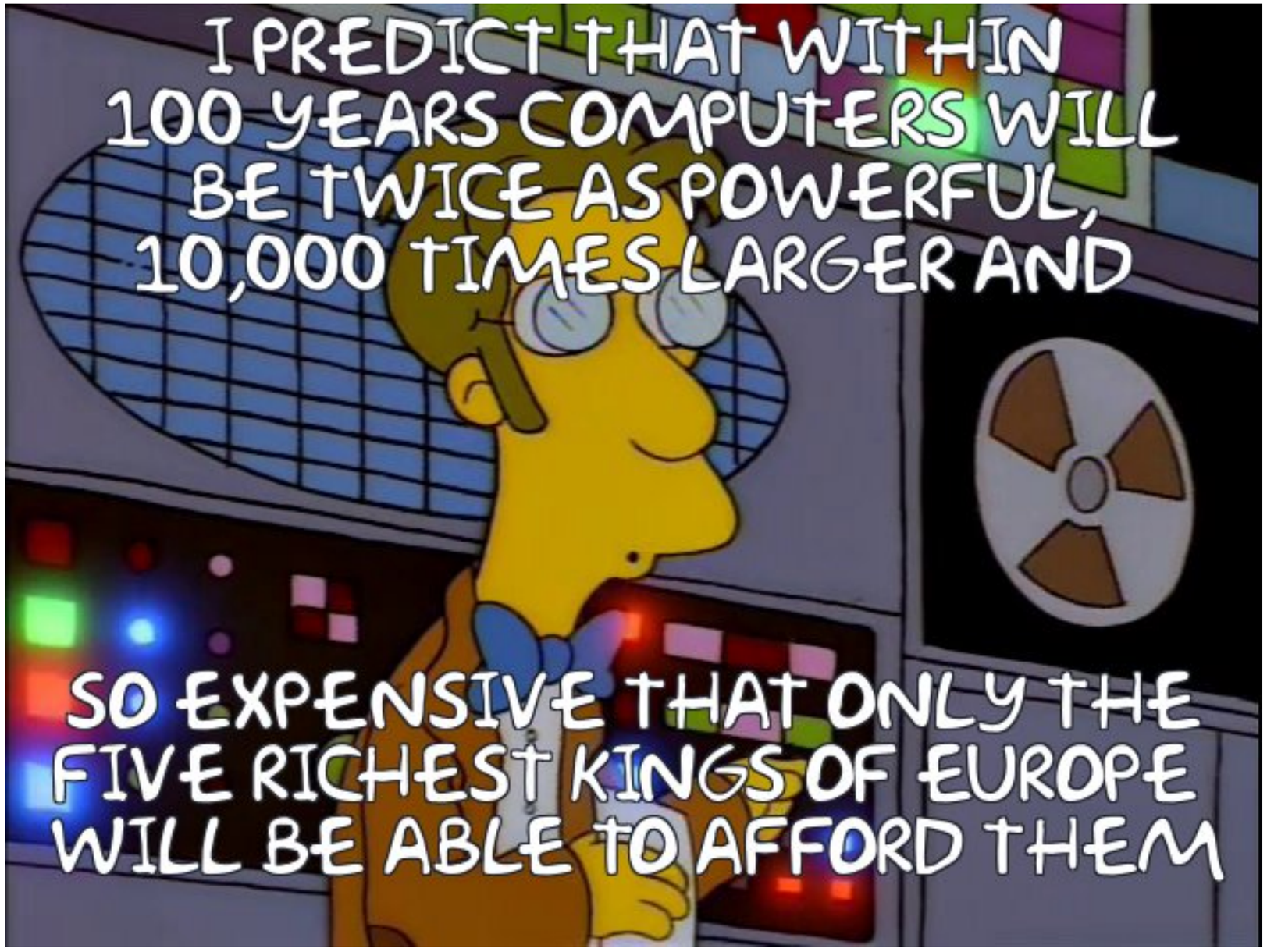


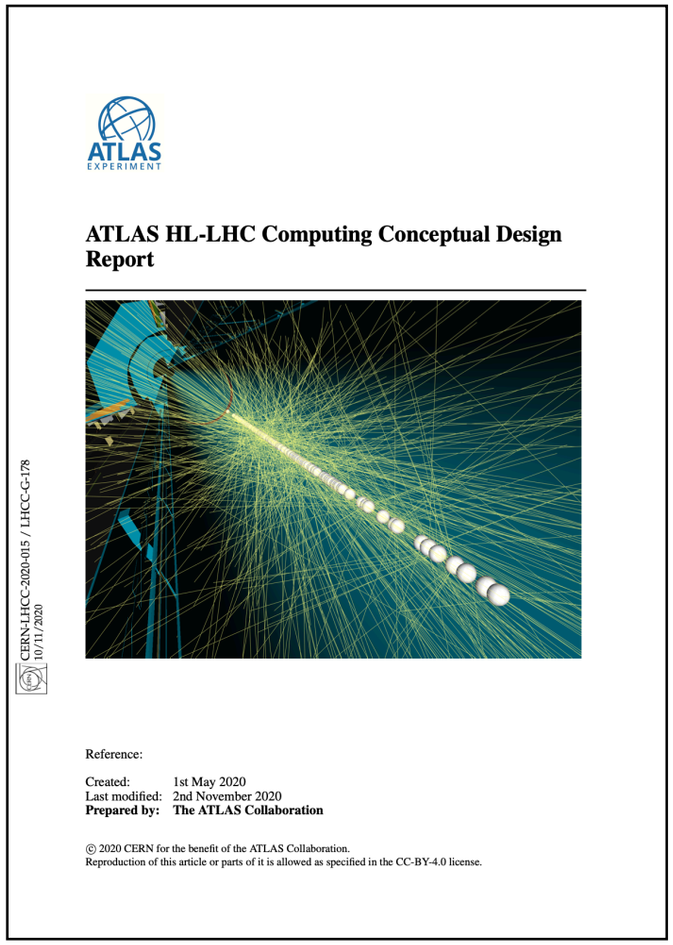
ATLAS R&D & the CCE

Zach Marshall (LBNL), HEP-CCE All-Hands, 20 April 2022
Thanks to David Cameron, John Chapman, Alessandro Di Girolamo, and Johannes Elmsheuser

Introduction and Context

- ATLAS has two documents describing our plans for the HL-LHC
  - The Conceptual Design Report (2020), laying out the broad challenges facing each area
  - The HL-LHC Roadmap (2022), layout out milestones and deliverables for reaching the HL-LHC successfully
- Both prepared in the context of LHCC reviews
- Most of the material I'll show here is drawn from those documents, with some extra opinions and ideas thrown in
  - They're great reads!

"Prediction is very difficult, especially about the future." – Maybe Bohr, Berra, Petersen, Twain…


ATLAS EXPERIMENT

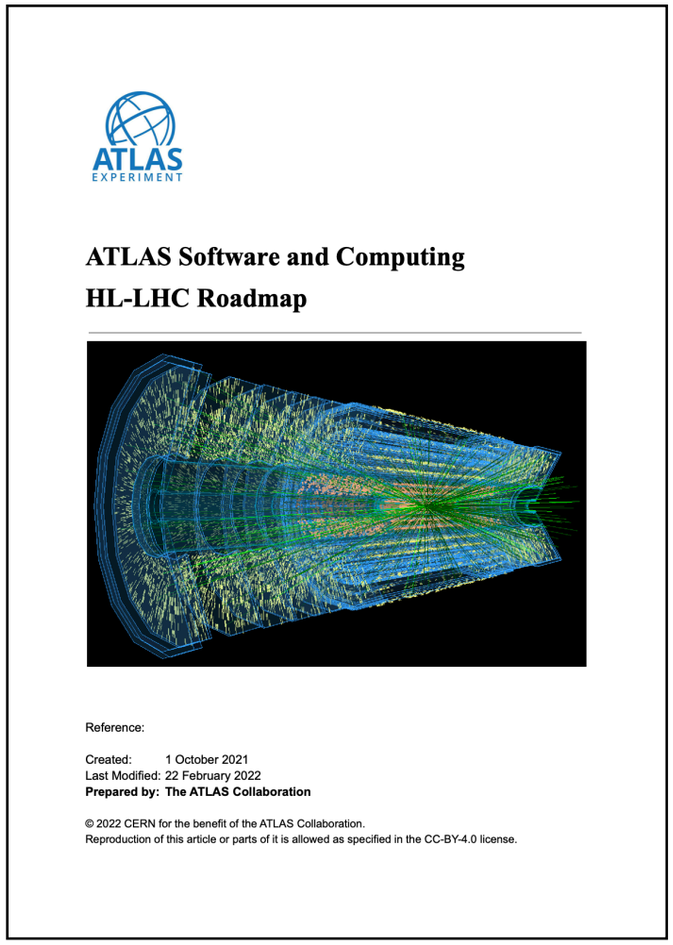
ATLAS HL-LHC Computing Conceptual Design Report

CERN-LHCC-2020-015 / LHCC-G-178
10/11/2020

Reference:

Created: 1st May 2020
Last modified: 2nd November 2020
Prepared by: The ATLAS Collaboration

© 2020 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.


ATLAS EXPERIMENT

ATLAS Software and Computing HL-LHC Roadmap

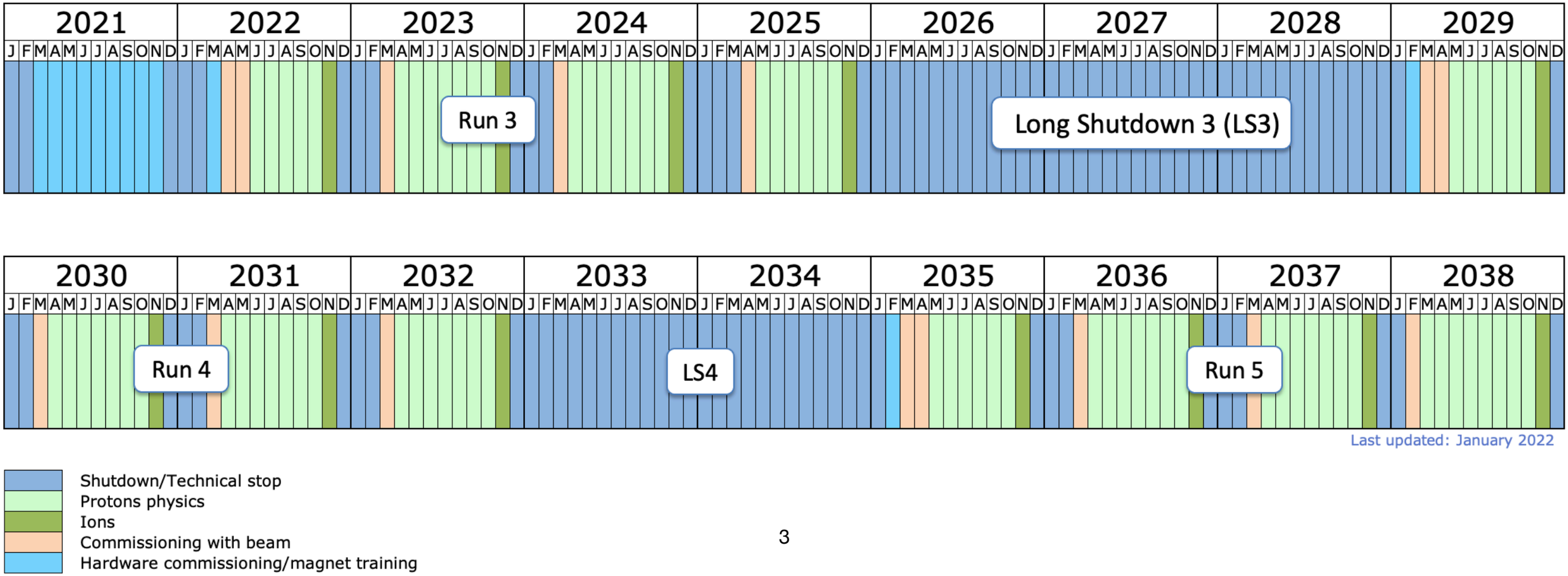
Reference:

Created: 1 October 2021
Last Modified: 22 February 2022
Prepared by: The ATLAS Collaboration

© 2022 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

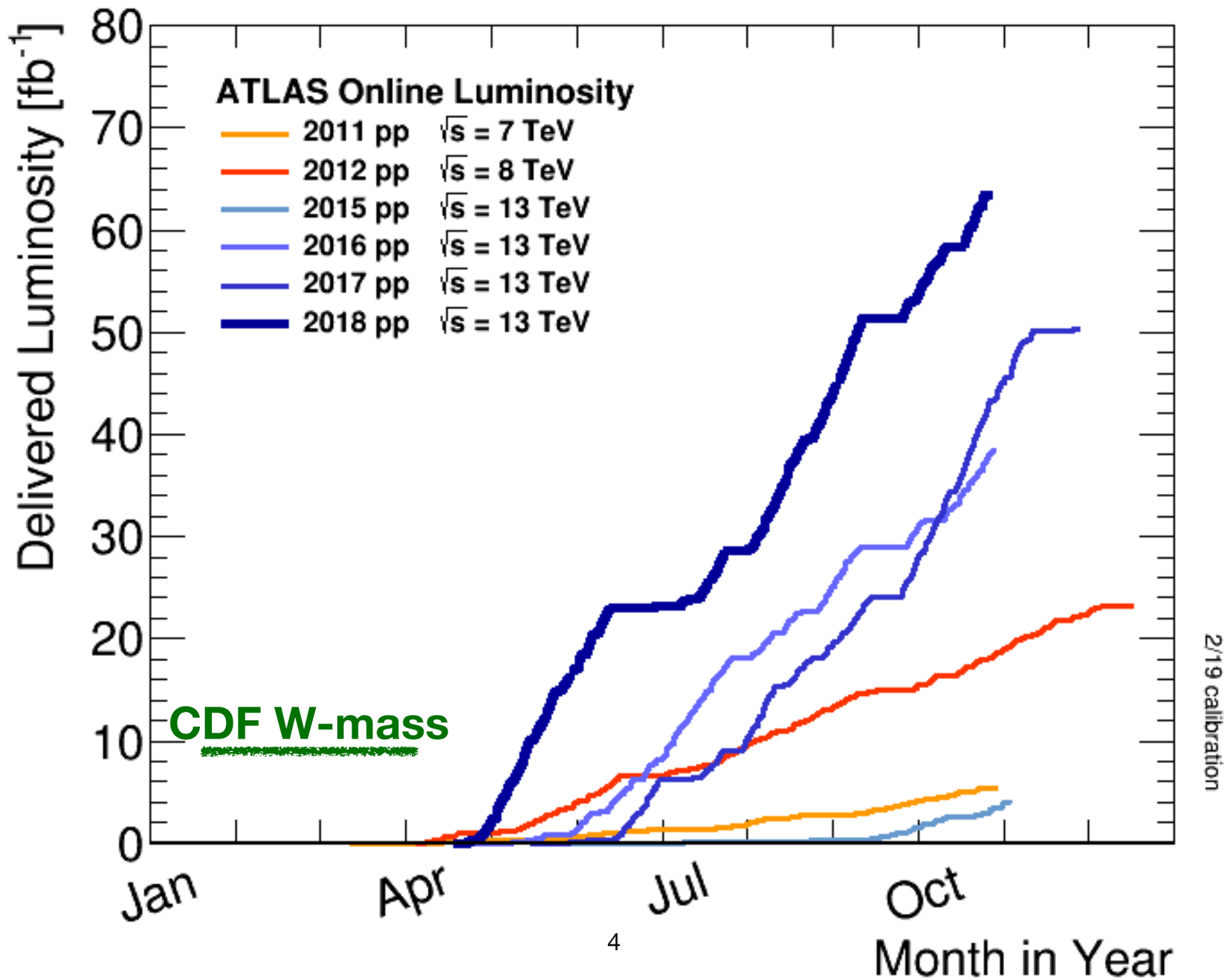
Schedule

- HL-LHC schedule recently adjusted to start in 2029
  - Essentially a two full-year delay w.r.t. CDR times
  - LS3 will be quite long – lots of time for upgrades!
- Nominal CERN plan includes a Run 6 in 2040-41 after LS5
  - This makes FCC scary: pushes machines at CERN to late 2040s



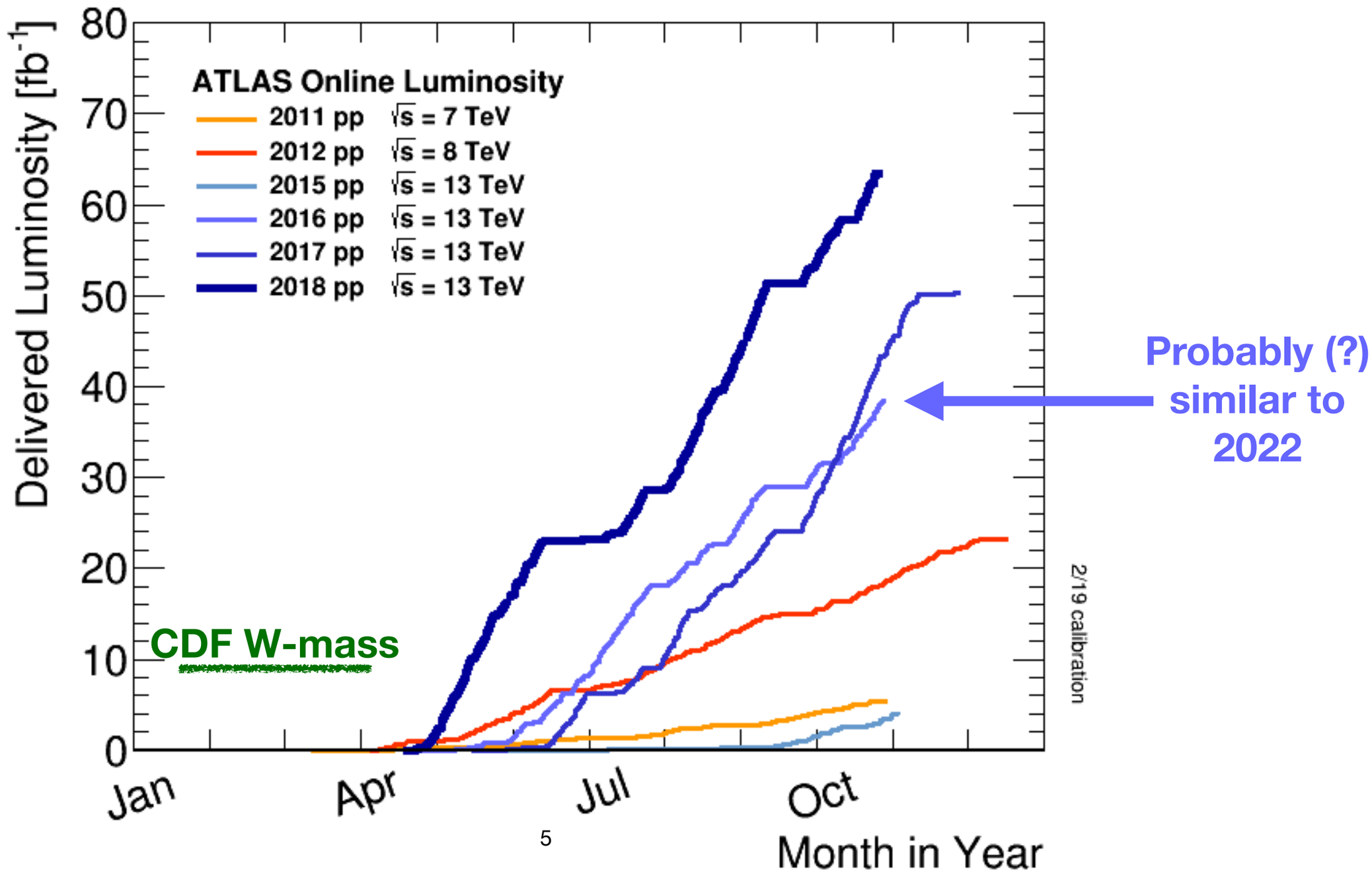
Physics Challenge

- Our data set size is measured in either events or inverse fb
- Here is what our data taking to date has looked like:



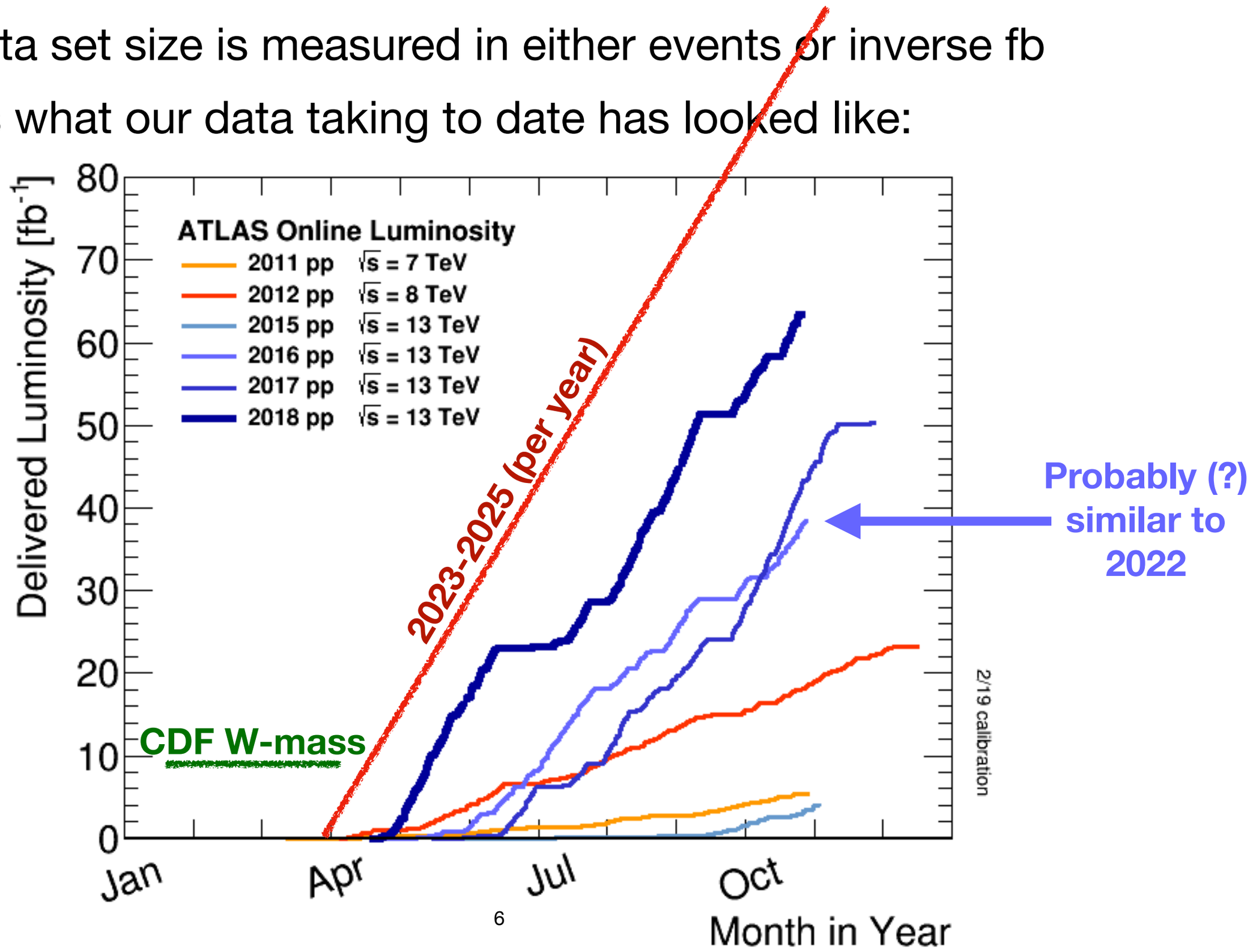
Physics Challenge

- Our data set size is measured in either events or inverse fb
- Here is what our data taking to date has looked like:



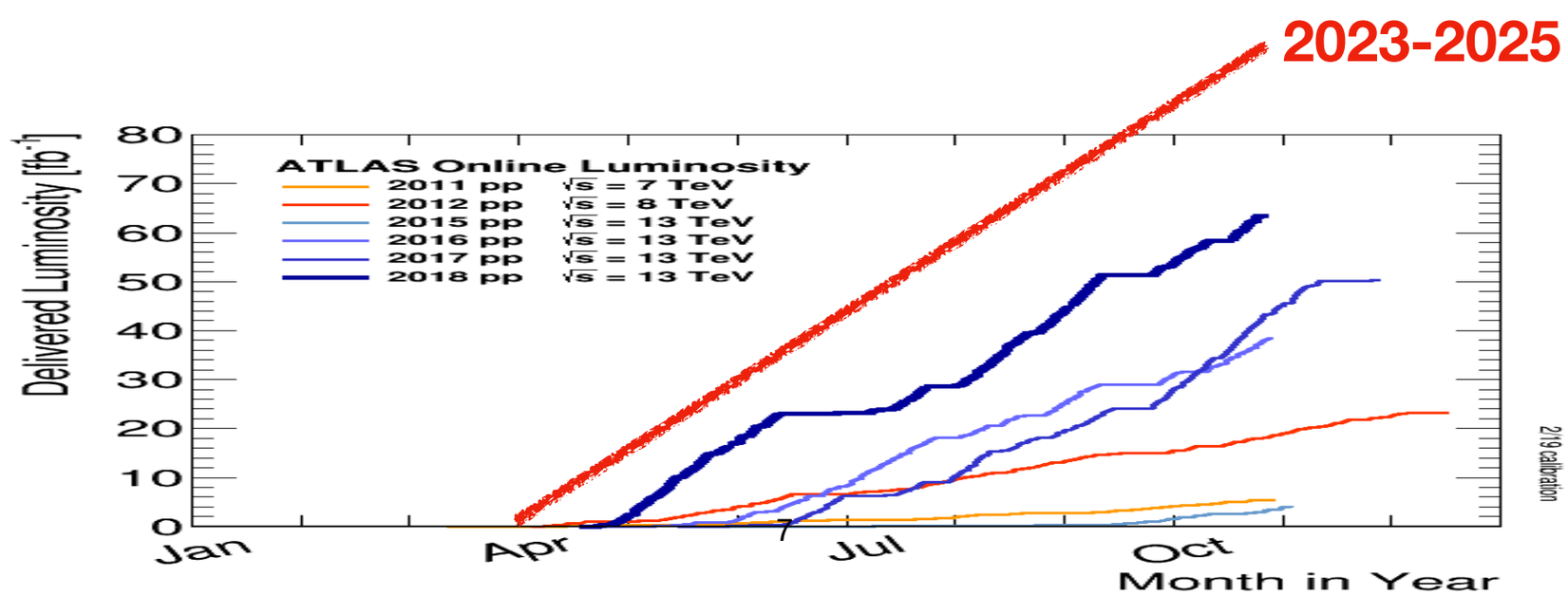
Physics Challenge

- Our data set size is measured in either events or inverse fb
- Here is what our data taking to date has looked like:



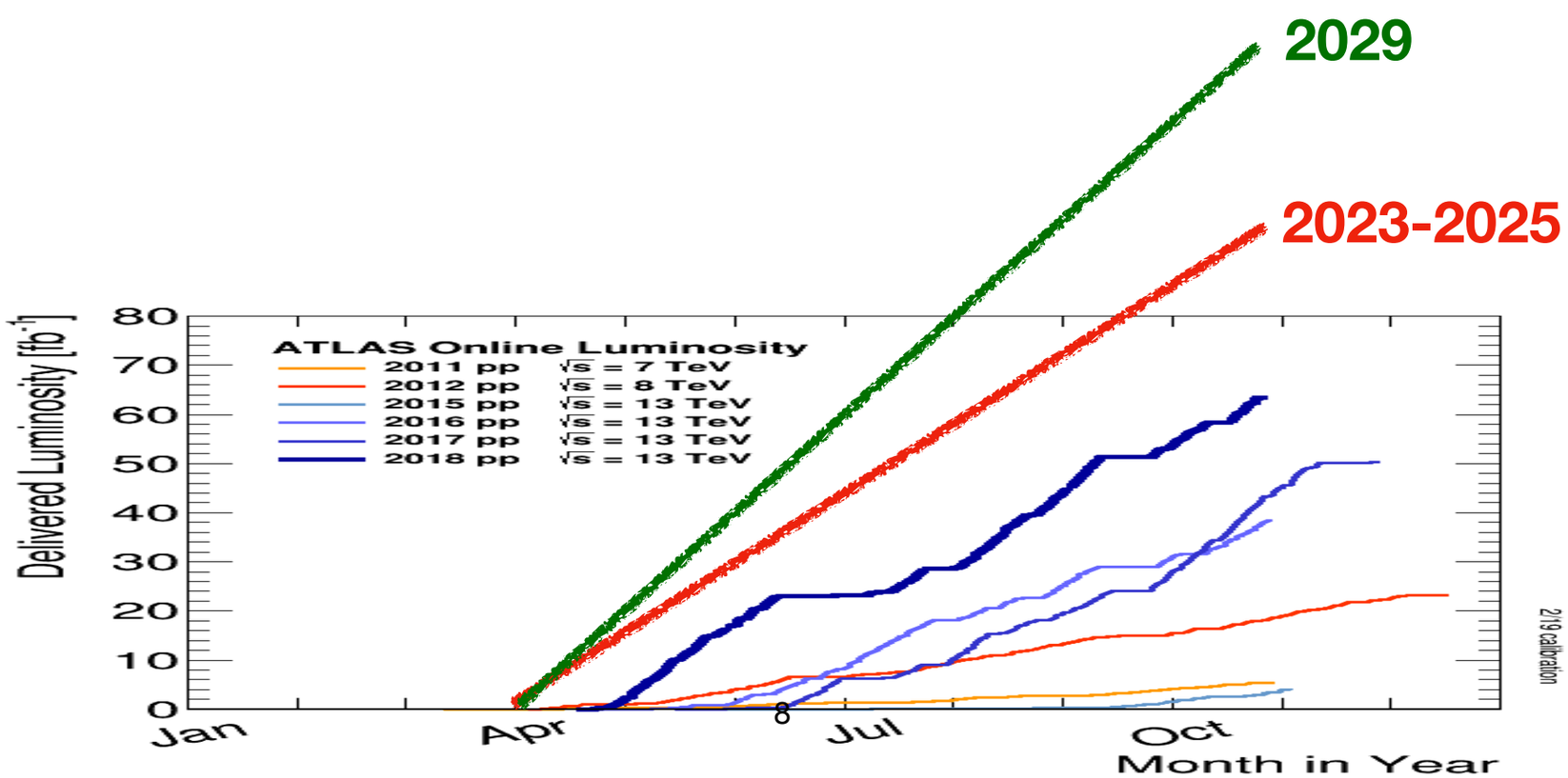
Physics Challenge

- Our data set size is measured in either events or inverse fb



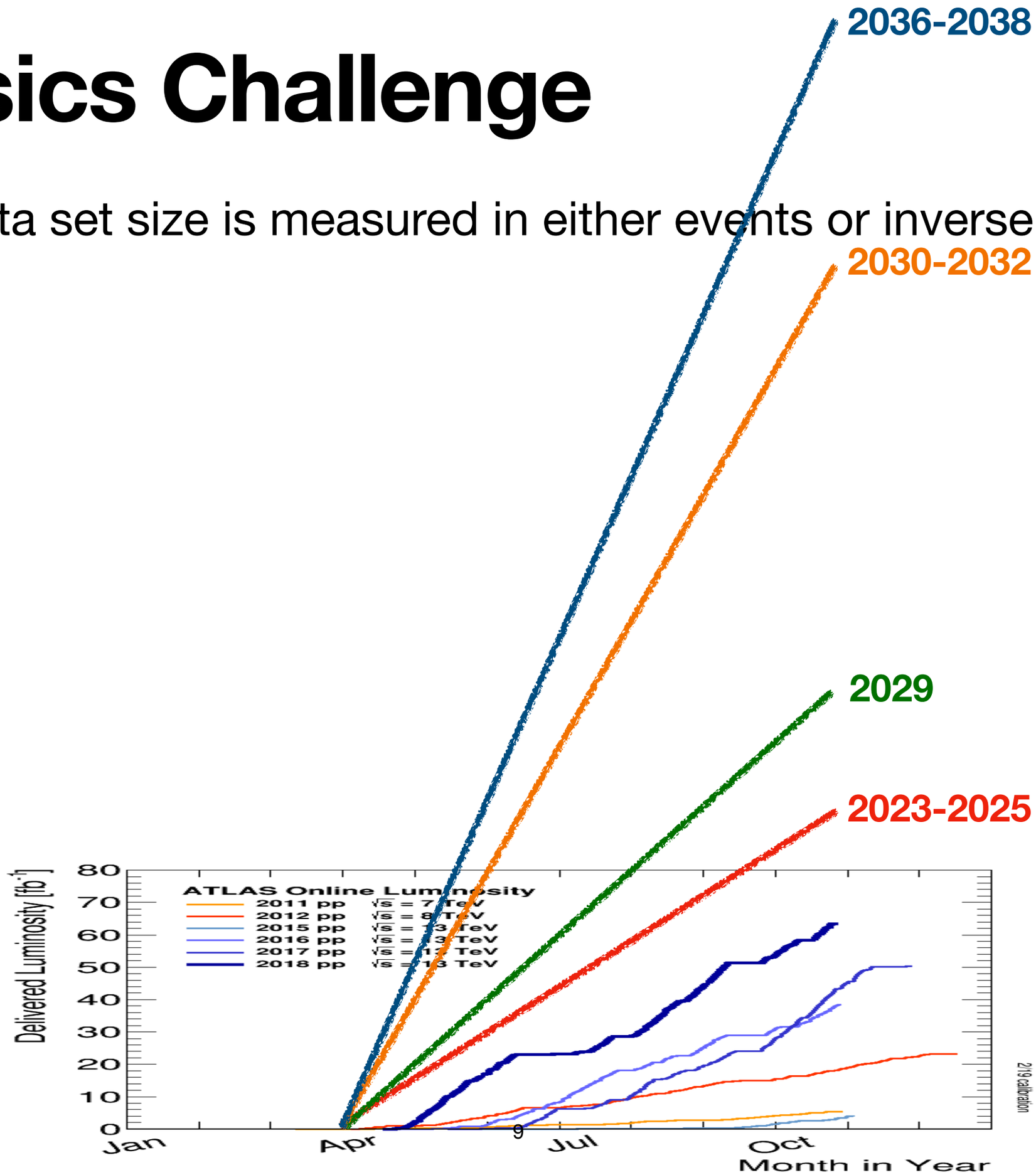
Physics Challenge

- Our data set size is measured in either events or inverse fb



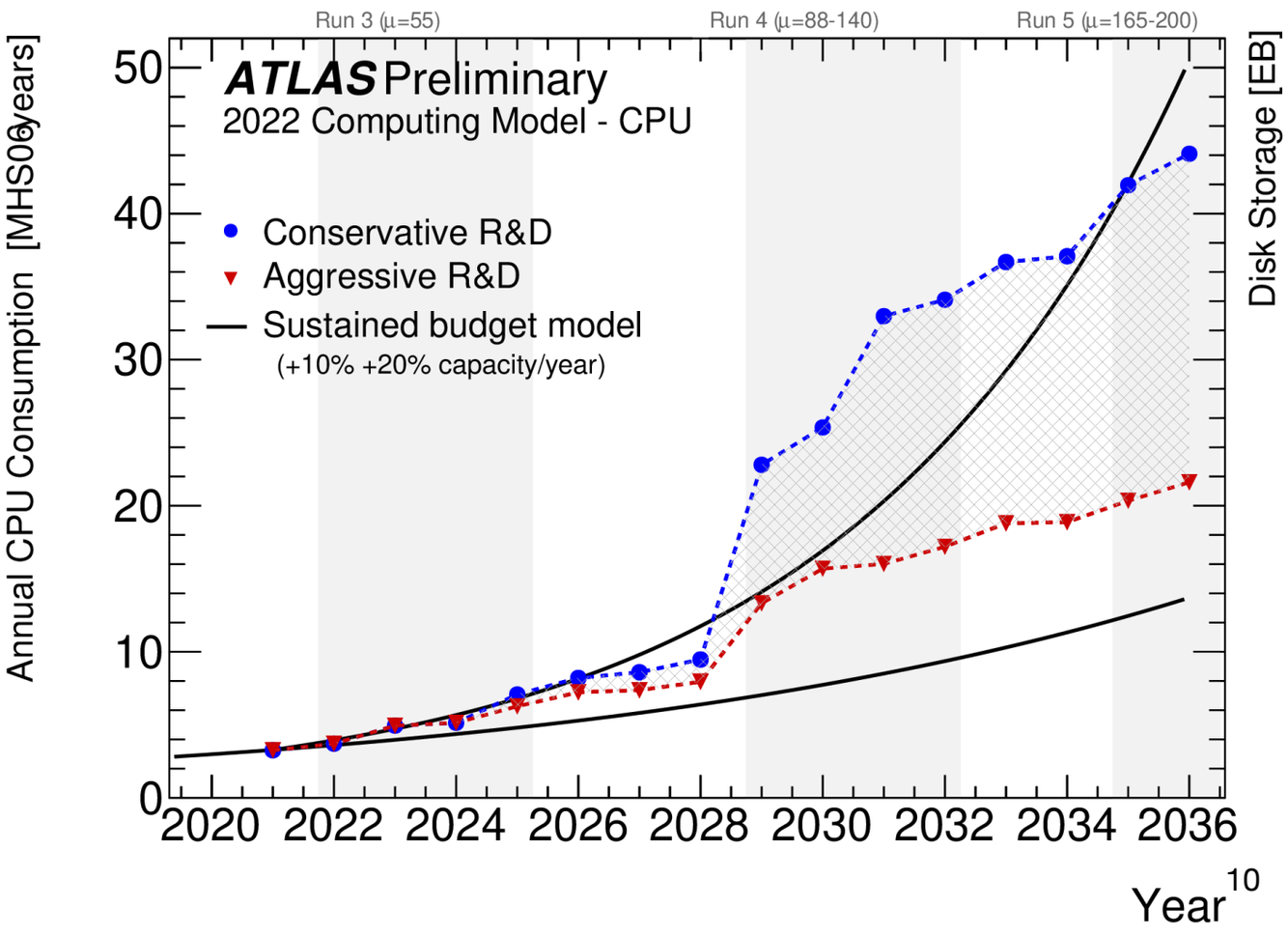
Physics Challenge

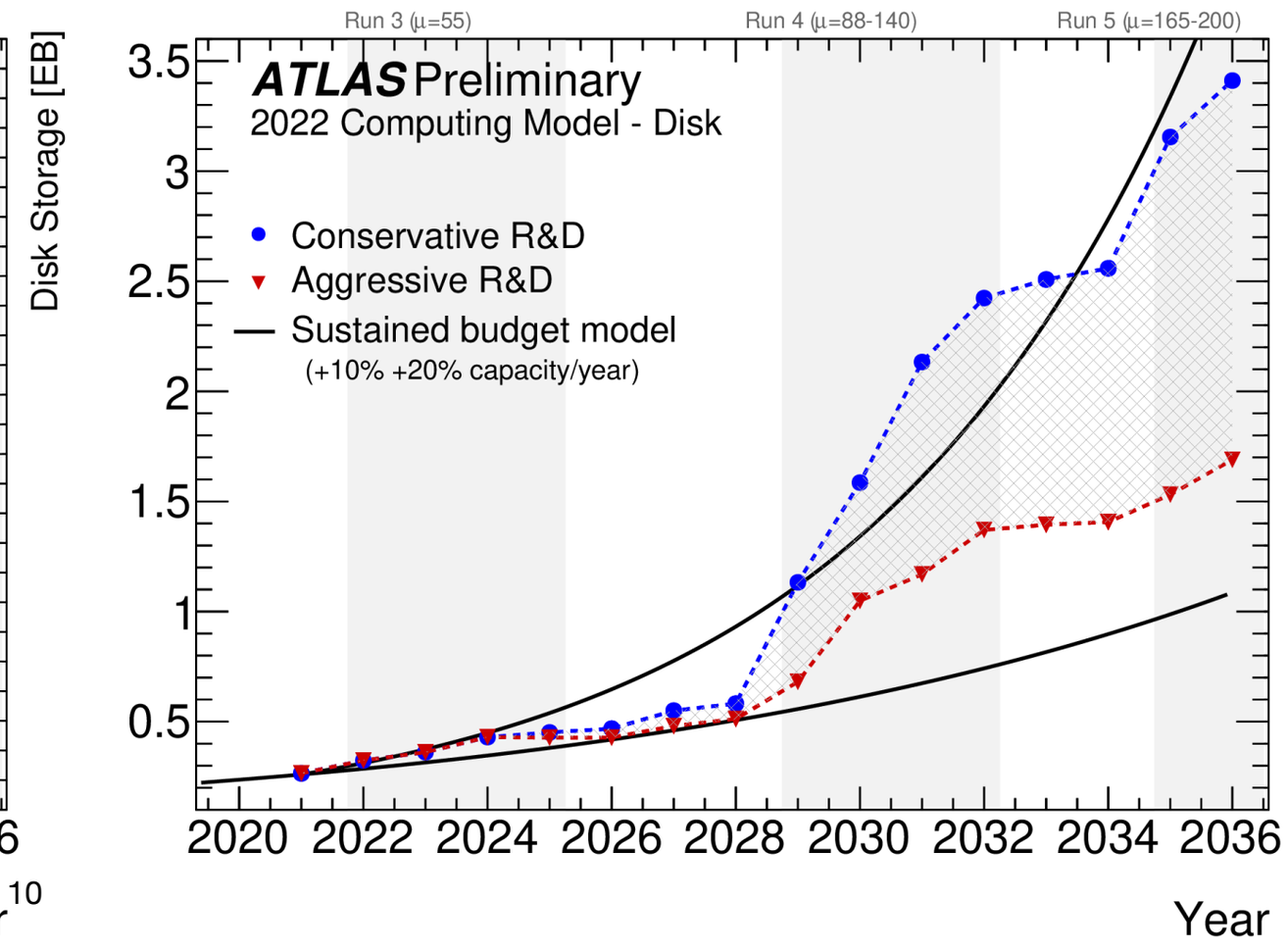
- Our data set size is measured in either events or inverse fb



Cutting to the chase…

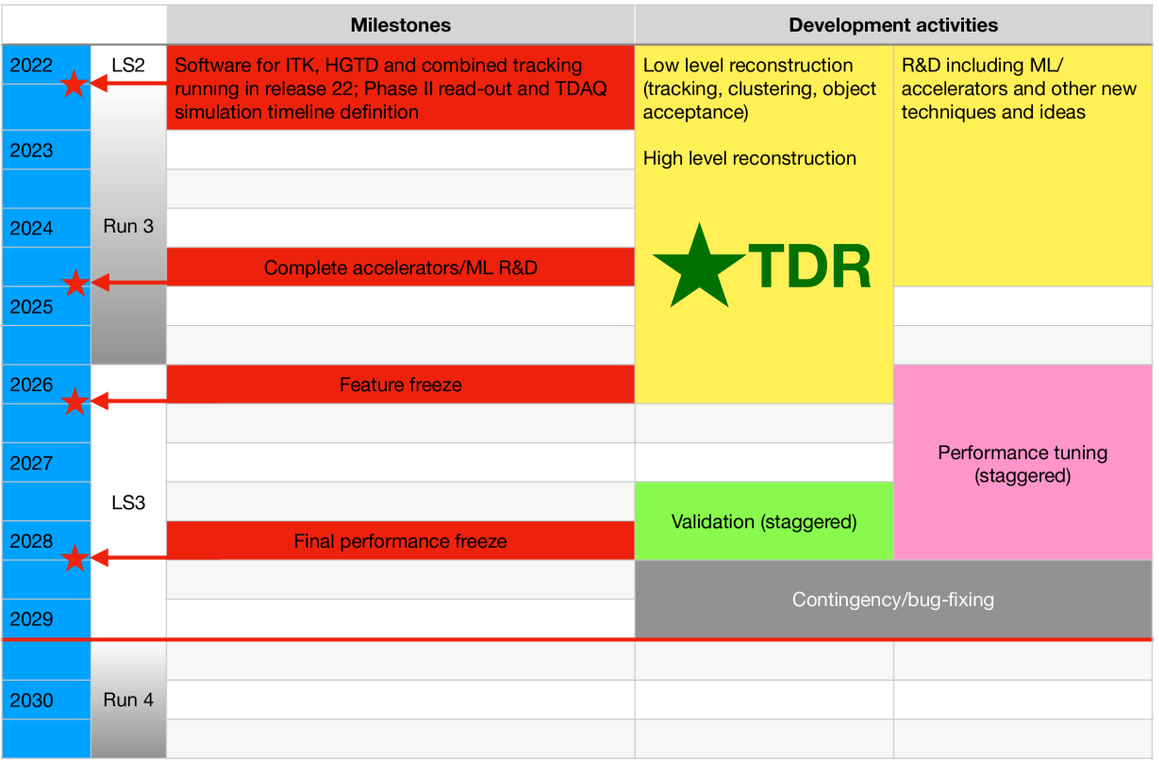
- We have defined **conservative** and **aggressive** R&D scenarios
  - Conservative should be achievable with today's effort and people
  - Aggressive requires more people, or some good fortune with R&D
  - Some projects not (yet?) included in aggressive (e.g. GPU usage)
- Comparing those to **'flat budget'** capacity increases (10-20%/yr)

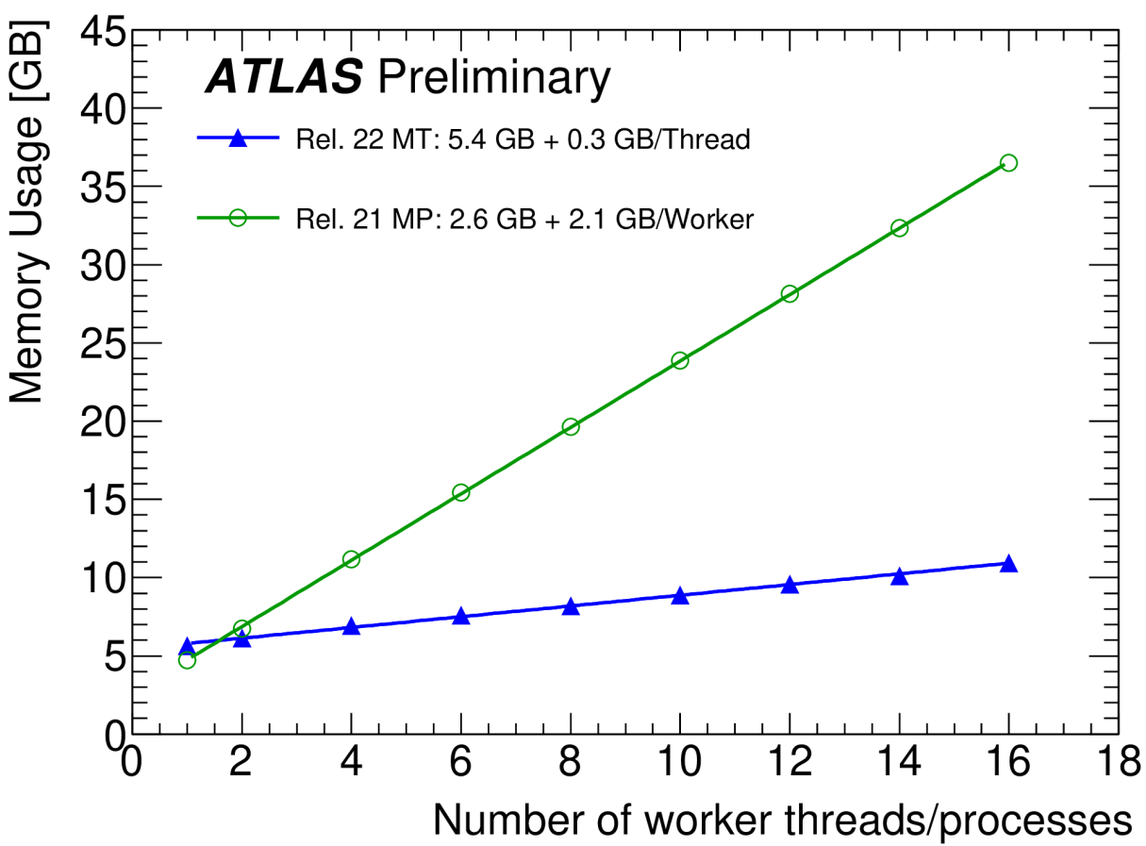




Time to Act

Combined performance milestones and activities





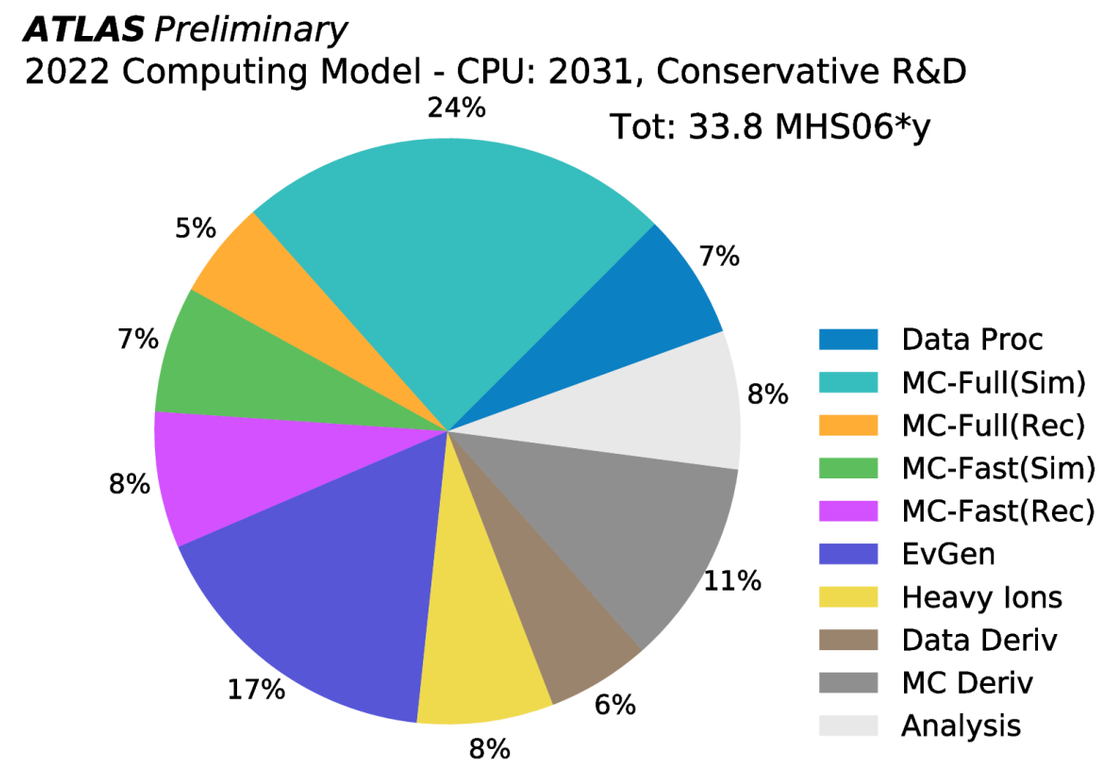
- Working backwards from the HL-LHC schedule, it is clear that R&D needs to be going on **now**
  - You're helping! Thanks!
- We will need **ample** time for integration and validation of solutions before data taking
  - We have some experience with this from the recent MT migration
  - Almost 10 years from ideas to validated and running reco!
  - Implies **late arriving R&D is risky**!
- Expecting that in 2025 we have a clear path to HL-LHC data taking, e.g. with or without accelerators

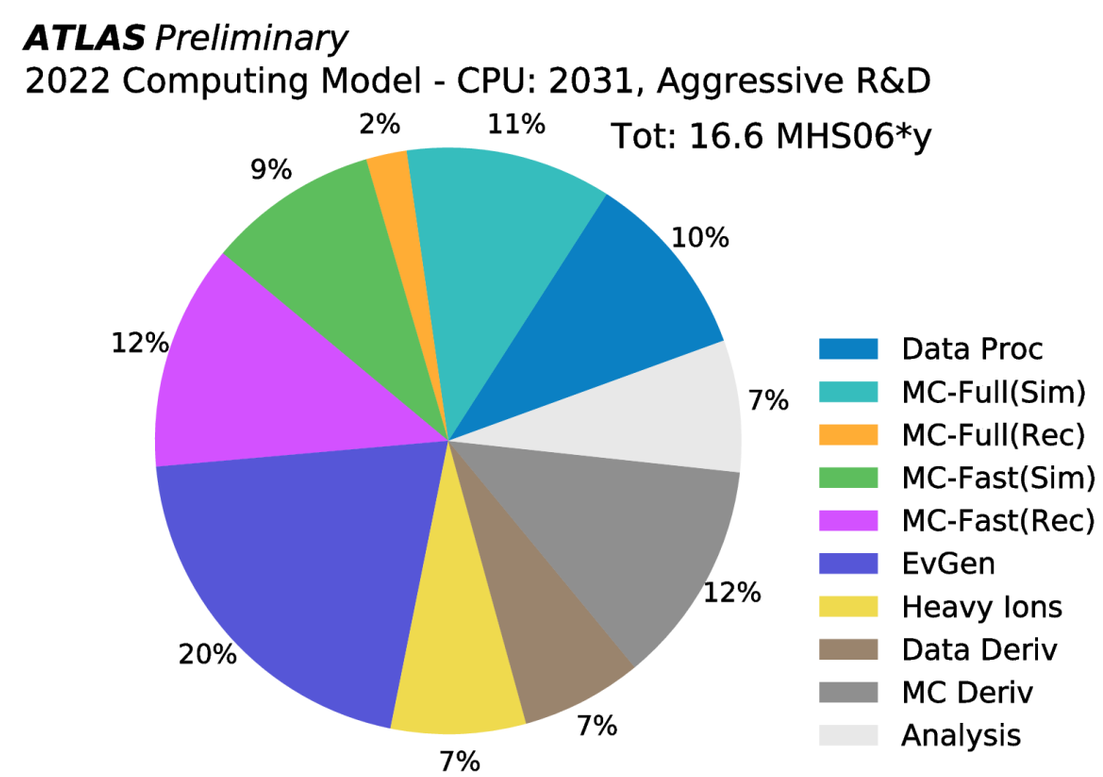
Lots of Business as Usual

- There is a *lot* of effort that needs to go into non-R&D work
  - Maintaining our current software and old-fashioned optimization
  - Updates to database infrastructure and usage
  - Improvements to metadata storage and usage
  - Upgrade detector geometry and corresponding digitization
  - I/O improvements (RNtuple ~between R&D and business as usual)
  - Re-tuning of fast and Geant4-based simulations for Run 4 detector
  - Re-tuning of reconstruction to the Run 4 detector and conditions
  - Distributed computing evolutions (tokens, OS, network, storage)
- R&D efforts are *on top of* this effort
- There's an art / balance in ensuring Run 3 and Run 4 success

ATLAS's Big CPU Problem

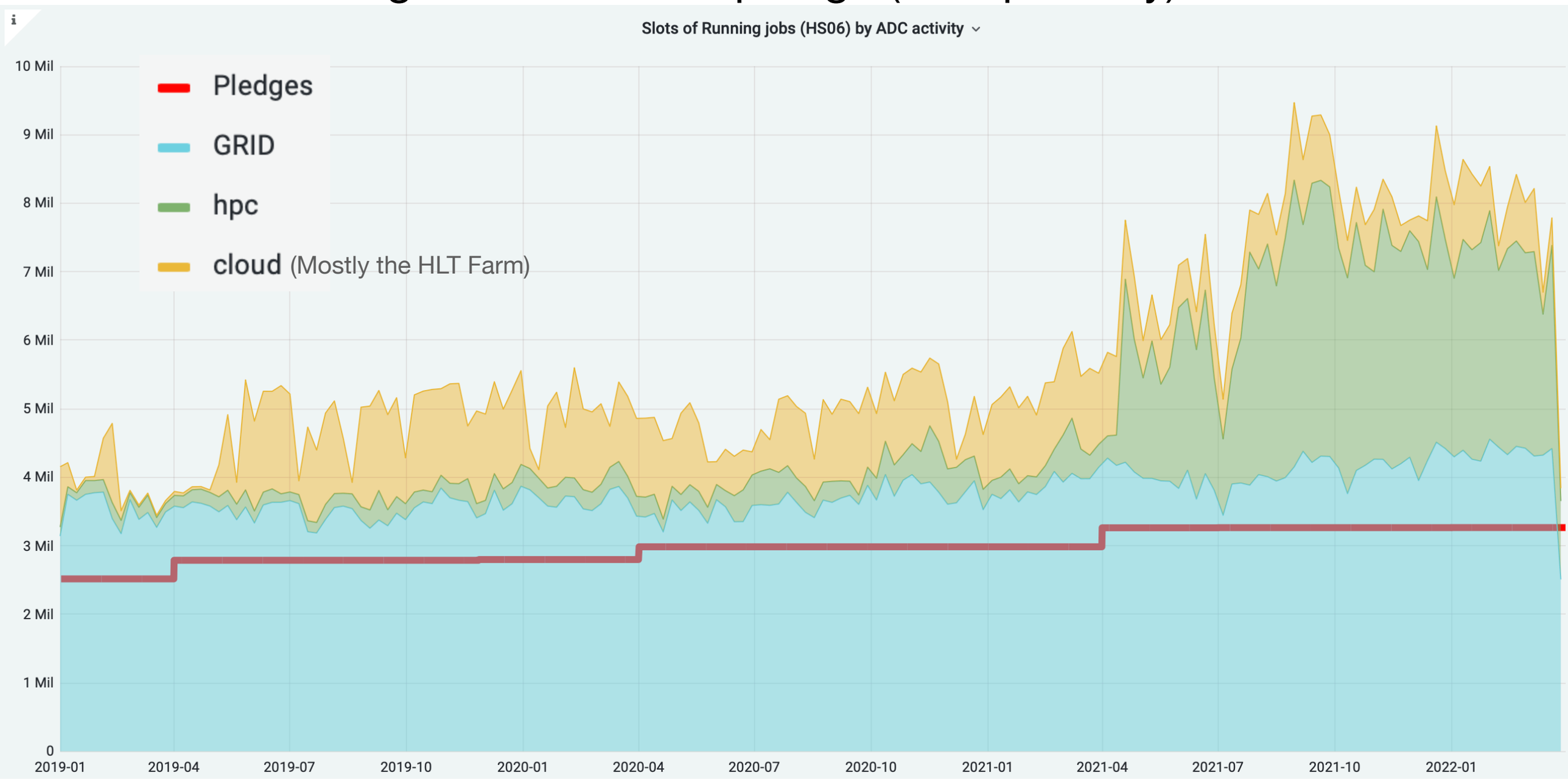
- Whether in the conservative or aggressive scenario, our biggest "problem" is that there is **not one dominant workload**
- We will have to work on everything to succeed
- This also means that we can afford to stumble without a serious risk to our physics program
- One additional possibility: that we get more CPU "for free"





HPCs and ATLAS

- We have been extremely successful with HPC resources lately
- Note: some grid sites deliver pledge (transparently) via HPCs!



HPCs and ATLAS (II)

- Most resources provided by:
  - Vega (large majority) and Karolina, both EuroHPC systems
  - Both run ~all workloads (no analysis)
- Why were these so successful?
  - Basically x86 machines (no major code adaptation; similar to a std grid site)
  - Insiders helping with edge services (and even machine setup)
  - Early adoption helping give large allocations (in before others)
- Biggest problem with HPCs in general: No associated **disk allocations!**
- Second biggest problem: **Single-year allocations**.

ATLAS's HPC Strategy

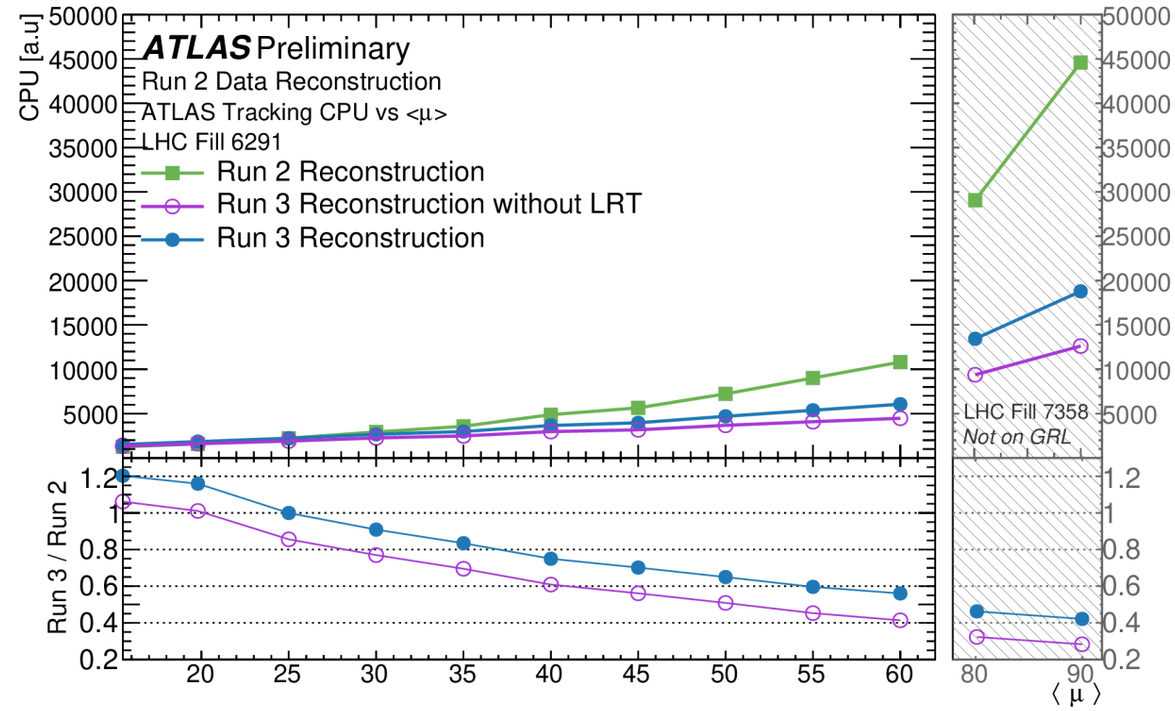
- Investment in proportion to opportunity
  - At the moment, e.g., we see ARM as a higher priority than Power
  - **Not** trying to get in on every machine!!
    - We don't need to – getting the easy ones gives ample CPU.
    - Without multi-year pledges with disk allocations, this is about it.
- We have a good (and growing) toolkit for edge services
  - With several available solutions, one is likely to fit a new machine
- Work towards several HPCs is ongoing
  - Perlmutter in the US is just about working for us
  - Hope to capture some of Fugaku(2)/CSCS with ARM development
  - Getting close to using Toubkal (largest HPC in Africa)
- Always watching for the next **low-hanging** fruit
  - While still finding "features", like the importance of I/O on HPCs

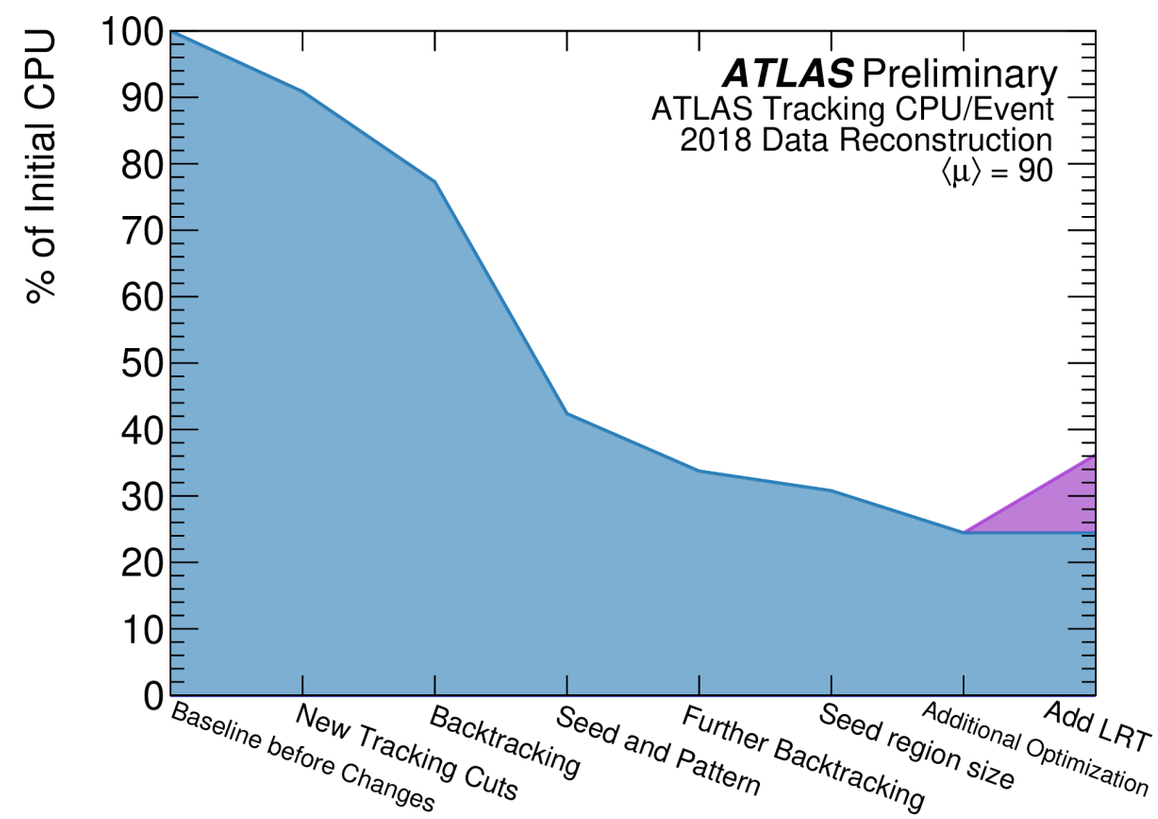
Compute R&D: All of the Above

- Trying to push R&D projects to define their scope and potential
  - The hardest part is defining success: in a many-dimensional space of “performance”, what is “better”?
  - The TDR (2024) is a good moment for a down-select: decide what projects will be a part of the HL-LHC baseline
- We are watching with great interest the Geant4 collaboration (+Adept and Celeritas) and Event Generator groups’ R&D efforts
  - Also trying to show our schedule around to make deadlines clear
- Note that not all our R&D will reduce our resource consumption!
  - I expect we will run a different mix of event generators in 2028
  - Major migrations often *increase* resource reqs, at least temporarily
  - Lots of R&D associated with reconstruction, today mostly focusing on improvements or extensions of what we already run

CPU R&D: A personal take

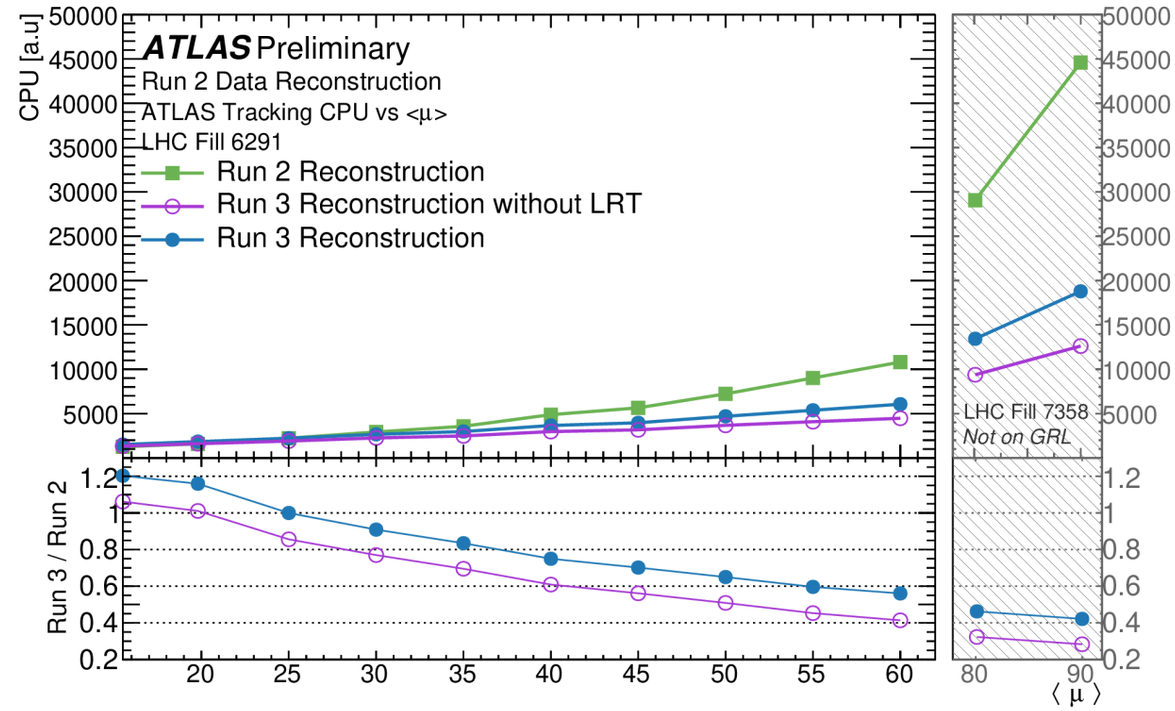
- Charged particle tracking is a great example of hot R&D
  - It is slow in the HL-LHC – it's always been one of the heavier parts of the reconstruction
- Lots of work was done in preparation for Run 3
  - The **vast** majority of the speed up came from **physics-driven optimizations** of the algorithms
  - We should be investing in some of these optimizations **now** to avoid wasted code optimizations
  - N.B. physicists will **find ways** to use the CPU again :)

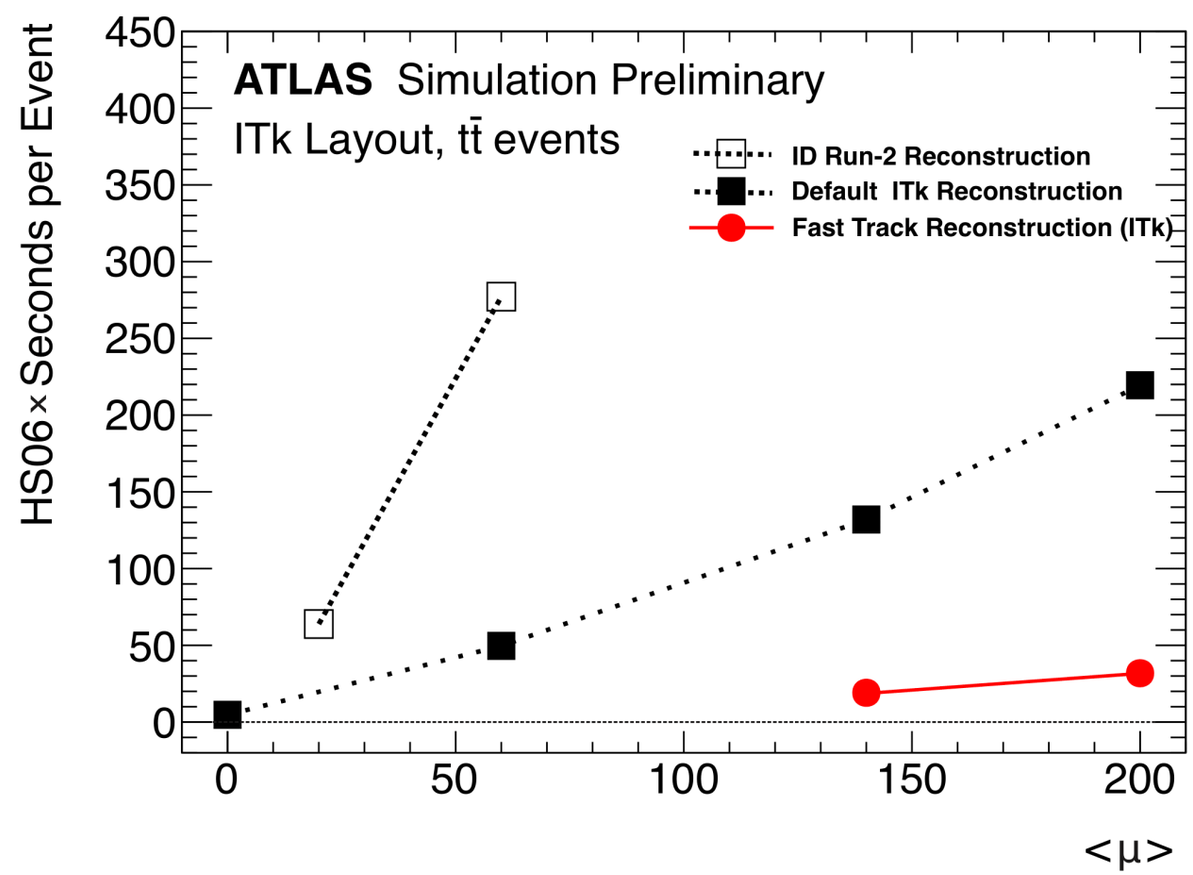




CPU R&D: A personal take

- Charged particle tracking is a great example of hot R&D
  - It is slow in the HL-LHC – it's always been one of the heavier parts of the reconstruction
- Lots of work was done in preparation for Run 3
  - The **vast** majority of the speed up came from **physics-driven optimizations** of the algorithms
  - We should be investing in some of these optimizations **now** to avoid wasted code optimizations
- Early/simple attempts already show this feature





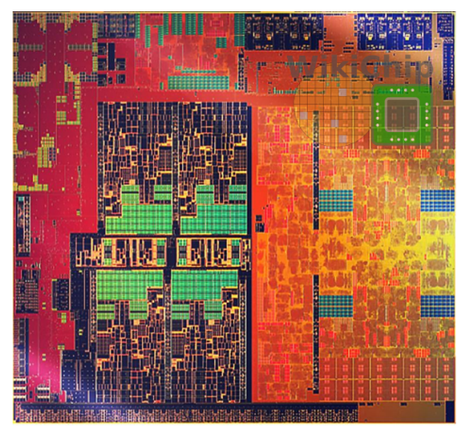
Compute R&D: Side-benefits

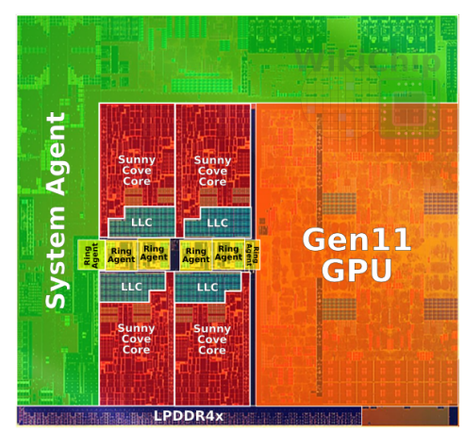
We choose to go to the Moon in this decade and do the other things, not because they are easy, but because they are hard; because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone

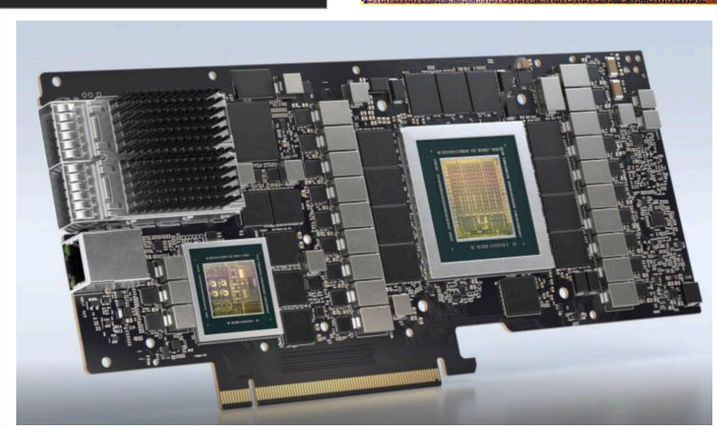
- Attempts at "revolutionary" R&D often have *significant* ancillary benefits
- Two examples from recent years:
  - GeantV: parallel detector simulation (closed)
  - ATLAS's heavy involvement in ACTS (ongoing)
- Even if we don't hit the "primary" target, many good things happen along the way
  - Detailed code reviews and clean-up campaigns (sometimes the first in many years)
  - Serious considerations of redevelopment (e.g. EDM changes, changes to virtualization…)
- Both these projects were about *toolkits* – and we can still use *some* of the tools!
- These have *major* benefits (independent of e.g. accelerator use)

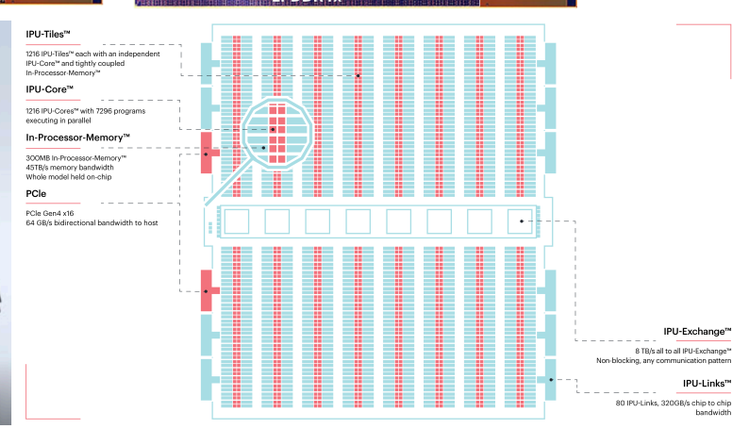
Accelerators: Why

- In terms of **Total** cost of ownership, accelerators often don't pay
  - Modern GPUs are *extremely* expensive
  - Good FPGA/GPU-based software engineering is *very* expensive and requires long-term expert support (v. CPU-based code)
- Accelerators do pay when you have a *fixed space or power budget* and need to expand your *total compute*
  - This is the case for e.g. our trigger farms









Accelerators: What, When, How

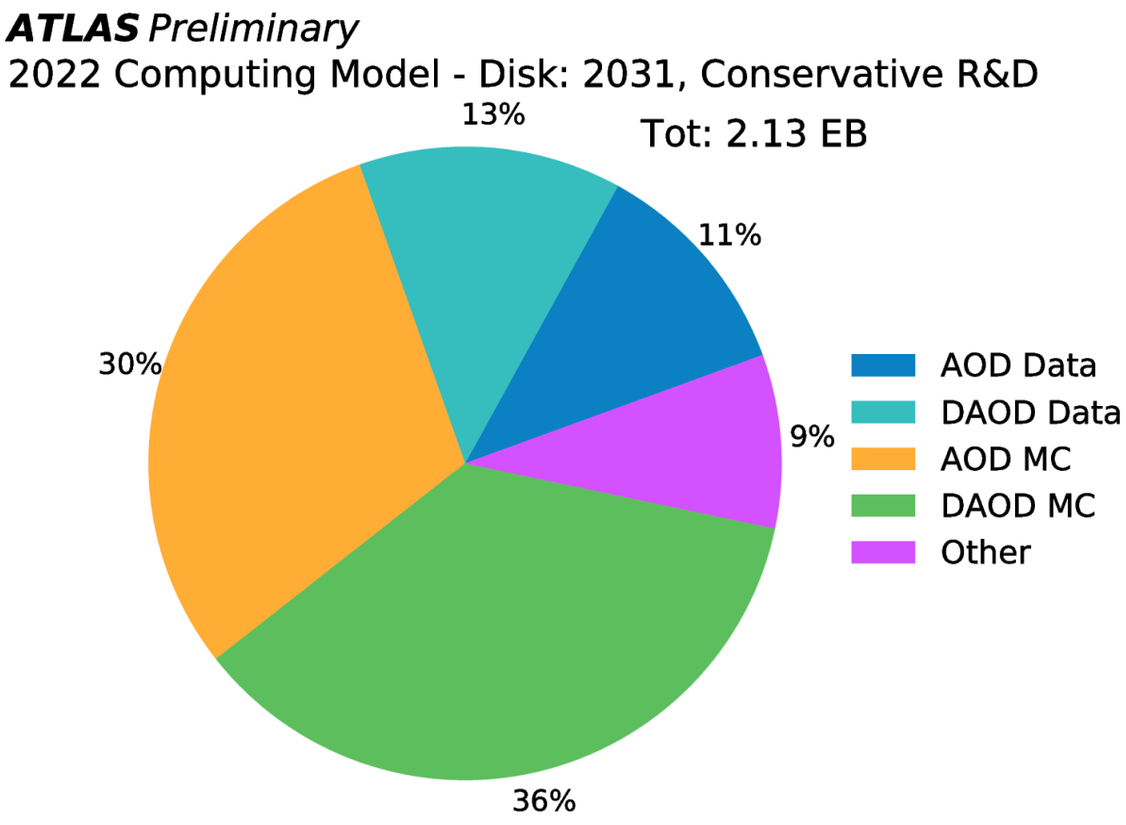
- ATLAS is trying to take on accelerators for our HL-LHC software-trigger tracking (aka “EF tracking”) – most focus on FPGAs
- The big decision point for accelerators comes in ~2025/6
  - Likely both online and offline will decide within a year of each other
  - May need to commit earlier to a limited set of technologies (e.g. decide in 2024 “CPU vs CUDA on GPUs”, then go/no-go in 2026)
- Lots of work meanwhile to prepare our infrastructure
  - Memory management (VecMem)
  - Evaluation of a more accelerator-friendly data model
  - Improvements of scheduling (including multi-node scheduling)
- We have some beautiful examples already for the infrastructure
  - FastCaloSim (thanks!) and calo energy clustering on GPUs
  - We look forward to playing with an ACTS/traccc demonstrator soon!

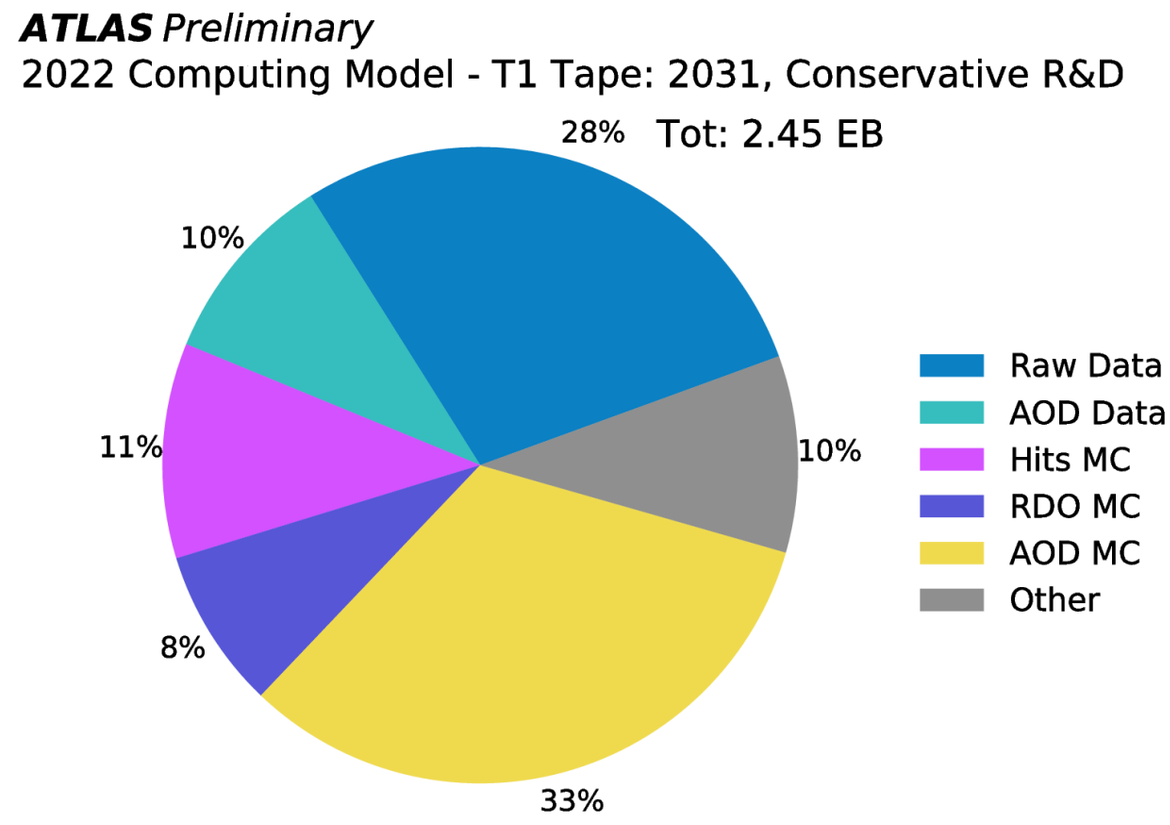
ATLAS's (Very) Big Disk Problem

- Our (real, RAW and downstream) dataset size in computing terms is *not* strongly correlated with our "physics" dataset size (in fb^{-1})
- Instead, it's a combination of three factors:
  - Event rate – going up from ~**3.3 kHz** today to ~10 kHz in HL-LHC.
  - Event size – RAW going up from 1.1 MB in Run 2 to 1.4 MB in Run 3 to 4.3 MB in Run 5.
  - Running seconds – constrained by the machine. About $6x10^6$/year.
- Our MC Simulation has similar factors but **HUGE** uncertainties.
  - Event rate – depends hugely on the outcome of the next 8 years of R&D, in event generation and simulation especially!
  - Event size – comparatively well under control.
  - Volume required – ~3x that of data, but depends on how many generators we use, versions, reproductions…

ATLAS's (Very) Big Disk Problem

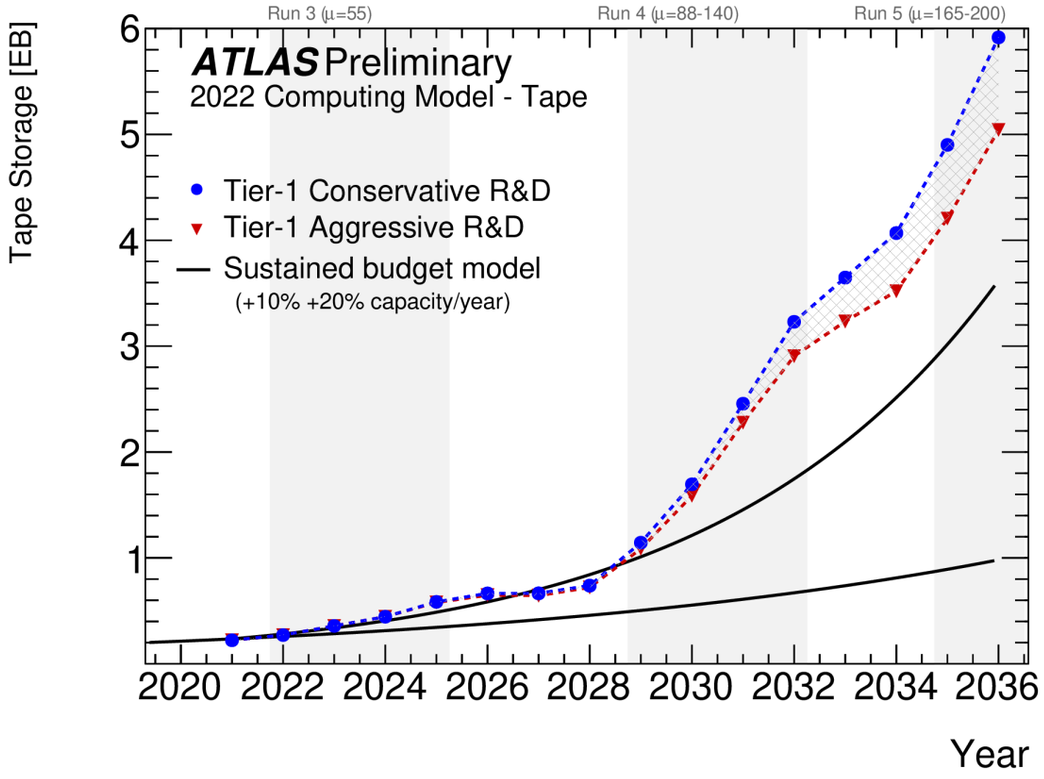
- Our storage isn't dominated by RAW data but by MC and downstream (analysis) data products
- These projections vary hugely depending on what analysis model you assume in the HL-LHC
- "Common" R&D (beyond RNTuple) is hard, but we are working…





Other Opportunities

- Some R&D items not included in “conservative” or “aggressive”
  - FastChain – a fast end-to-end simulation. If successful, could choose to save less intermediate data, which (if widely used) would affect both our disk and our tape projections.
  - Disruptive R&D, like an ML model to produce (simulate) all pileup or a full GAN fast simulation
  - CPU / disk trade-offs by (risking) re-running samples
- Probably still space for Geant4 optimization towards Run 4
- At least one space not yet mentioned: analysis
  - Many analysis resources are off-grid (dark) CPU/disk
  - It can be very hard to predict what users will want in 10 years
  - But let’s talk about it anyway…

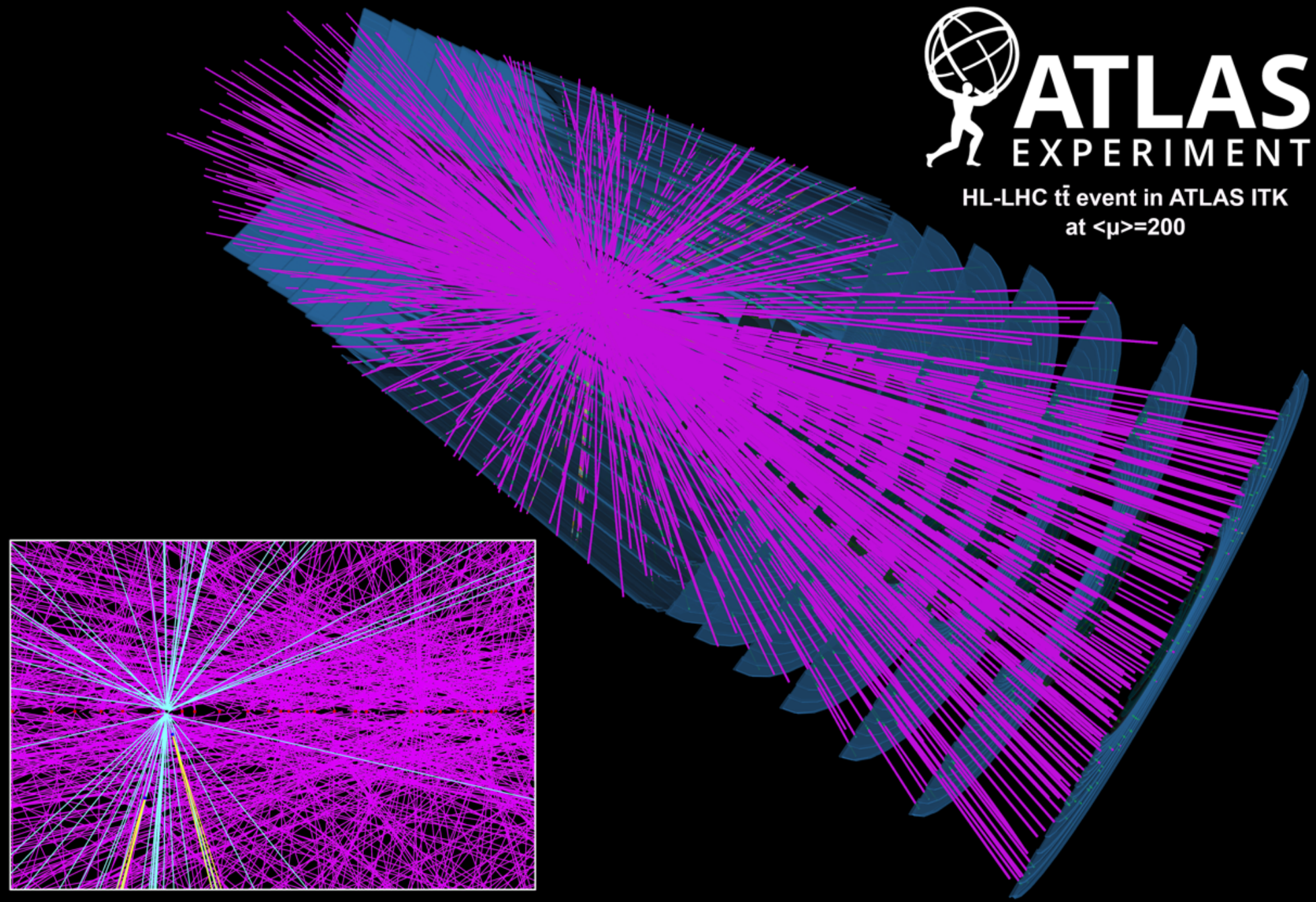


Analysis (& Facilities)

- Our analysis model involves using a lighter format in Run 4
  - What analyzers do with that format is still an item of hot debate – not clear if “columnar analysis” is going to dominate in ATLAS
    - This is not to say that *some people* won’t use it!
  - This includes some *very* important issues (e.g. systematics) to which we are still discussing new approaches
- Analysis facilities are extremely hard to define, but we have some facilities that look like analysis facilities today
  - Universally accessible, offering interactive and batch resources at scale, Jupyter support, GPUs: *this is lxplus at CERN!*
  - We have to be wary of scaling up resources that aren’t a match to requirements (SWAN?).
  - We hope that the community can **work together** to solve the real and interesting problems of scaling and access!

Summary

- ATLAS is facing interesting, difficult, but solvable software and computing challenges for the HL-LHC
  - One of the biggest challenges not mentioned here is supporting and retaining skilled developers – your help is always welcome!
- We appreciate the collaboration with HEP-CCE thus far, and look forward to continuing to work together
  - Keeping close collaboration is critical – high risk of trying to solve a non-problem, and we don't have the effort to lose
- One of the most important items is **validation**
  - We can't integrate R&D dropped off the day before the run starts; even purely technical changes sometimes get tripped up
  - Your help in validation – early and frequent iteration with your ATLAS correspondents – is paramount!



Thank you!

Extras

AI/ML

- Lots of R&D ideas around ML approaches to specific problems
  - ML for “jet tagging” (what is the origin of this spray of particles) is extremely successful
  - ML for tracking is a popular R&D topic
- I believe it’s unlikely that ML approaches will replace parts of our reconstruction en masse, but I think it’s very likely ML approaches will replace specific pieces of our reconstruction
  - ML for energy calibration? Yes! ML for jet reconstruction? Eh.
- ATLAS has demonstrated itself to be more resistant to ML approaches in analysis than other collaborations

Another

A.I. they’ve developed, called Delphi, takes an ethical approach. Delphi has analyzed ethical judgments made by crowdworkers, and has learned to say which of two actions is more morally acceptable; it comes to commonsense conclusions seventy-eight per cent of the time. Killing a bear? Wrong. Killing a bear to save your child? O.K. Detonating a nuclear bomb to save your child? Wrong. A stabbing “with” a cheeseburger, Delphi has said, is morally preferable to a stabbing “over” a cheeseburger.