higgs



Click for animation: https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsHIG/HZZ4l_date_animated.gif