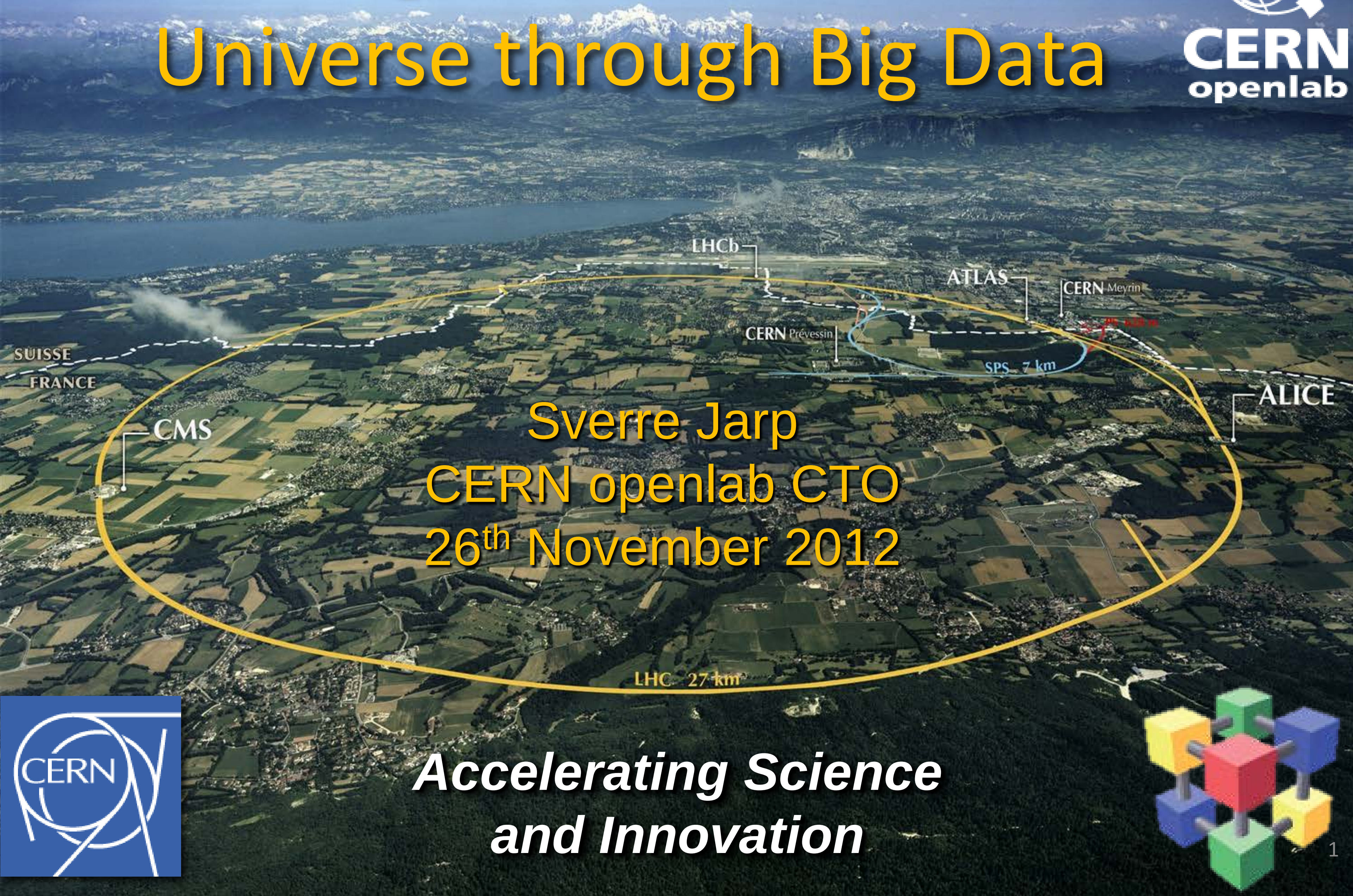
Solving the Riddles of the Universe through Big Data

Sverre Jarp
CERN openlab CTO
26th November 2012



Accelerating Science and Innovation

What is CERN

- The European Particle Physics Laboratory based in Geneva, Switzerland
  - Current accelerator: The Large hadron Collider (LHC)
- Founded in 1954 by 12 countries for fundamental physics research in a post-war Europe
- Today, it is a global effort of 20 member countries and scientists from 110 nationalities, working on the world's most ambitious physics experiments
- ~2'300 personnel, > 10'000 users
- ~1 billion CHF yearly budget

CERN openlab

- **A unique research partnership between CERN and the industry**
- Objective: The advancement of cutting-edge computing solutions to be used by the worldwide LHC community

CERN openlab
www.cern.ch/openlab

PARTNERS
hp invent
intel
ORACLE
SIEMENS

CONTRIBUTOR
HUAWEI

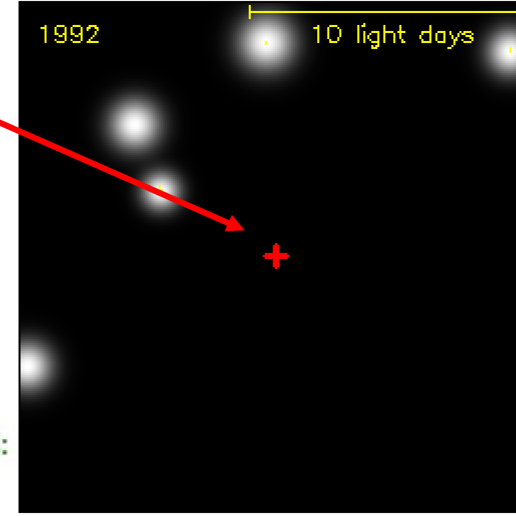
CERN: The Mecca of the Particle Physics Community

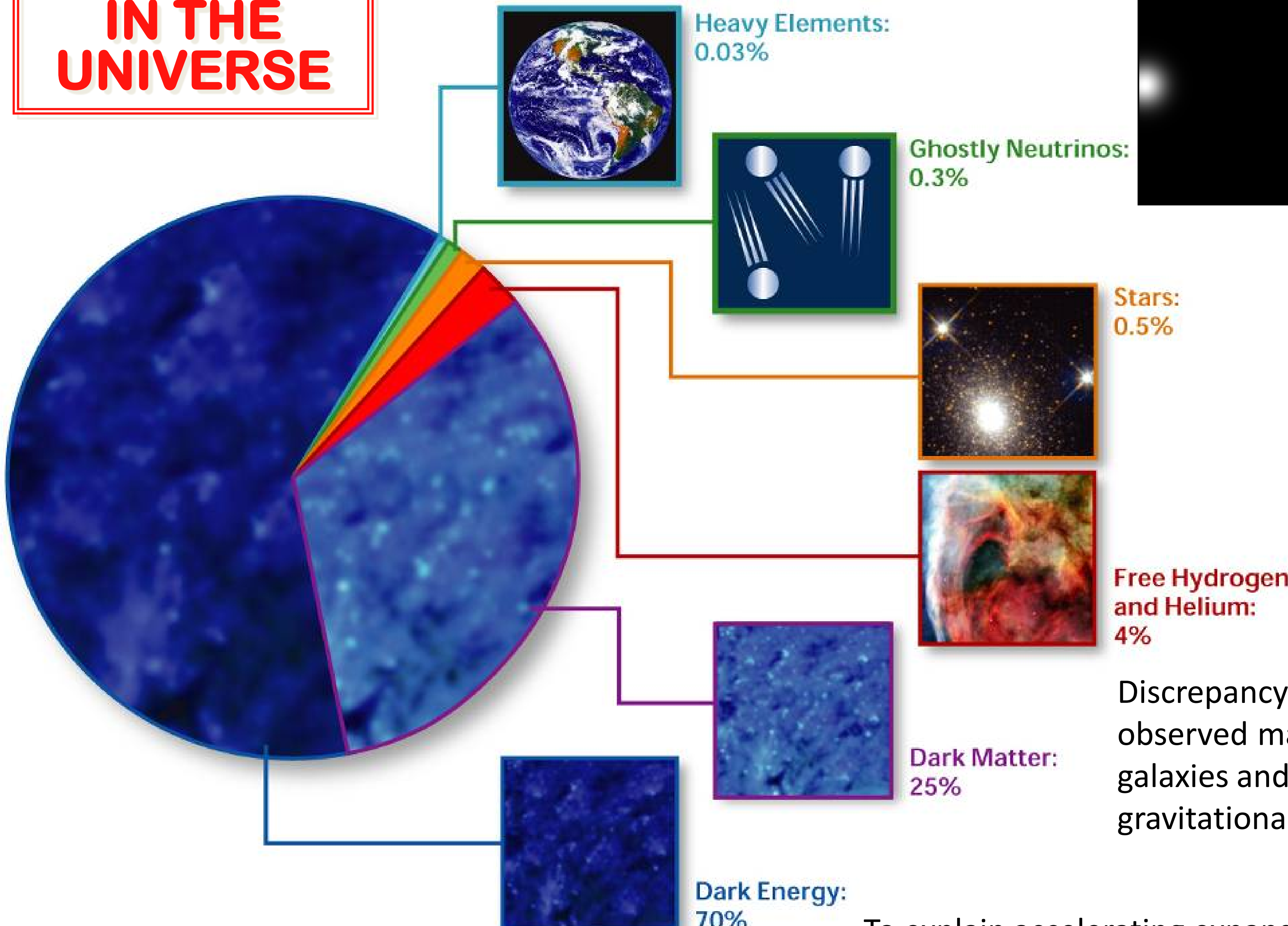
… bringing the world together

WHY “CERN”?

> 95% UNKNOWN STUFF IN THE UNIVERSE

Black hole





Discrepancy between observed mass of stars & galaxies and their gravitational motion

To explain accelerating expansion of the universe

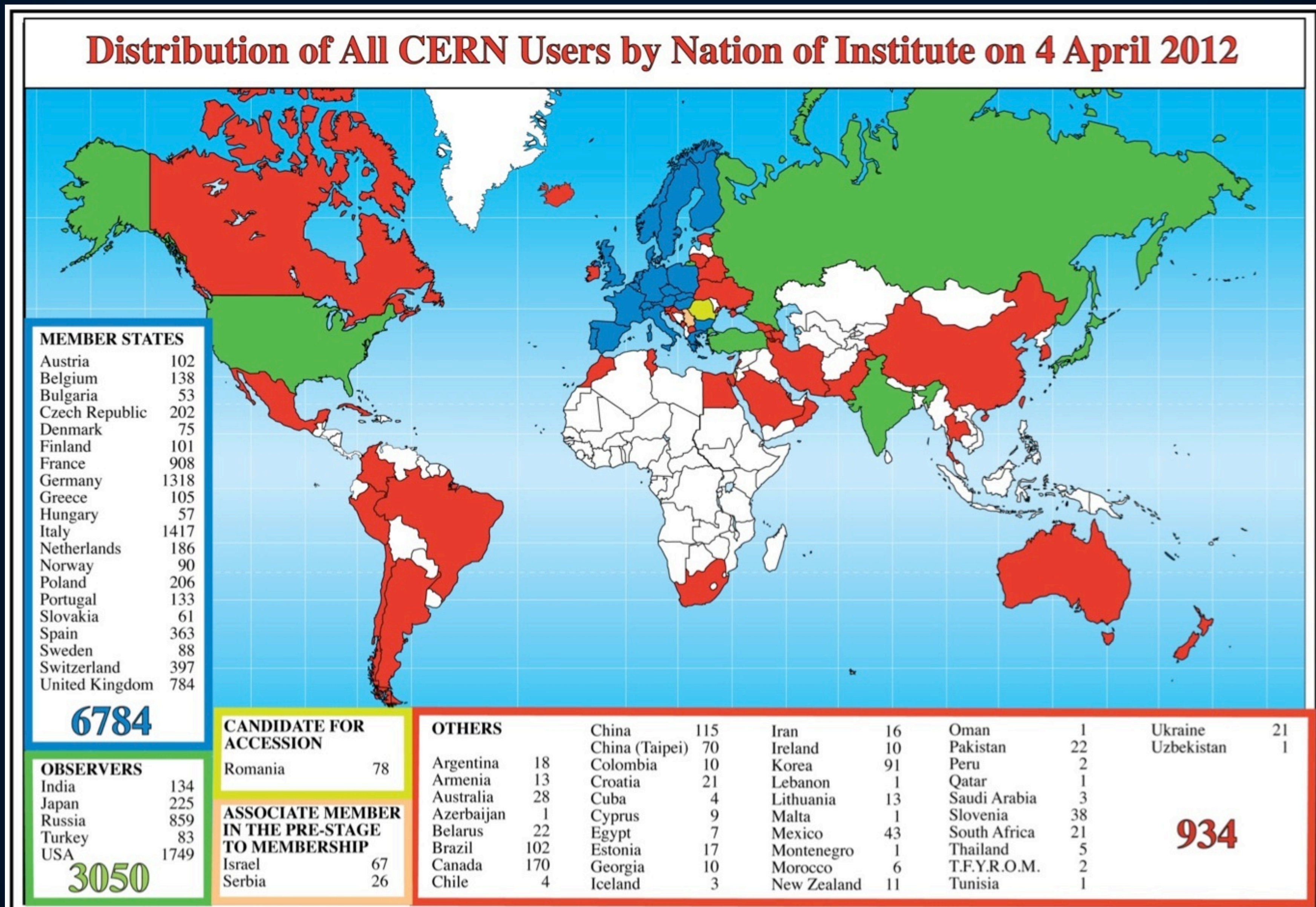
Fundamental Physics Questions

- Why do particles have mass?
  - Newton could not explain it - and neither can we…
- What is 95% of the Universe made of?
  - We only observe a fraction! What is the rest?
- Why is there no antimatter left in the Universe?
  - Nature should be symmetrical, or not?
- What was matter like during the first second of the Universe, right after the "Big Bang"?
  - A journey towards the beginning of the Universe gives us deeper insight

The Large Hadron Collider (LHC), built at CERN, enables us to look at microscopic big bangs to understand the fundamental behaviour of nature

CERN

Science is more and more global



So, how do you get from this ←

to this →

Higgs boson-like particle discovery claimed at LHC

COMMENTS (1665)

By Paul Rincon
Science editor, BBC News website, Geneva

The moment when Cern director Rolf Heuer confirmed the Higgs results

Cern scientists reporting from the Large Hadron Collider (LHC) have claimed the discovery of a new particle consistent with the Higgs boson.

Relat

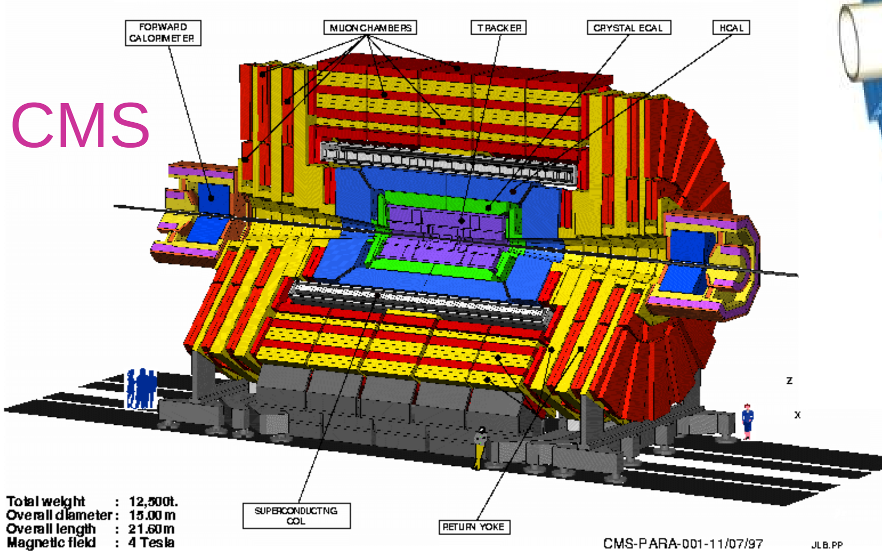
Some facts about the LHC

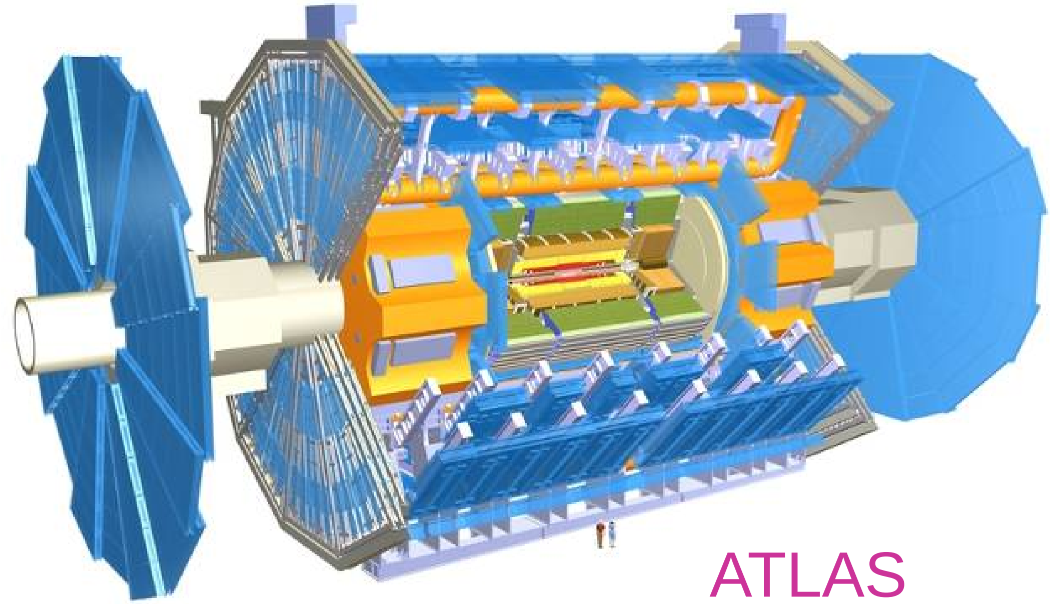
- Biggest accelerator (largest machine) in the world
  - 27 km circumference, 9300 magnets
- Fastest racetrack on Earth
  - Protons circulate 11245 times/s (99.9999991% the speed of light)
- Emptiest place in the solar system – high vacuum inside the magnets:
  - Pressure 10^{-13} atm (10x less than pressure on the moon)
- World's largest refrigerator (need only 1/8 of LHC magnets to qualify): -271.3 °C (1.9K)
- Hottest spot in the galaxy
  - During Lead ion collisions create temperatures 100 000x hotter than the heart of the sun; new record 5.5 Trillion K
- World's biggest and most sophisticated detectors
  - 150 Million "pixels"
- Most data of any scientific experiment
  - 15-30 PB per year (as of today we have about 60 PB)

Scale of ATLAS and CMS?

ATLAS superimposed to a CERN 5-storey building

CMS





ATLAS

	ATLAS	CMS
Overall weight (tons)	*7000*	*12500*
Diameter	*22 m*	*15 m*
Length	*46 m*	*22 m*
Solenoid field	*2 T*	*4 T*

Some history of scale…

Date	Collaboration sizes	Data volume, archive technology
Late 1950's	2-3	Kilobits, paper notebooks
1960's	10-15	KB, punchcards
1970's	~35	MB, tape
1980's	~100	GB, tape, disk
1990's	700-800	TB, tape, disk
2010's	~3000	PB → EB, tape, disk

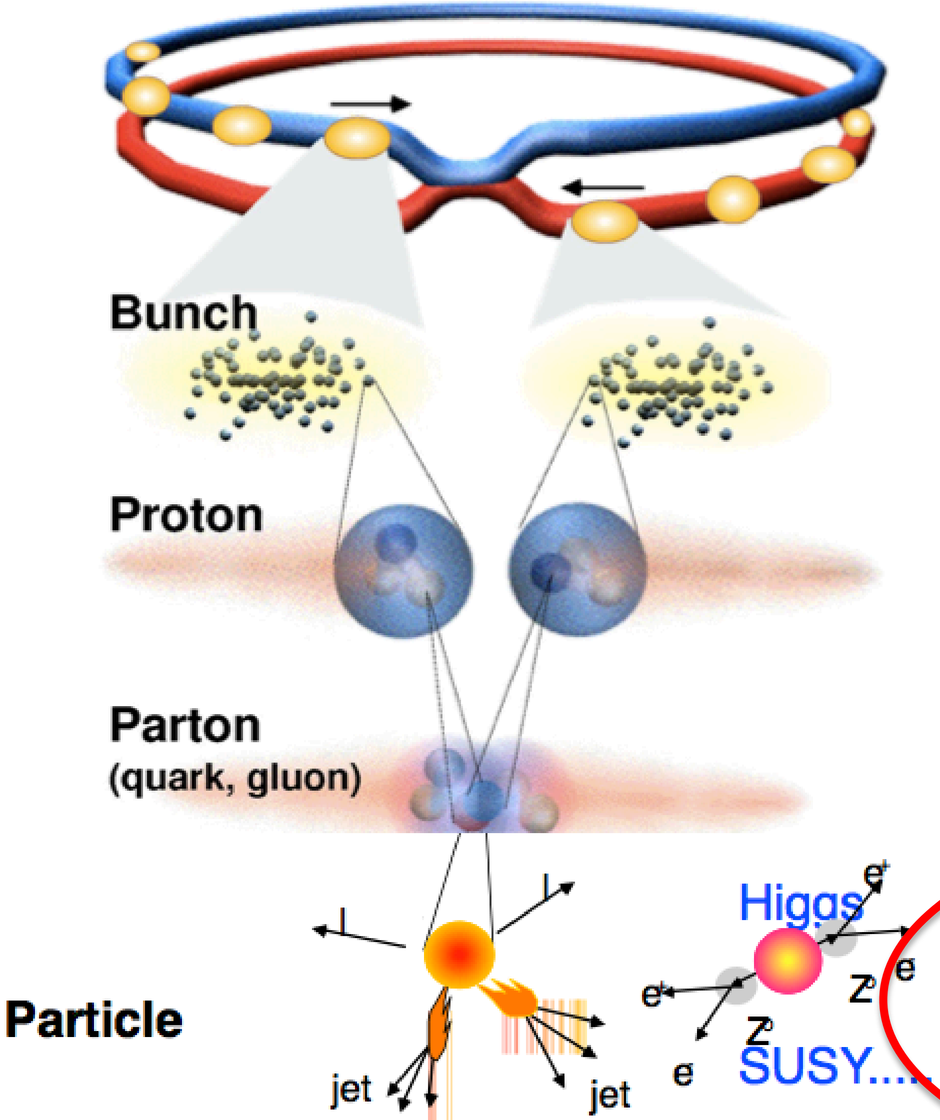
For comparison:
1990's: Total LEP data set ~few TB
Would fit on 1 tape today

Today: 1 year of LHC data ~30 PB

CERN has about 60,000 physical disks to provide about 20 PB of reliable storage

Why do we have to produce so much data ?

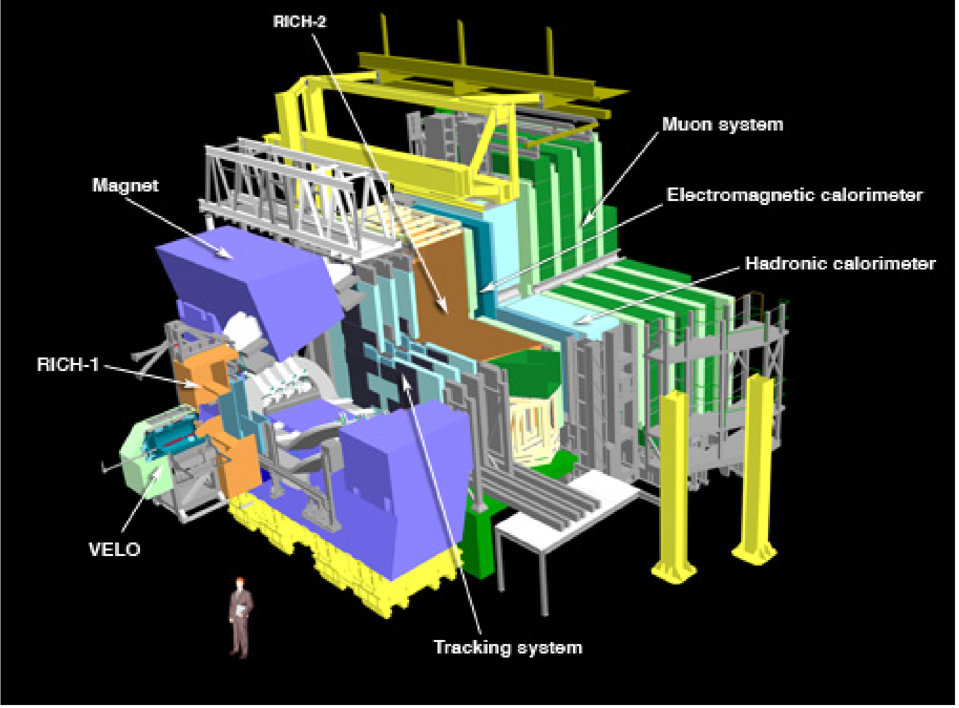
Collisions at the LHC: summary

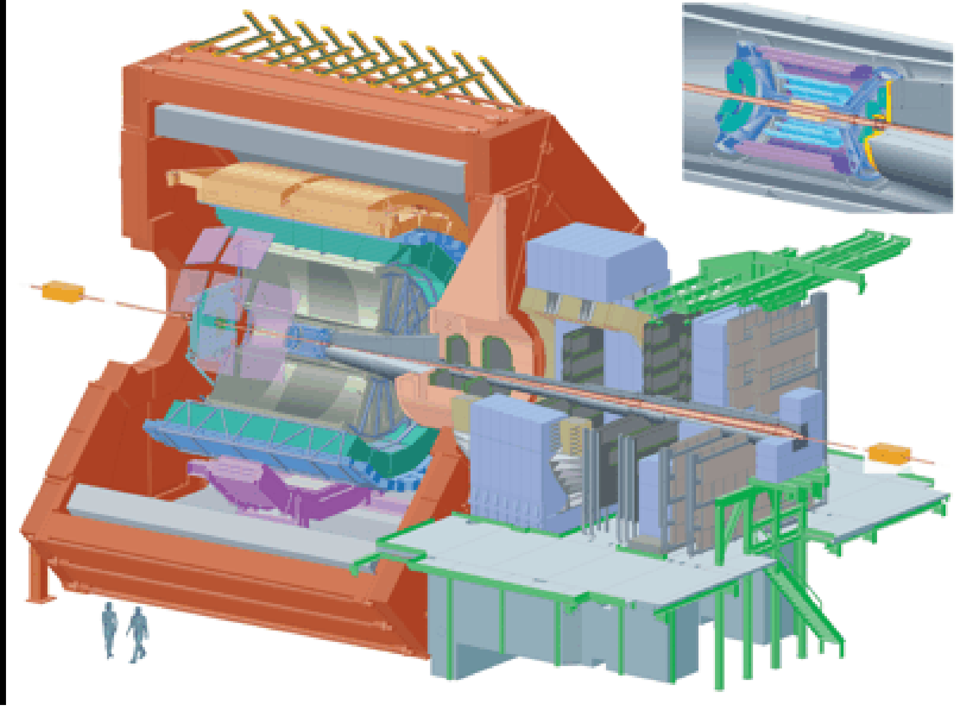


Proton - Proton	2808 bunch/beam
Protons/bunch	10^{11}
Beam energy	7 TeV ($7x10^{12}$ eV)
Luminosity	$10^{34}cm^{-2}s^{-1}$
Crossing rate	40 MHz
Collision rate ≈	10^7-10^9

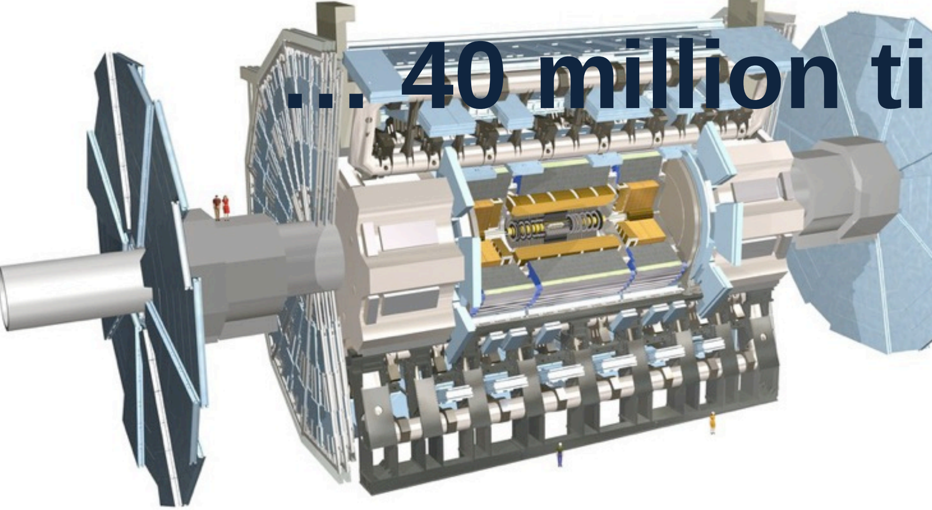
New physics rate ≈ .00001 Hz

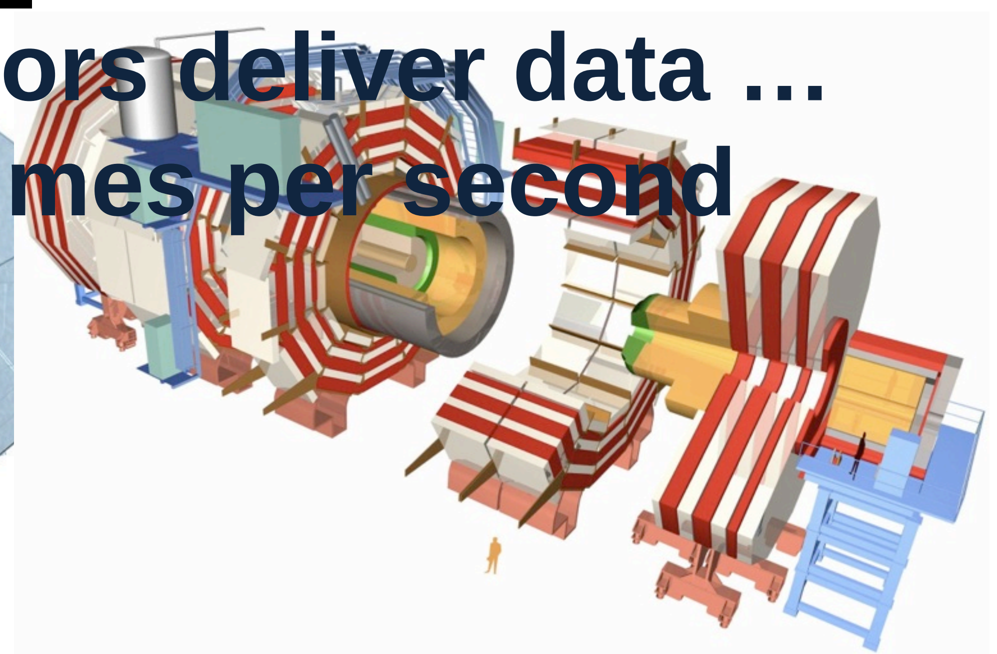
Event selection:
1 in 10,000,000,000,000



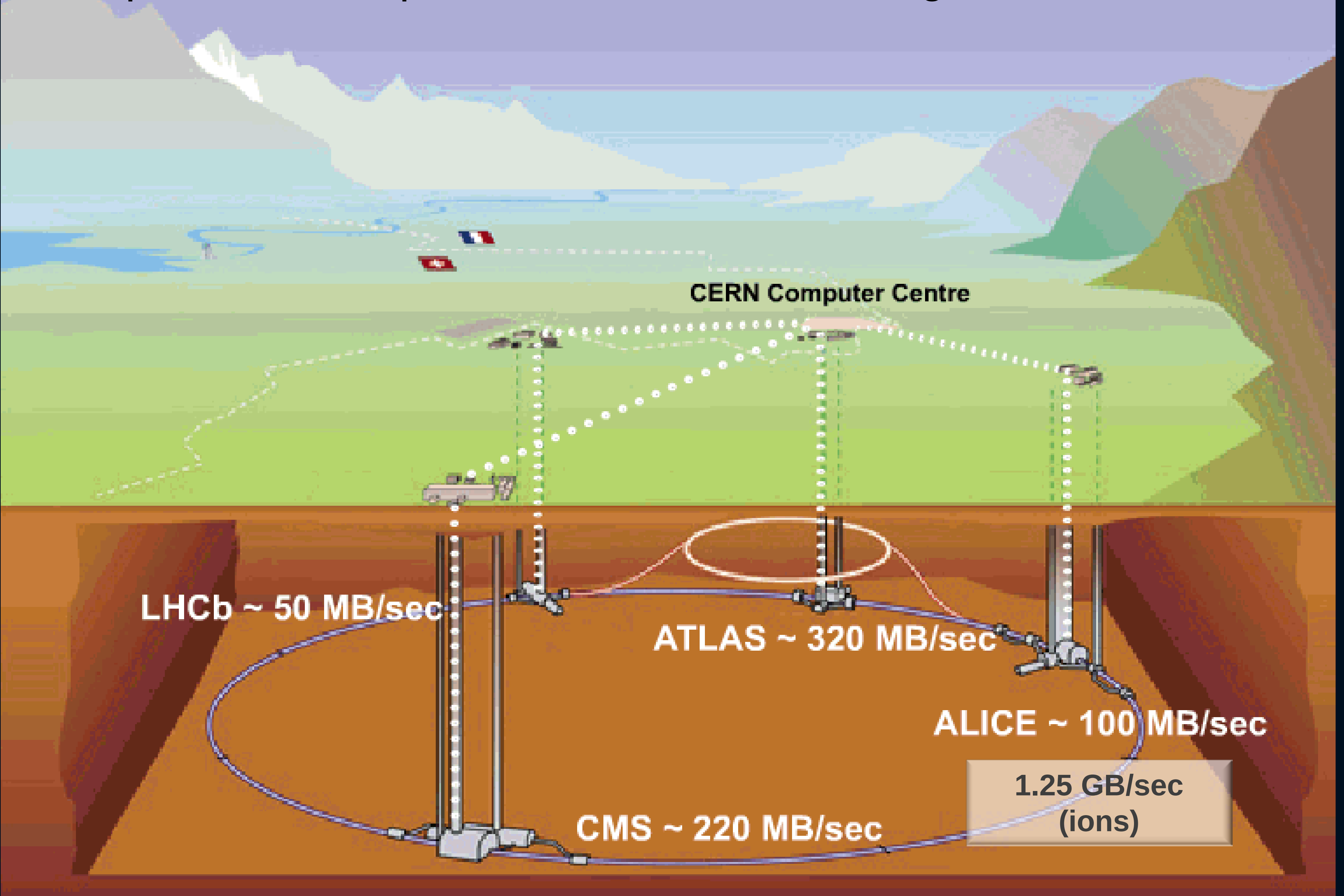


150 million sensors deliver data …
… 40 million times per second





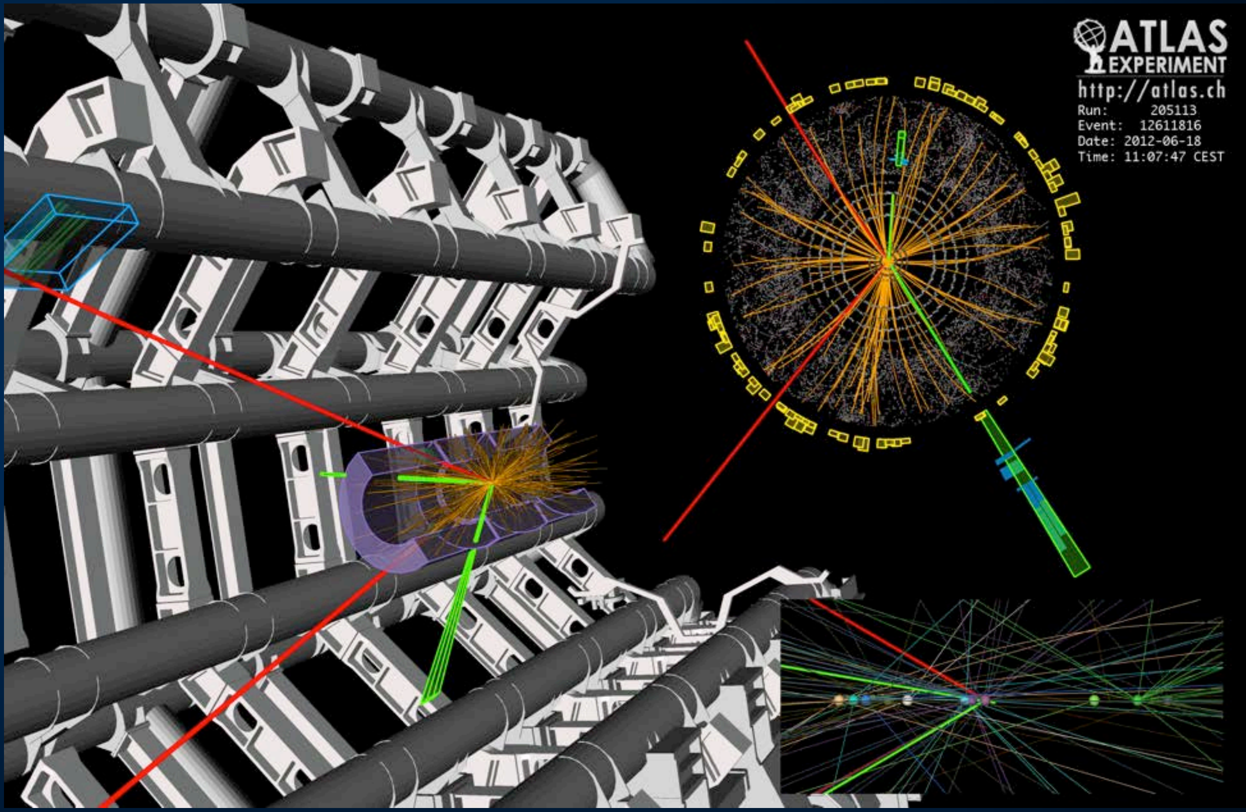
Tier 0 at CERN:
Acquisition, First pass reconstruction, Storage & Distribution



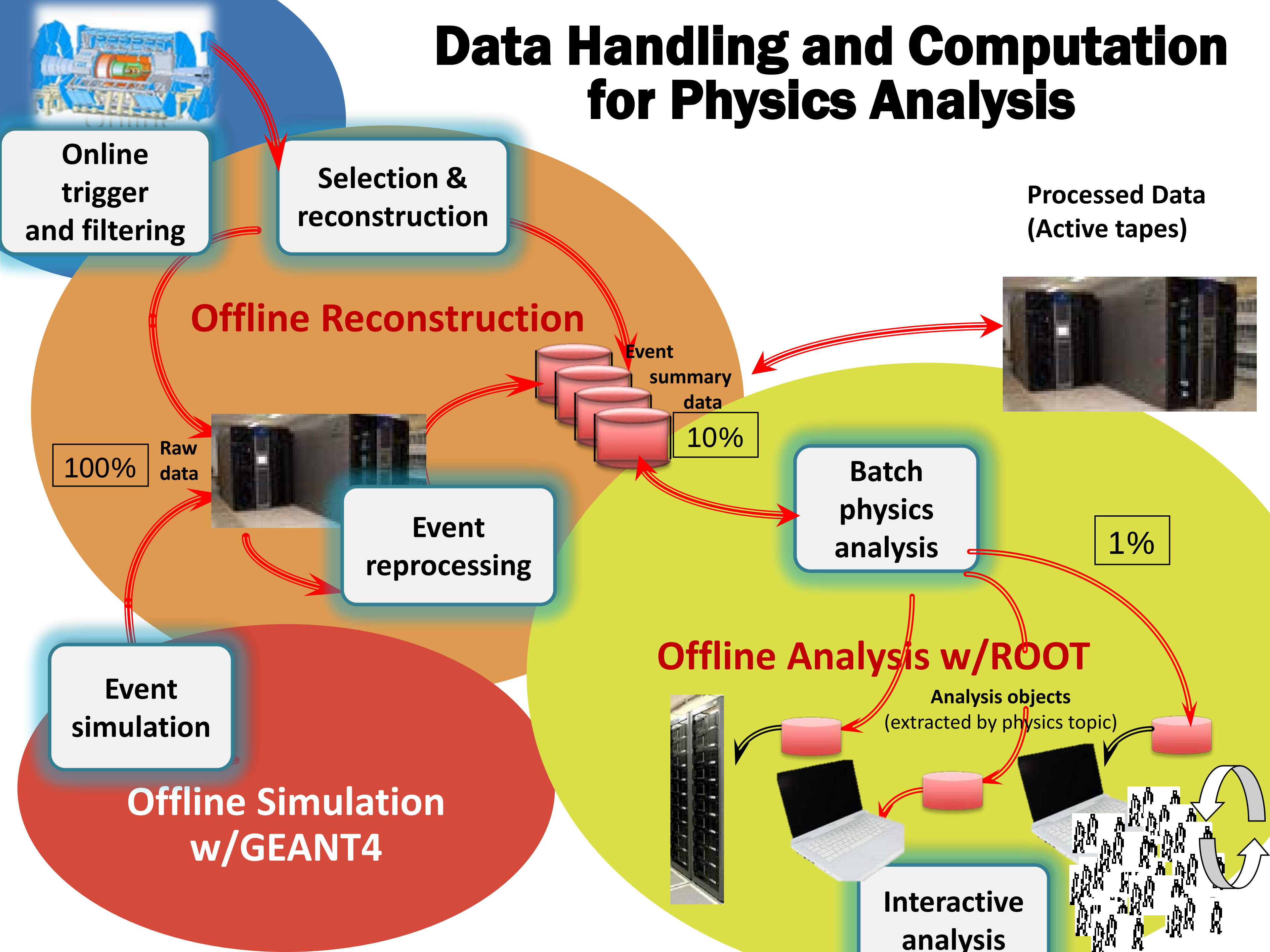
What is this data?

- Raw data:
  - Was a detector element hit?
  - How much energy?
  - What time?

- Reconstructed data:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information

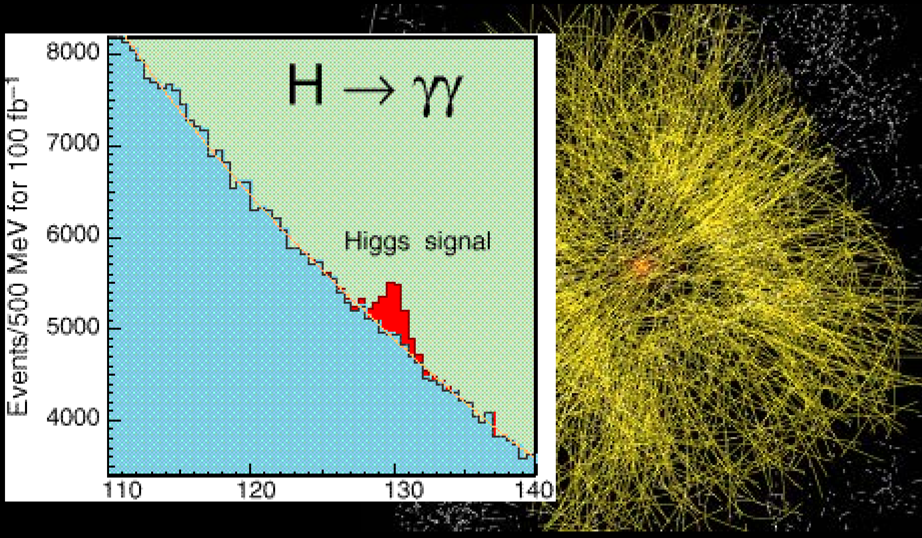


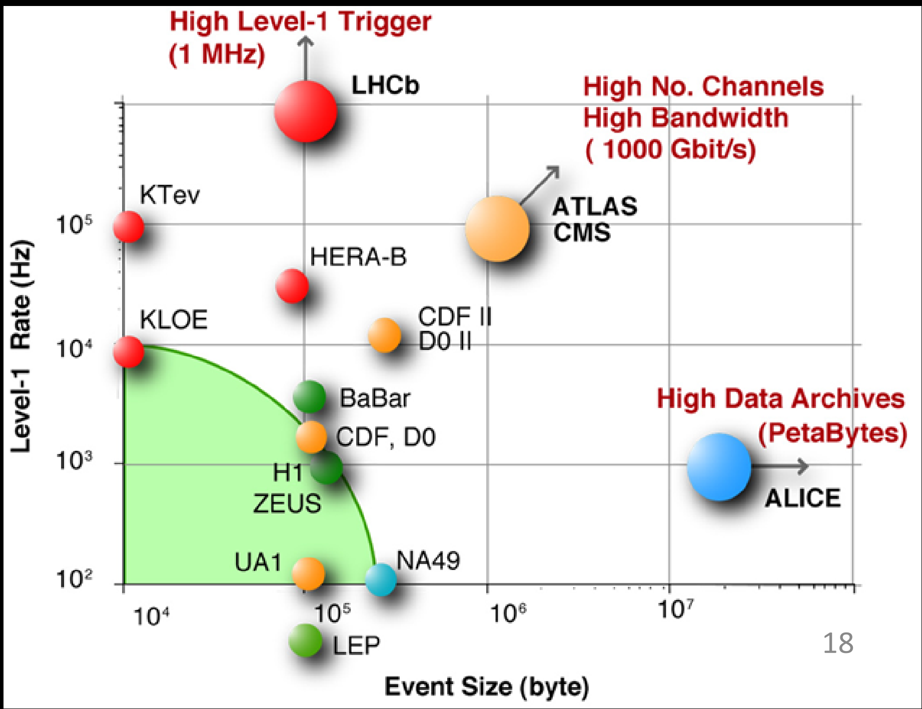
Data Handling and Computation for Physics Analysis



The LHC Computing Challenge

- Signal/Noise: 10^{-13} (10^{-9} offline)
- Data volume
  - High rate * large number of channels * 4 experiments
  - ➔ **15 PetaBytes of new data each year** → 30 PB in 2012
- Overall compute power
  - Event complexity * Nb. events * thousands users
  - ➔ **200 k cores** → 250 k cores
  - ➔ **45 PB of disk storage** → 150 PB
- Worldwide analysis & funding
  - Computing funding locally in major regions & countries
  - Efficient analysis
  - ➔ **GRID technology**





WLCG Worldwide LHC Computing Grid

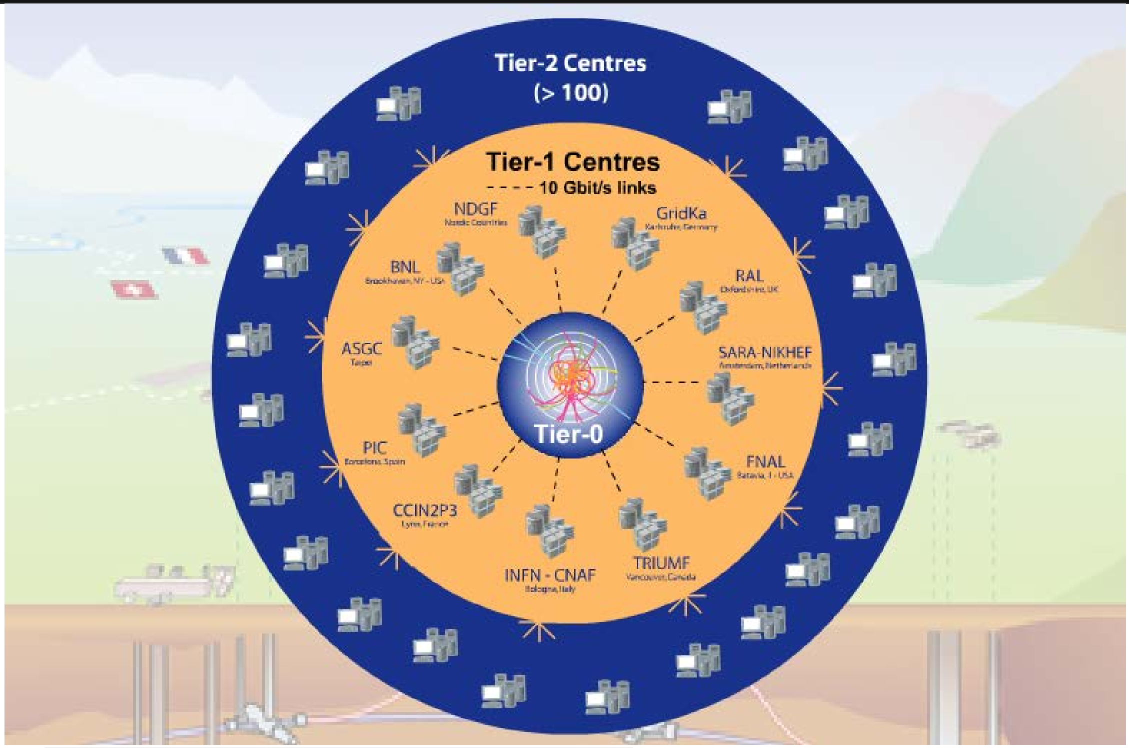
LCG

So, what are our issues with Big Data?

Credit: A. Pace

World-wide LHC Computing Grid

- A distributed computing infrastructure to provide the production and analysis environments for the LHC experiments
- Managed and operated by a worldwide collaboration between the experiments and the participating computer centres
- The resources are distributed – for funding and sociological reasons
- Our task was to make use of the resources available to us – no matter where they are located



Tier-0 (CERN):
- Data recording
- Permanent storage
- Initial data reconstruction
- Data distribution

Tier-1 (11 centres):
- Permanent storage
- Re-processing
- Analysis

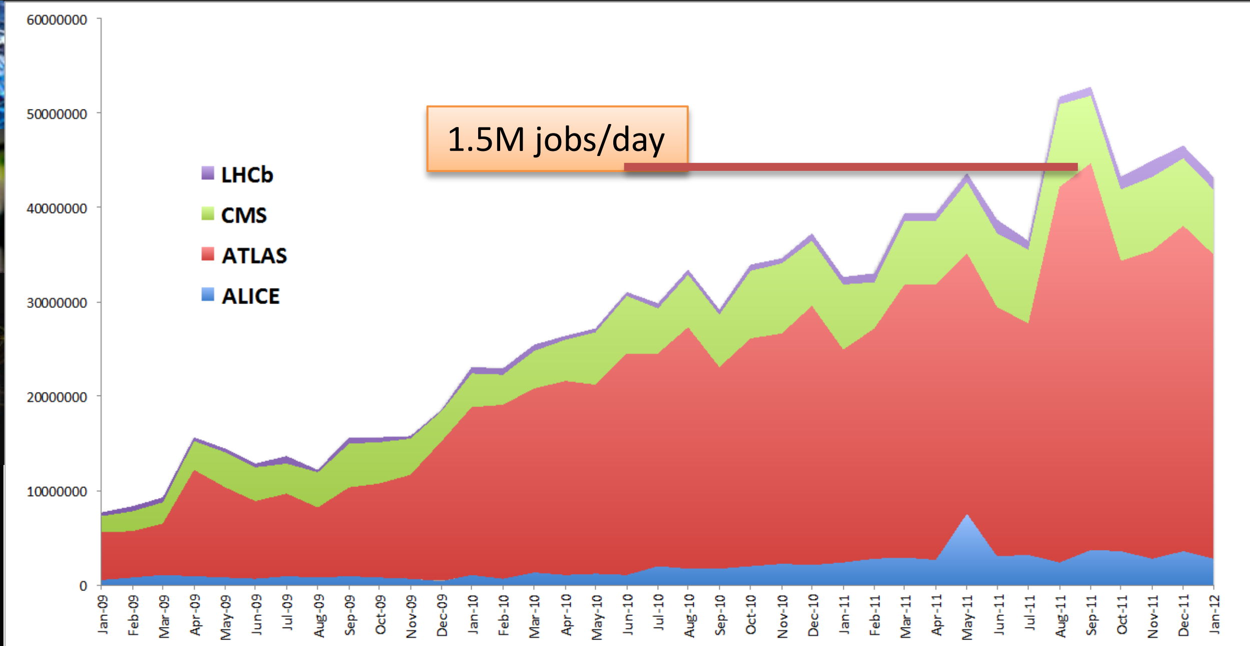
Tier-2 (~130 centres):
- Simulation
- End-user analysis

WLCG Grid Sites

- Today >140 sites
- >250k x86 PC cores
- >150 PB disk

Tier 0 Tier 1 Tier 2

Processing on the grid

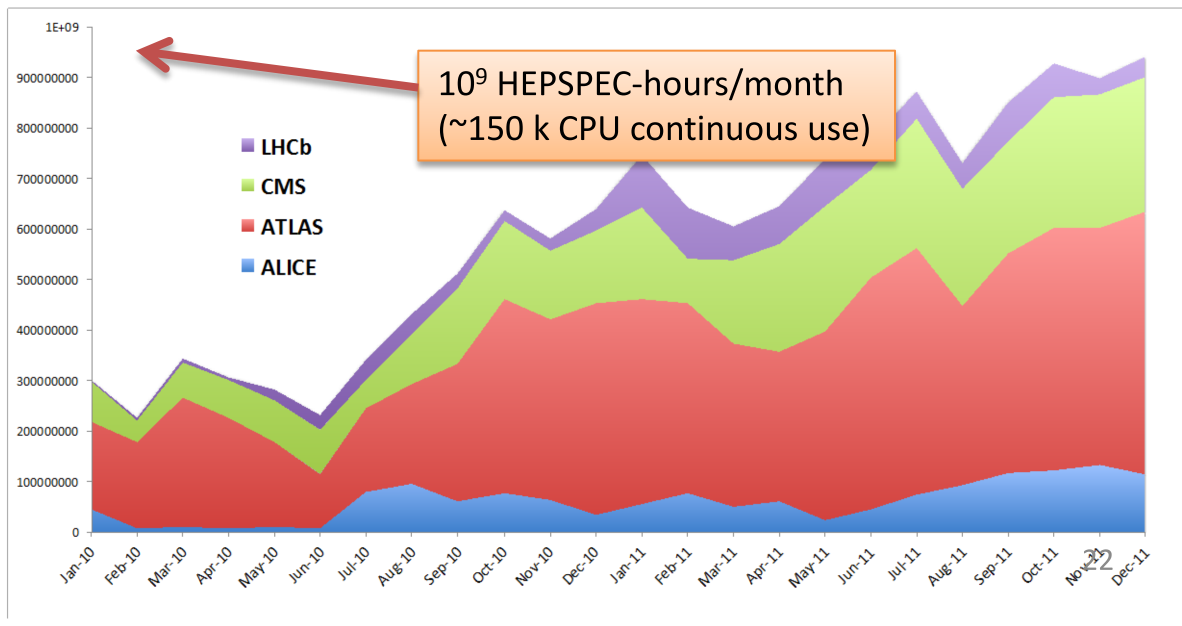


Usage continues to grow…

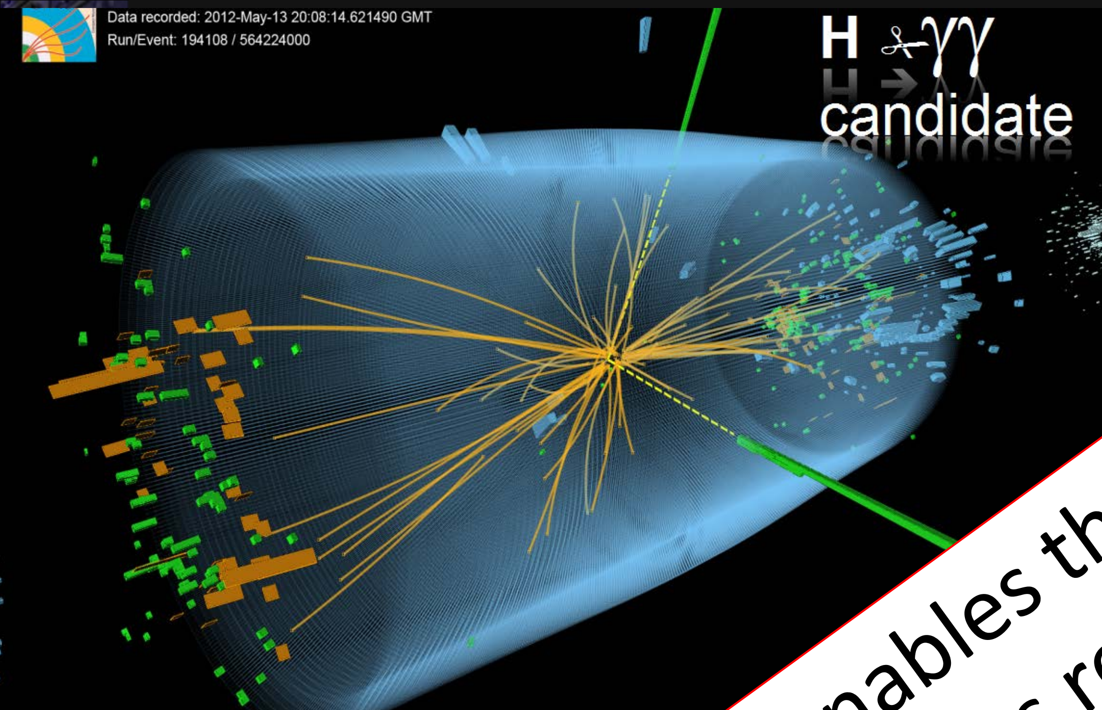
- # jobs/day
- CPU usage

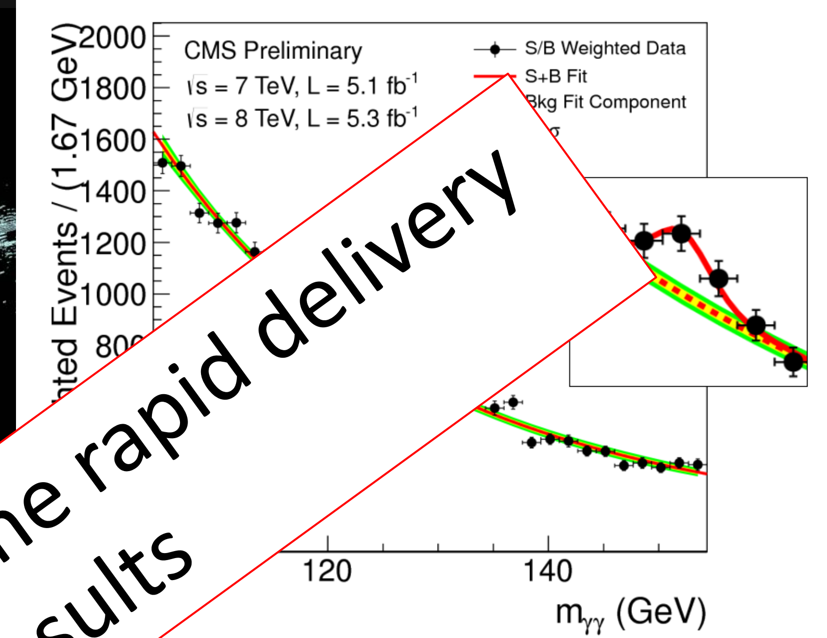
~ 150,000 years of CPU delivered each year

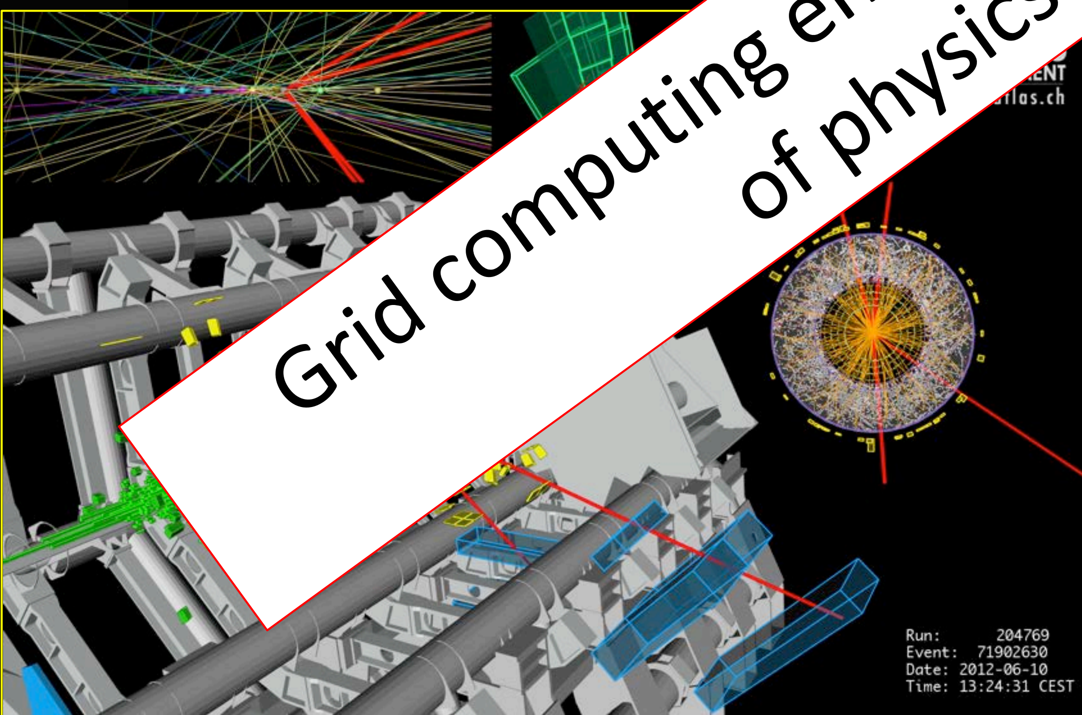
This is close to full capacity
We always need more!

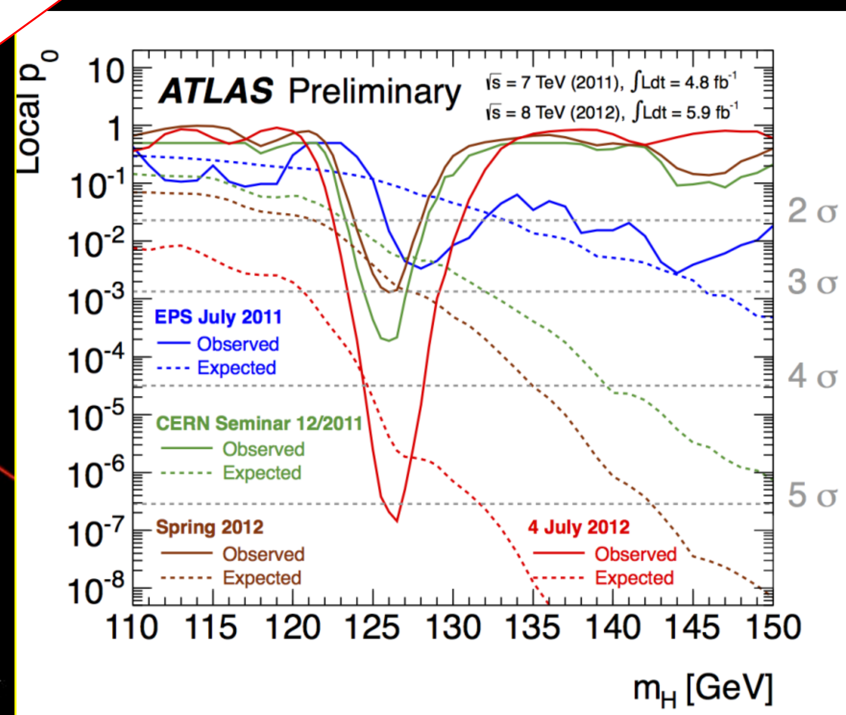


Grid computing enables the rapid delivery of physics results

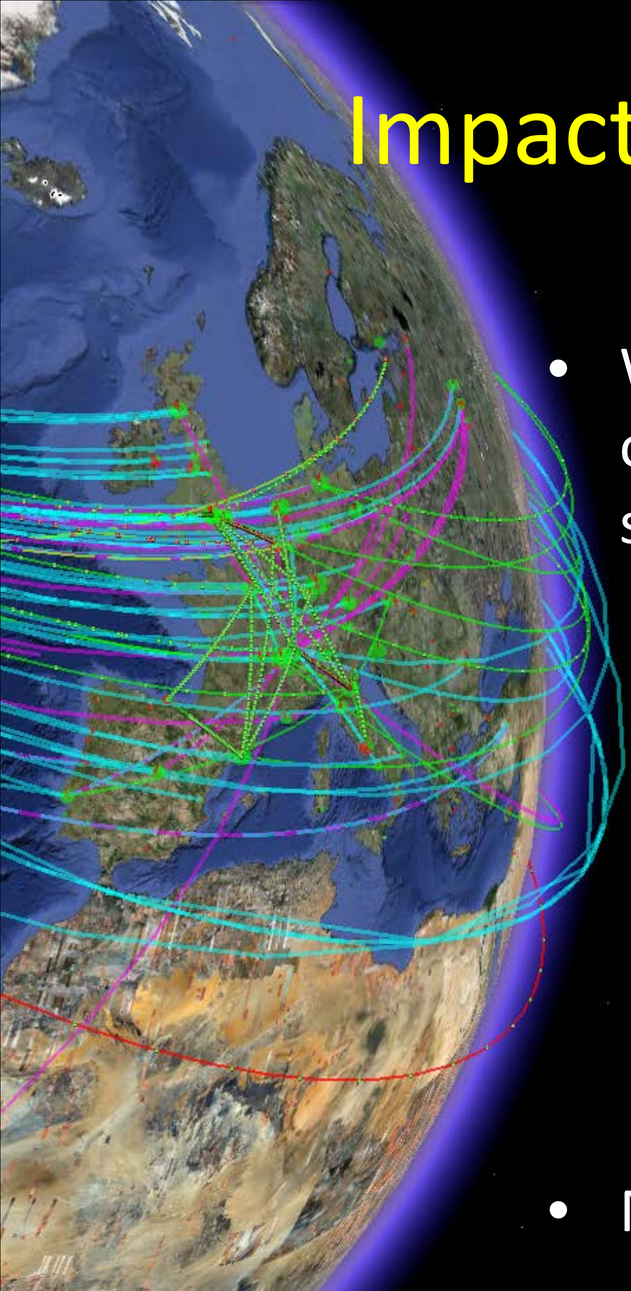








Impact of the LHC Computing Grid



- W-LCG has been leveraged on both sides of the Atlantic, to benefit the wider scientific community
  - Europe:
    - Enabling Grids for E-sciencE (EGEE) 2004-2010
    - European Grid Infrastructure (EGI) 2010--
  - USA:
    - Open Science Grid (OSG) 2006-2012 (+ extension?)
- Many scientific applications ➔

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

WLCG Worldwide LHC Computing Grid

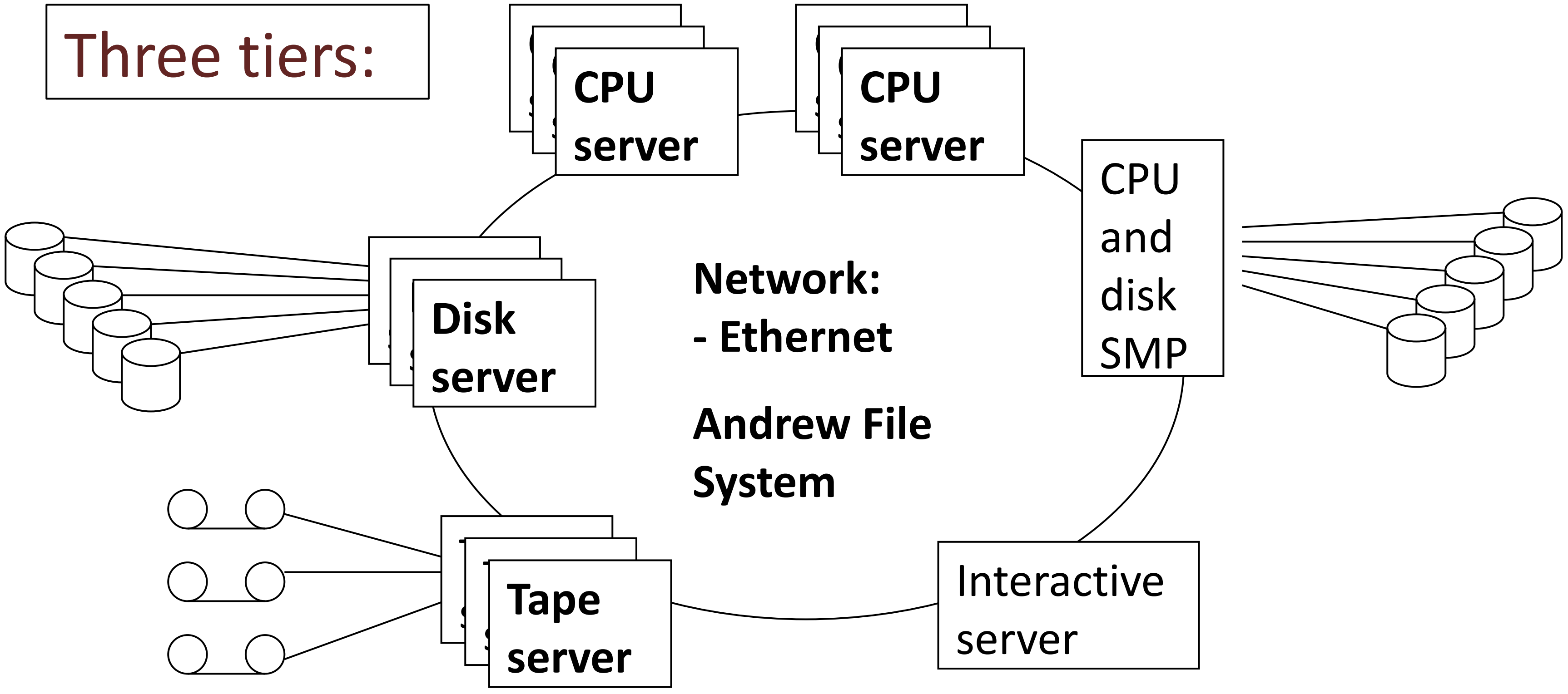
Back to our issues with Big Data

Credit: A. Pace

SHIFT architecture

(Scalable Heterogeneous Integrated Facility

- In 2001 the architecture won the 21st Century Achievement Award issued by Computerworld



Tier-0: Central Data Management

- **Hierarchical Storage Management: CASTOR**
  - Rich set of features:
    - Tape pools, disk pools, service classes, instances, file classes, file replication, scheduled transfers (etc.)
  - DB-centric architecture

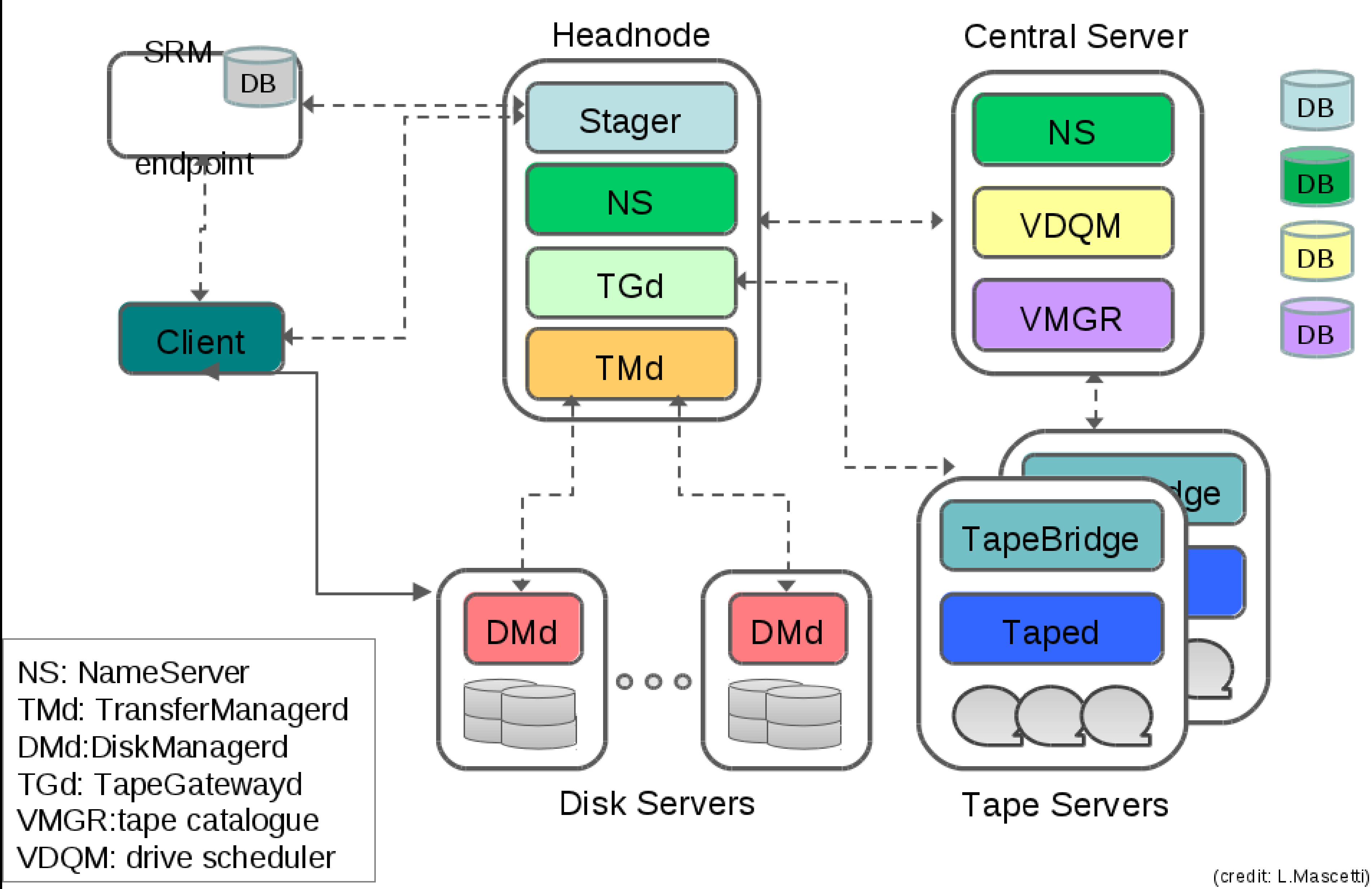
- **Disk-only storage system: EOS**
  - Easy-to-use, stand-alone, disk-only for user and group data with in-memory namespace
    - Low latency (few ms for read/write open)
    - Focusing on end-user analysis with chaotic access
    - Adopting ideas from other modern file systems (Hadoop, Lustre, etc.)
    - Running on low-cost hardware (JBOD and SW RAID)

Active tapes

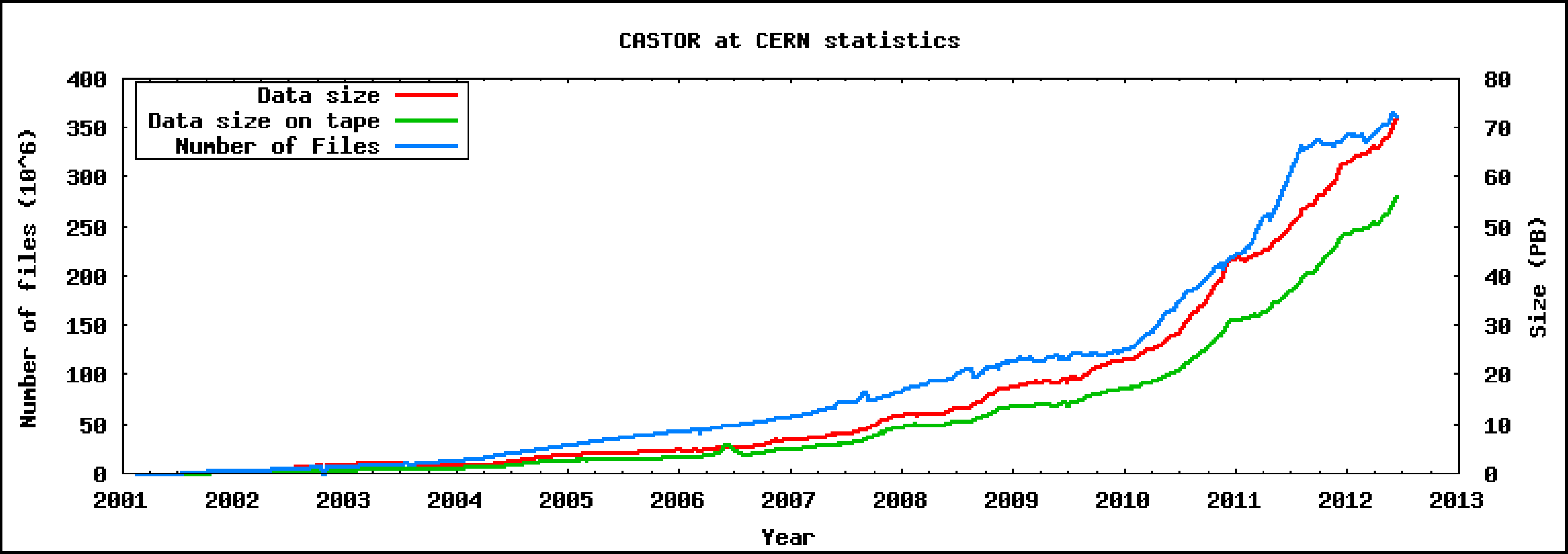
- Inside a huge storage hierarchy tapes may be advantageous!

We use tape storage products from multiple vendors

CASTOR architecture



CASTOR current status

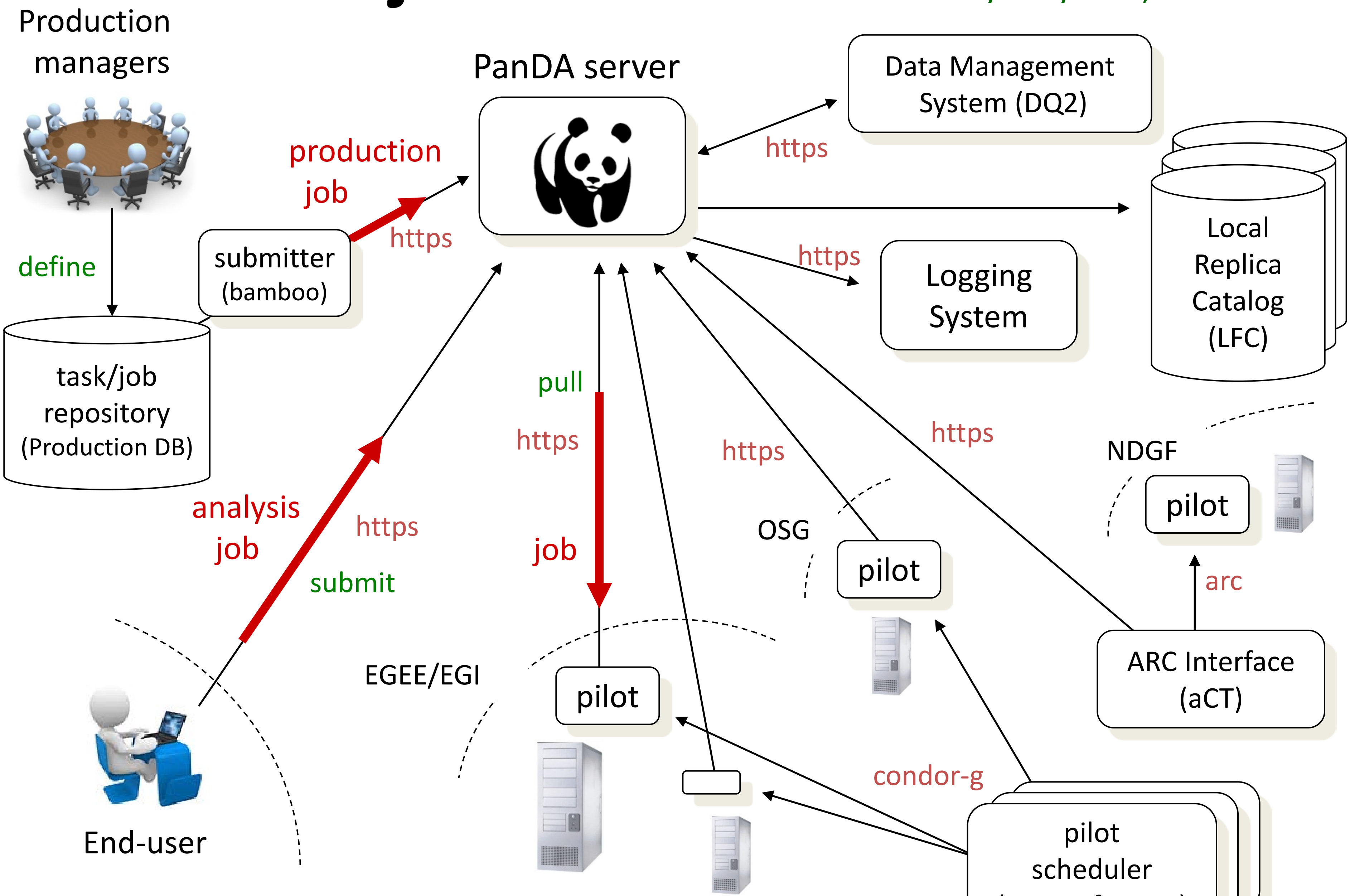


66 petabytes across 362 million files

Credit: J. Iven

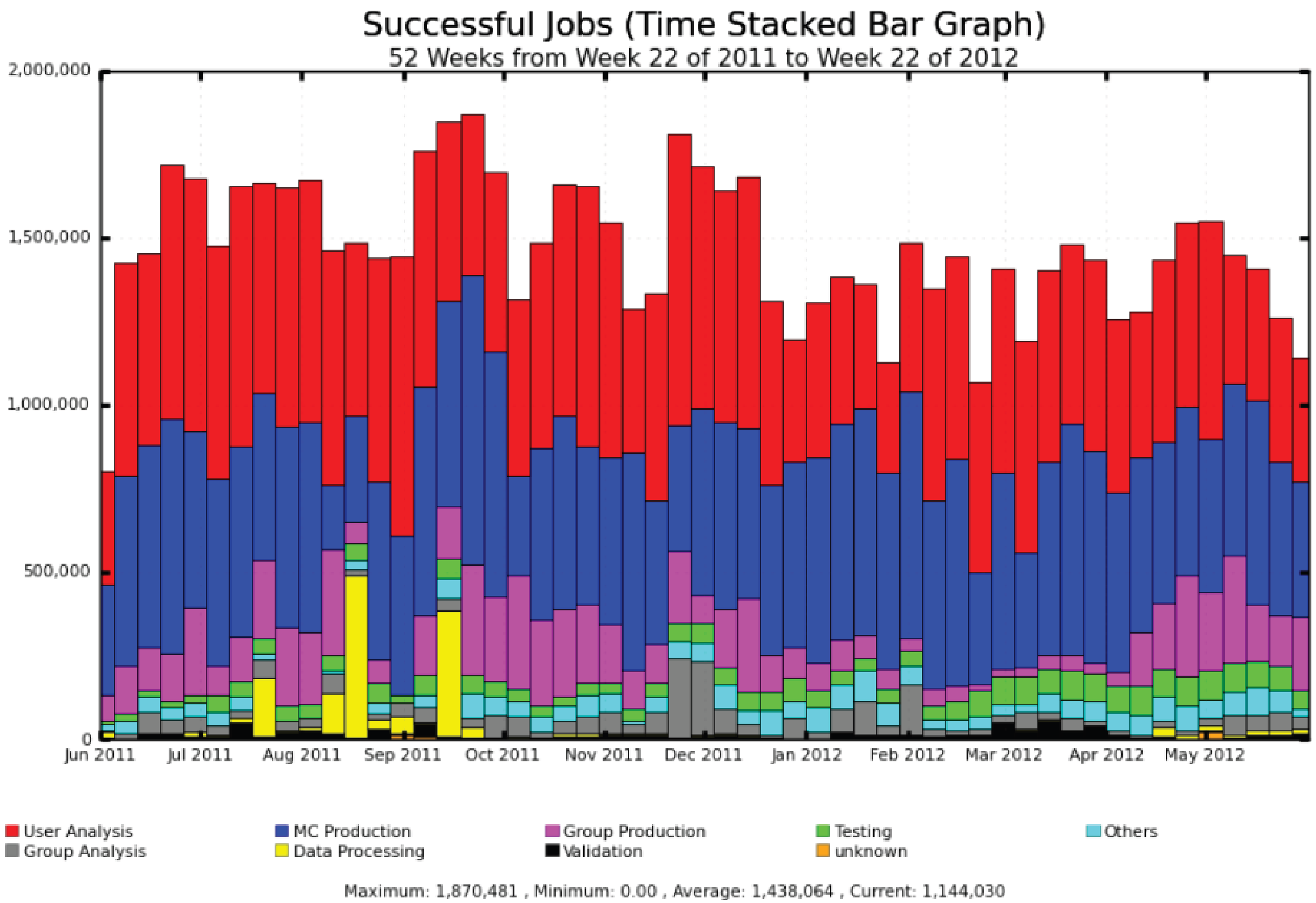
ATLAS job control

(Production and Distributed Analysis System)



Credit: T. Maeno

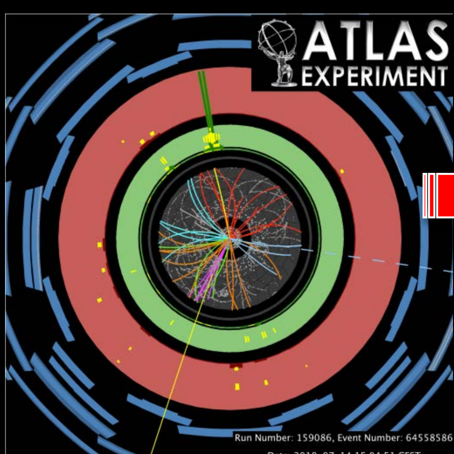
ATLAS User Analysis

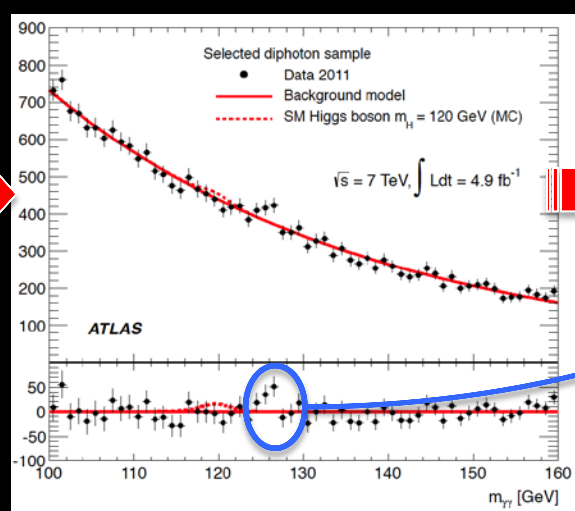


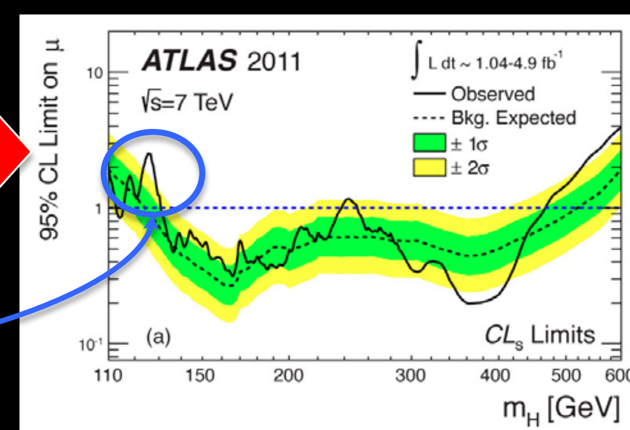
4

Data Analytics

- Huge quantity of data collected, but most of events are simply reflecting well-known physics processes
  - New physics effects expected in a tiny fraction of the total events:
    - "The needle in the haystack"
- Crucial to have a good discrimination between interesting events and the rest, i.e. different species
  - Complex data analysis techniques play a crucial role



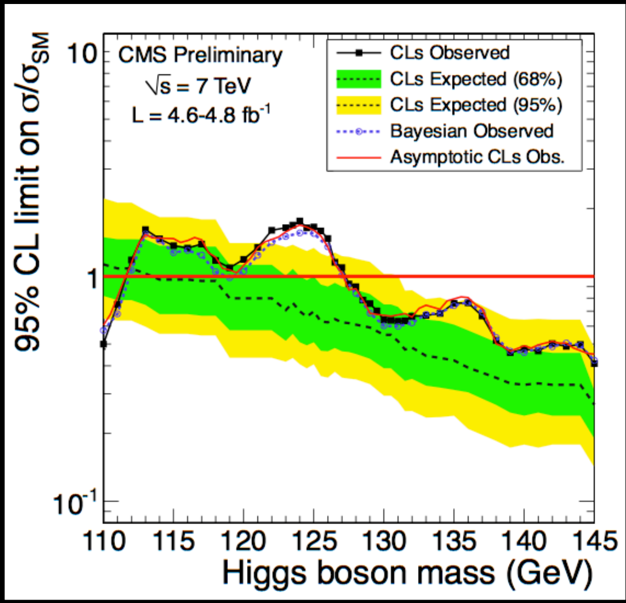


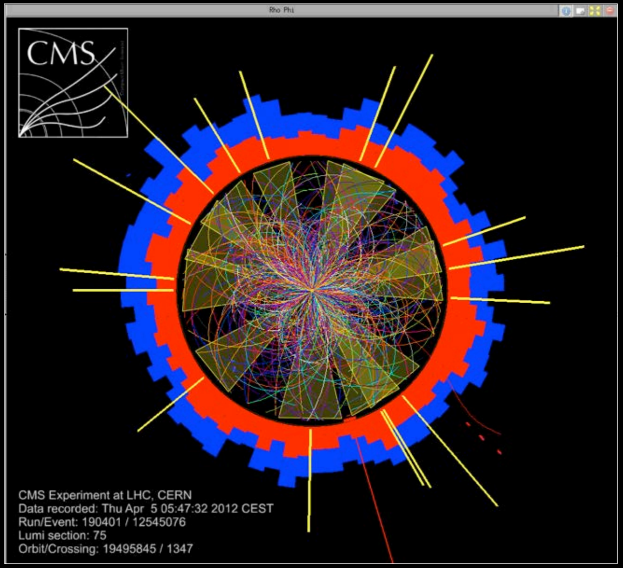


Credit: A. Lazzaro

ROOT Object-Oriented toolkit

- Data Analysis toolkit
  - Written in C++ (millions of lines)
  - Open source
  - Integrated interpreter
  - File formats
  - I/O handling, graphics, plotting, math, histogram binning, event display, geometric navigation
  - Powerful fitting (RooFit) and statistical (RooStats) packages on top
  - In use by all our collaborations

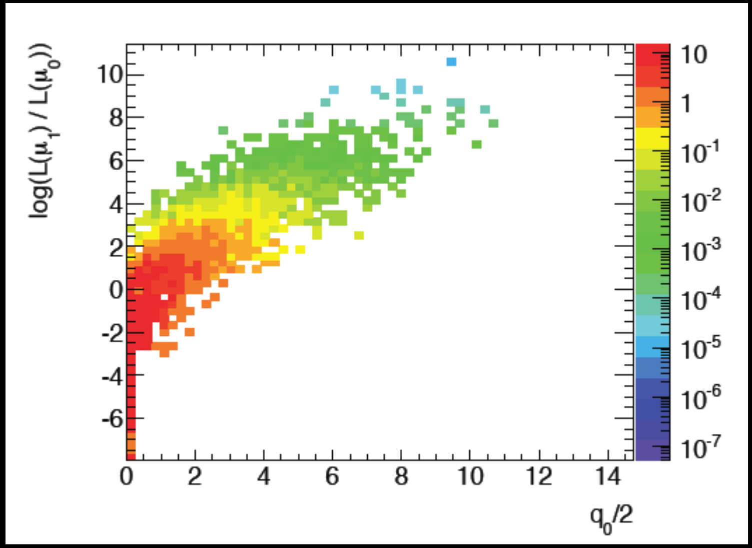


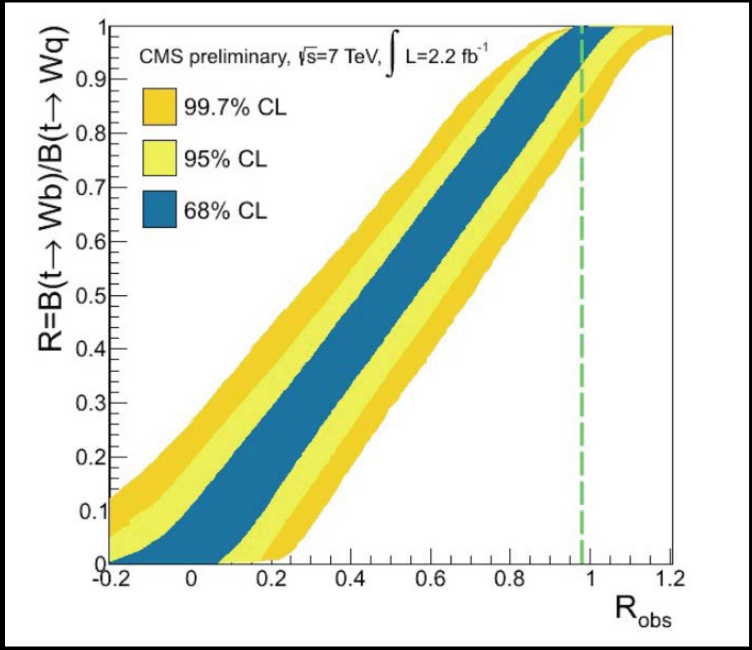


Credit: F. Rademakers

RooFit/RooStats

- Standard tool for producing physics results at LHC
  - Parameter estimation (fitting)
  - Interval estimation (e.g limit results for new particle searches)
  - Discovery significance (quantifying excess of events)
- Implementation of several statistical methods (Bayesian, Frequentist, Asymptotic)
- New tools added for model creation and combinations
  - Histfactory: make RooFit models (RooWorkspace) from input histograms





Credit: L. Moneta

ROOT files

- Default format for all our data
- Organised as Trees with Branches
  - Sophisticated formatting for optimal analysis of data
    - Parallelism, prefetching and caching
    - Compression, splitting and merging

Over 100 PB stored in this format (All over the world)

Big Data at CERN

- All of this needs to be covered!

Credit: A. Pace

Conclusions

- Big Data Management and Analytics require a solid organisational structure at all levels
- Must avoid "Big Headaches"
  - Enormous files sizes and/or enormous file counts
  - Data movement, placement, access pattern, life cycle
  - Replicas, Backup copies, etc.
- Big Data also implies Big Transactions/Transaction rates
- Corporate culture: our community started preparing more than a decade before real physics data arrived
  - Now, the situation is well under control
  - But, data rates will continue to increase (dramatically) for years to come: Big Data in the size of Exabytes!

There is no time to rest!

References and credits

- http://www.cern.ch/
- http://wlcg.web.cern.ch/
- http://root.cern.ch/
- http://eos.cern.ch/
- http://castor.cern.ch/
- http://panda.cern.ch/
- http://www.atlas.ch/

I am indebted to several of my colleagues at CERN for this material, in particular:
Ian Bird, WLCG project Leader
Alberto Pace, Manager of the Data Services Group at CERN and the members of his group

Q & A

CERN openlab