Big Fast Data in High-Energy Particle Physics

Andrew John Lowe

3 June 2015

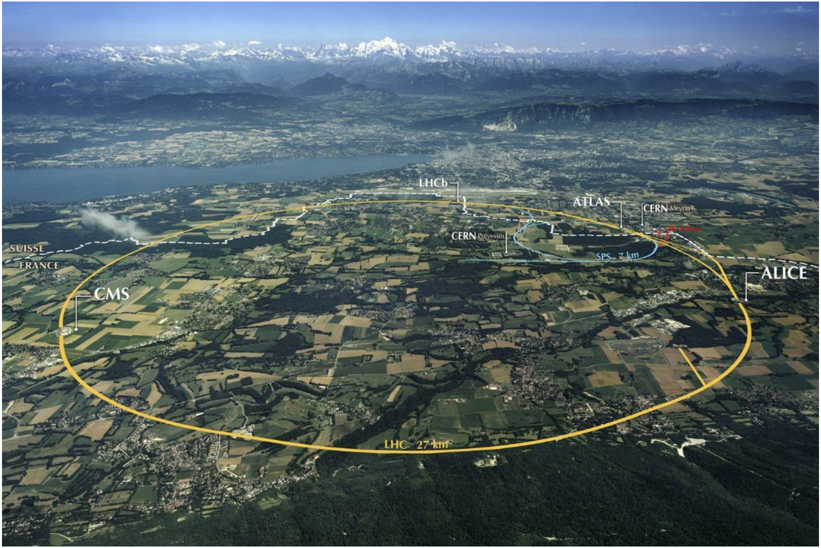
About CERN

- CERN is the European Organization for Nuclear Research, in Geneva, Switzerland
- Worldwide community
- 21 members states (+ 2 incoming members)
- Observers: Turkey, Russia, Japan, USA, India
- About 2300 staff and 10,000 users (about 5,000 on-site)
- Budget (2014) ~1000 MCHF
- Birthplace of the World Wide Web

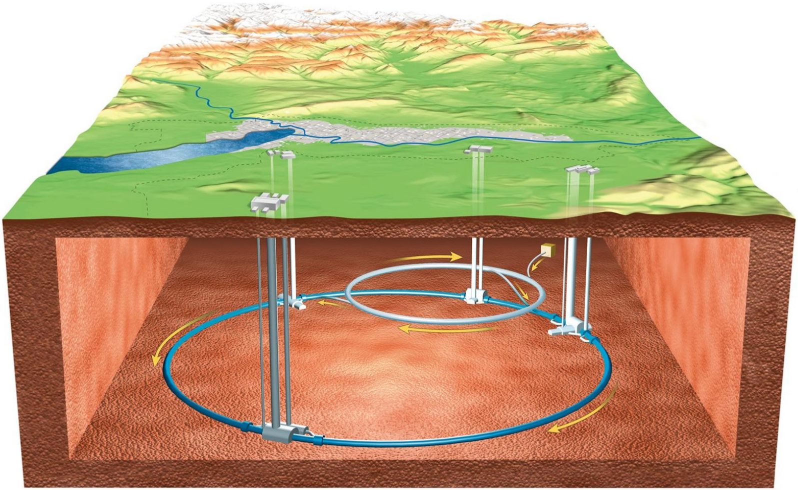
Particle physics

- Particle physics is the study of subatomic particles and the fundamental forces that act between them
- Present-day particle physics research represents man's most ambitious and organised effort to answer the question: *What is the universe made of?*
- We found the Higgs boson (after a 40-year search) but many questions remain unanswered:
  - Our present theoretical model doesn't explain gravity, identity of dark matter, neutrino oscillations, matter/antimatter asymmetry of universe ...
- To find the Higgs boson and attempt to unlock some of these other mysteries, we needed to:
  - Build the world's most powerful particle accelerator and largest machine ever constructed by humankind
  - Construct incredibly sophisticated particle detectors
  - **Collect an enormous amount of data**

The Large Hadron Collider (LHC)

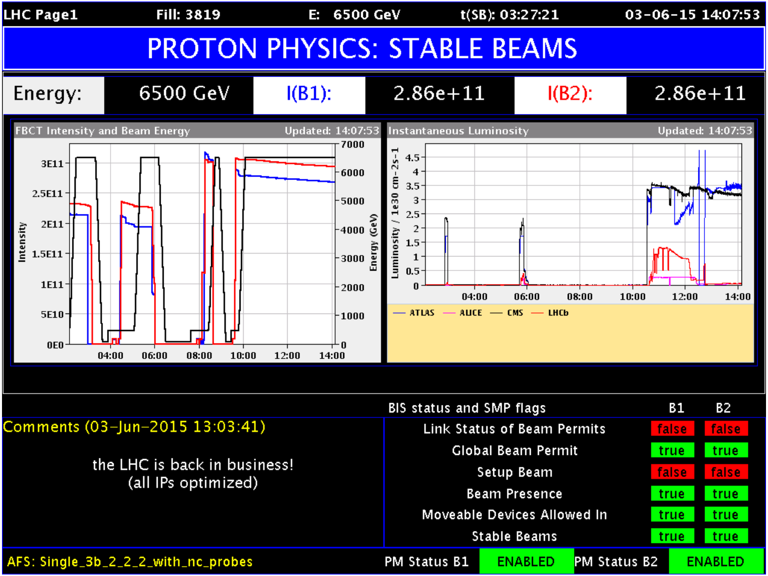


LHC underground structure



- 27 km circumference, 50–175 m underground
- Now running at 13 TeV after a two-year break for upgrades

LHC Run 2: first physics at 13 TeV starts **TODAY!**



Machine back on at 10:40 this morning!

BBC | Sign in | News | Sport | Weather | Shop | Earth | Trav

NEWS

Home | Video | World | UK | Business | Tech | Science | Magazine | Entertainment & Arts

Science & Environment

Hadron collider turns on data tap

The Large Hadron Collider has re-started scientific investigations after a two-year pause.

2 hours ago

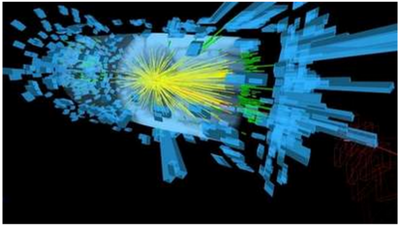
Science & Environment

LHC smashes collision energy record

LHC restart sees first collisions

Hadron Collider restarts after pause

What next for the Large Hadron Collider?



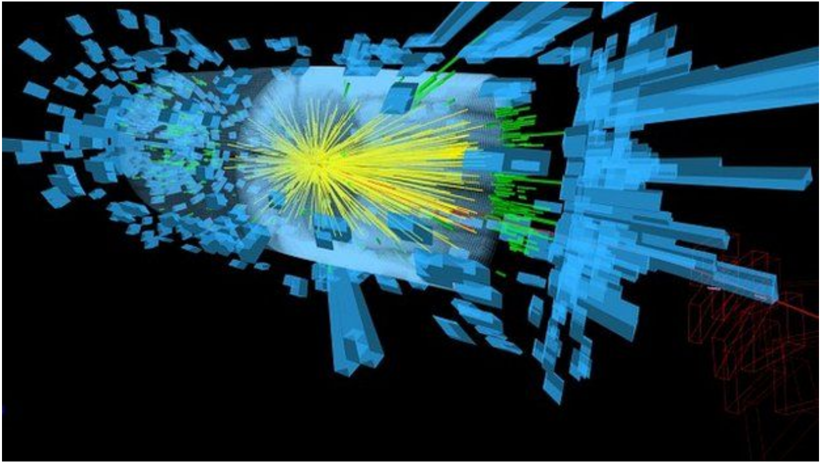
We are taking data at record energy **NOW!**

CMS Control Room

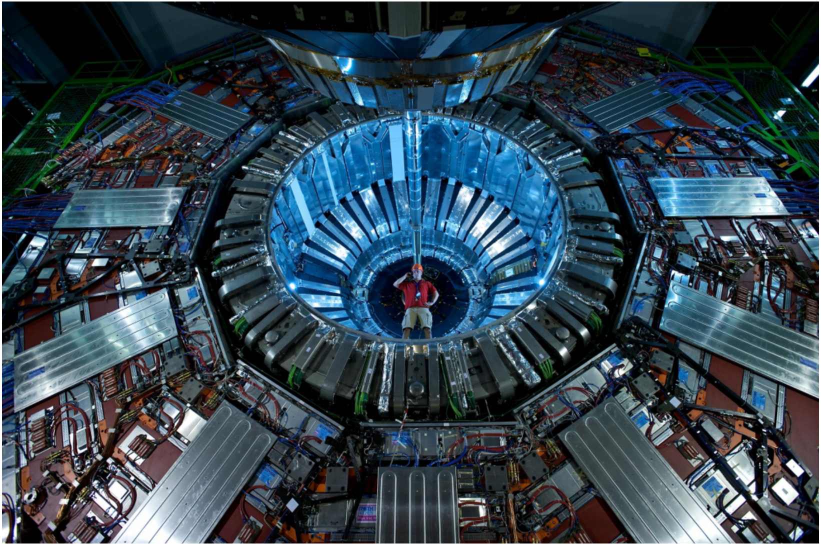
ATLAS Control Room

LHC Control Room

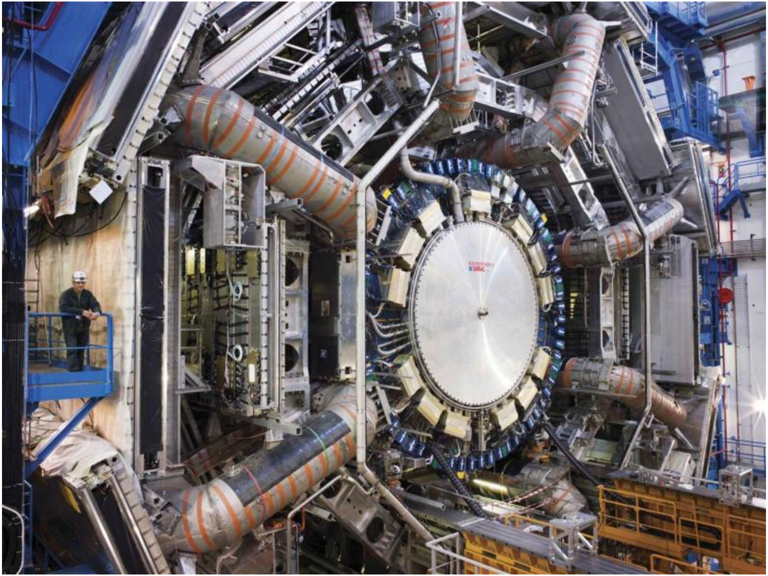
New data!



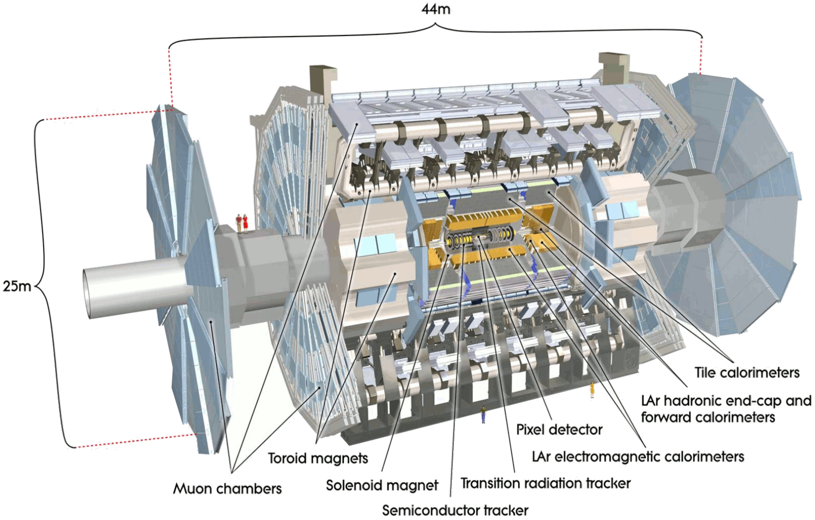
LHC detector experiment: CMS



LHC detector experiment: ATLAS



Example detector at the LHC: the ATLAS detector



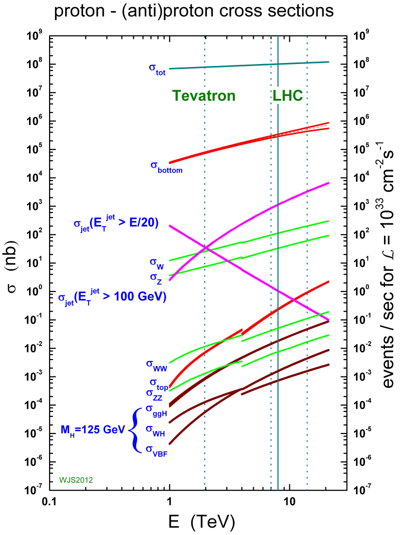
Weight: 7000 tonnes • 3000 km of cables • 100 million electronic channels — *a big digital camera* — focus on this experiment now

Big machines and big data

- Why is the LHC so big?
  - Need to collide particles with enough energy to manifest new particles which (if they exist) have masses beyond those accessible with previous machines ($E = mc^2$)
  - Exploration of the energy frontier
- Why are the LHC detector experiments so big?
  - Large decay length of some particles; require decays happen *inside* detector volume where they can be recorded
- Why is our data so big?
  - How likely a given collision event occurs depends entirely on quantum mechanics and is a property intrinsic to that specific type of event
  - However, the *rate* depends on experimental variables (like beam intensity) that can be controlled
  - **Require huge data throughput**
- Parameters, architectural decisions and technology choices are driven by the physics

Rates for different physics processes at the LHC

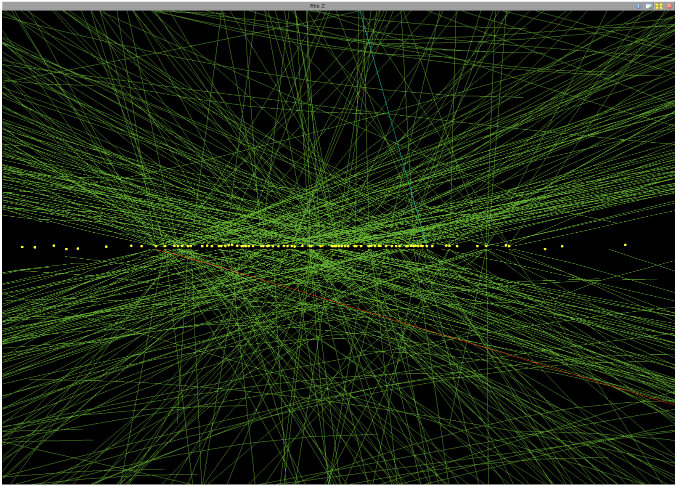
(Rare processes at bottom, frequent processes at top)



Data challenges for ATLAS

- We want to study extremely rare processes
  - For example, the production rate of Higgs bosons at the LHC is 10^{-11} that of the total proton-proton interaction rate
- A high collision rate (and long runs to collect *lots* of data) increases our chances of observing rare processes
- Beams are composed of "trains" of proton bunches that cross in the LHC detectors every 25 ns (rate = 40 MHz)
  - Bunches travelling close the speed of light $\rightarrow$ bunch separation is 7.5 m $\rightarrow$ before you've read out a single electronic channel from the 1$^{\text{st}}$ collision, the 2$^{\text{nd}}$ pair of colliding bunches are already in the detector, with the 3$^{\text{rd}}$ pair about to enter!
- There are a (Poisson) average of 23 proton-proton collisions per bunch crossing
  - These collision events are superposed; that is, **piled-up** one upon another!
- Full (zero-suppressed) event size of ATLAS is 1.5 MB
  - **This would result in a data rate of 60 TB/s!**

Example of collision event with "pile-up" (side view)

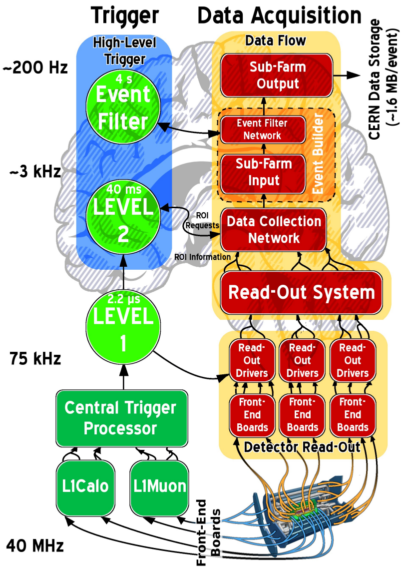


- There are 78 superposed proton-proton collisions in this single bunch-crossing event – very messy, but not uncommon!
- Is one of these collisions interesting enough to trigger the read out of the detector? *Must decide quickly!*

The ATLAS Trigger: processing Big Fast Data

- *Triggering* is the process whereby the detector's read-out system is triggered to record the data for a collision event that has been identified as interesting
  - Throwing away data in an unrecoverable way; focus on fast rejection
- The *Trigger* is a real-time multi-stage cascade classifier composed of three levels; each refine the trigger decision:
  1. Radiation-hard electronics, latency: 2 µs, output rate: 75 kHz,
  2. Software-based, latency: 10 ms, output rate: 3 kHz,
  3. Software-based, latency: ~1 s, output rate: 200 Hz, write-out to offline storage at 300 MB/ s, **expect to store a few PB/year**
- Level-2 and Level-3 run in PC farm (~17000 CPU cores)

ATLAS Trigger Architecture



ATLAS Level-1 Trigger

- Hardware based, radiation tolerant
- Mounted on, or near, the detector
  - Cable propagation delays limit the time available for processing
- Coarse granularity ("low pixel resolution") detector data
- Uses only fastest subdetector systems
- During processing, data for multiple bunch-crossings are held in pipelined memories
- Identifies *Regions of Interest* – locations in the detector of objects passing trigger thresholds

ATLAS Level-2 and Level-3 Triggers

- Software based
- Access to full-precision detector data
- Basic idea: *seeded* and *stepwise* reconstruction
- Regions of Interest from Level-1 seed processing
  - Means only ~2% of the data needs to be transferred to Level-2
- The trigger software has four main components:
  - The Algorithms which process the event data
  - The Steering which guides and steers the algorithmic processing of events and is responsible for the trigger decision
  - The Data Manager which handles the event data during the trigger processing
  - The Event Data Model which specifies the objectified representation of the event data to be used by the algorithms

Algorithmic processing in the ATLAS Trigger software

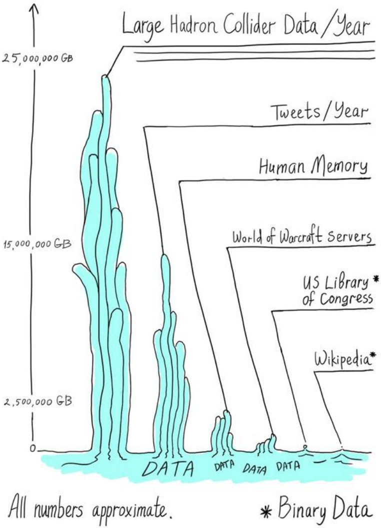
- There are two types of trigger algorithm:
  - Feature extraction algorithms process the event data and produce abstract physics objects ("features") that represent candidates for electrons, muons, jets, and so on. FEX algorithms operate on features and produce new ones, thereby refining the event information.
  - Hypothesis algorithms perform a task similar to particle identification; a Hypothesis algorithm tests whether a previously created feature agrees with the hypothesis of an assumed physics object by applying selection cuts on the feature's properties. It can then flag the hypothesis as valid or invalid.
- Algorithm sequencing is driven by a static configuration that informs the Steering which Algorithm must be executed in the case that a particular (dynamic) trigger condition is active
- Configuration: menu of trigger signatures we're interested in
- Chain of algorithms can be stopped at any validation step
- Reach end of algorithm chain: read out data for offline storage

ATLAS Level-2 and Level-3 processor farm

How big is our data?

- LHC experiments produced ~30 PB of data per year in Run 1
- Run 2 (now!) ~50 PB/year
- By 2023: 400 PB/year
- A typical LHC experiment dataset has a size of tens of TB
  - On my own experiment, sizes are sometimes hundreds of TB
  - Simulated 3.5 PB of Monte-Carlo data with combined running time of 18811 years
- Over the past 20 years, the CERN Computer Centre has recorded 130 PB or data – about 100 PB in the last five years
- Bulk of data is stored on magnetic tape
- Frequently-accessed (hot) data stored in disk pool system, cold data on tape; stage-in data to disk from tape on demand

Data size comparison

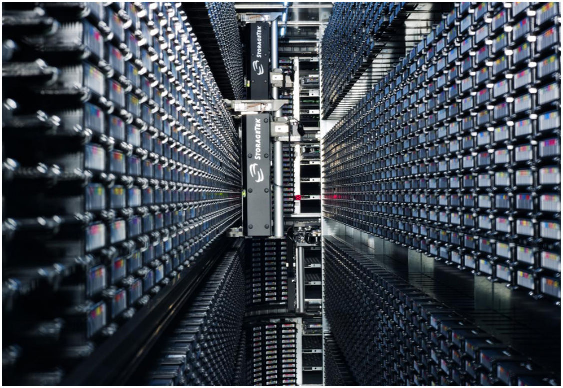


From "Particle physics tames big data", Symmetry, August 2012

Physics data handling — CERN Computer Centre

- CERN Computer Centre hosts 11,000 servers with 110,000 processor cores, 120 PB raw disk space, consumes 3.5 MW of power, processes about 1 PB per day

Tape storage



- 106 PB on tape • 25,000 tape cartridges • 1–5.5 TB each
- Cheap, compact and long-lasting; reliably read 30 years later
- If a tape snaps, it can be spliced back together
  - CERN looses only a few hundred MB of data on tape per year
- Don't need power to preserve the data held on them
- Safe from hackers

Data analysis on the Grid

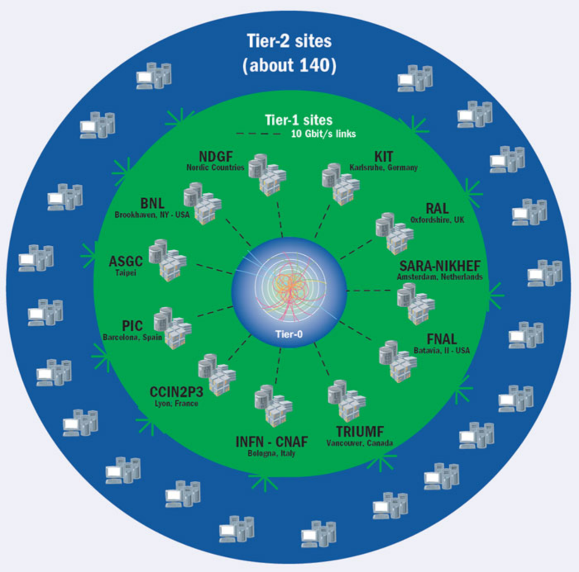
- The Worldwide LHC Computing Grid consists of some 200,000 processing cores and 150 petabytes of disk space, distributed across 36 countries through leased data lines
- These computer centres are arranged in "Tiers"
  - **Tier-0:** This is the CERN Data Centre, which is located in Geneva, Switzerland and also at the **Wigner Research Centre for Physics in Budapest, Hungary**. First copy, first pass reconstruction, distribution of data to Tier-1s (by 10 Gbps optical fibre private network)
  - **Tier-1:** 13 computer centres located worldwide. Storage of a proportional share of data, large-scale reprocessing, distribution of data to Tier-2s,
  - **Tier-2:** Around 160 sites, typically universities and scientific institutes. End-user analysis and proportional share of data simulation and reconstruction.
- Users send analysis jobs to the data, job runs, get back results
- Every day WLCG processes more than two million jobs, corresponding to a single PC running for more than 600 years

The Wigner Data Centre

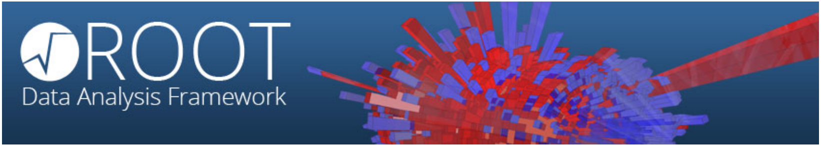
- Inaugurated in June 2013
- The Wigner Data Centre acts as a remote Tier-0 and an extension to the CERN Data Centre
- Also ensures full business continuity for the critical systems in case of a major problem on CERN's site
- 2700 servers, 43,000 computing cores, and 72 PB of storage
  - Installed capacity will eventually be increased to a level similar to that at CERN
- Long distance network connection to CERN: two independent 100 Gbps circuits
  - Bandwidth equivalent to the entire Hungarian domestic internet traffic
- The Wigner Data Centre was chosen after a tender open to all 20 CERN Member States

CERN-Wigner high-bandwidth connections

Architecture of Worldwide LHC Computing Grid



Tier-0: CERN (Geneva) + Wigner RCP (Budapest)

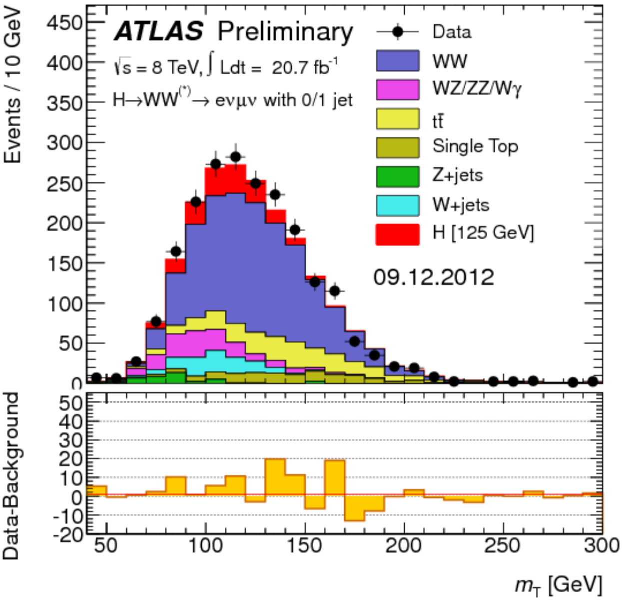


- For experimental particle physics, *ROOT* is the ubiquitous data analysis tool, and has been for the last 20 years old
- Command language: CINT ("interpreted C++") or Python
  - Small data: work interactively or run macros
- Data format optimised for large data sets
- Data in ROOT "tree" (like a hierarchical database)
- An entry represents an event (*i.e.*, a collison)
  - "Branches" (electrons, muons, photons, *etc.*)
    - "Leaves" (energy, momentum, mass, *etc.*)
- Basic idea: don't need all of the data all of the time
- Trees in many different files can be merged into one "chain"
- Access data in chain as if it was a tree in a single file
- Big data: build application with ROOT libraries, run on Grid

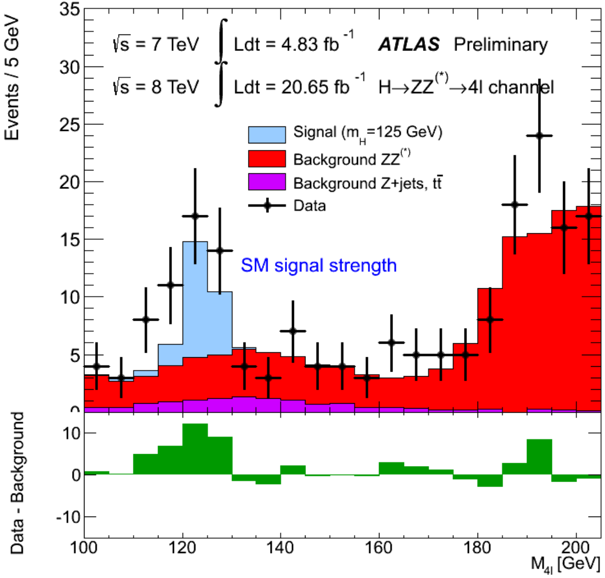
LHC data flow

1. Detected by LHC experiment
2. Online multi-level filtering (hardware and software)
3. Transferred to CERN and Wigner Tier-0, archived and reconstructed
4. Transferred to Tier-1 sites, archived, reconstructed and skimmed
5. Transferred to Tier-2 sites, reconstructed, skimmed, filtered and analysed
6. Written to locally-analysable files, put on PCs
7. Turned into plot in a paper

Higgs boson → WW signal in 2011 and 2012 data



Higgs boson → 4-leptons signal in 2011 and 2012 data



More information

- Data science @ LHC2015 Workshop
  - Workshop to help foster long-term connections between the data science and particle physics communities
- A mailing list `HEP-data-science@googlegroups.com` has just been created to deal with anything concerning both particle physics and data science, in particular machine learning
  - Announcement/discussion about workshops, challenges, papers, tools, *etc.*
  - Open to all, subscription by sending a mail to `HEP-data-science+subscribe@googlegroups.com`
- Explore the CERN experiments with Google Streetview
- Explore CERN's Computer Centre with Google Streetview
- "Processing LHC data" (short film)

Thanks!

https://www.linkedin.com/in/andrewjohnlowe

Bonus slides. . .

Data Centre statistics (2 June 2015)