UiO : University of Oslo

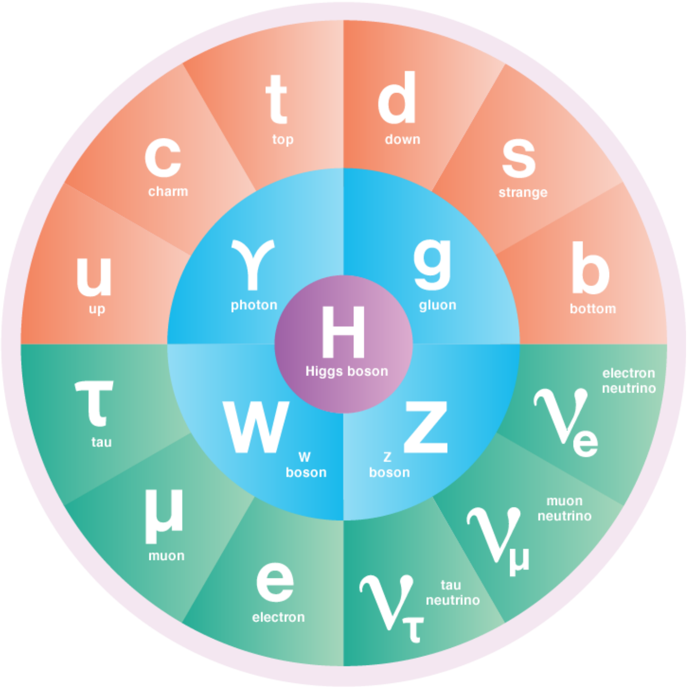
LHC status and future

(with a particular emphasis on Software and Computing)

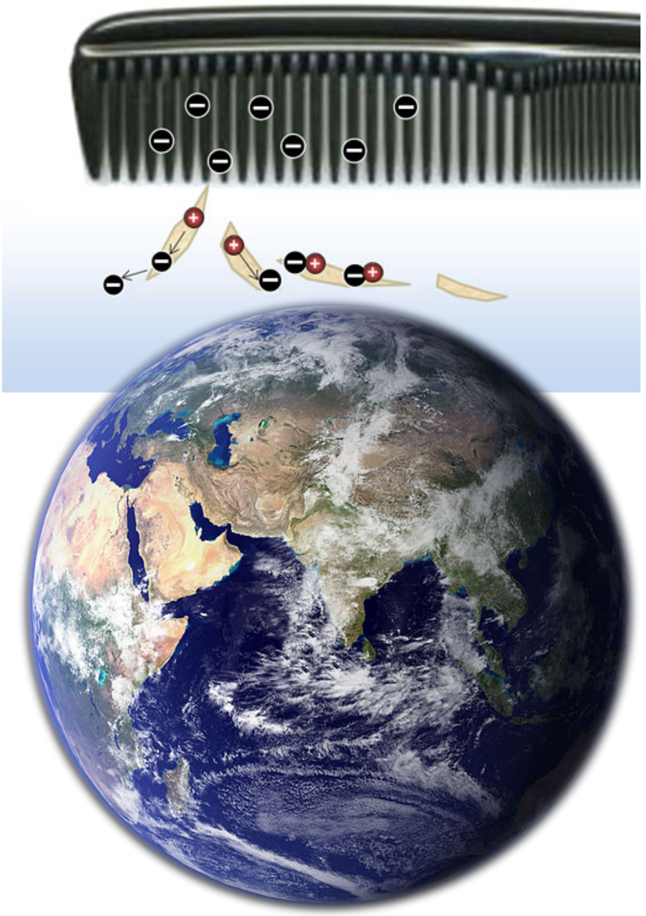
James Catmore
University of Oslo

Taking a step back: why are we here?

- Understanding the basic constituents of the Universe and the forces that govern them
  - Are there additional symmetries of Nature (+ new particles) other than those described by the Standard Model? SUSY? Something else?
  - Are quarks and leptons the smallest object that exist, or do they have internal degrees of freedom?

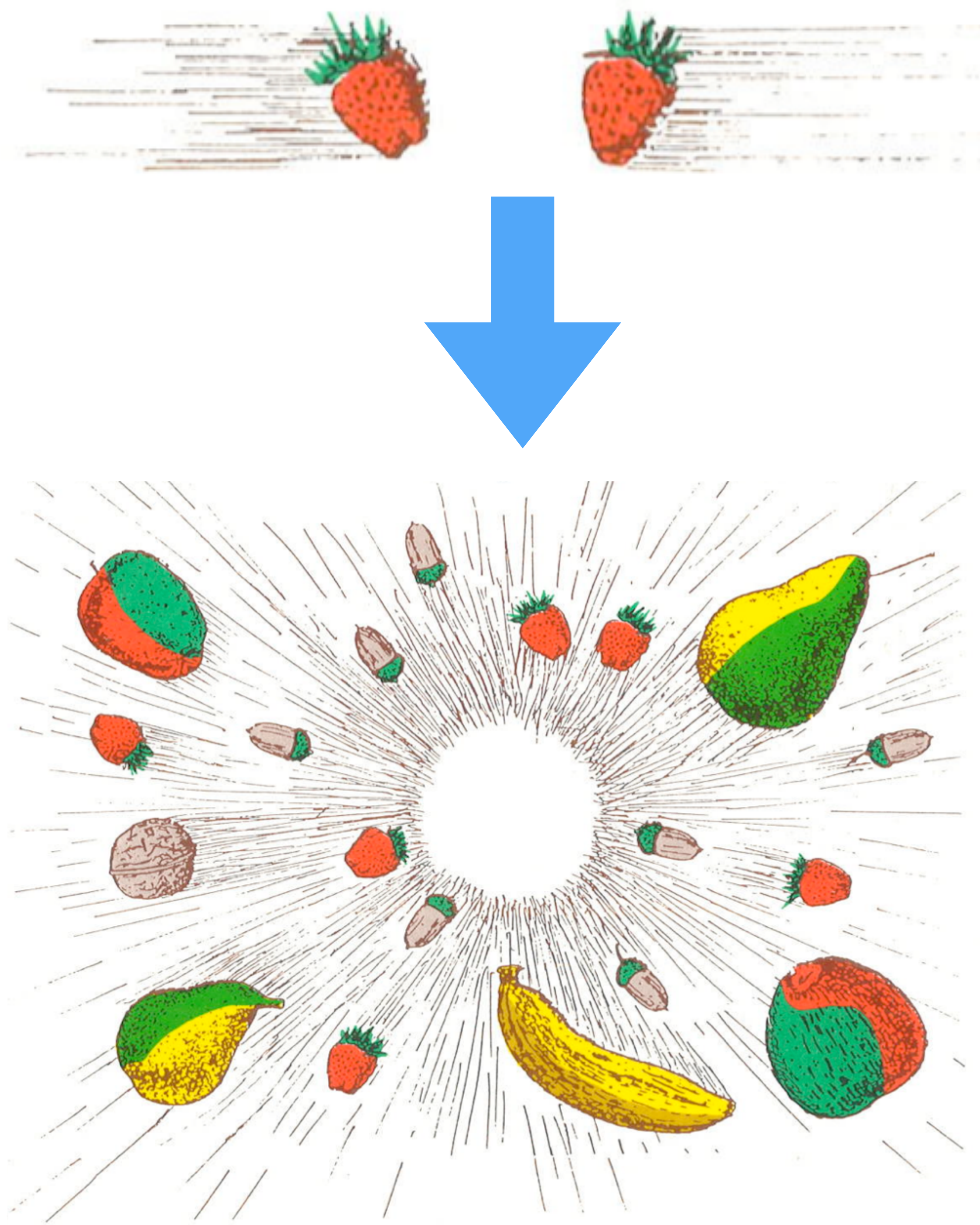


- Understanding how cosmology and fundamental physics interact
  - What is gravitation? Why is it so weak?
  - What is Dark Matter? And Dark Energy?
  - How did matter behave just after the Big Bang?
  - Where did all the anti-matter go after the Big Bang?

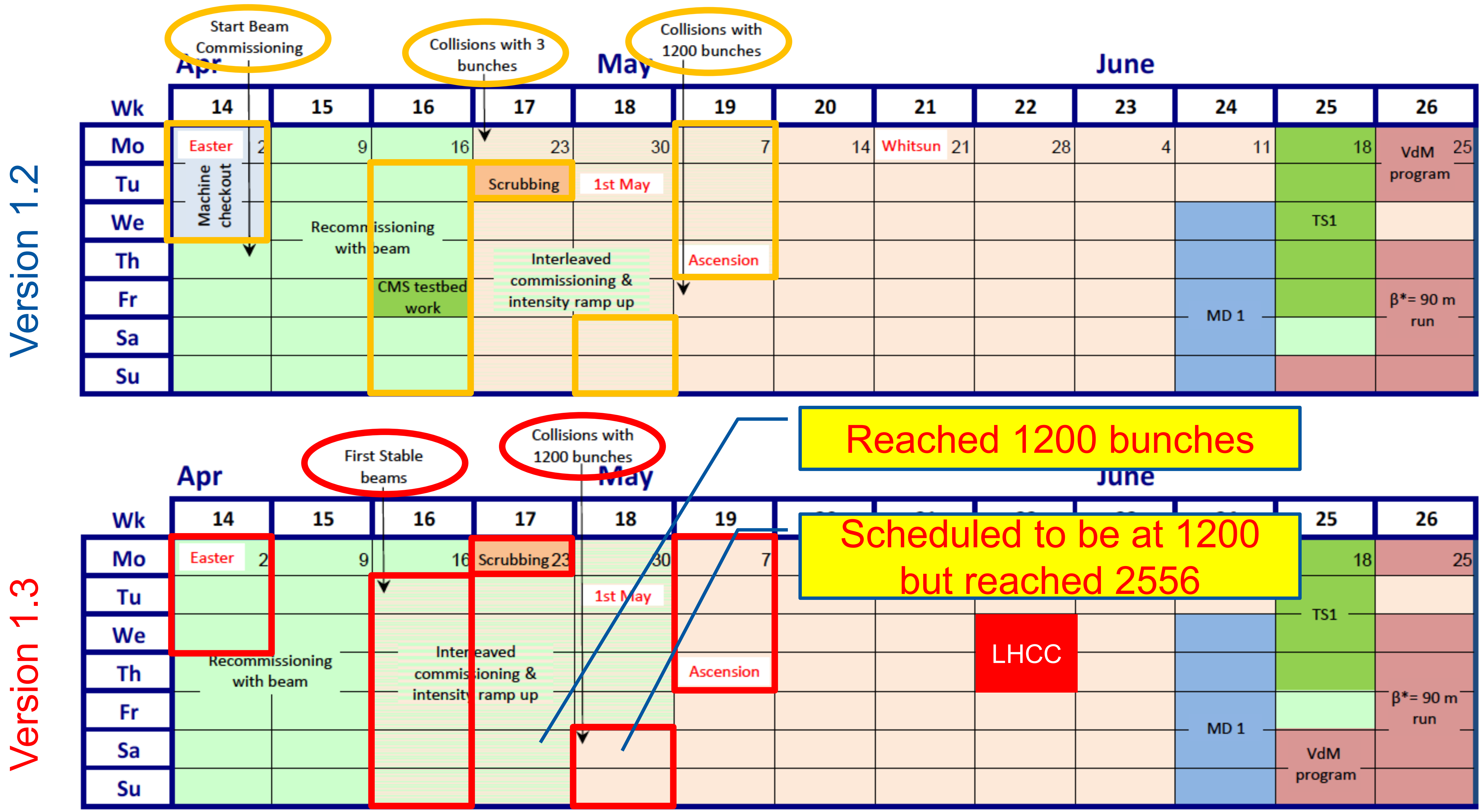


How to discover New Physics

- **DIRECT** searches: looking for evidence of new particles
  - SUSY, exotica, microscopic black holes
- **INDIRECT** searches: making high-precision measurements of known particles to observe deviations from SM predictions
  - SM bosons: Higgs, W, Z, γ
  - Top quarks and B-hadrons
  - Heavy Ion collisions

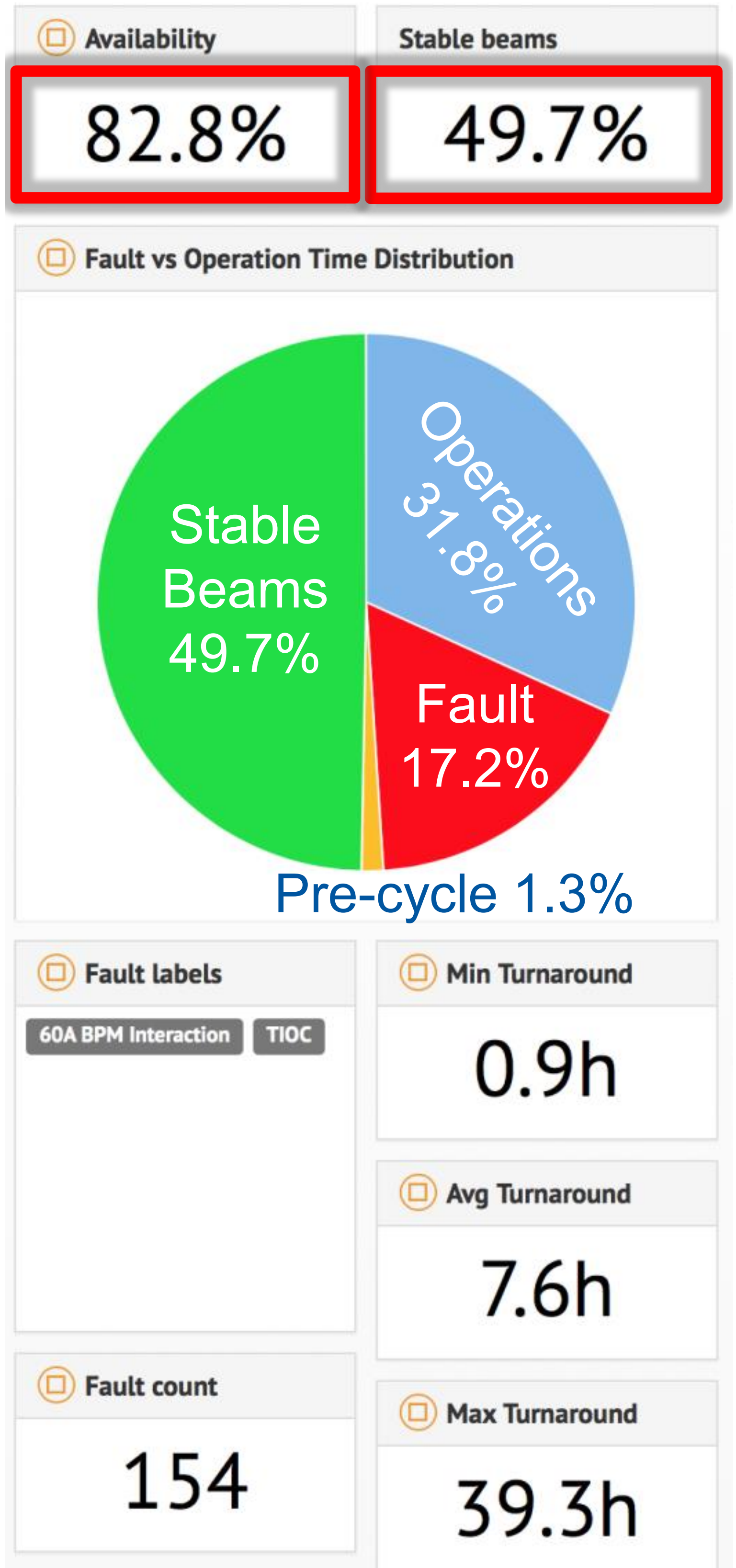


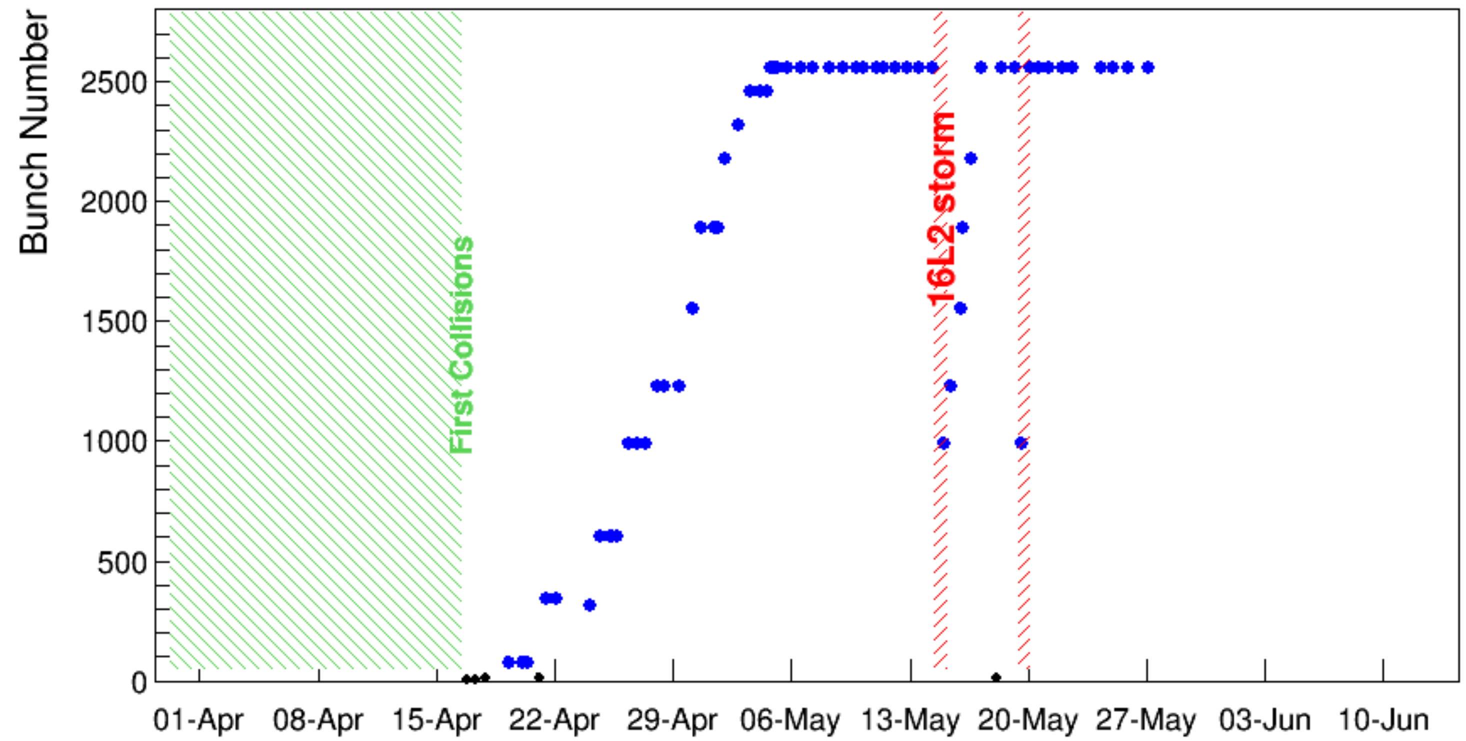
2018 LHC Schedule – Q2

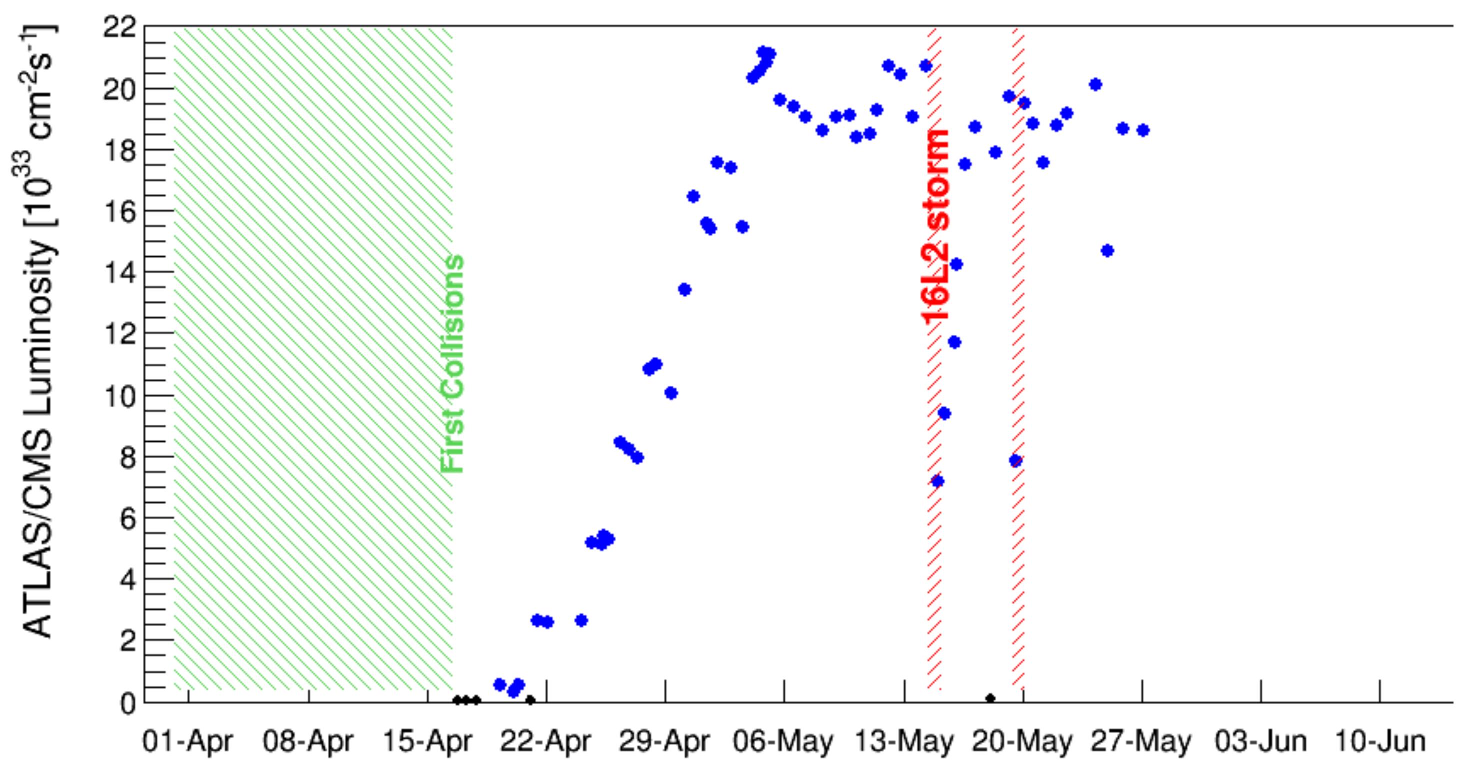


Full machine 13 days ahead of schedule

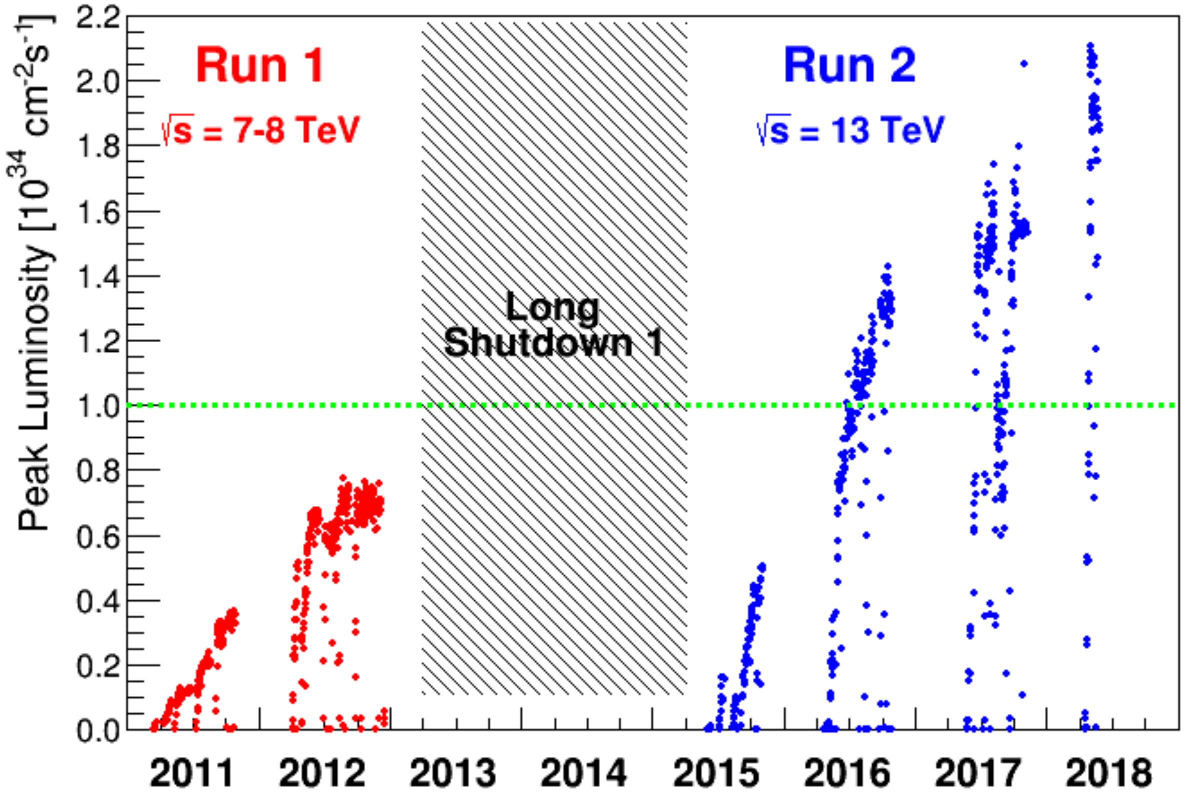
LHC status

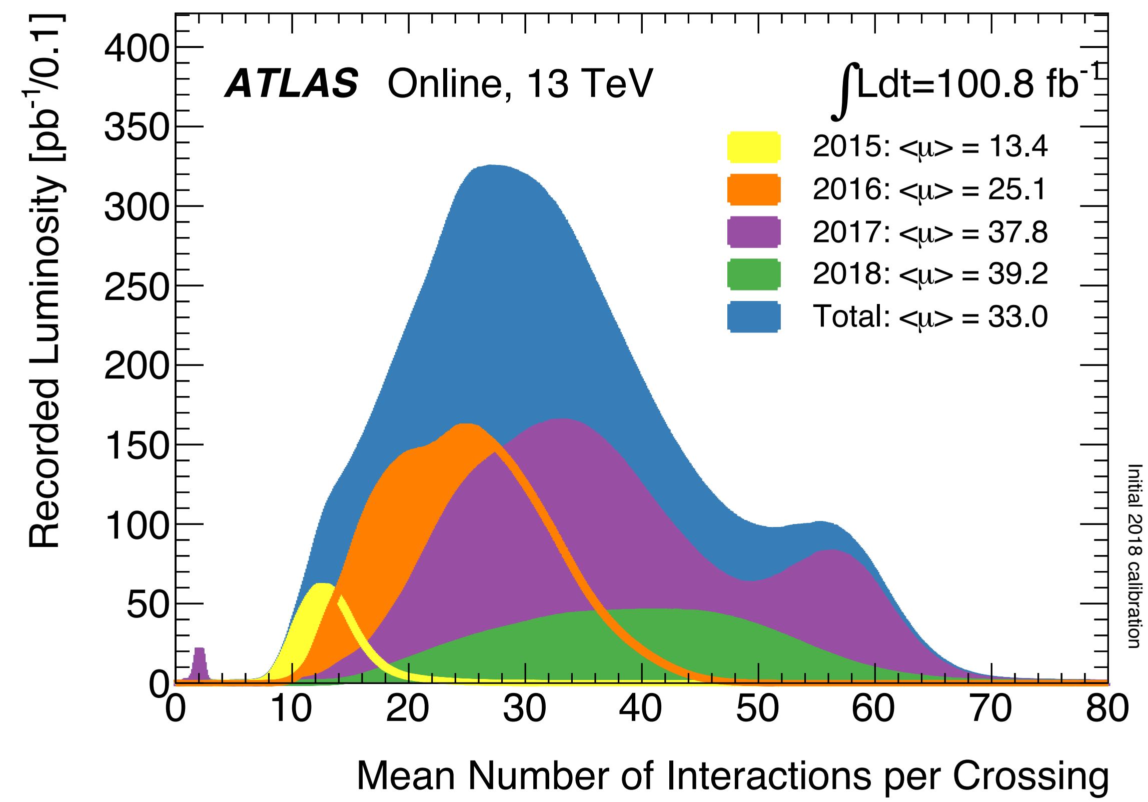


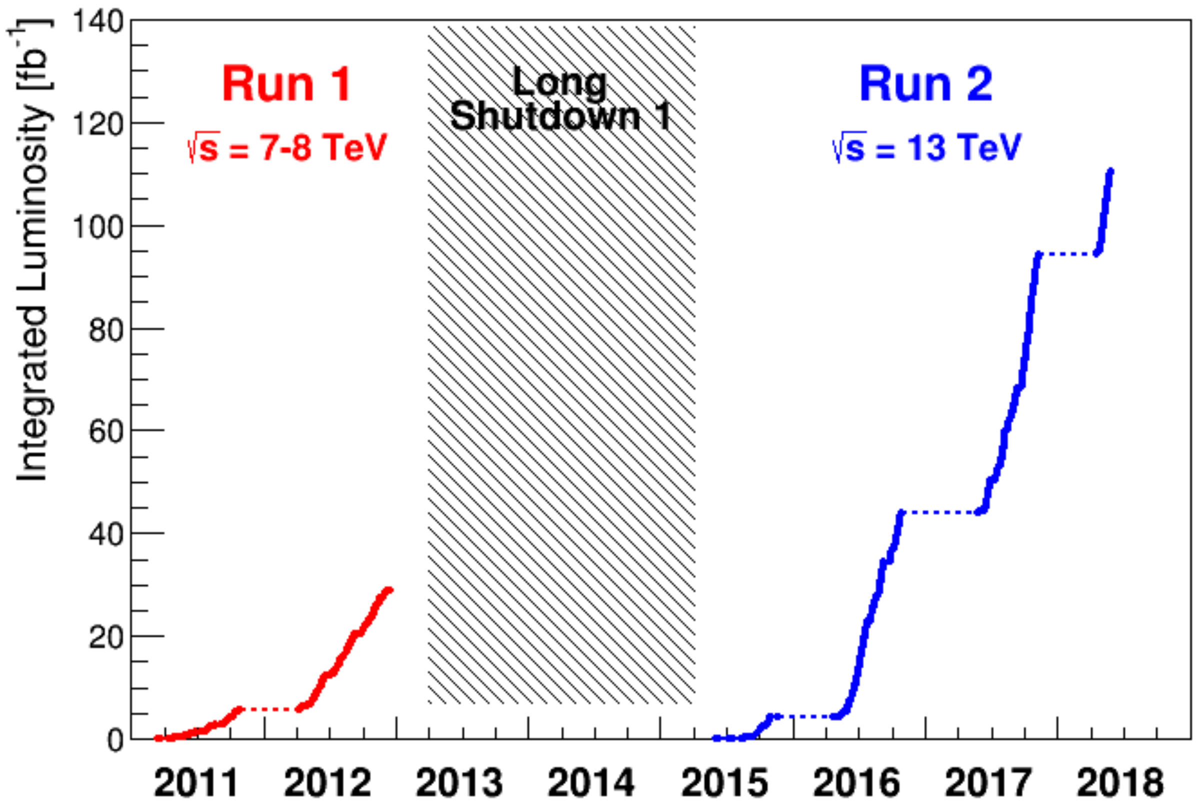


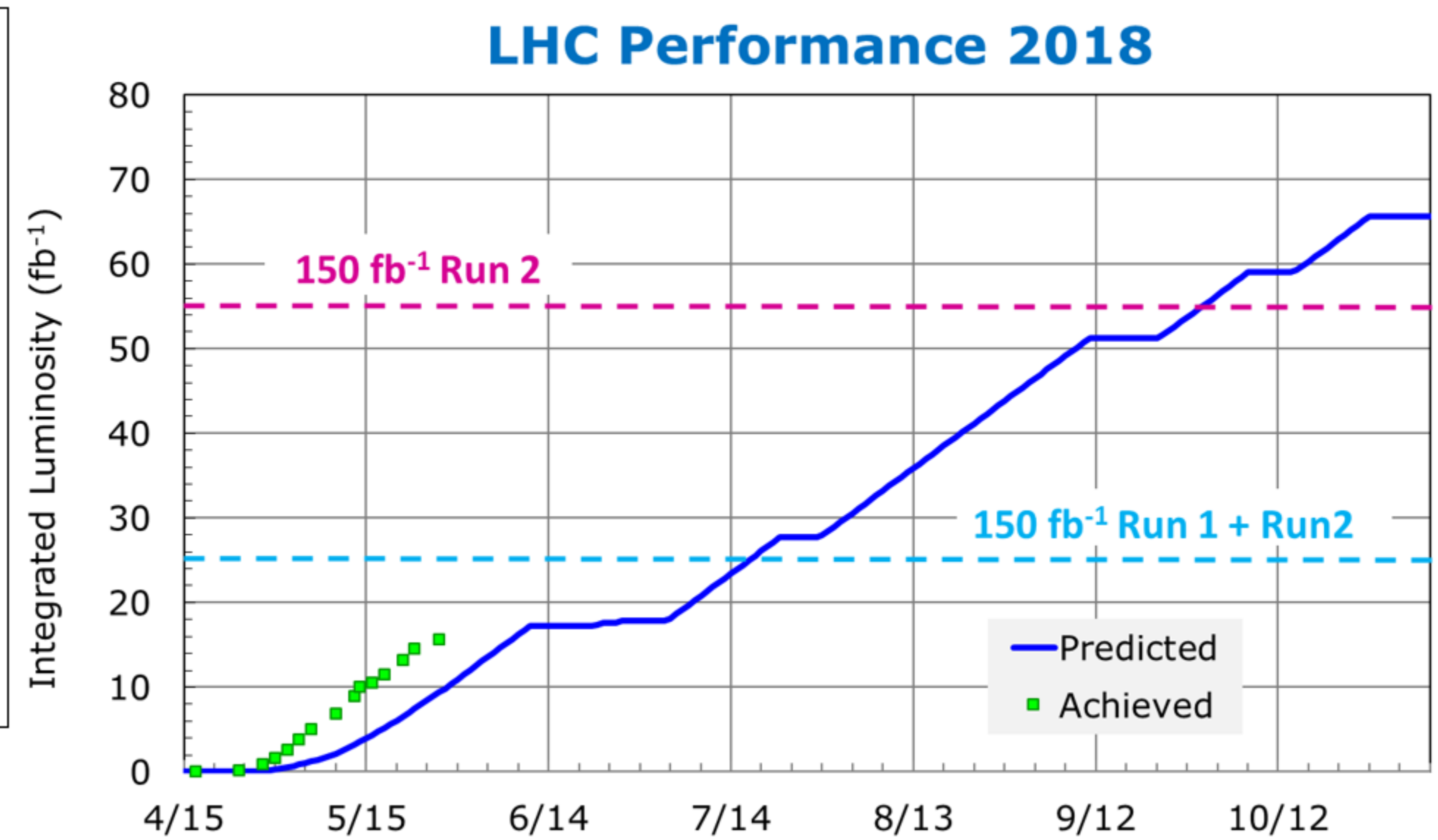


Rende Sterneberg, LHCC #134



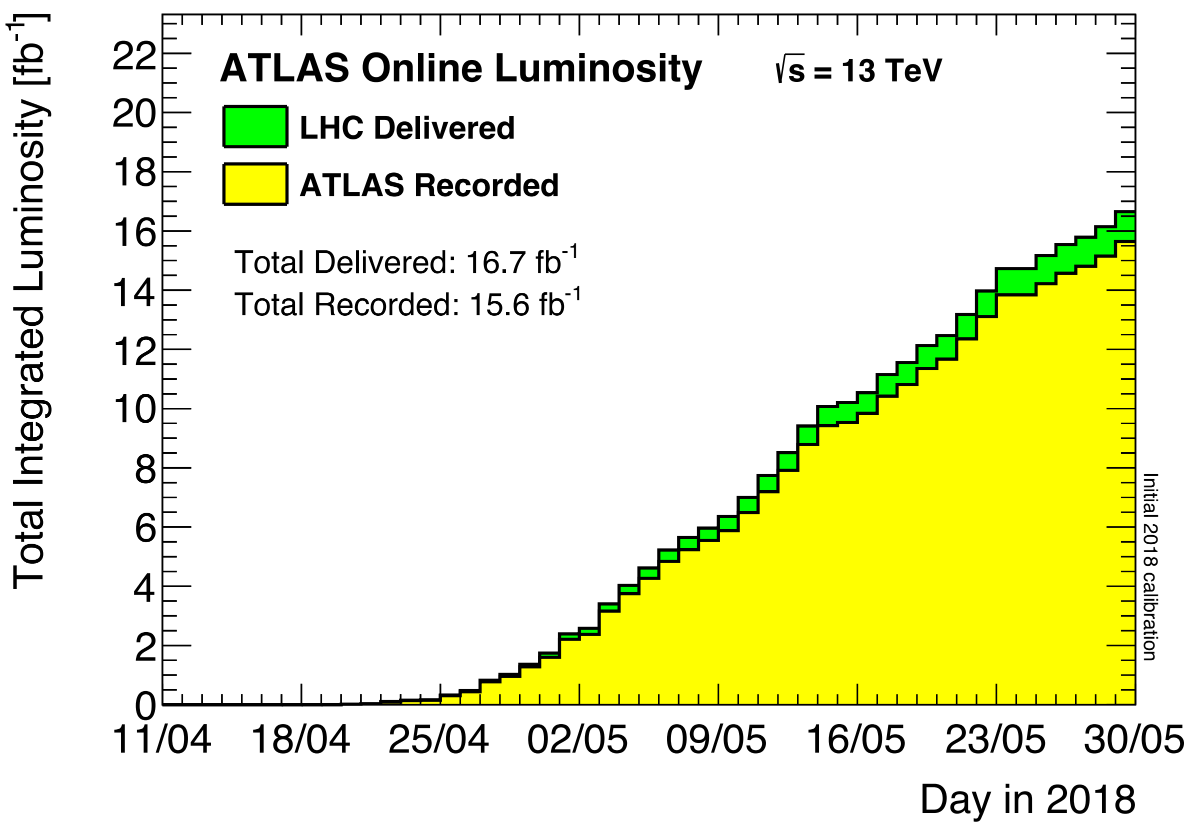


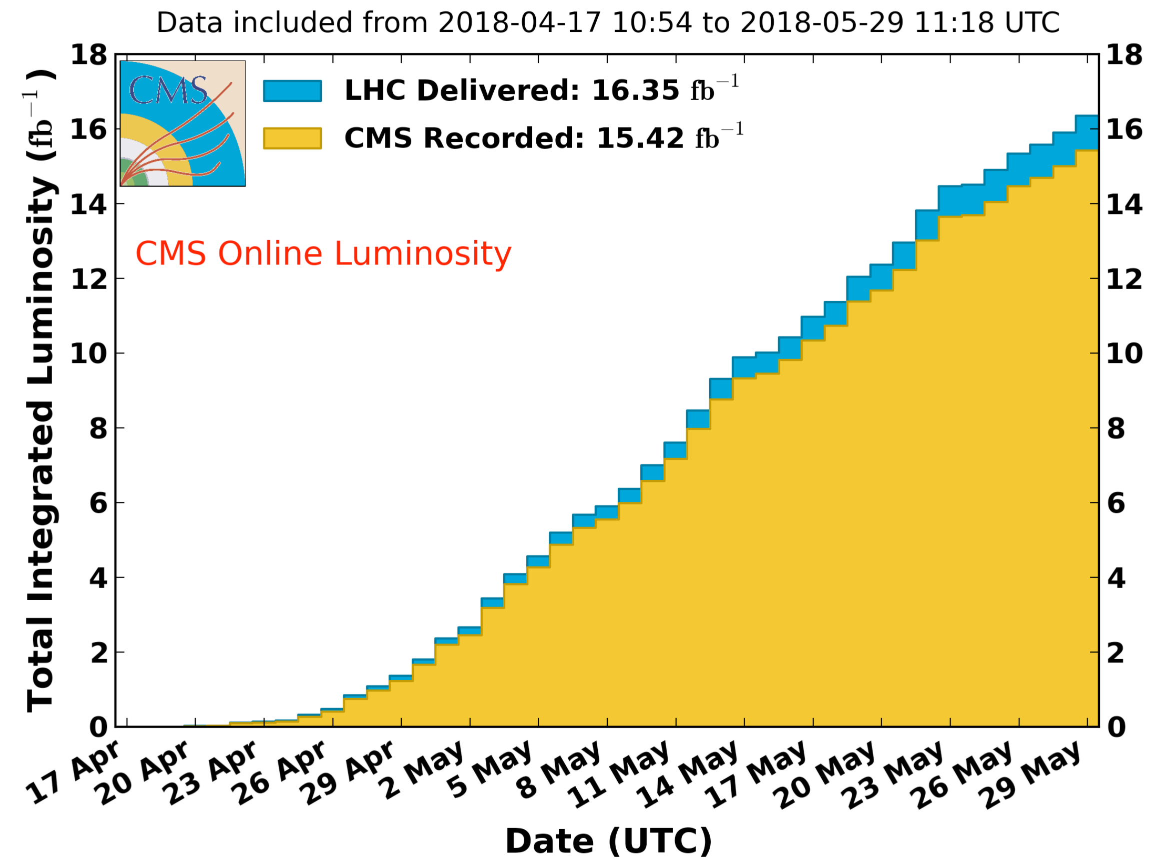


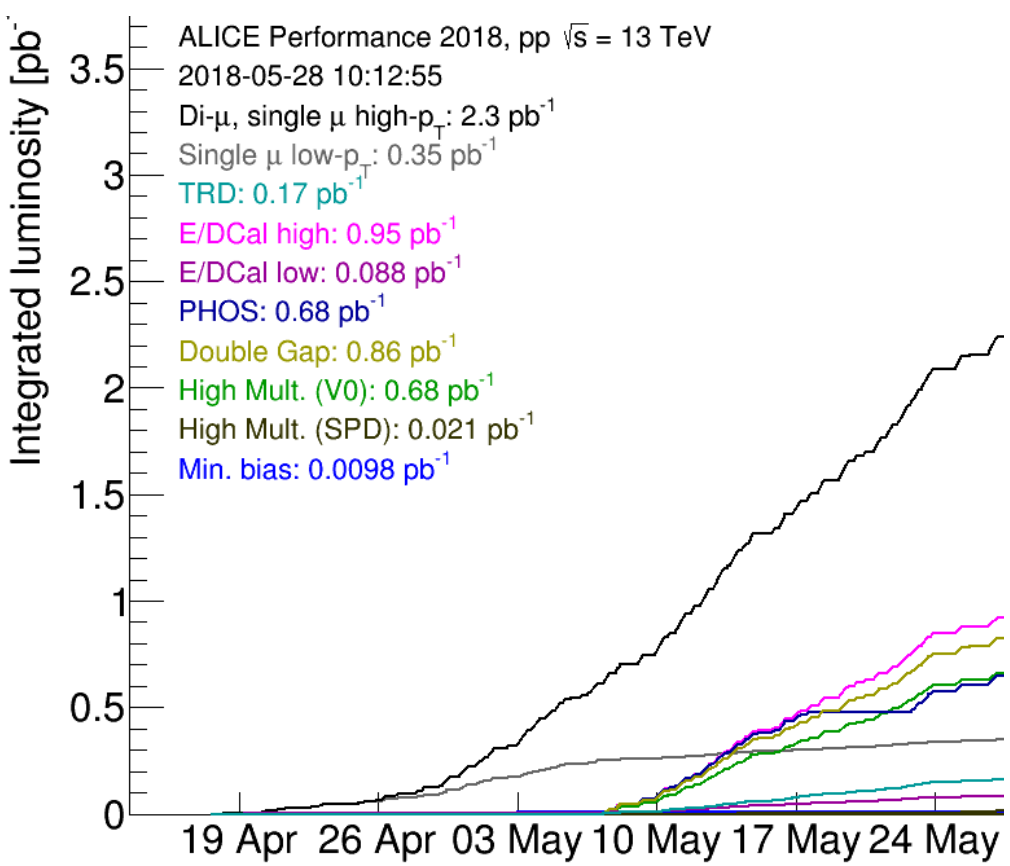


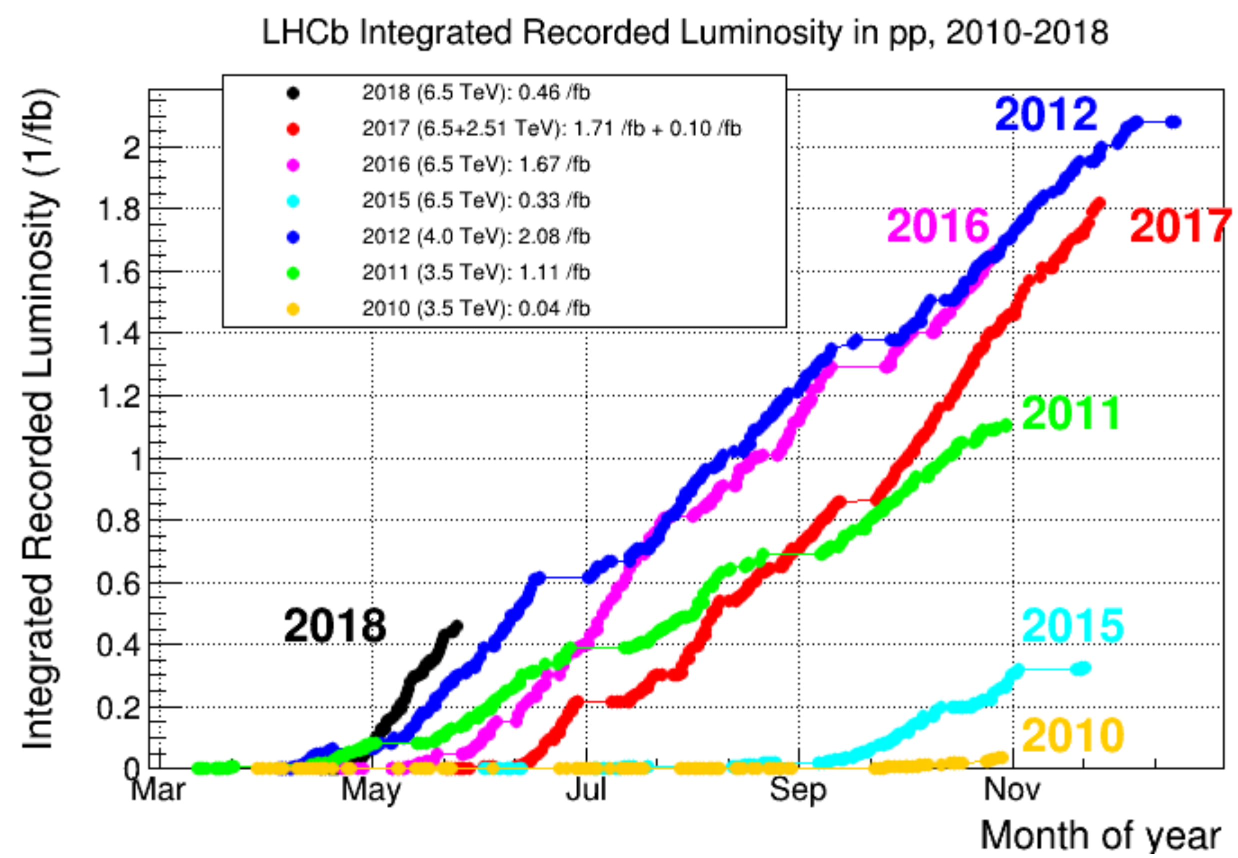
Rende Steerenberg, LHCC #134

Status of the experiments

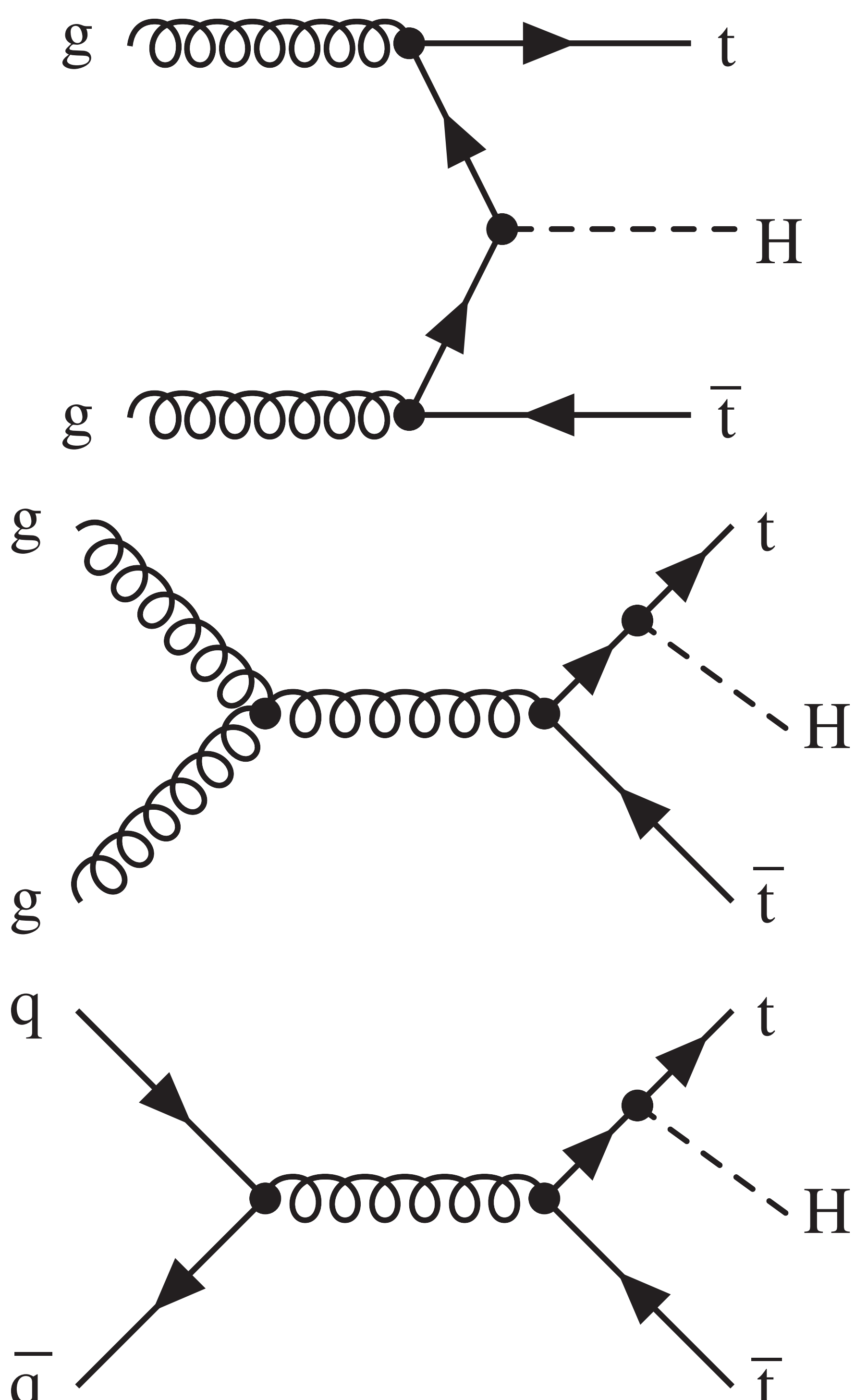






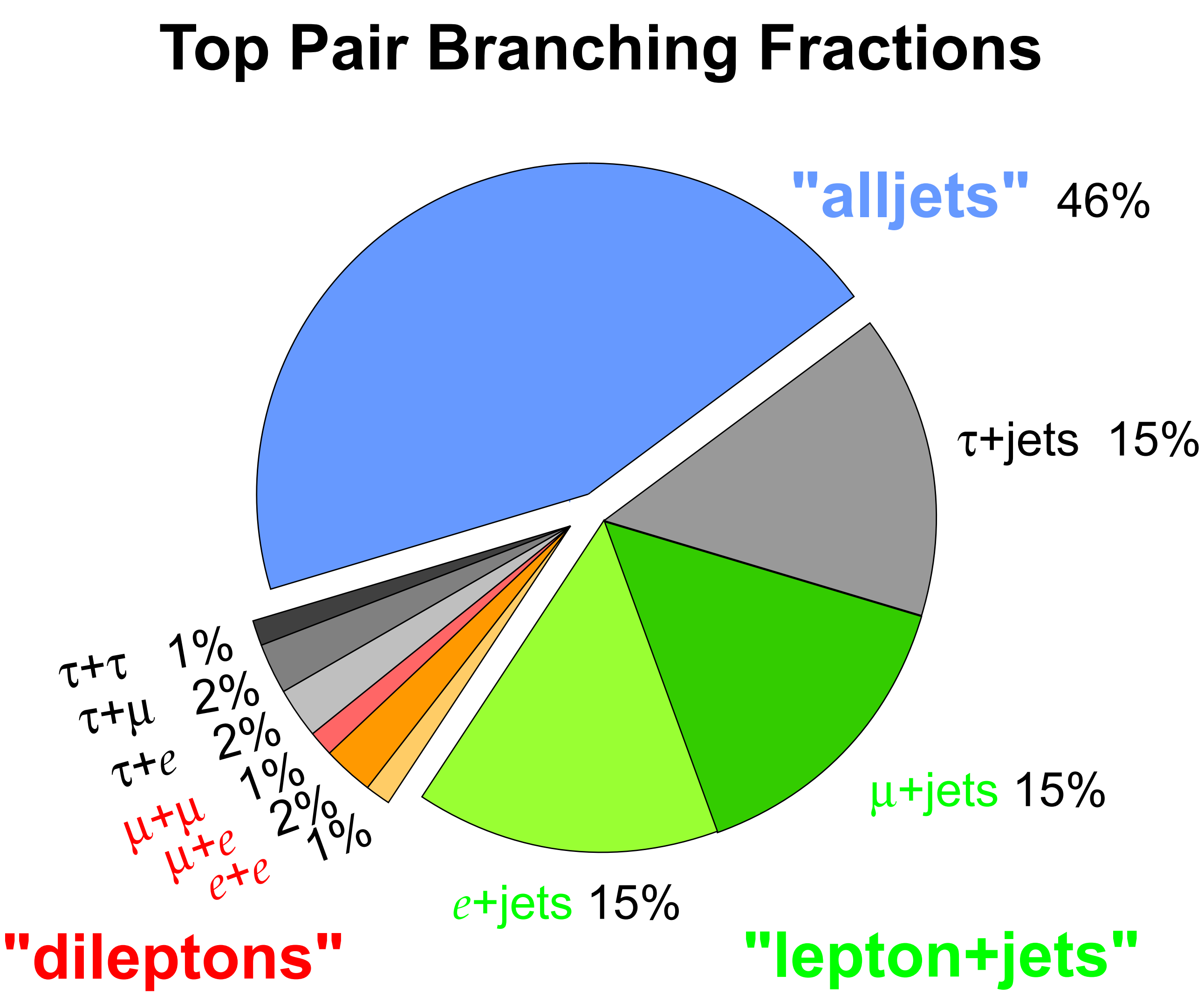


Physics: Associated production of Higgs bosons and top quark pairs ***Breaking news!***

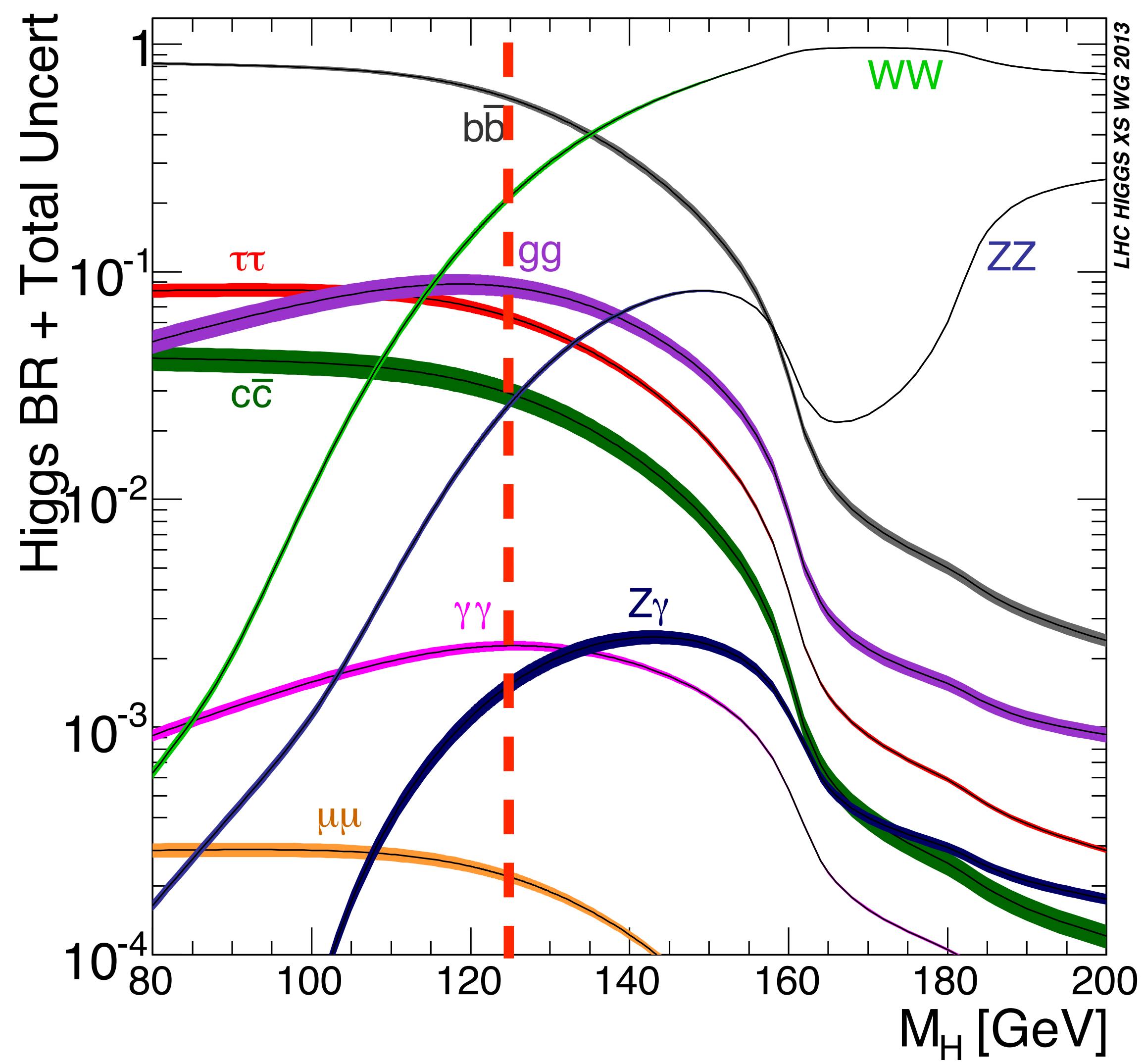


- Hugely important result
  - Extremely important to measure directly the top quark Yukawa coupling (close to 1 in the SM)
  - If it were to deviate from unity, it would be clear evidence of New Physics, so studying this vertex is hugely important
  - First direct observation of how the Higgs couples to a quark
- Difficult measurement
  - Both Higgs Bosons and Top Quarks can decay into a large variety of final states
  - Both have challenging backgrounds
  - And then you are looking both both in the same event!

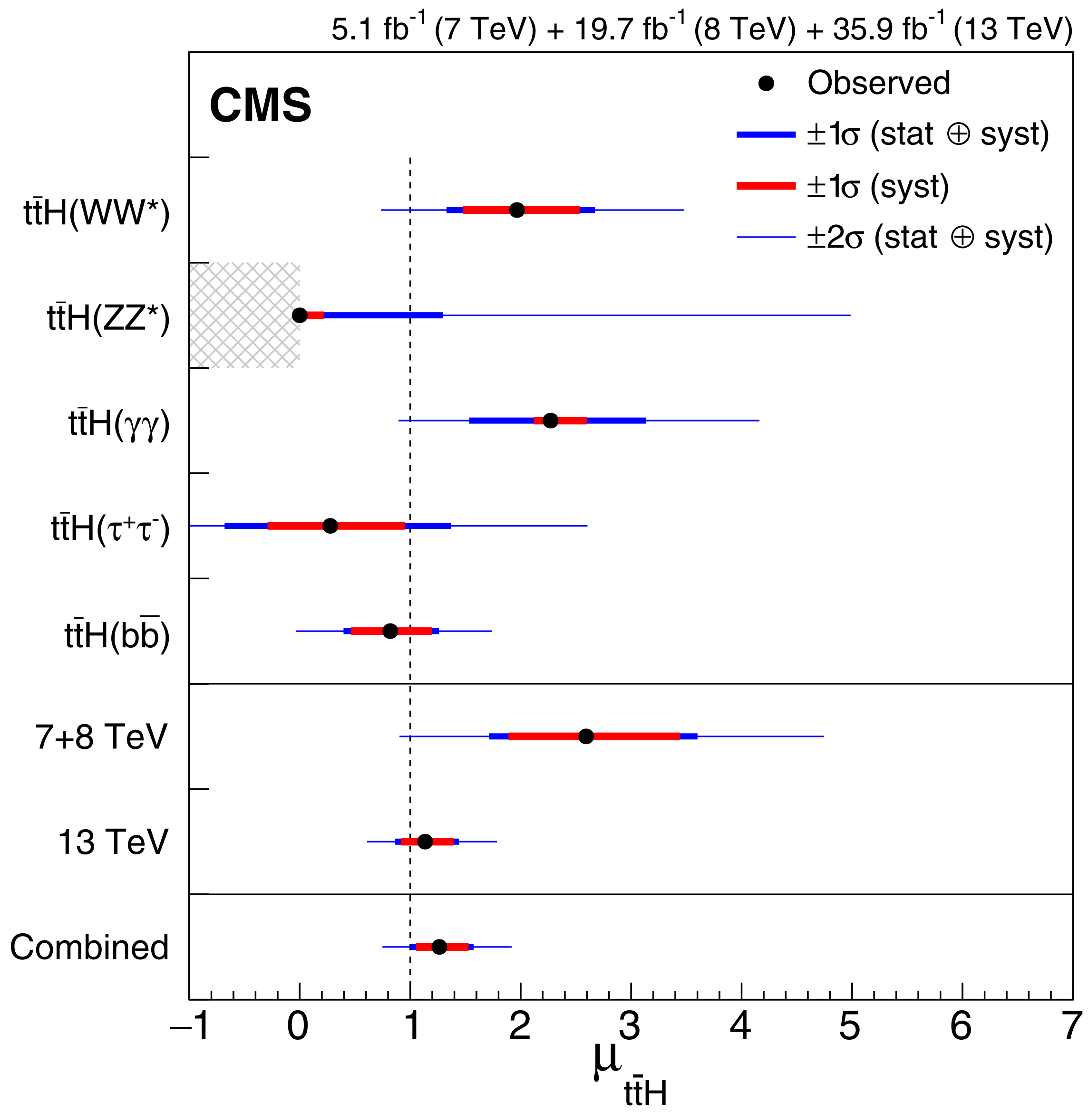
Physics: Associated production of Higgs bosons and top quark pairs *Breaking news!*

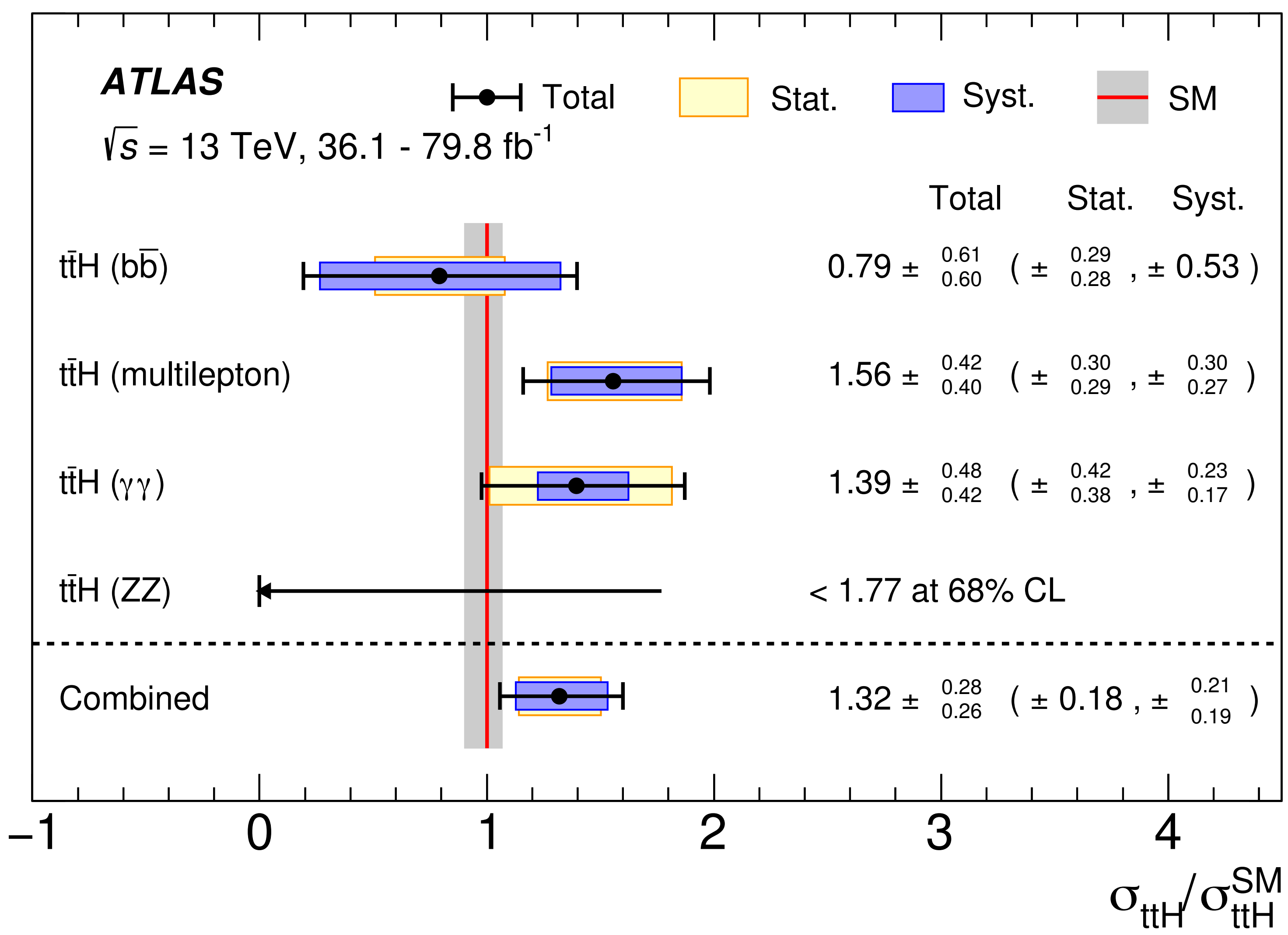


$\otimes$



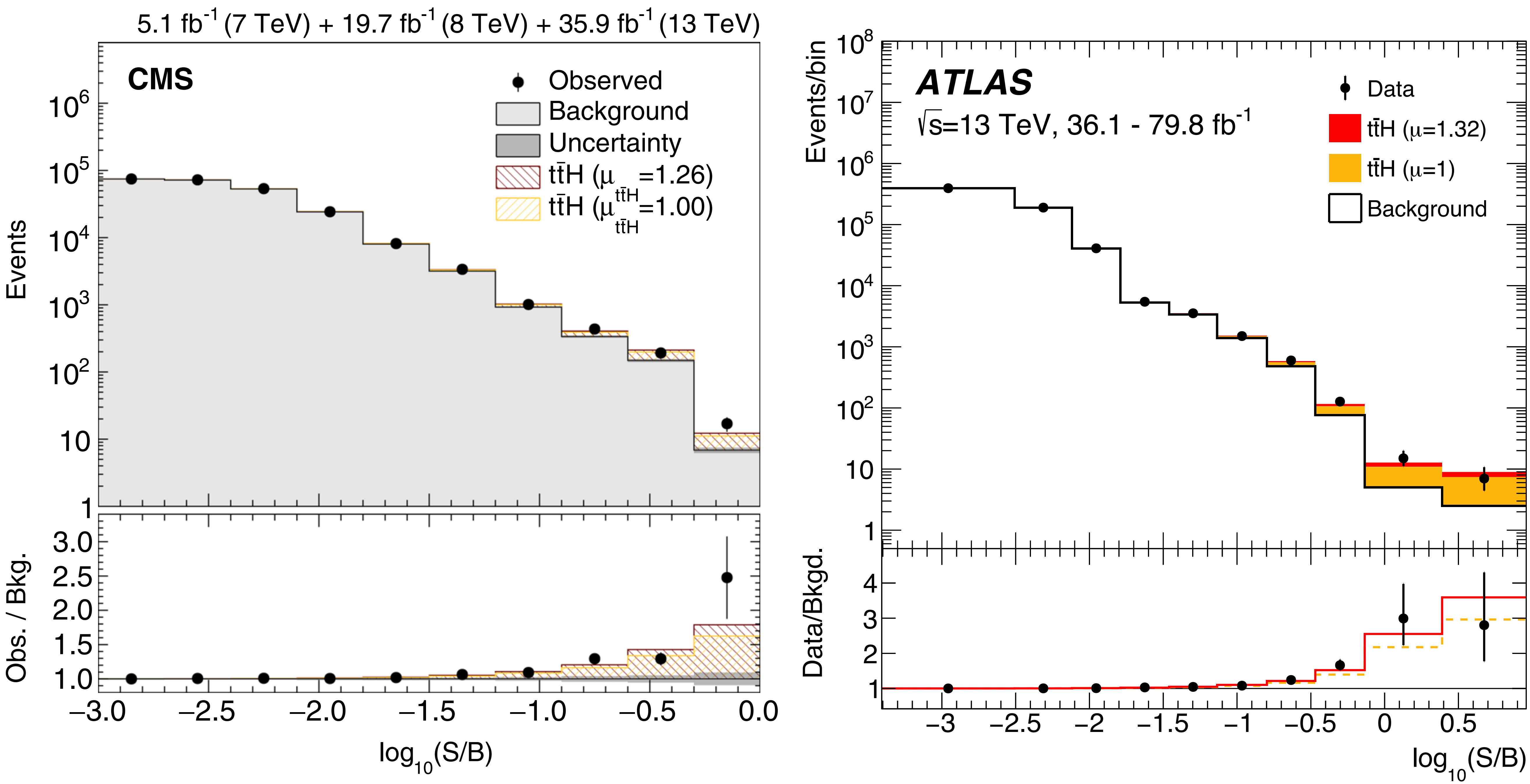
Physics: Associated production of Higgs bosons and top quark pairs *Breaking news!*





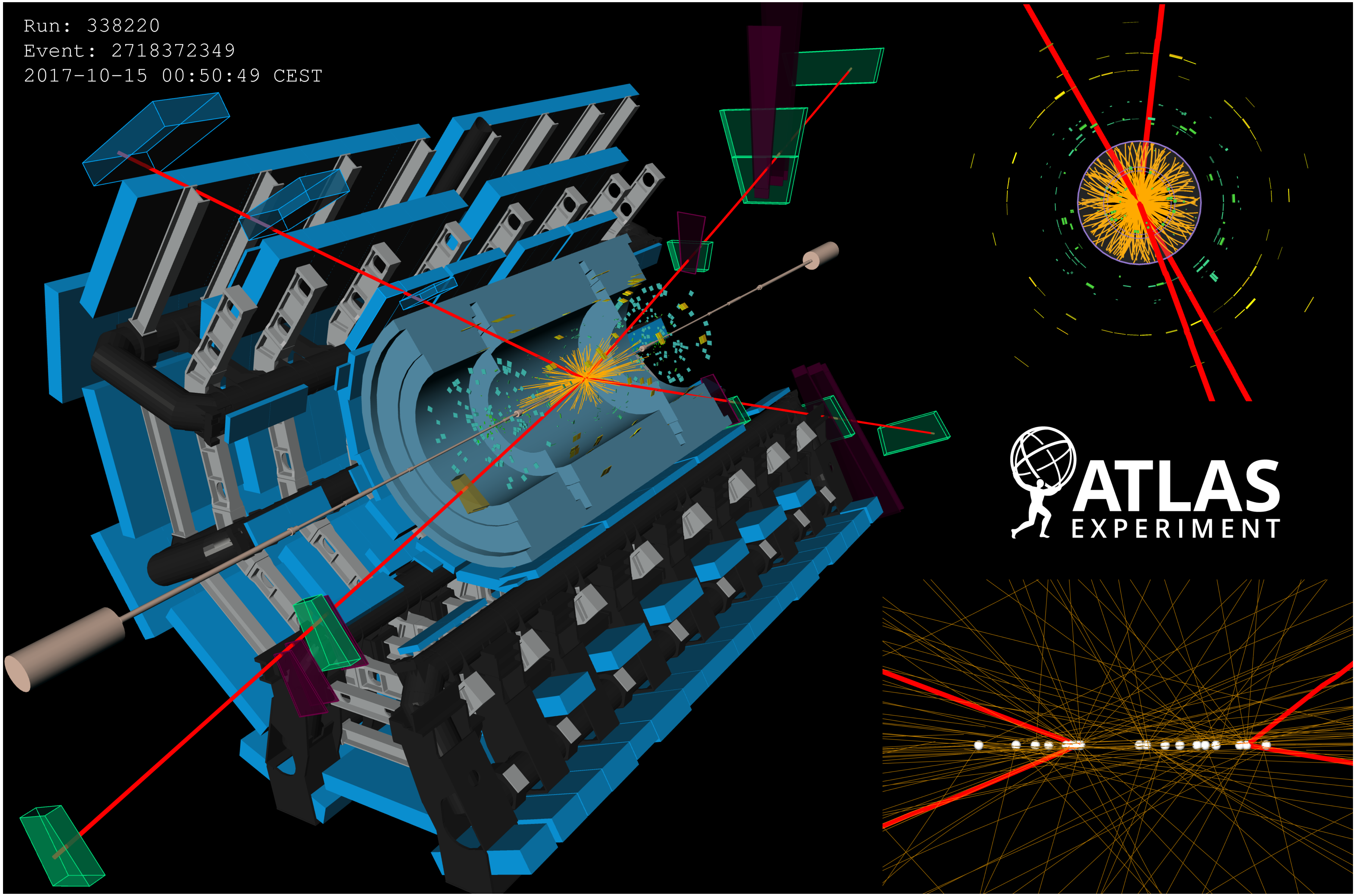
Physics: Associated production of Higgs bosons and top quark pairs

Breaking news!



Observation of events containing two hard-scatter processes at high pileup

- 2 $Z \to \mu\mu$ candidates from different pp interactions, but in the same bunch-crossing, observed in 2017 data
  - their production vertices are separated by 67 mm ATL-PHYS-PUB-2018-007



Observation of the $\chi_b(3P)$ states

For the first time the two states $\chi_{b1}(3P)$ and $\chi_{b2}(3P)$, corresponding to J=1,2, are resolved

The mass difference is measured to be:

$\Delta M = 10.60 \pm 0.64(stat.) \pm 0.17(syst.)$ MeV

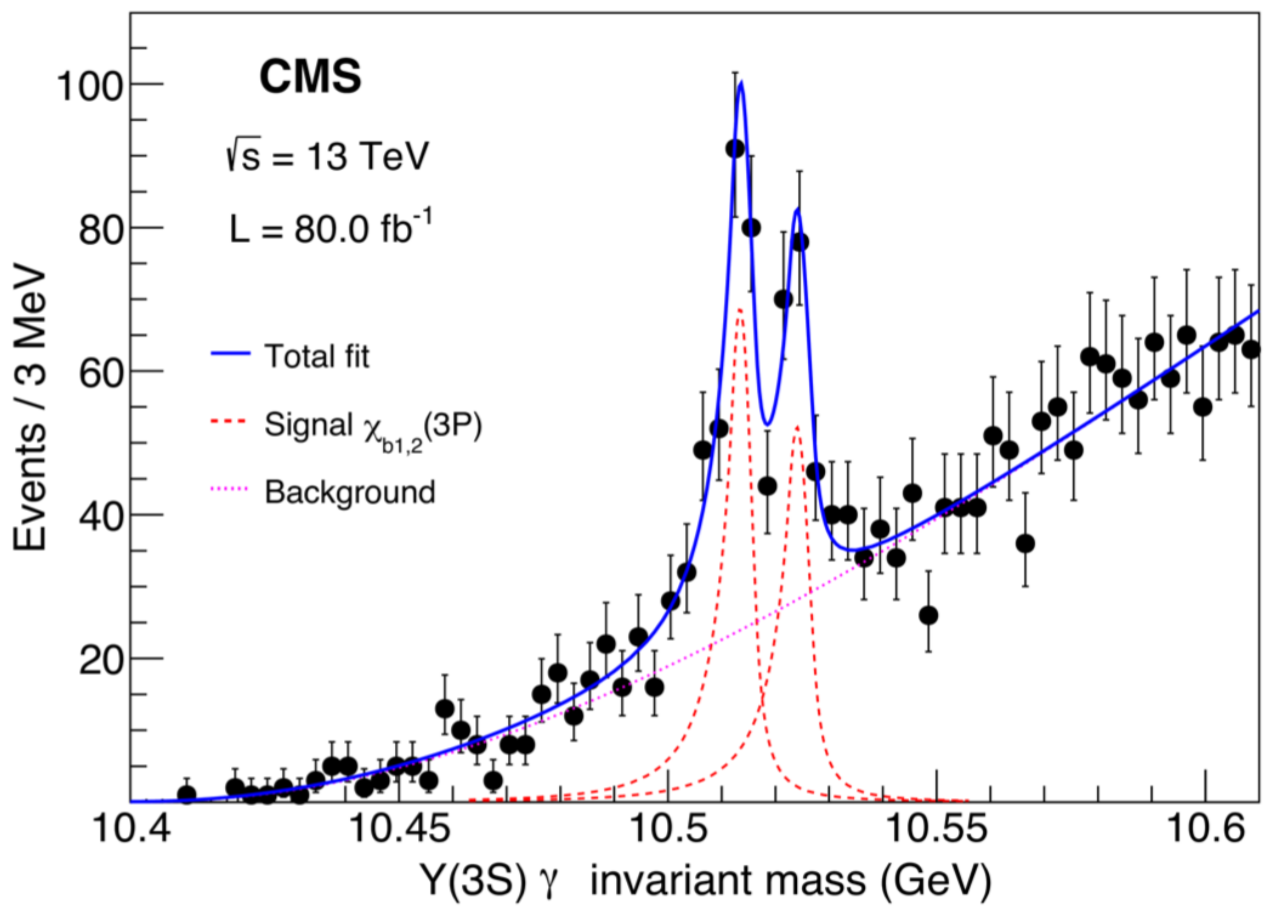
- Most predictions from non-perturbative QCD range from 8 to 18 MeV
- One predicts -2 MeV reflecting the coupling with the open-beauty threshold

The masses of the two states are

$M_1 = 10513.42 \pm 0.41(stat.) \pm 0.18(syst.)$ MeV

$M_2 = 10524.02 \pm 0.57(stat.) \pm 0.18(syst.)$ MeV

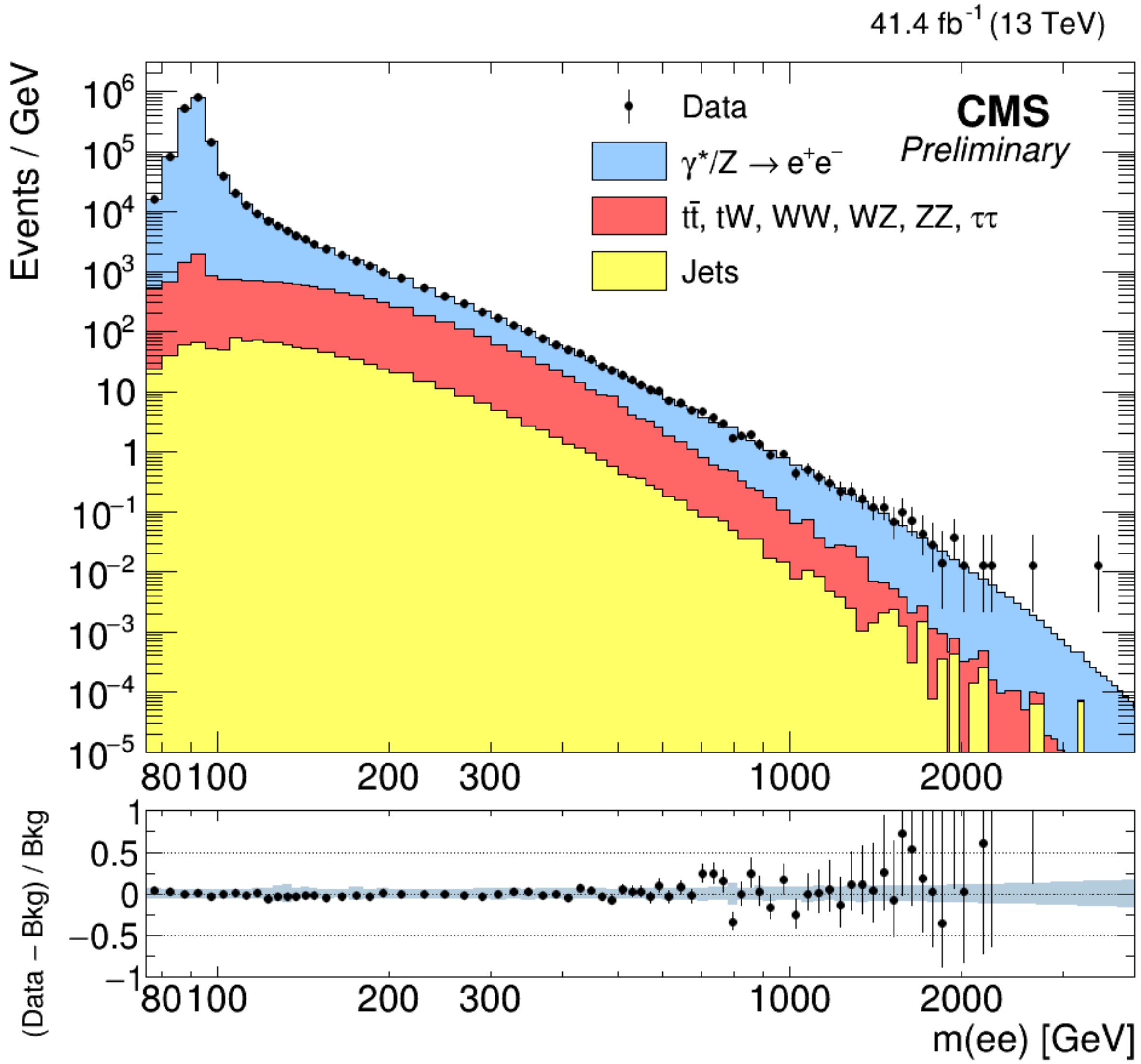
$\chi_b(3P)$ mass resolution 2.2MeV

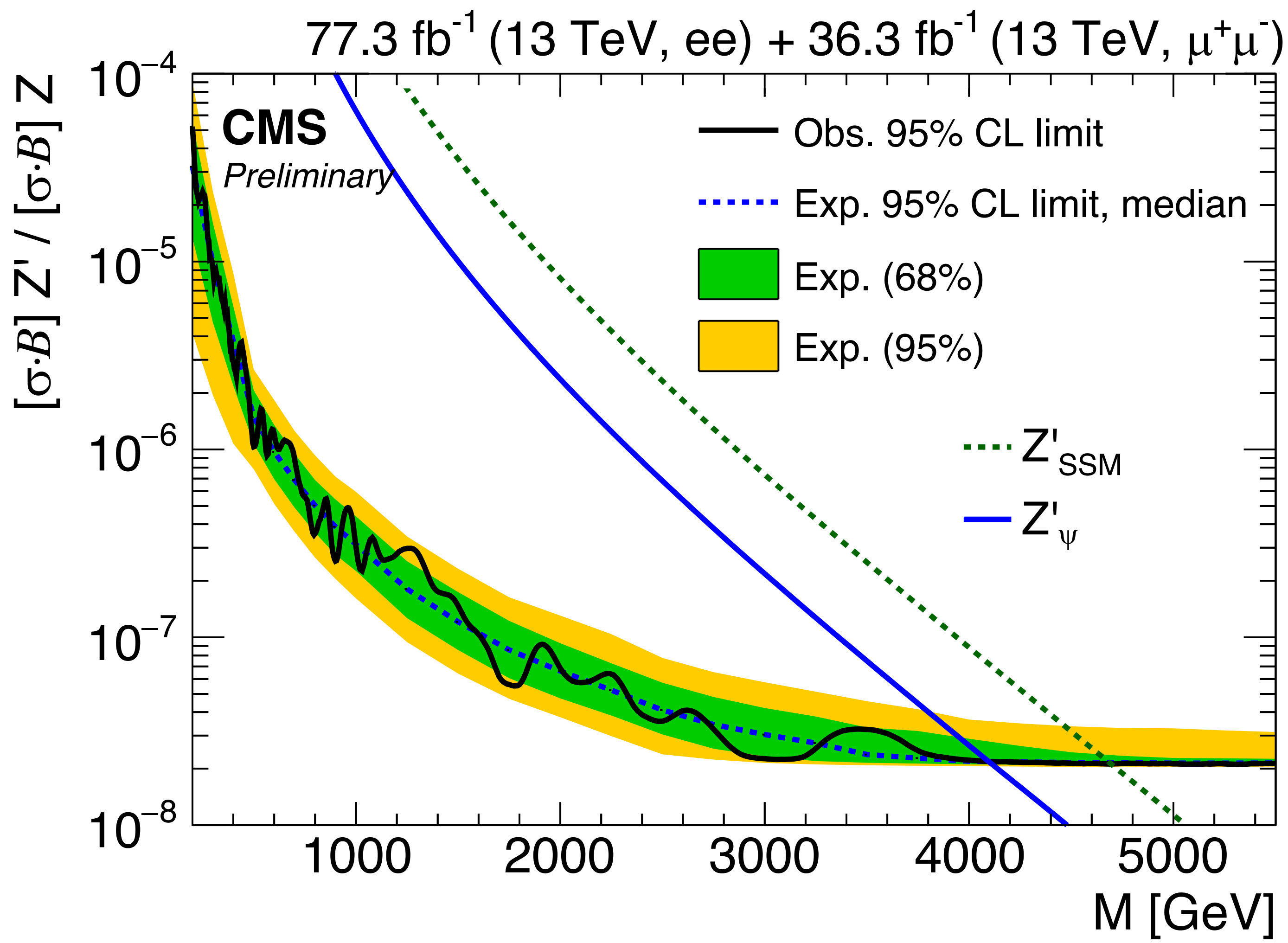


arXiv1805.11192

Searches for high-mass di-lepton resonances

First 2017 analysis





Limits for high mass searches extending beyond 4 TeV

Channel	Model	Obs. limit (TeV)	Exp. limit (TeV)
ee (2017)	Z'_{SSM}	4.10	4.15
	Z'_{ψ}	3.35	3.55
ee (2016 and 2017) + $\mu\mu$ (2016)	Z'_{SSM}	4.7	4.7
	Z'_{ψ}	4.1	4.1

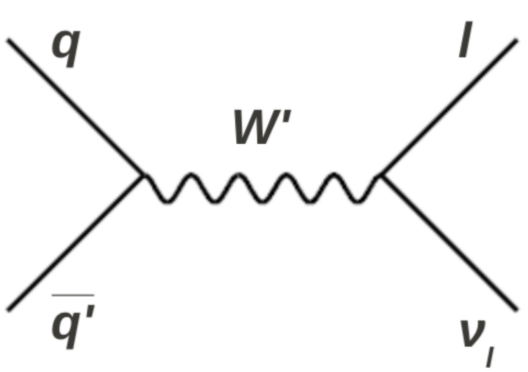
CMS-PAS-EXO-18-006

Searches for W' with 80 fb^{-1}

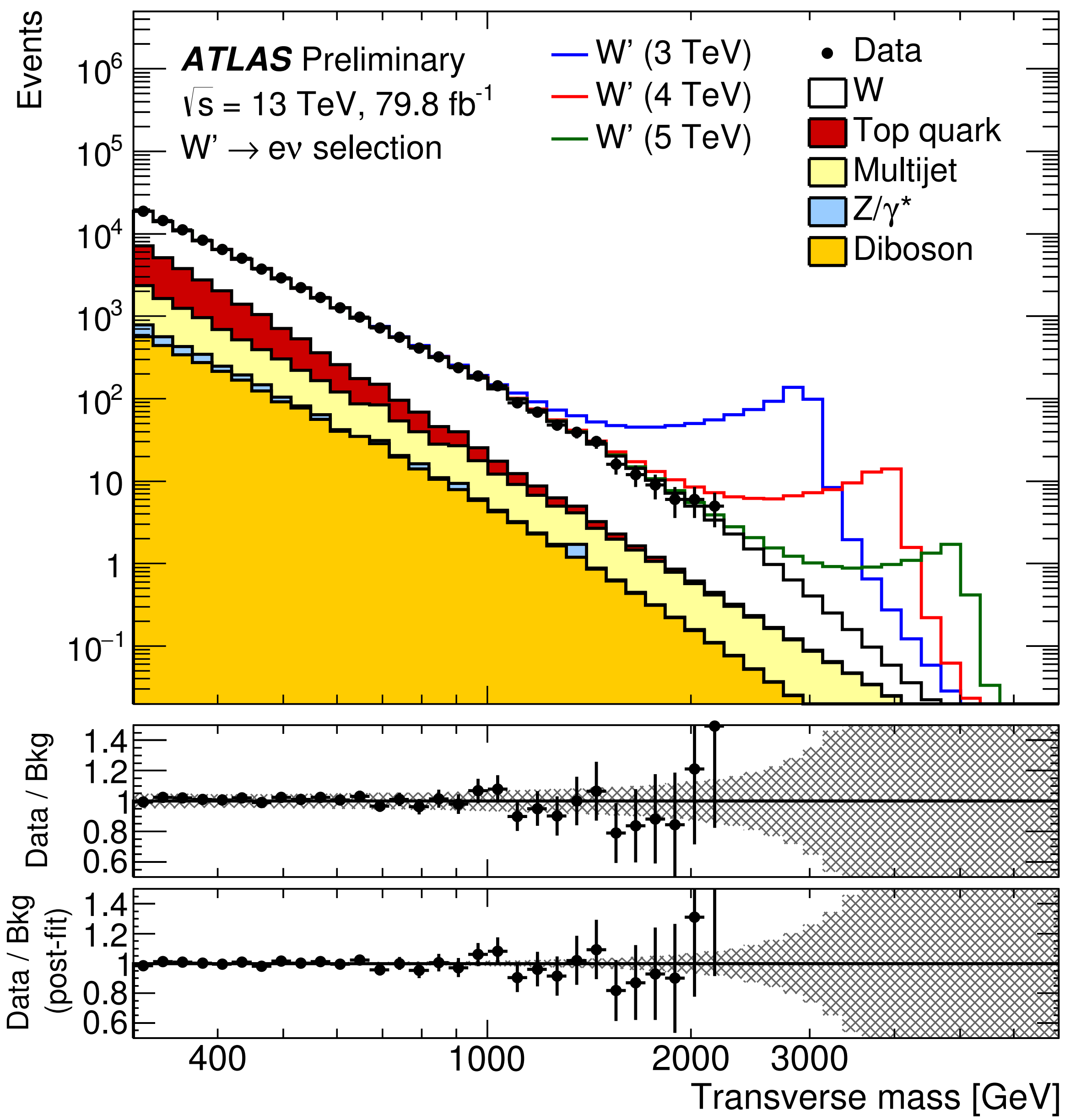
80 fb^{-1}

- search for $W' \to \ell\nu$ updated with **80 fb^{-1}**

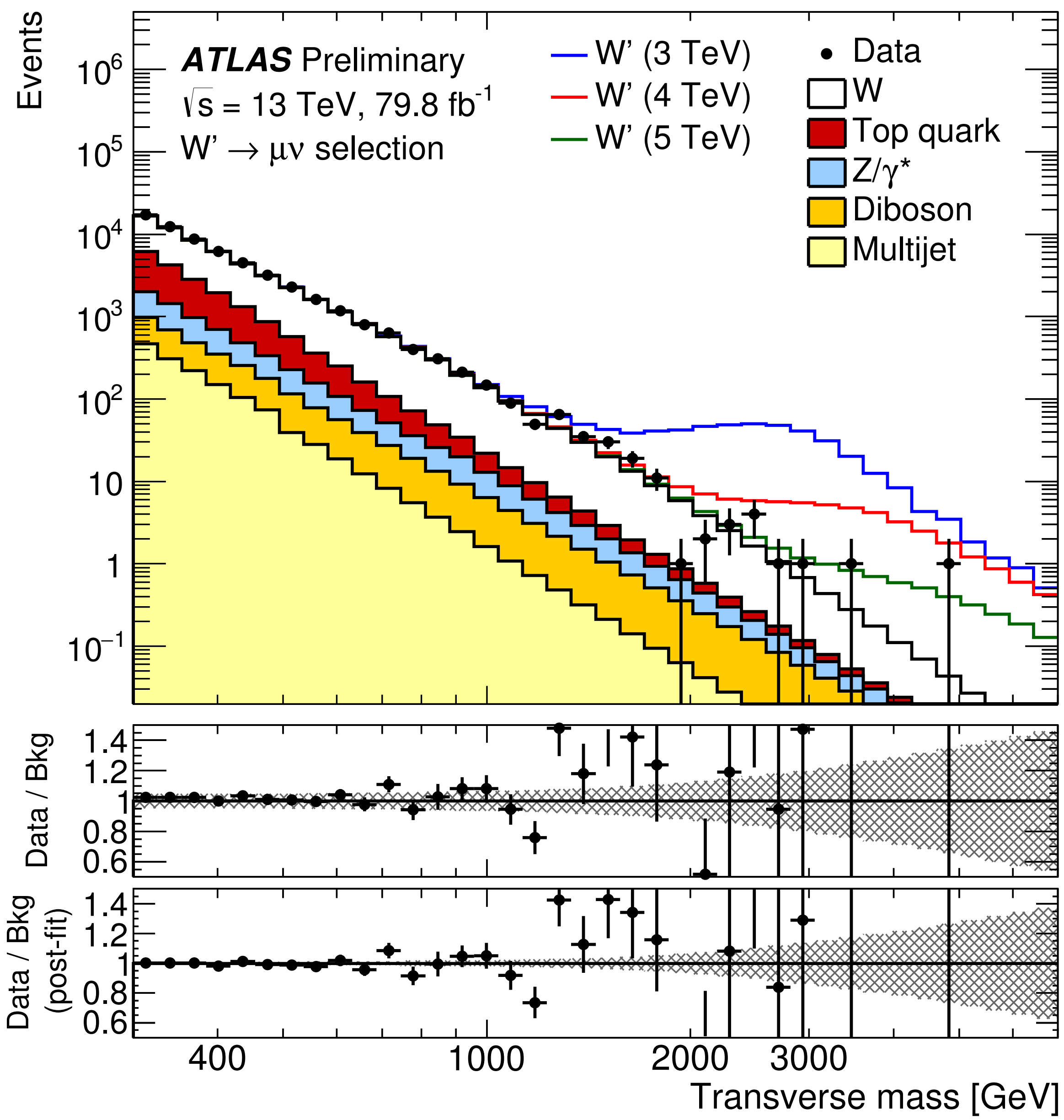
  $\Longrightarrow$ **W'_{SSM} masses below 5.6 TeV are excluded**



e channel



μ channel

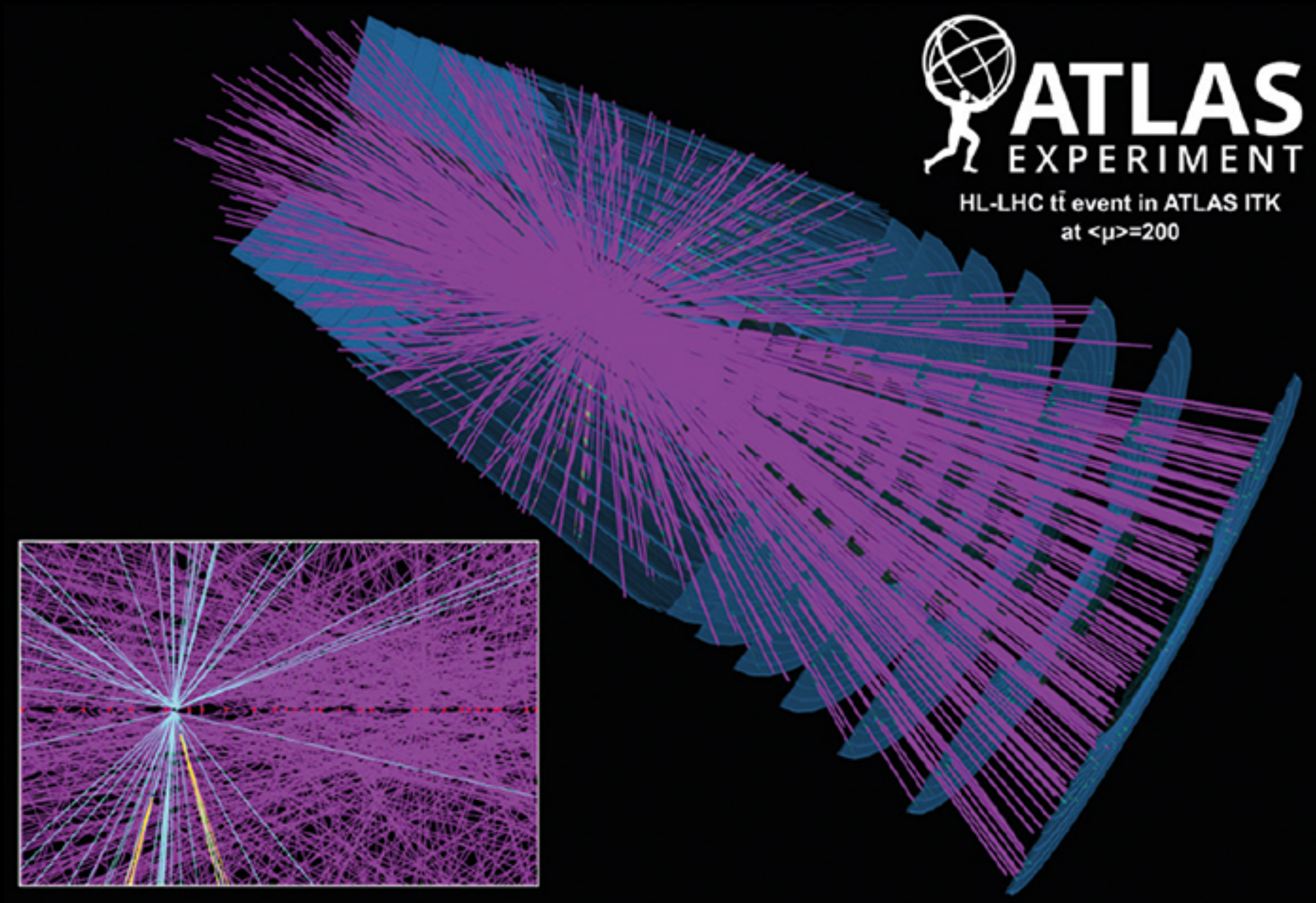


And there's so much more…

- LHCC reports (last week)
  - https://indico.cern.ch/event/726320/
- LHCP conference (ongoing)
  - http://lhcp2018.bo.infn.it
- Tantalising signs of flavour anomalies
- Precise measurements
  - Multi-boson production
  - Higgs boson production
  - CP violation in new processes
  - New B-hadrons and quarkonia species
- Searches - sadly no discoveries but lots of limits…

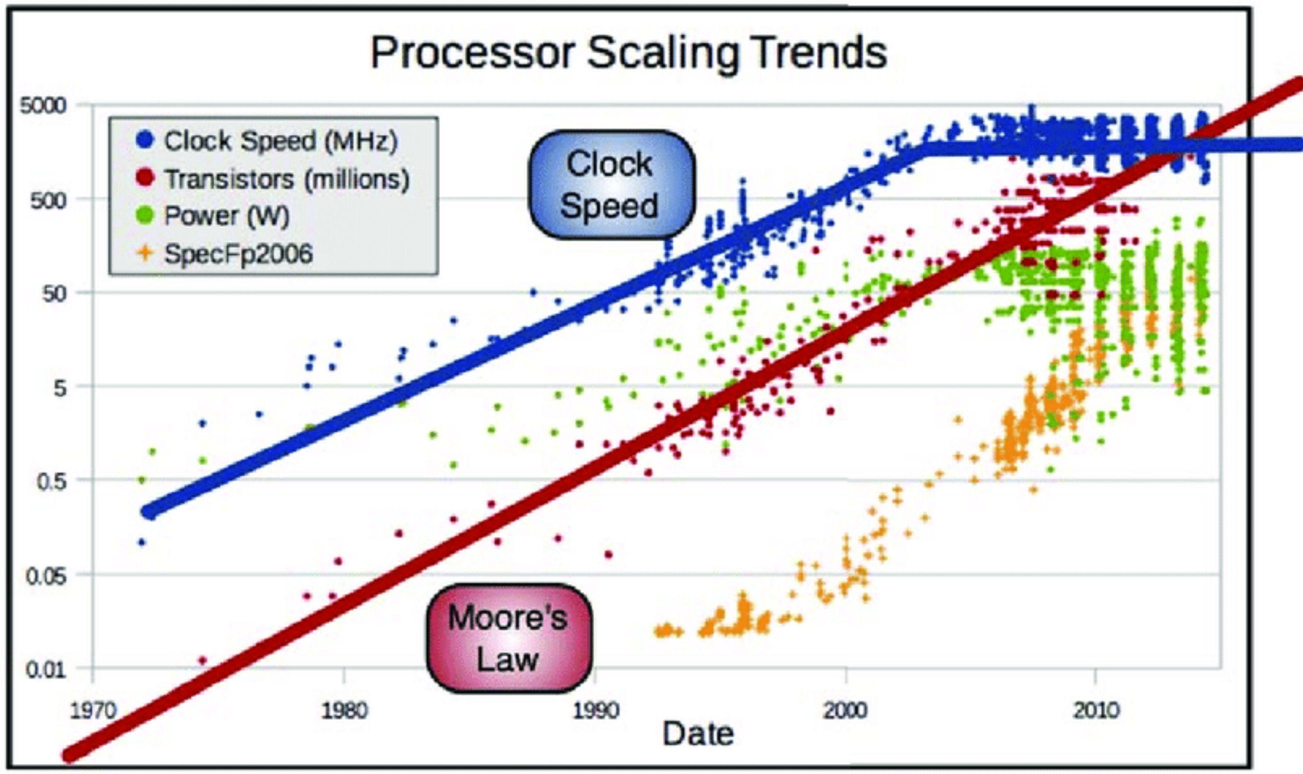
Challenges and opportunities for Software and Computing in LHC Runs 3 and 4

Slides largely inspired by Graeme Stewart's talk @ Spåtind 2018 conference

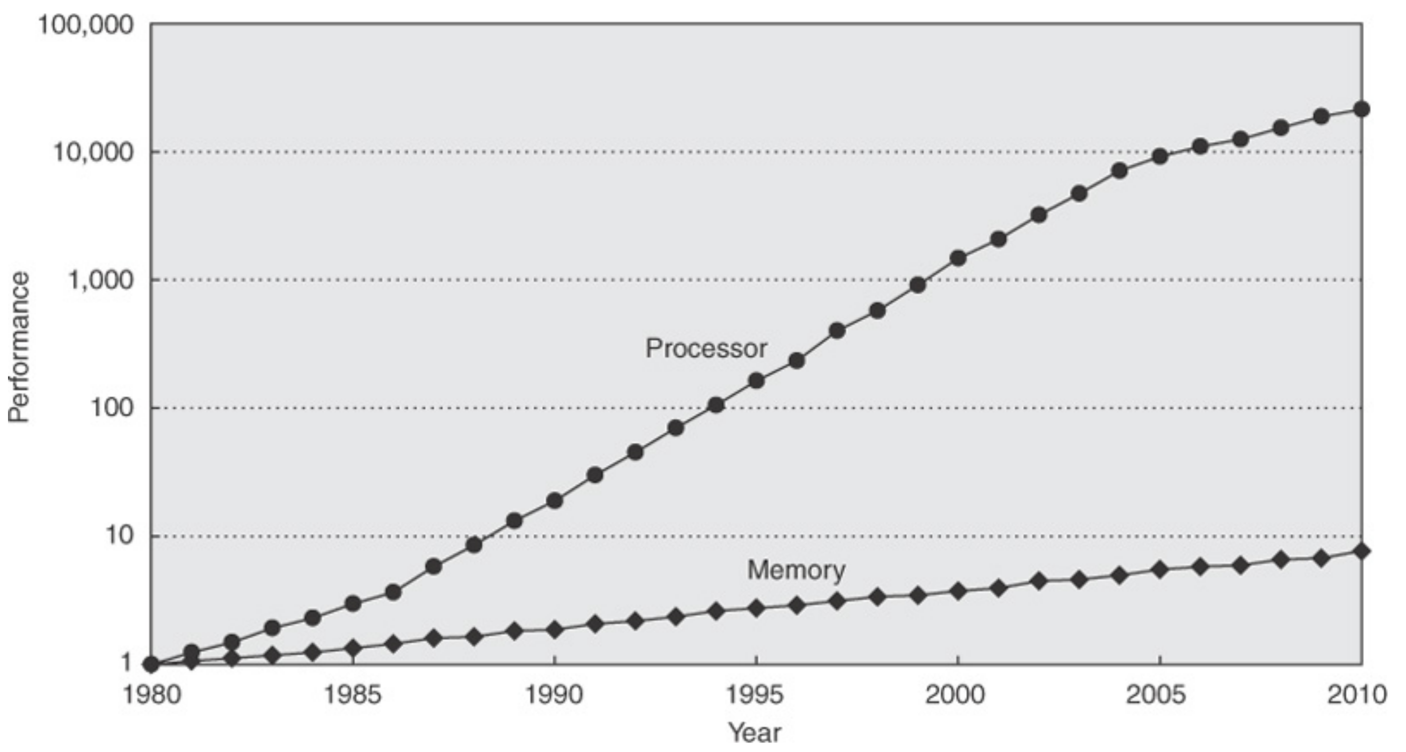


Technology evolution

- CPUs are not getting *faster*, but they are getting *wider*
- To continue to ride the wave of Moore's Law we must make use of *multi-threading* and *vectorisation*
- Co-processors such as GPUs can deliver improved throughput but code must be re-cast to make optimal use of each architecture
- Other issues:
  - Deeper hierarchy of CPU caches means that cache misses are very costly as data is hauled up through the layers
  - I/O performance not keeping up with storage capacities
- Network capacities continue to grow impressively



C. Leggett, LBNL



J. Hruska

What is driving technology evolution?

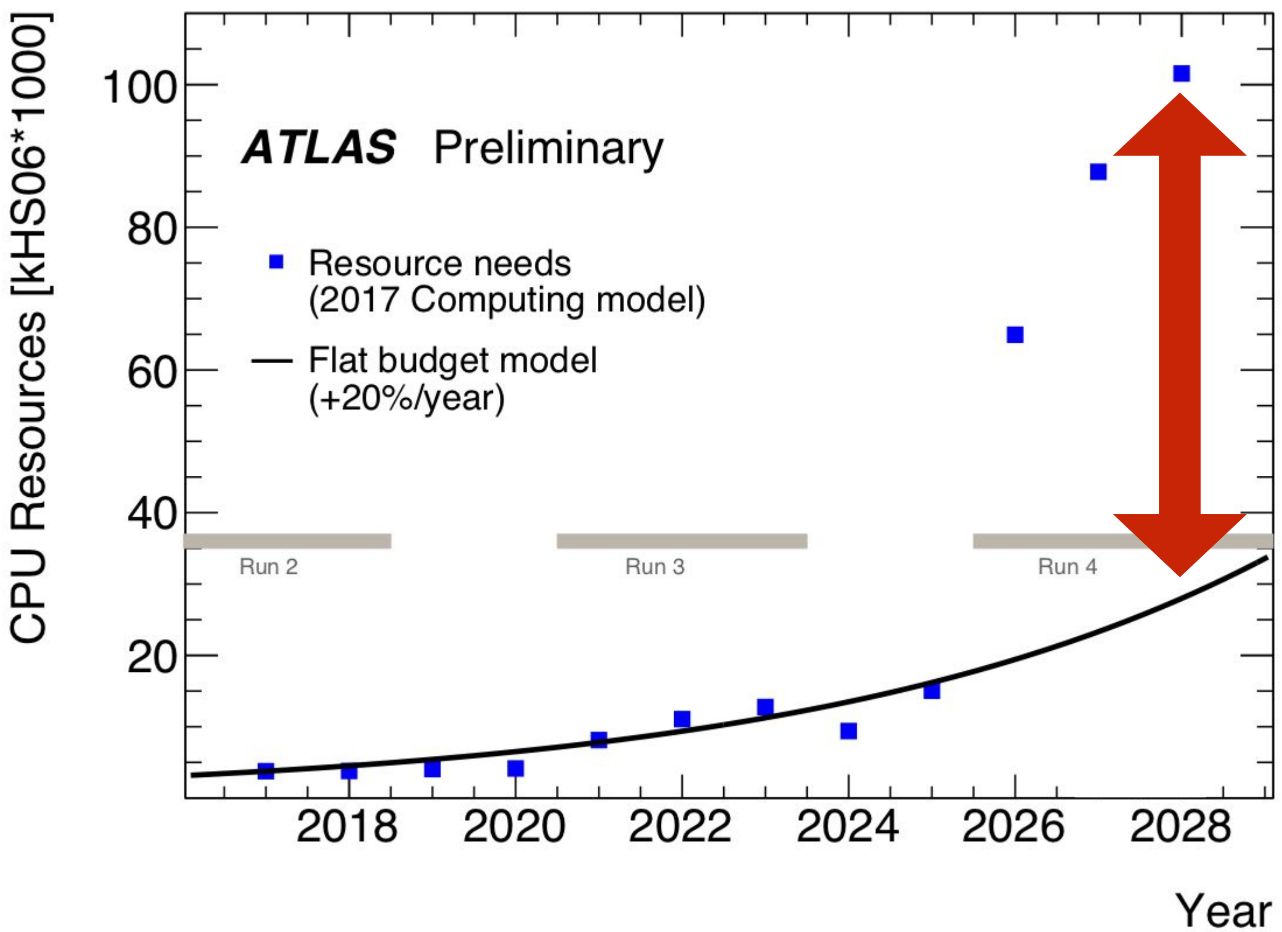
- Not experimental particle physics, for sure!
- Power consumption
  - Mobile devices
  - Internet of things
- Machine learning (= lots of matrix manipulations)
  - Specialist architectures (neural/tensor processing units) now appearing to support the efficient running of these applications
- Computing is moving to *clouds* as organisations try to reduce their in-house resources and large technology companies offer more and more internet-based consumer services

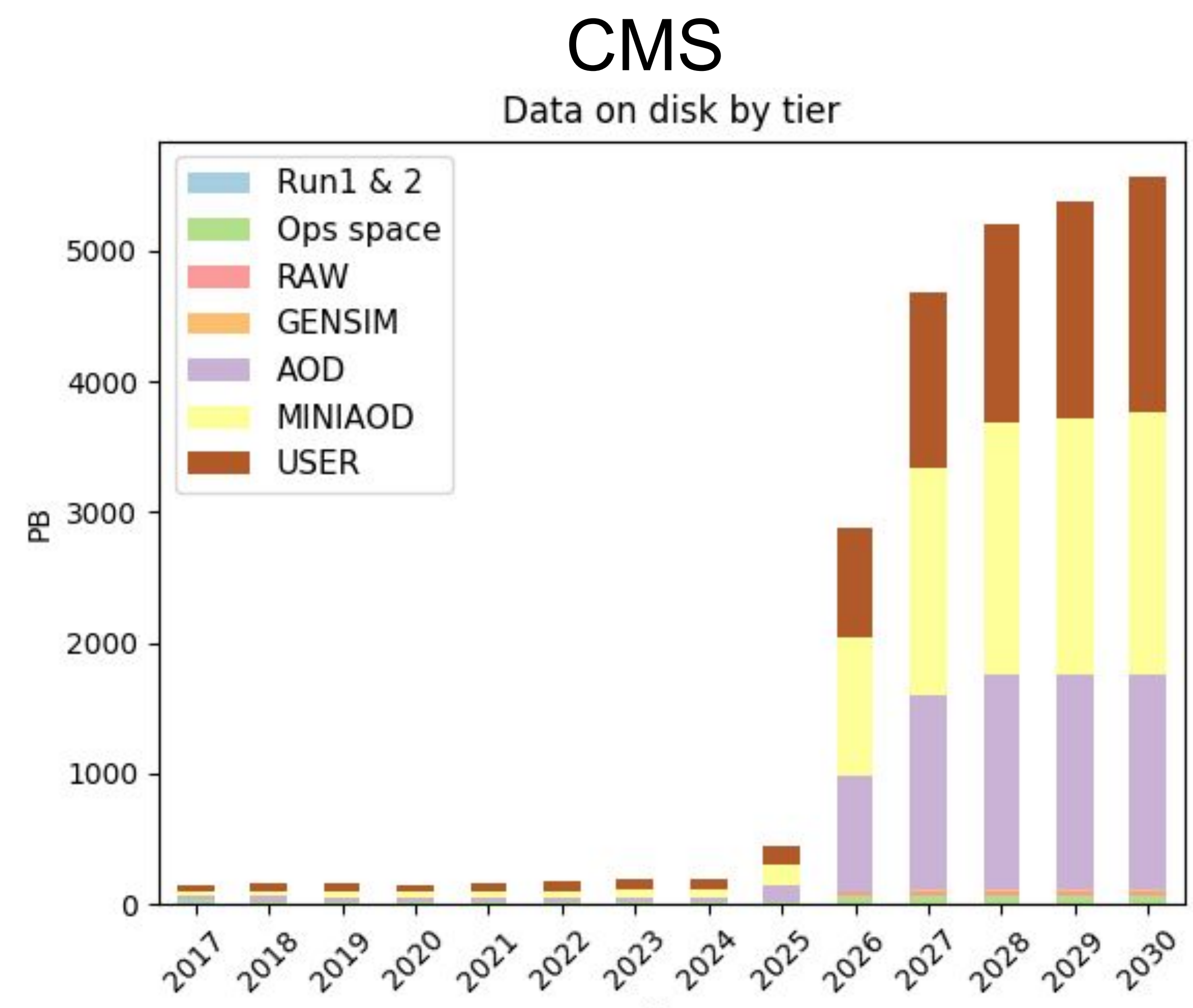
LHC Runs 3 and 4

	Run 2	Run 3	Run 4-5 (HL-LHC)
Year	2015-2018	2021-2023	2026++
CoM energy (TeV)	13	14	14
Lumi/Nominal	1.0 (2015-2016) 2.5 (2017-2018)	2.5	5-7
∫L at end of run (fb^{-1})	150	300	3000

Increasing data volumes, rate and event complexity

Increasing complexity and rate





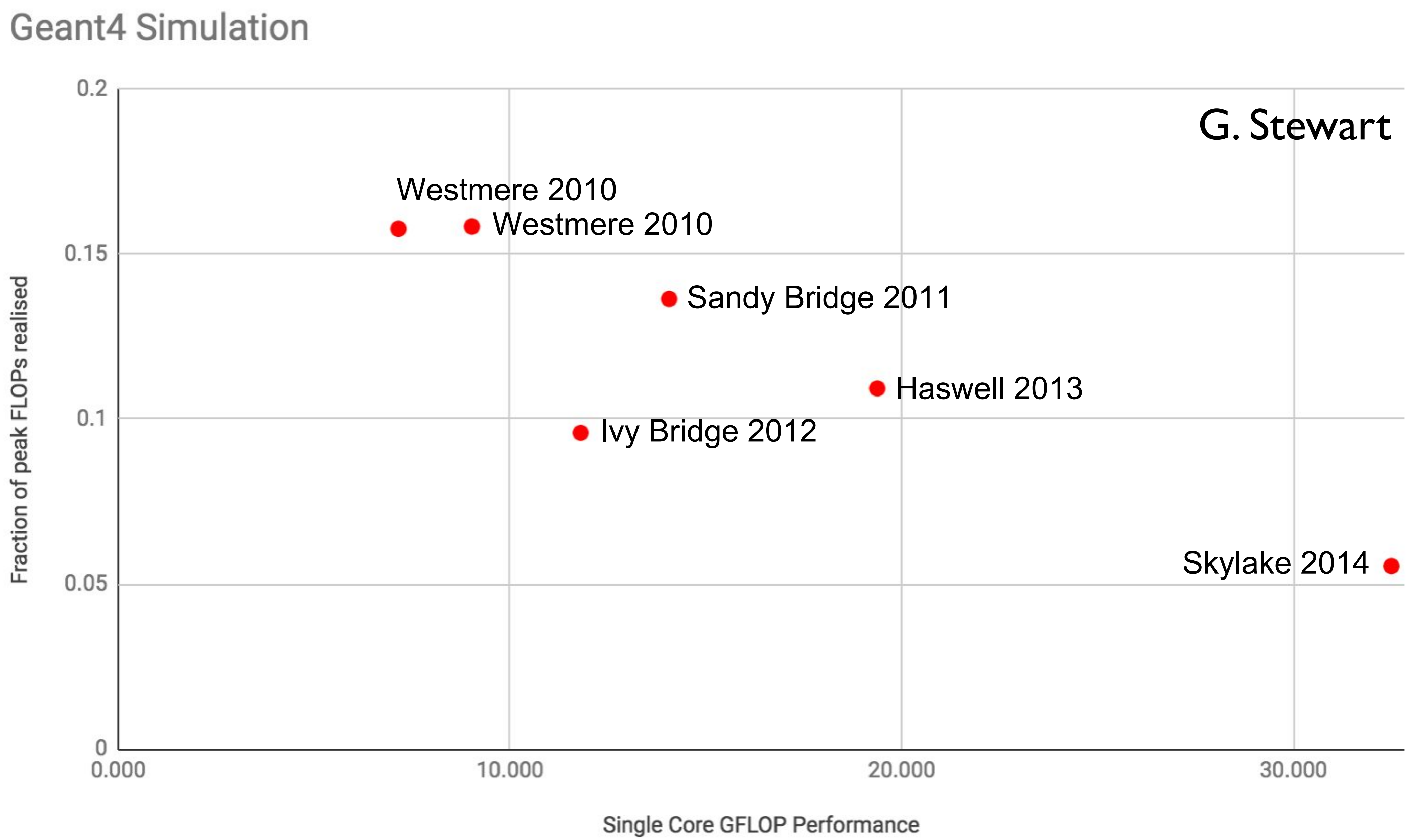
Rest of run 2 and run 3: smart & efficient use of existing model will see us through

Run 4: <μ> ~ 200 and much higher rate - need to do things very differently → we must be able to make full use of evolving technology to have a hope of keeping up with the HL-LHC

This means that our software has to change radically

How are we doing?

Could do better…



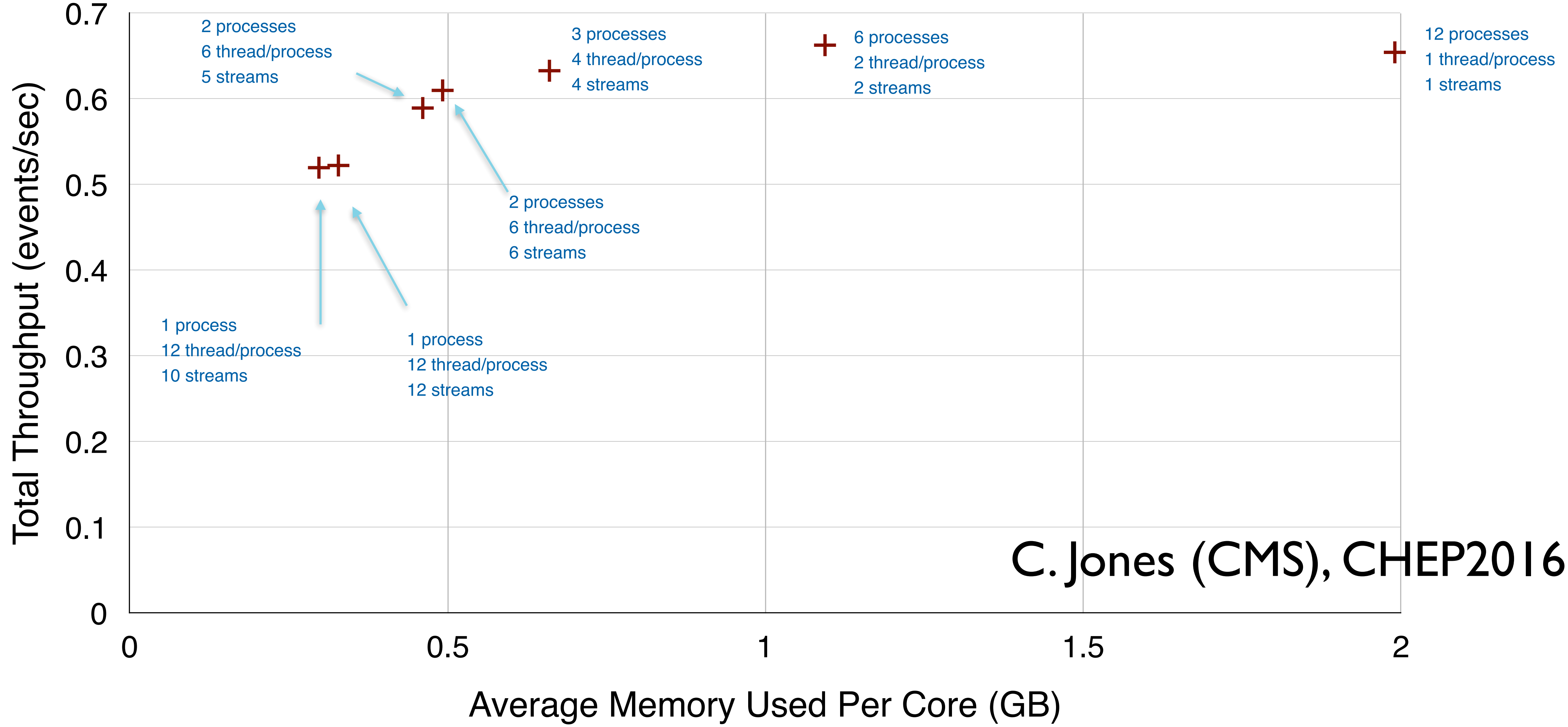
Wider vector registers that we aren't fully using…

Expensive cache misses due to deeper hierarchies…

Making use of frameworks

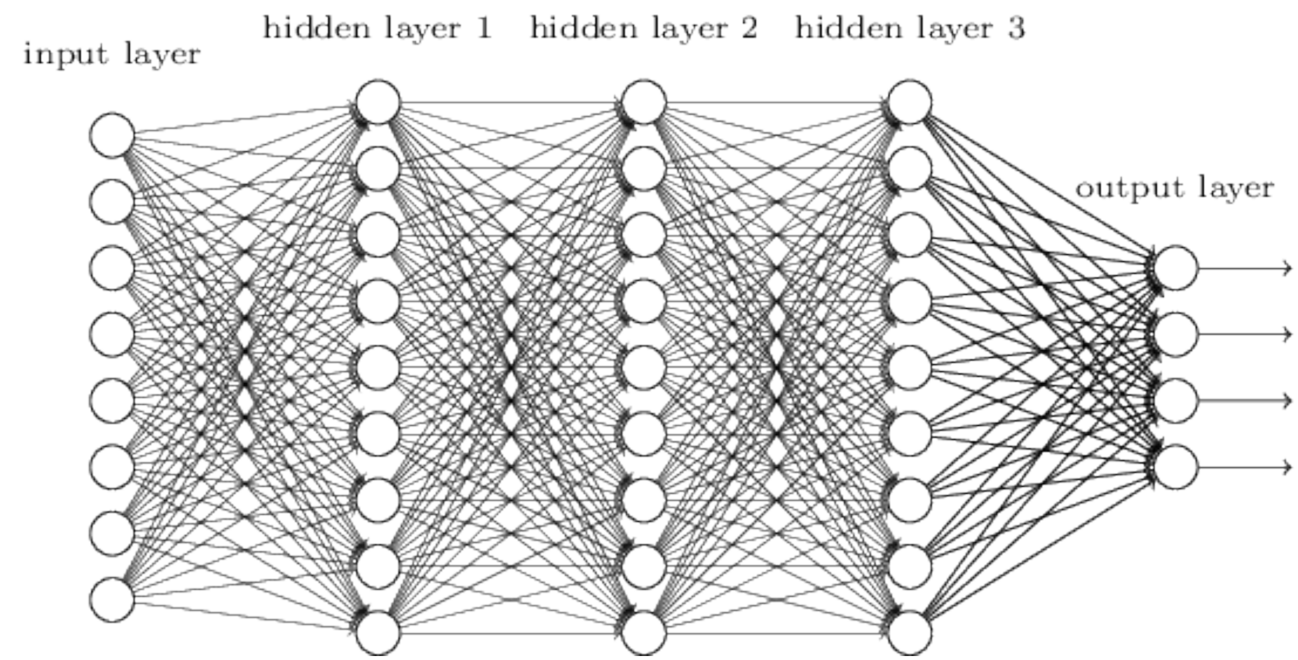
- We can't re-write everything
- Making optimal use of concurrency and vectorisation can be frustrating, even for experts and even with abstraction libraries
- The software *frameworks* used by the experiments are our friends in this regard
  - These provide the basis of the algorithmic processing, persistent/transient layers, services, tools etc
  - Incorporate as much of the concurrency and vectorisation into the framework and shield those writing algorithmic code from the ugly details
  - But we still need people to implement the core software - and the migration is still not zero-work for the clients
- This approach has been used in the experiments: CMS is already multi-threaded; ATLAS has a multi-process framework in use (AthenaMP) and plans to be multi-threaded in time for Run-3 (AthenaMT)
- The same approach can be used for dealing with increasing inhomogeneity of resources, e.g. from volunteer computing to grid sites to HPCs to commercial clouds
- Some elements of the frameworks are shared between experiments: further sharing could reduce the workload (see later)

Total Throughput vs Memory for Fully Loaded Machine

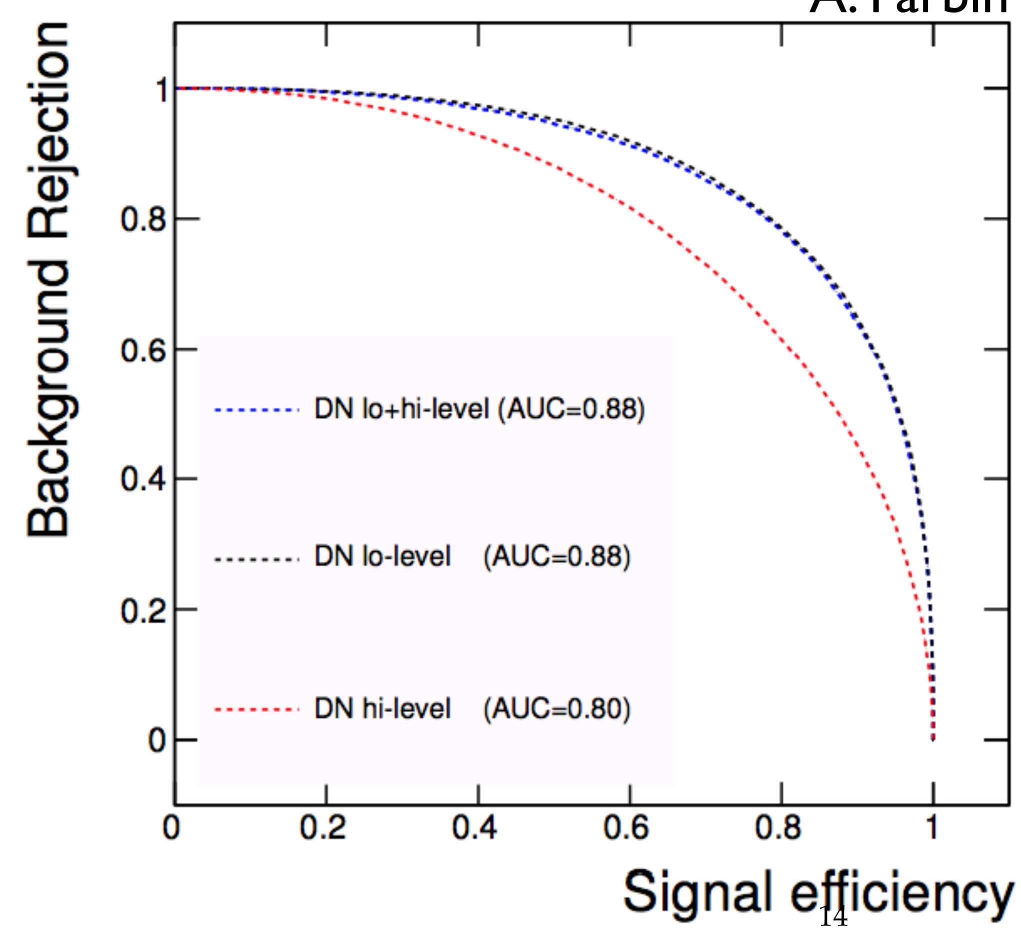


Multithreaded framework allows use of low per-core memory whilst maintaining throughput

- Long history in HEP: BDTs and shallow neural networks used since the 1980s
    - Instrumental in discovery of single top quark production and Higgs measurements
    - We called it "multivariate analysis"
- More powerful computers and bigger training datasets have led to the growth of "deep learning" and an accompanying cluster of high-performance open source tools from outside HEP
- Possibly this has revolutionary potential at all levels of the field
    - Simulation, reconstruction, physics analysis, automation of shift work, optimisation of computing resources, anomaly detection
- Need good links with academic and commercial experts
- Challenges our current data processing model
- Fortunately, there is a very strong and wide interest in ML across the field



A. Farbin



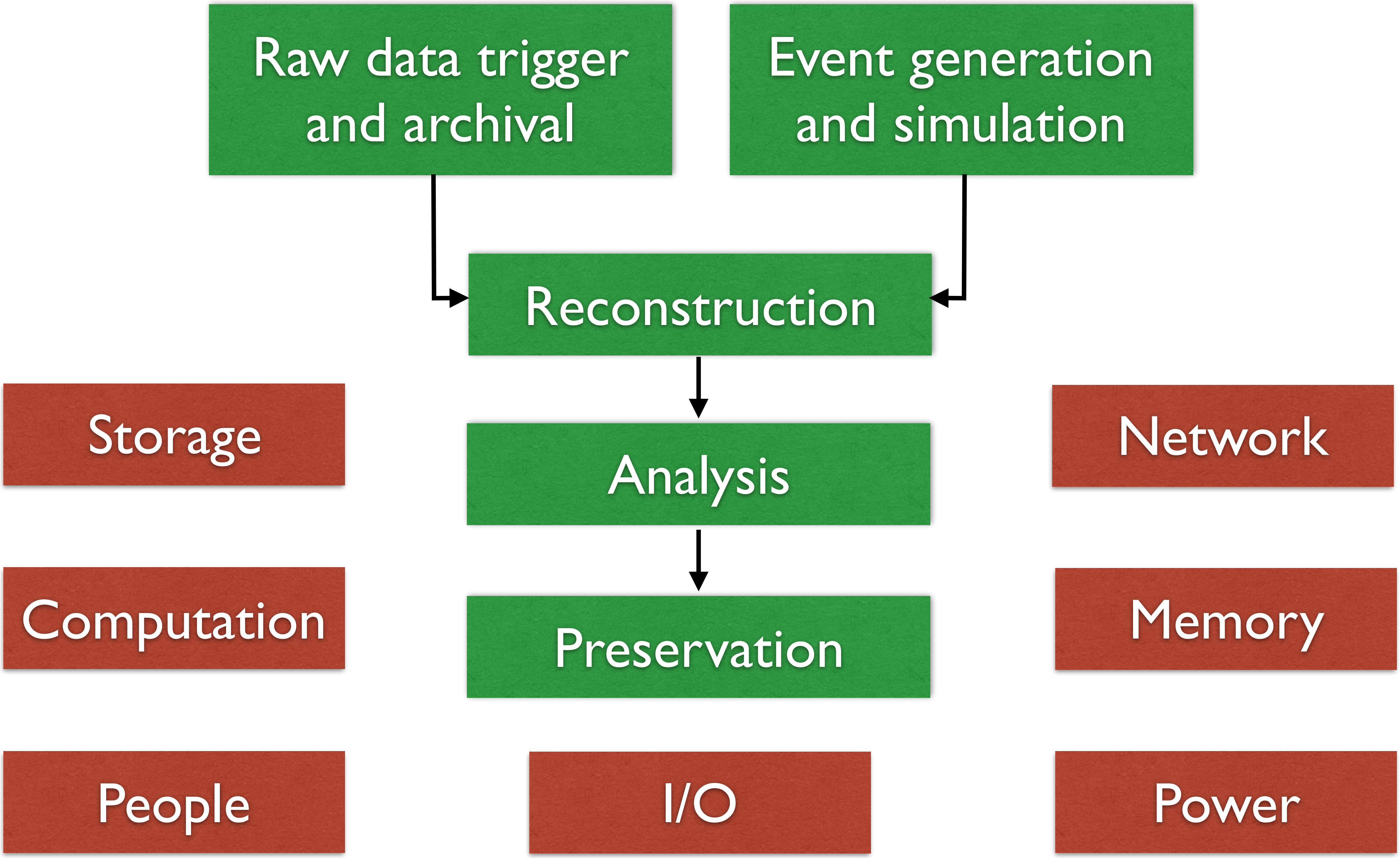
Facilities and data management

- Storage and computing are overwhelmingly from WLCG resources
  - This is expected to continue into Run 4, but with other resources in the mix
- Volunteer computing, commercial clouds, HPCs
  - Need to ensure our software can work seamlessly in these environments
- How to take fullest advantage of major improvements in network capacity?
- Data lakes?
- Strengthen links with other big data sciences especially with regards to sharing network resources
- Storage is a major challenge for HL-LHC
  - Sheer volume
  - How to support fast access for analysis and machine learning?
  - High granularity access for using opportunistic resources…
  - Technologies: more SSD, less disk? Relative costs of tape?

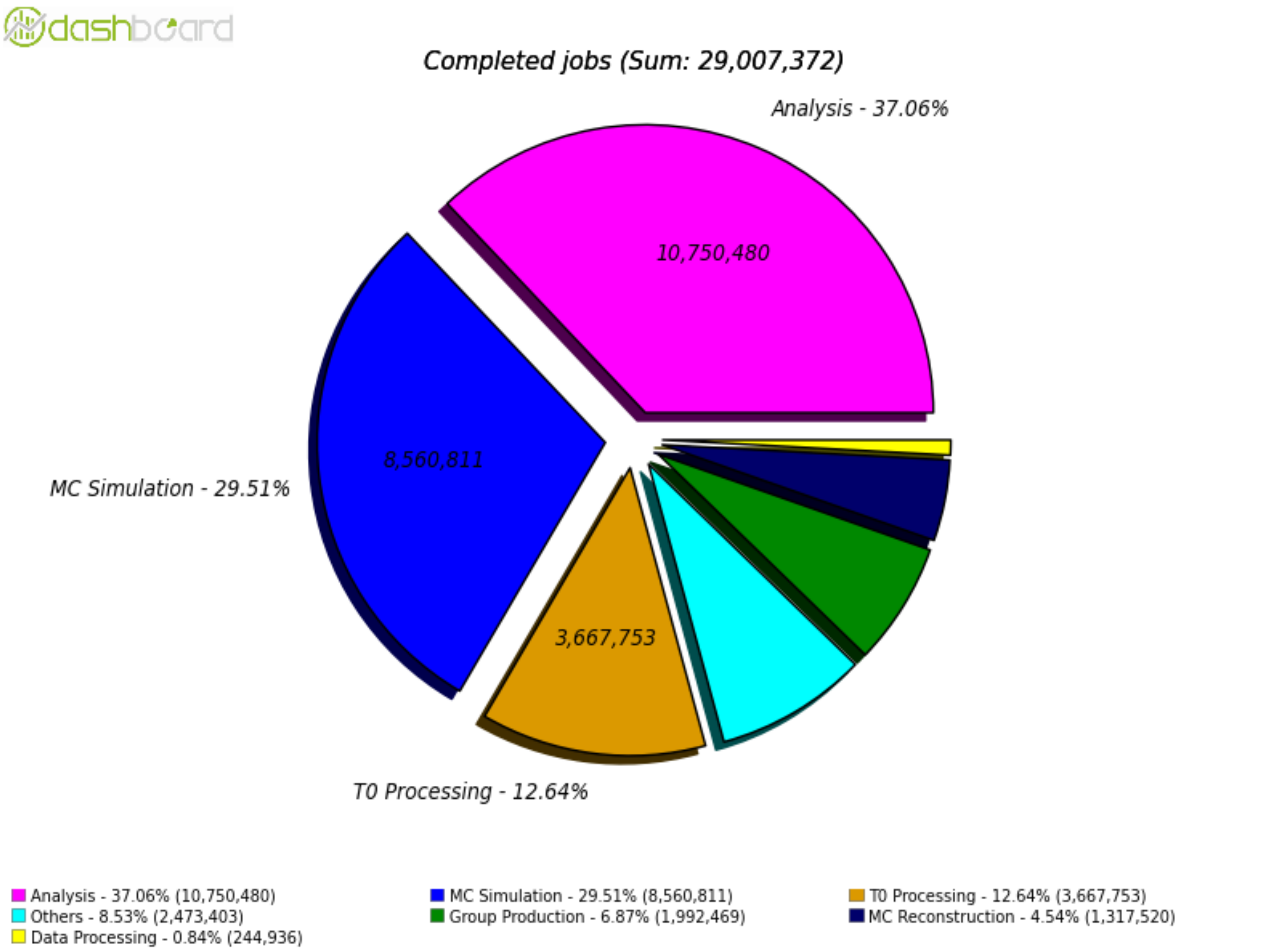
Software development, management and packaging

- Very significant improvements in this regard in the past years
  - Widespread adoption of Git, CMake, code review via GitHub/Lab, continuous integration techniques
  - Software is more portable than in the past
    - CVMFS
    - Slimmed-down builds for laptop analysis
    - Container-based analysis
- Debugging and optimisation is still very tough: complicated frameworks often don’t play nicely with standard tools
  - `std::cout << “Got here 79.5” << std::endl;`

Activities and Limitations



We should aim to minimise the limitations that computing and software impose on our ability to do physics research


Completed jobs (Sum: 29,007,372)
Analysis - 37.06%
10,750,480
8,560,811
MC Simulation - 29.51%
3,667,753
T0 Processing - 12.64%
Analysis - 37.06% (10,750,480)
MC Simulation - 29.51% (8,560,811)
T0 Processing - 12.64% (3,667,753)
Others - 8.53% (2,473,403)
Group Production - 6.87% (1,992,469)
MC Reconstruction - 4.54% (1,317,520)
Data Processing - 0.84% (244,936)

Event generation

- Event generation begins the data processing chain with a physics simulation of the proton-proton collision, creation of initial states and evolution into final states that interact with the detector
  - As our knowledge of the Standard Model improves and more precise measurements are needed, we require higher-precision generation
  - Leading order has modest CPU requirements, next-to-leading is less trivial, but the HL-LHC will widely require NNLO: serious computing requirement
- Generators are written and maintained in the theory community, and maintenance of common software (e.g. HepMC, Rivet, LHAPDF) needs to be maintained
  - Some widely-used generators are not thread-safe
- Side point: event generation is particularly suited to closed HPCs since it has no input and small output

Detector simulation

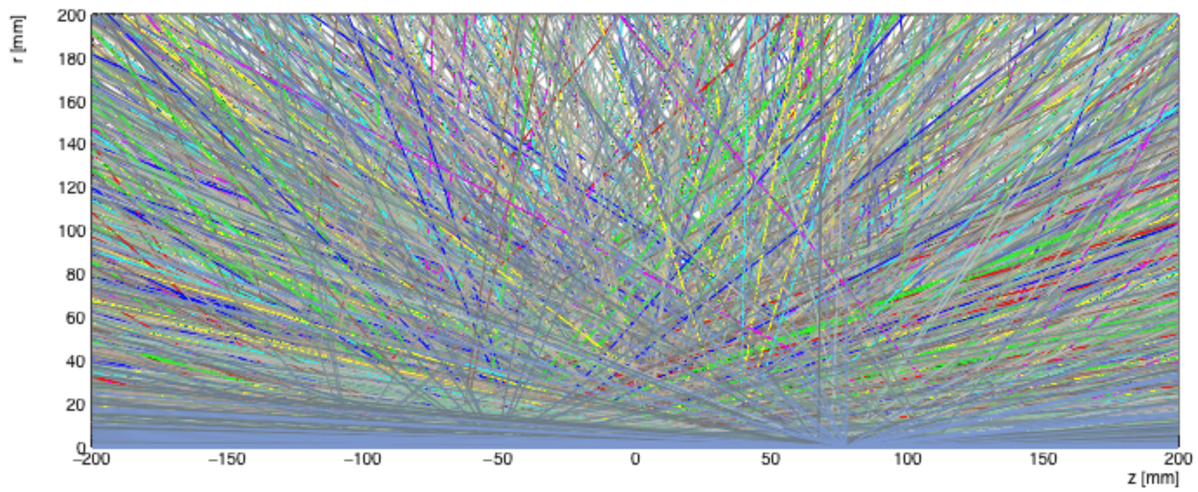
- This is our biggest consumer of CPU and it will only get bigger
- We do know how to speed it up already
  - More use of parameterisation
  - More use of partial simulation (e.g. only simulating regions around particles of interest)
  - Main challenge is validation
- Improved physics models
- Adapting to new computing architectures
- Advances in geometry modelling
- Machine learning?
  - Generative-adversarial systems?
  - Validation will be *tough*…

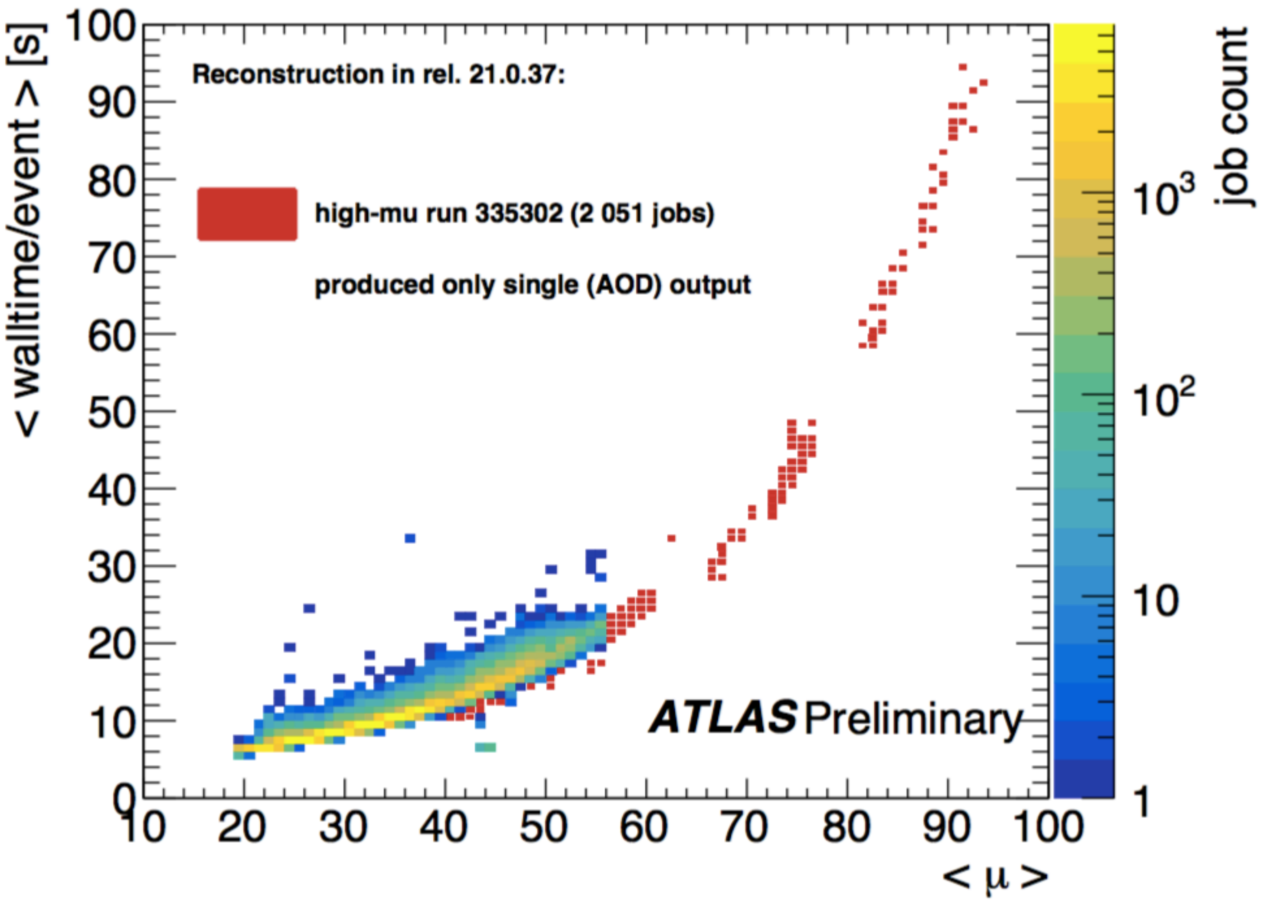
arXiv:1511.06434v2

Reconstruction

- Reconstruction (especially tracking) is particularly vulnerable to high pile-up and high track density
  - Especially at the clustering/pattern recognition step
  - Not clear our current physics performance can be maintained at high <μ>, especially at low momentum, with current algorithms
- This is a key area for use of concurrency and vectorisation
- Maybe machine learning can help?
  - Kaggle Challenge
- Beautiful presentation by Andi Salzburger @ Spåtind 2018 describing the challenges in detail

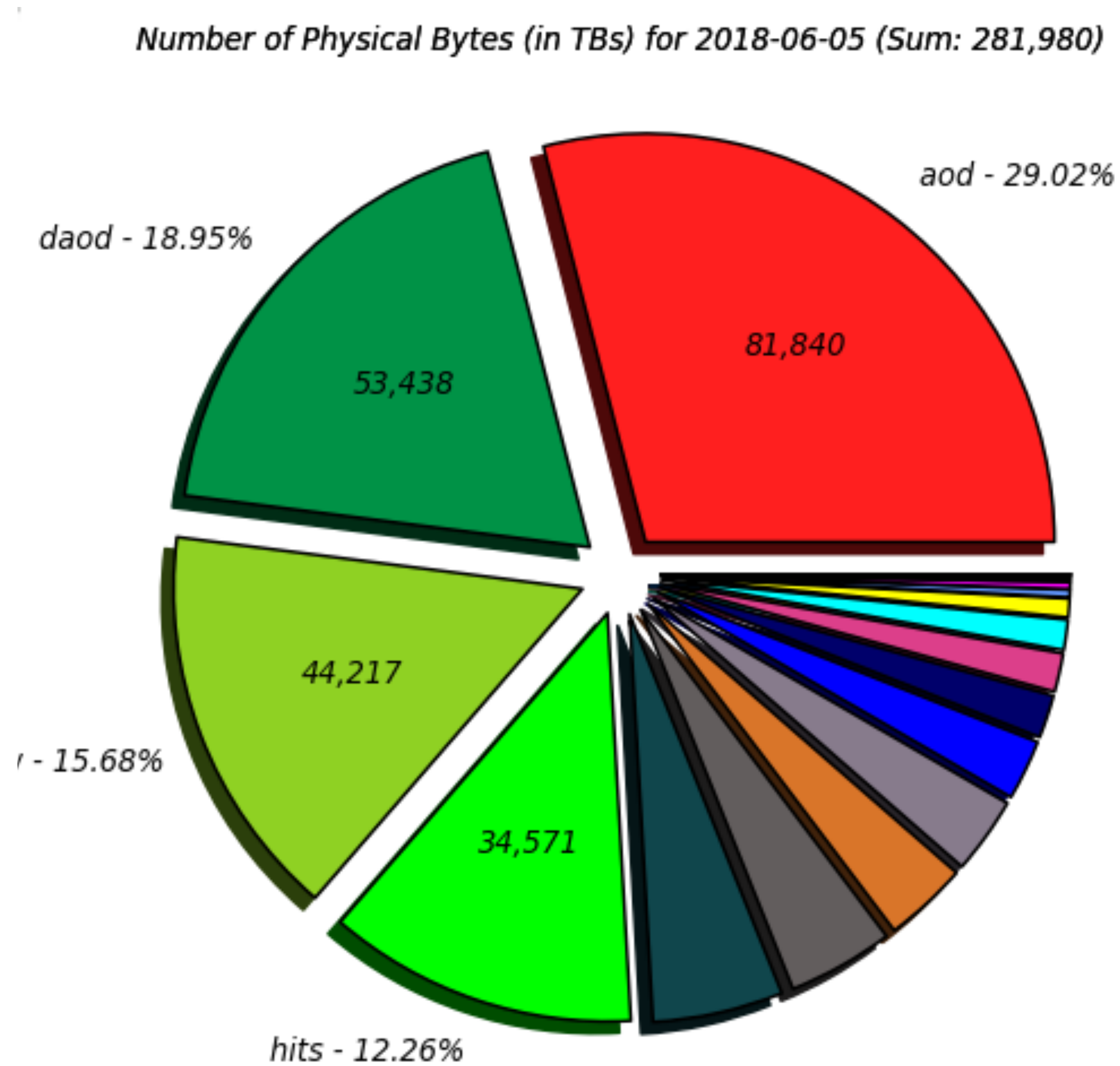
200 p-p collisions
HL-LHC conditions

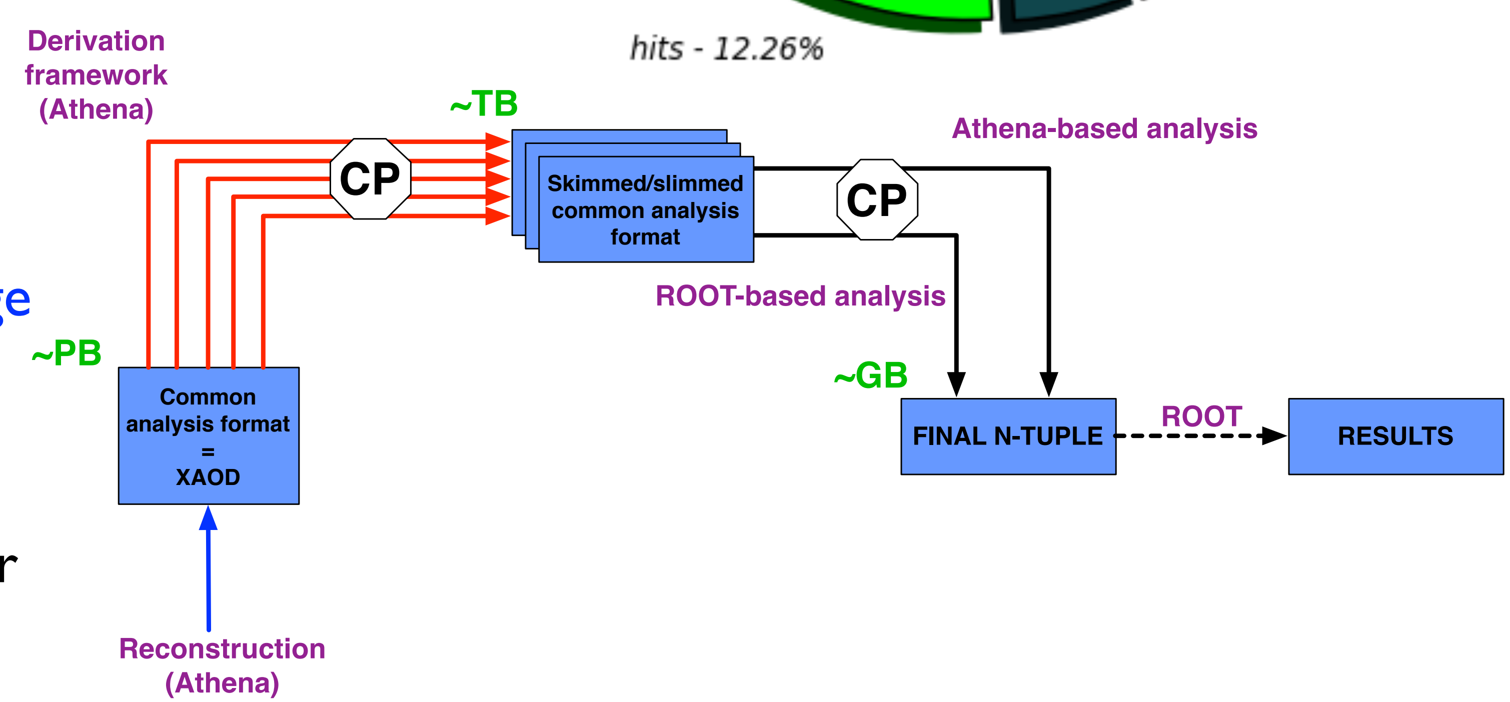




Analysis and analysis models

- How to enable analysts to do their work efficiently?
- Tensions
  - Flexibility vs resources vs uniformity vs imagination vs exceptions
  - Central processing vs user processing
  - Local analysis vs distributed analysis vs cloud analysis
  - Many formats vs few formats
  - Python vs C++
- How to account for changing calibrations within a restricted resources envelope?
- Experiments' analysis models differ widely
  - No single correct answer: requires continual review and willingness to change
- How to preserve physics data for the future (massive topic by itself…)
- Will we ever do our analysis by means other than looping over TB-sized ROOT files?





Other topics

- Conditions data access (opportunities for joint projects)
- Security
- Visualisation
- Analytics (yes, this is useful for physicists as well…)
- I/O and layout of data in memory; compression

People and training

- Not controversial to say that most experiments have a desperate shortage of people willing and able to work on software and computing
  - This comes at a time when we can't rely on technology improvements to keep us afloat
- We need to support
  - People who eventually plan to go into industry → need to stay relevant and provide training in modern technologies that are transferrable to the commercial or public sectors
  - People who want long-term careers in HEP → need to recognise SW&C work as equal to detector and physics analysis work
    - As things become more complicated we need physics leaders with strong interests in SW&C
- We need to improve our citation and publication record
- We must invest in training our community at different levels from basic analysis to advanced software engineering
- Collaboration with those from other academic fields and industry is important in this regard

A Roadmap for HEP Software and Computing R&D for the 2020s: https://arxiv.org/abs/1712.06982

Advancing from here

- Community White Paper process has been a success
  - Engaged more than 250 people and produced more than 300 pages of detailed description in many areas
- Summary roadmap lays out a path forward and identifies the main areas we need to invest in for the future for our software upgrade
  - Supporting the HL-LHC Computing TDRs and NSF S2I2 strategic plan
  - You can still sign :-)
- HEP Software Foundation has proved its worth in delivering this CWP Roadmap
  - Achieving a *useful* community consensus is not an easy process
  - Sign up to our forum to keep in touch and get involved (hep-sf-forum@googlegroups.com)
- We now need to marshal the R&D efforts in the community, refocusing our current effort and helping to attract new investment in critical areas
  - The challenges are formidable, working together will be the most efficacious way to succeed
  - HSF will play a vital role in spreading knowledge of new initiatives, encouraging collaboration and monitoring progress
  - Next HSF workshop in March, shared with WLCG, should start to put our ideas into practice:
    - C++ Concurrency, Workload Management and Frameworks, Facilities Evolution, Analysis Facilities, Training, …