

# Tackling the Computing Challenges @ CERN

CERN OpenLab Summer Student lecture

2th July 2024

Alessandro Di Girolamo - CERN IT



# ... a brief “random walk” through CERN

# What does **CERN** stand for?

**C**onseil

**E**uropéen pour la

**R**echerche

**N**ucléaire

1954



# Nuclear?



# CERN

## Who is it?

- The Council
  - Member states
- The Collaborations
  - experiments

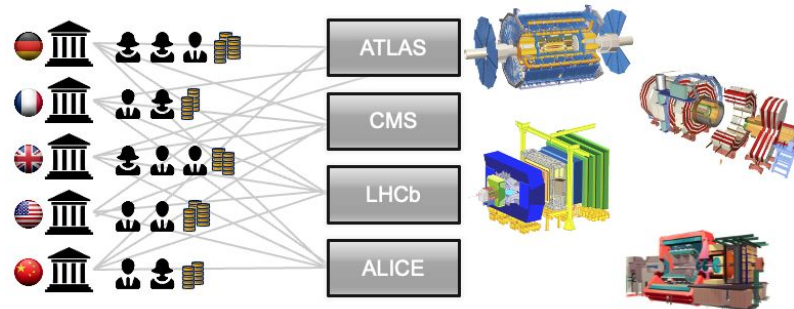
## Member states



~1 B CHF



## Collaborations

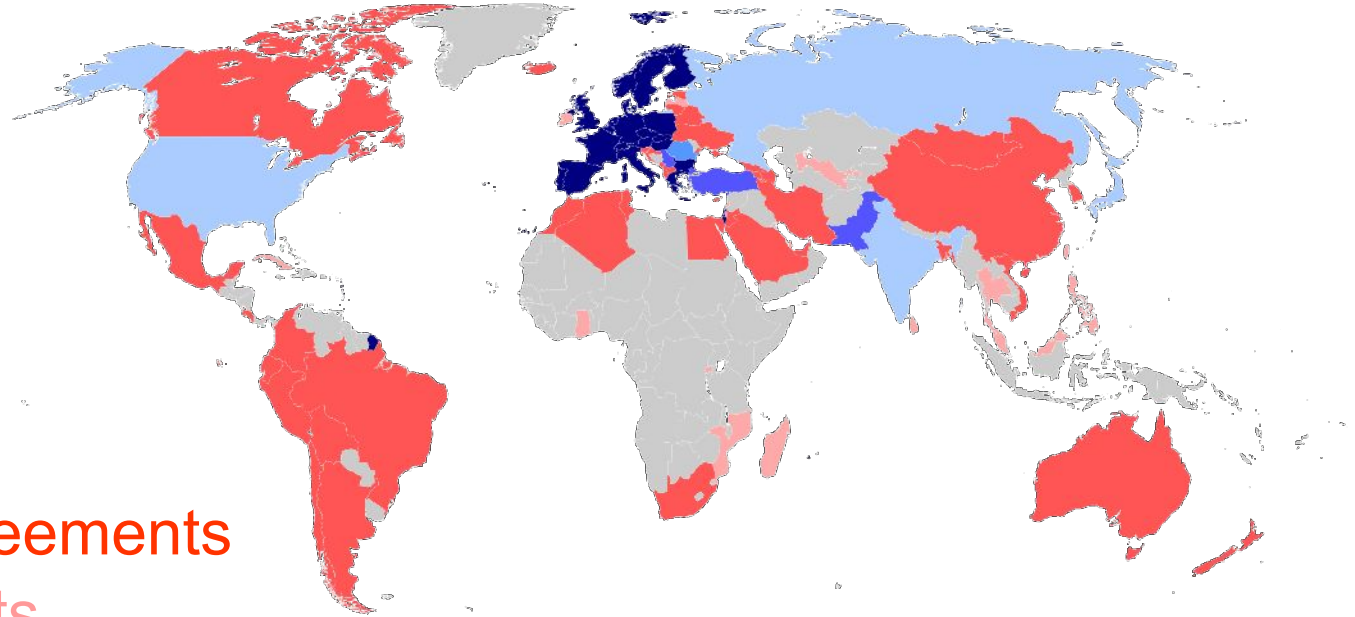


# **CERN: a worldwide collaboration**

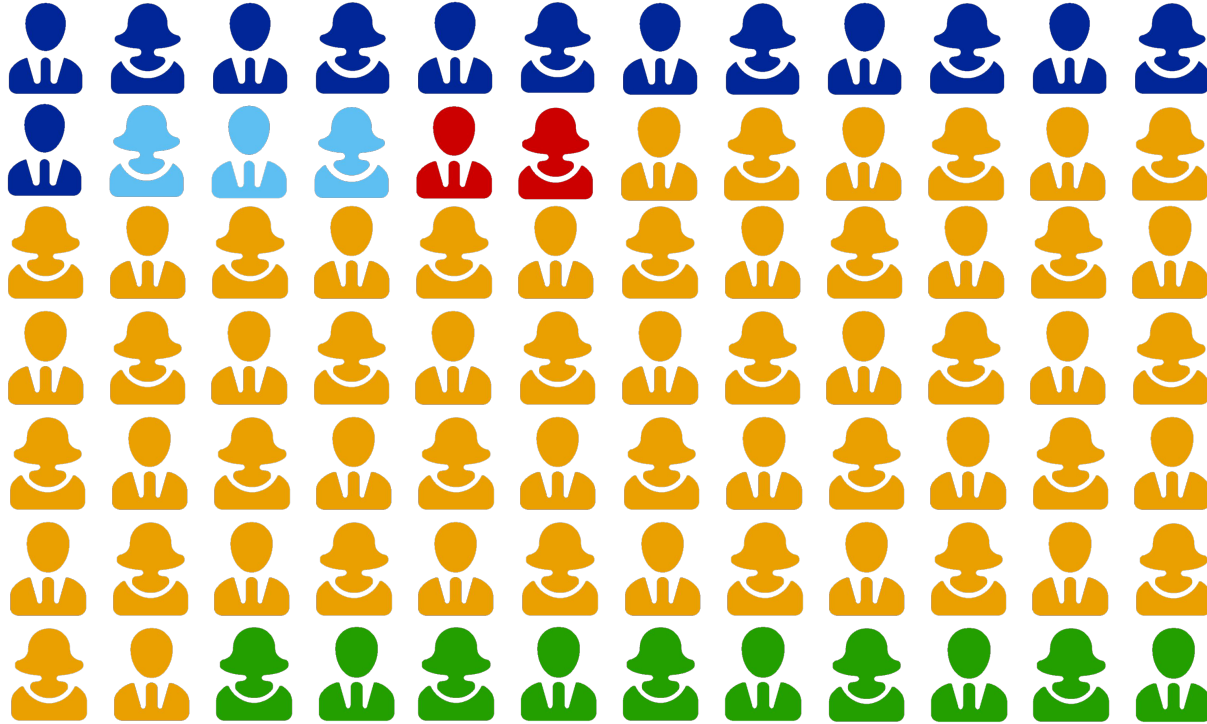
23 members  
+ Associates

+ Observers

Cooperation agreements  
Scientific contacts



# How many persons? **15000+!**



2500 staff

600 fellows & apprentices

500 students

10000+ users

2000 external companies

# Our Mission

## The CERN Mission

- perform world-class research in fundamental physics.
- provide a unique range of particle accelerator facilities that enable research at the forefront of human knowledge, in an environmentally responsible and sustainable way.
- unite people from all over the world to push the frontiers of science and technology, for the benefit of all.
- train new generations of physicists, engineers and technicians, and engage all citizens in research and in the values of science



# CERN

## *How does it work ?*

# Accelerating and colliding



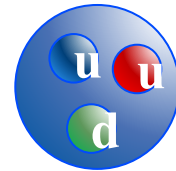
# Accelerator chain



# Incredible levels of energy

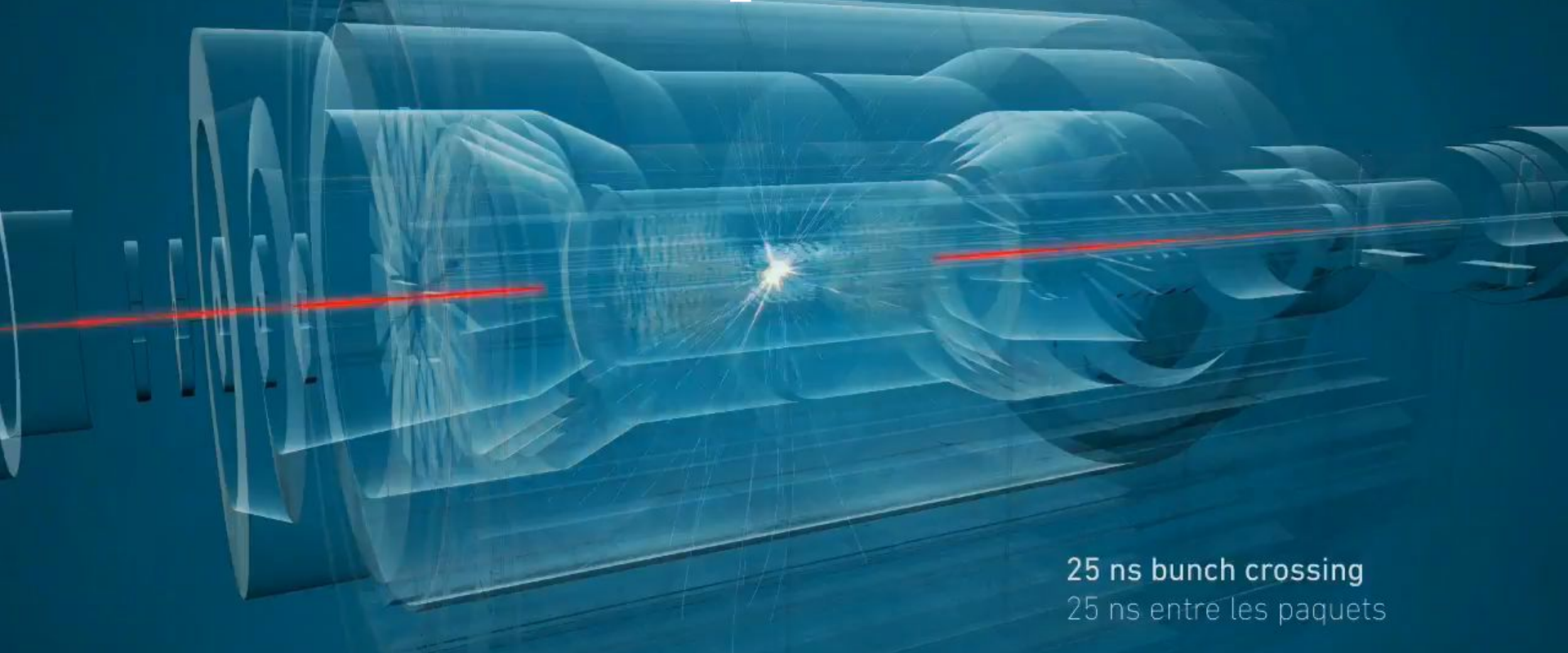
100'000'000'000'000'000'000'000

7+7 TeV





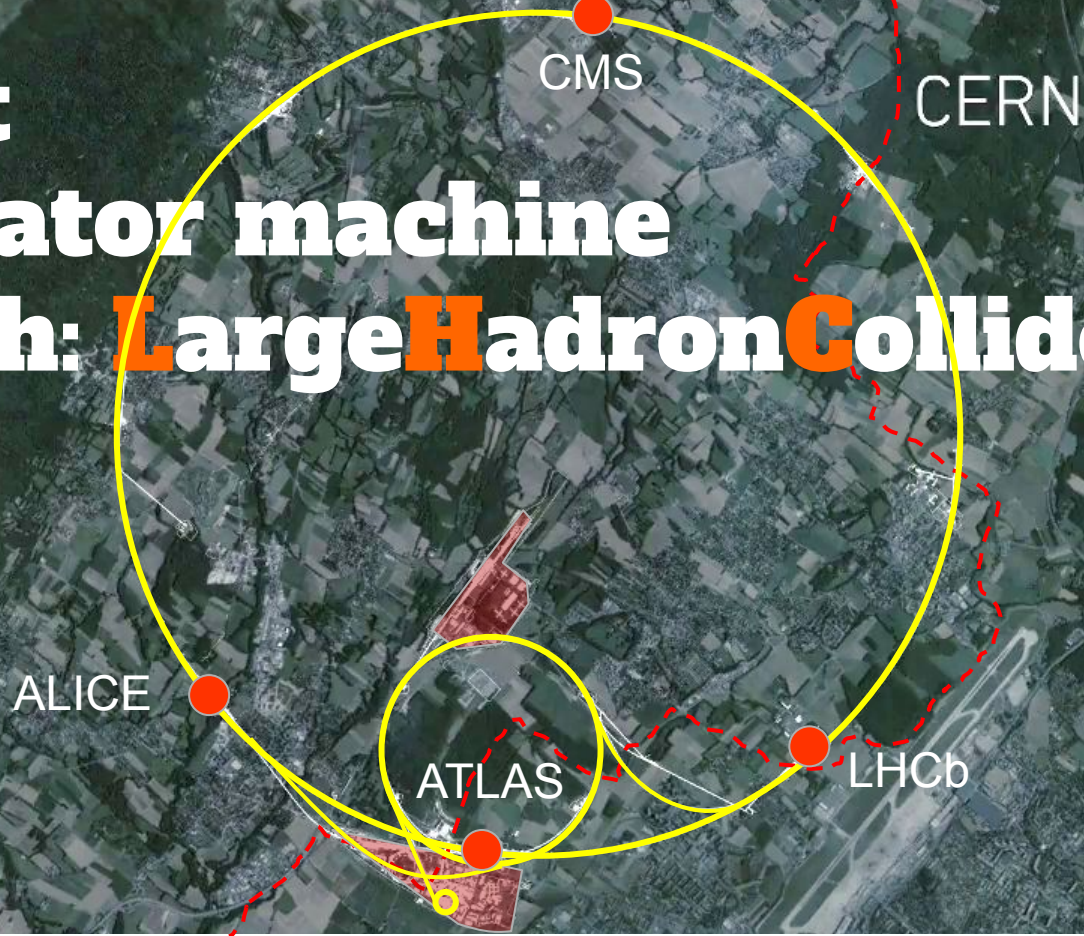
# Millions of collisions per second



25 ns bunch crossing  
25 ns entre les paquets



# Largest accelerator machine on earth: **L**arge **H**adron **C**ollider



# ATLAS and CMS

- Two general-purpose detectors cross-confirm discoveries, such as the Higgs boson.



46m long, 25m diameter  
weights 7'000 tonnes  
100 million electronic channels , 3 000 km of cables



22m long, 15m diameter  
weights 14'000 tonnes  
Most powerful superconducting solenoid ever built



# ALICE and LHCb

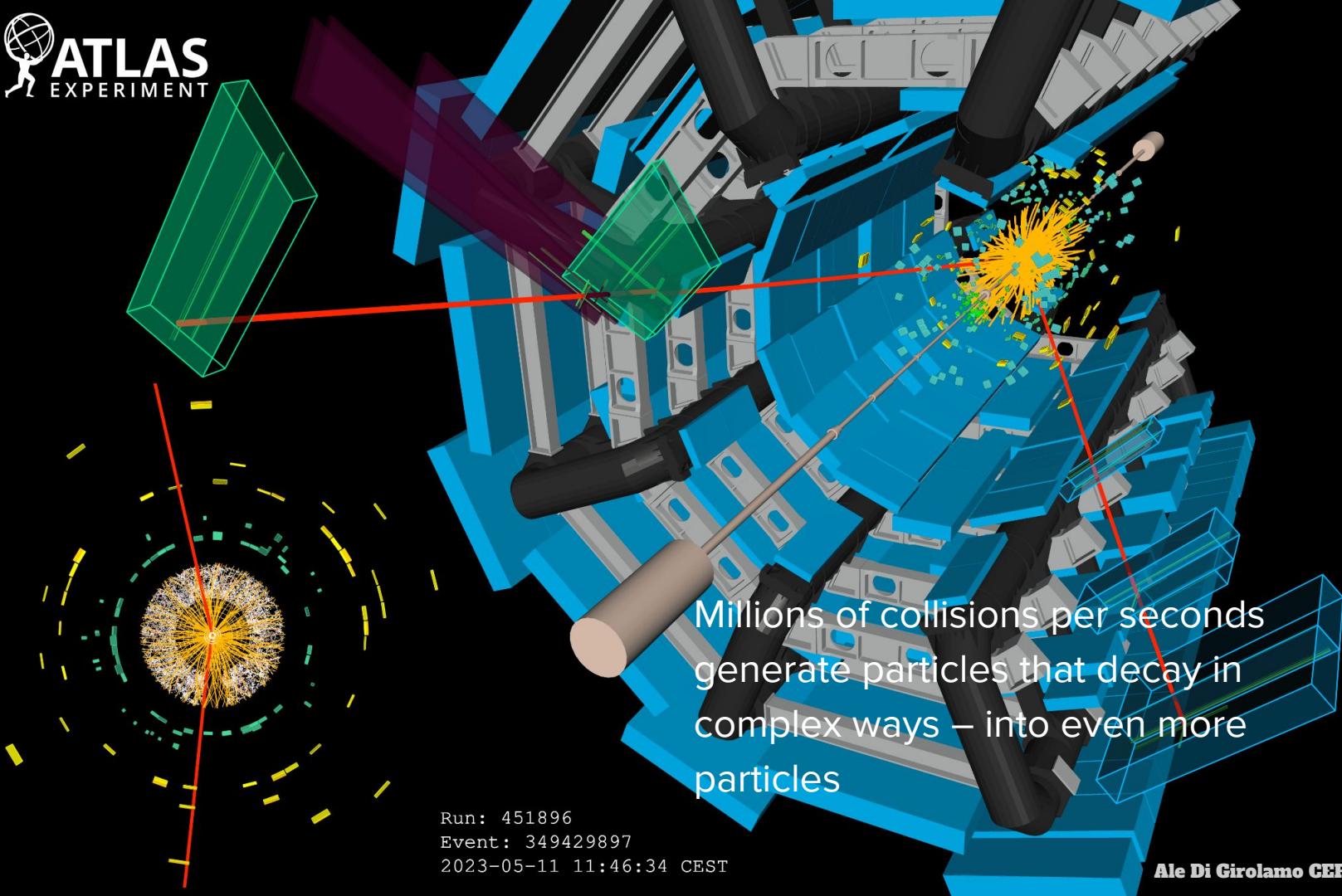
- ALICE and LHCb experiments have detectors specialised on studying specific phenomena.



Studies the «Quark Gluon Plasma», state of matter which existed moments after the Big Bang.



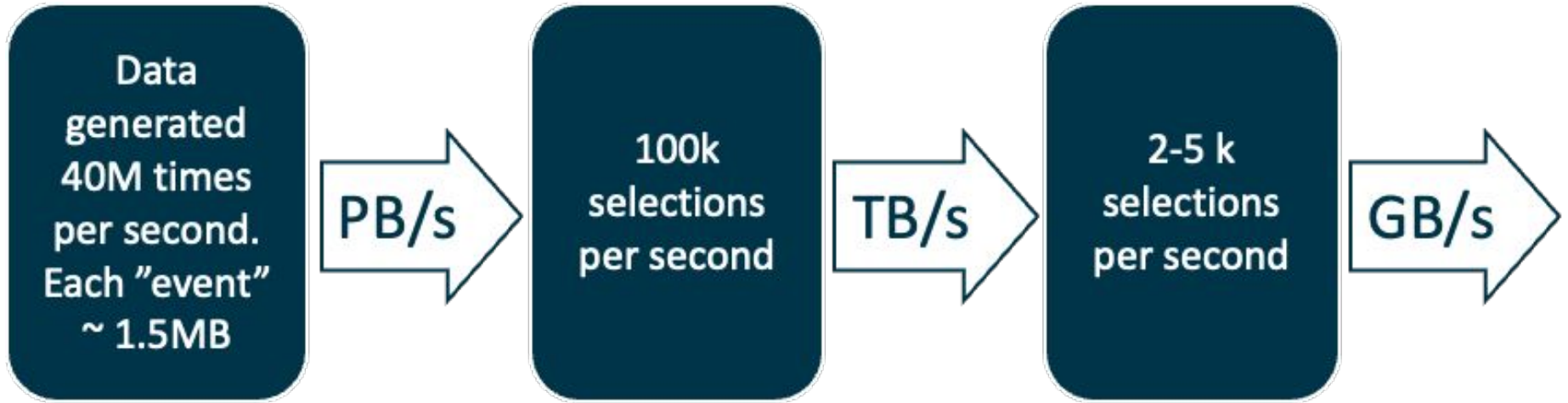
Studies the behaviour difference between the b quark and the anti-b quark to explain the matter-antimatter asymmetry in the Universe.



Millions of collisions per seconds  
generate particles that decay in  
complex ways – into even more  
particles

Run: 451896  
Event: 349429897  
2023-05-11 11:46:34 CEST

# Data filtering (Trigger)

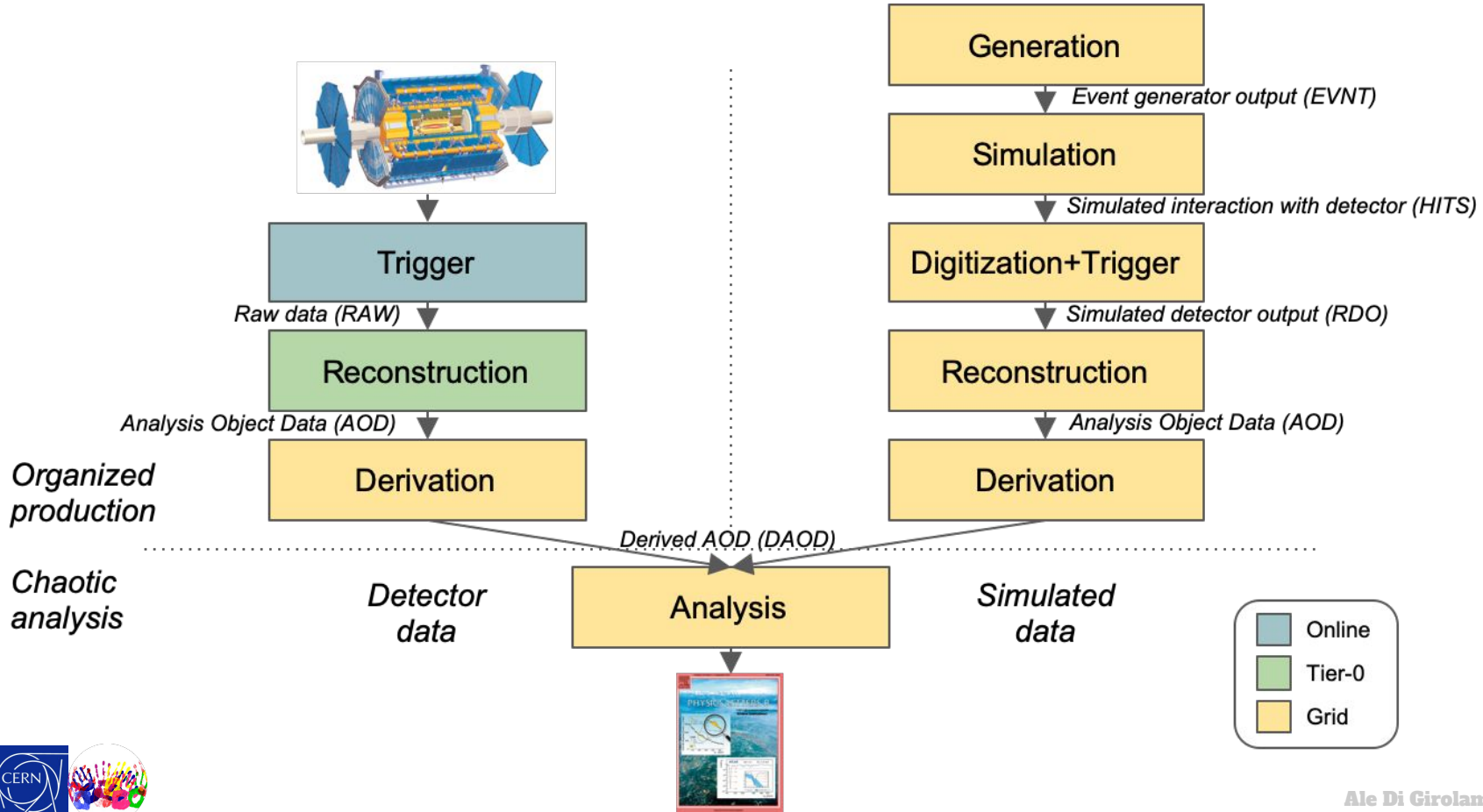


PetaBytes of data per second:

The "trigger" is filtering data in real time, selecting potentially interesting events

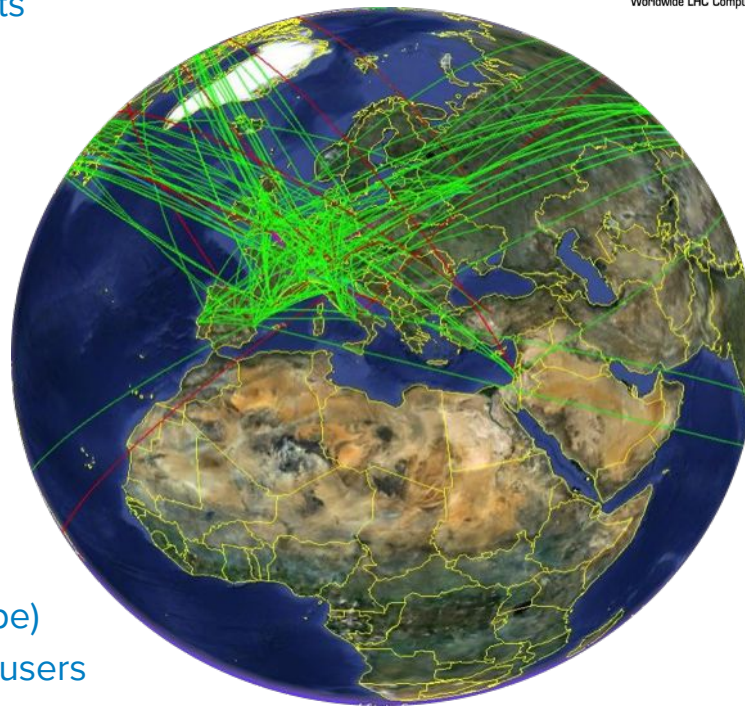


# From Data to Paper(s): data processing chain



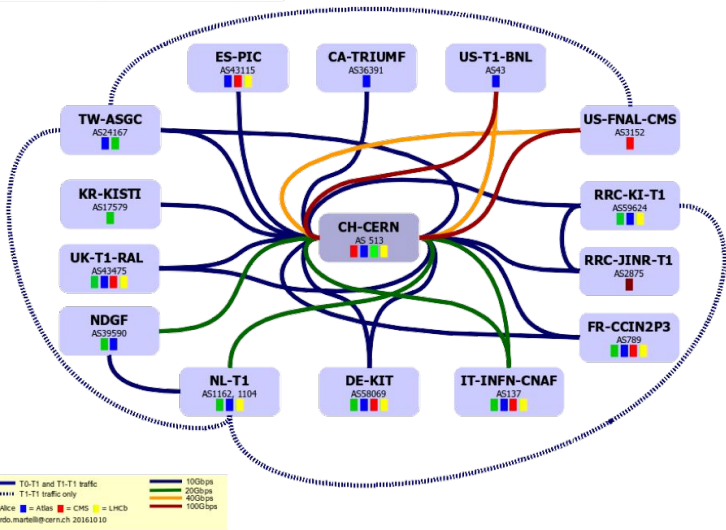
# The Worldwide LHC Computing Grid - WLCG

- International collaboration to distribute and analyse LHC data
- Integrates computing centres worldwide that provide **computing** and **storage** resource into a single infrastructure accessible by all LHC physicists
- Global **network** connectivity
  - **Tier-0 (at CERN):** Large resource, recording and custodial archival of collision data, prompt reconstruction
  - **Tier-1s:** Memory and CPU intensive tasks second tape copy of detector data, occasional Tier-0 overspilling
  - **Tier-2s:** Processing centres, nowadays many are similar to the Tier-1s
- Around 1.5 million CPU cores fully occupied 24/7
- Around 1.5 EB data (~600 PB on disk and >800 PB on tape)
- More than 100 PB moved every month, accessed by 10k users



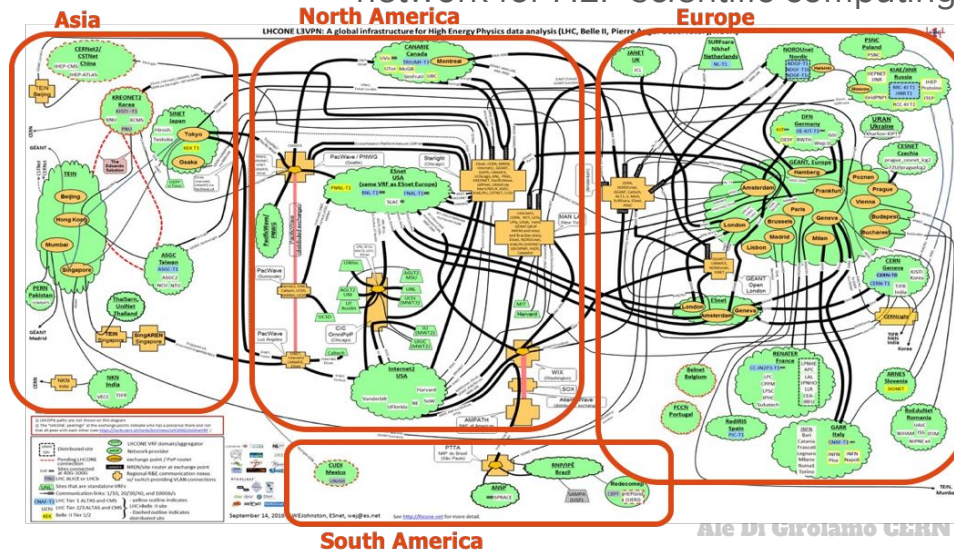
# Worldwide Networking

## LHCOPN



WLCG relies on powerful network infrastructure which connects the whole Grid

- LHCOne: overlay network
  - Enable NREN to manage general public network for HEP scientific computing

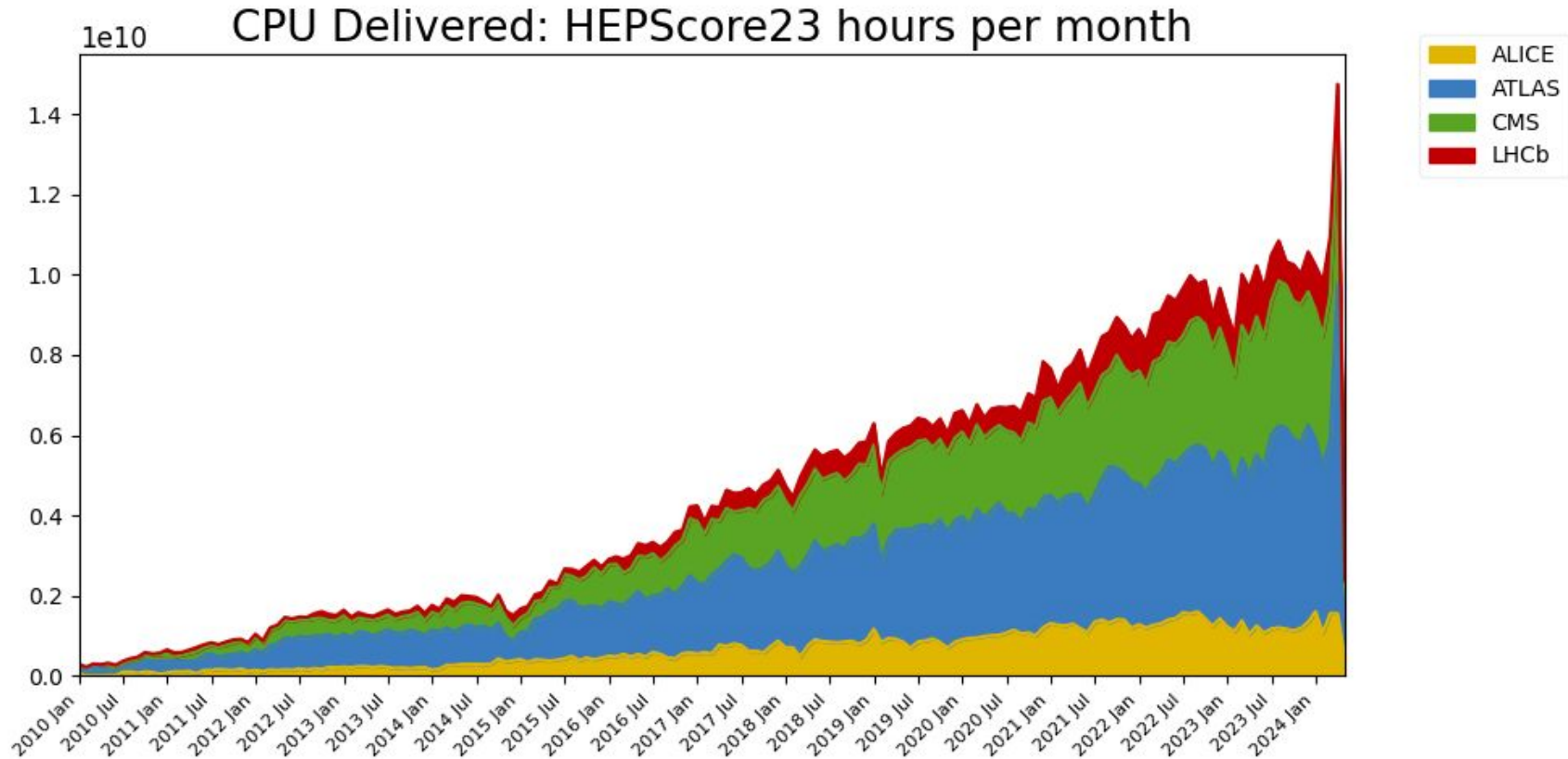


- LHCOPN: LHC Optical Private Network

- For Tier0 ↔ Tier1s traffic,
- managed by Tier0 and Tier1s



# WLCG resources: compute

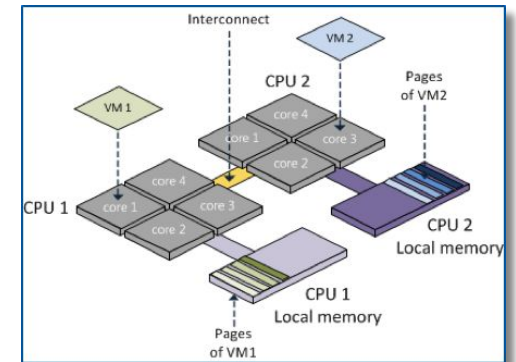
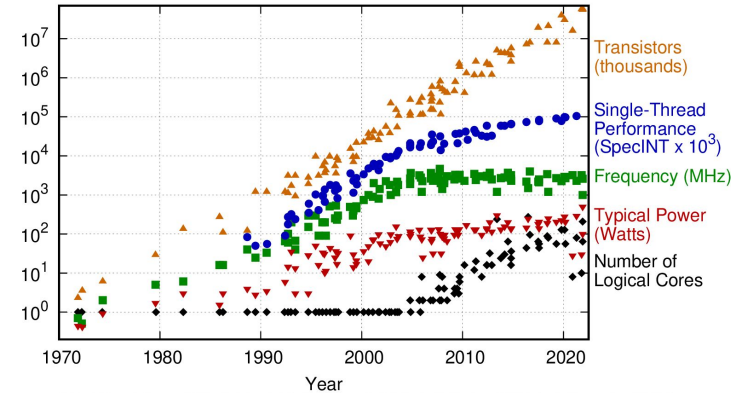


# Hardware evolution

- A challenge not only for the HEP computing:
  - Transistors go into many cores, co-processors, vector units...
  - Clock speed stalled since ~2005
    - Single core performance is essentially also stalled
  - Mobile devices and data centres are the key volume markets
  - Memory consumption is a huge driver now
  - Memory access and I/O paths also become problematic
- Modernizing code will play a vital role for the upgrades of experiments and LHC
- Increase software performance by adopting modern coding techniques and tools
- Fully exploit the features offered by modern HW architectures
  - Many-cores GPU platforms, acceleration co-processors and innovative hybrid combinations of CPUs and FPGAs
- Getting performant software requires significant investment in programming skills

Also, it does not happen overnight...

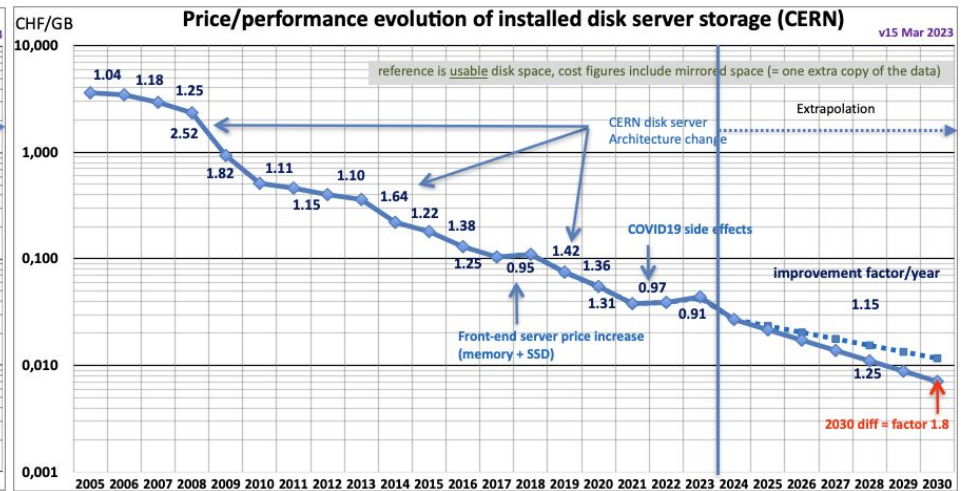
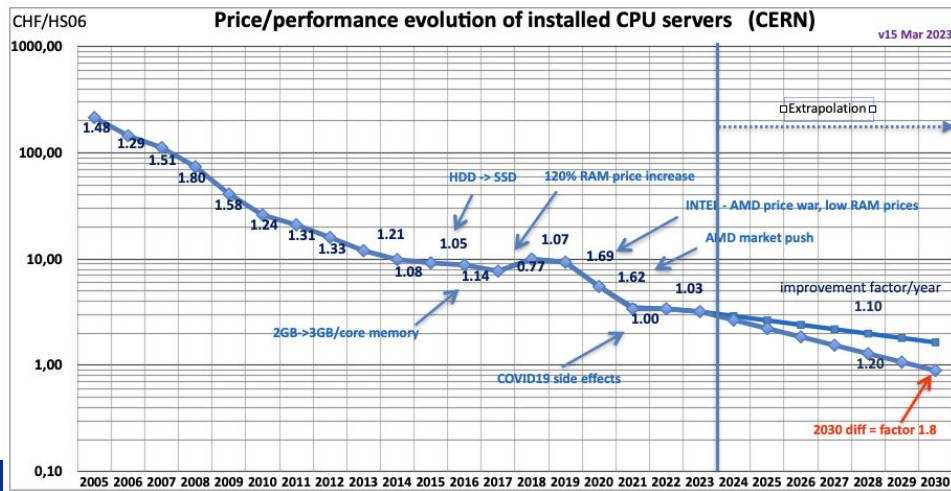
50 Years of Microprocessor Trend Data





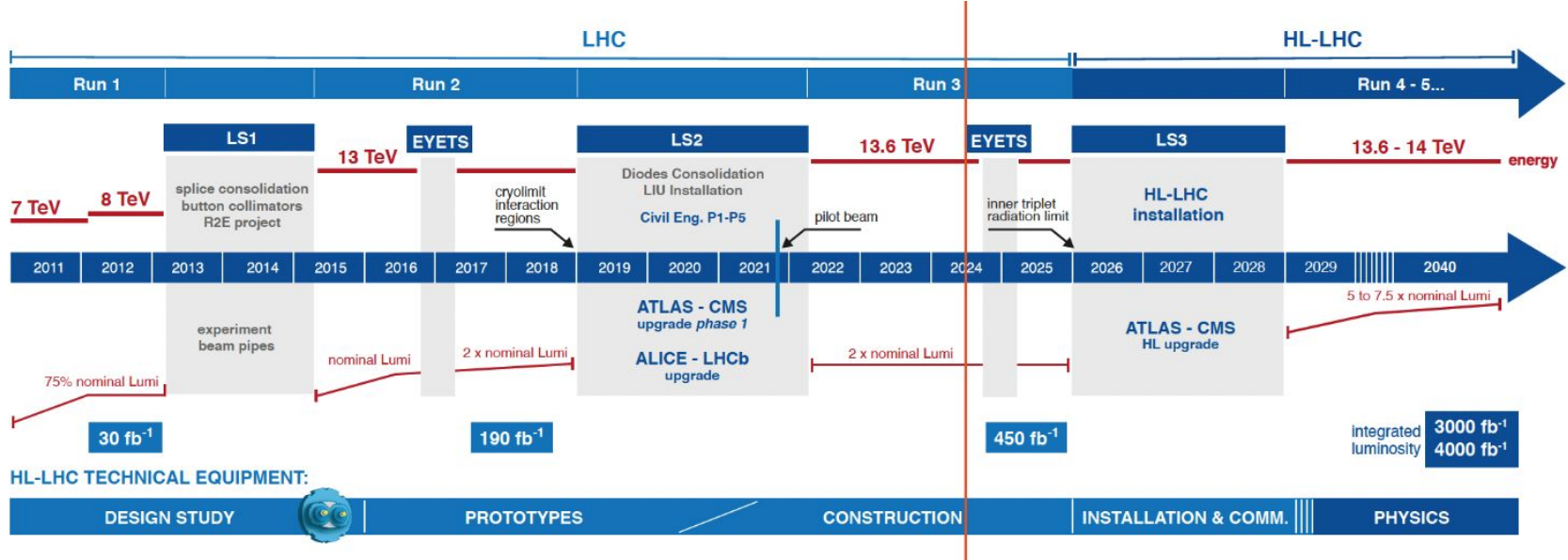
# Hardware cost evolution

- Hardware cost is more and more dominated by market trends rather than technology
  - science has no influence on these markets!
- For the budget we have been assuming +20%/year in storage and CPU capacity, but we have to deal with large fluctuations and a “flat” budget constraint



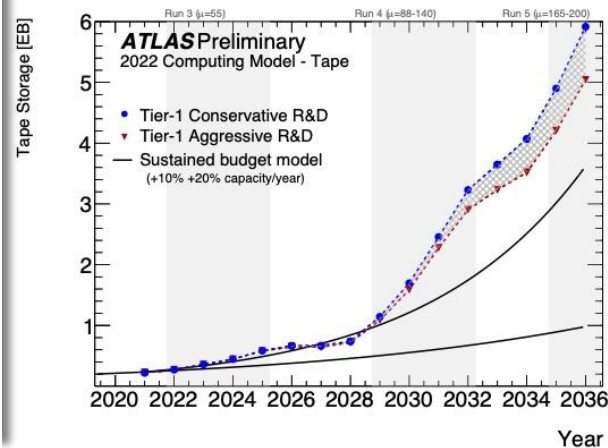
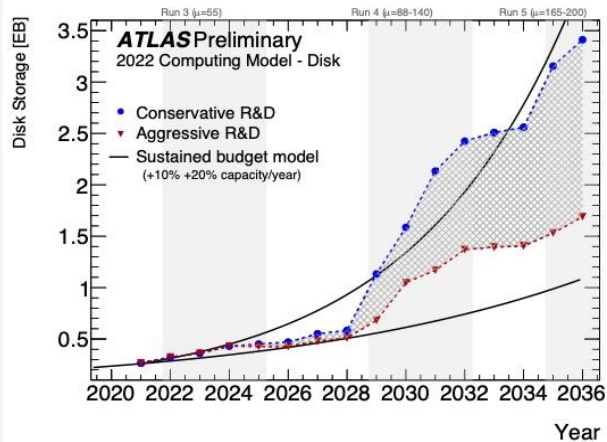
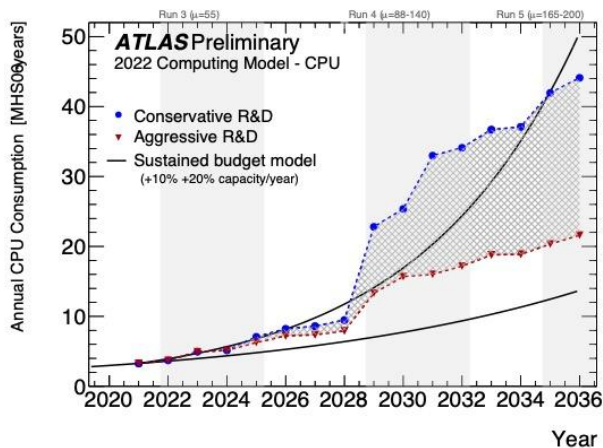
# HL-LHC plans

- Expected to be operational from 2029
  - Objective is to increase the integrated luminosity by a factor of 10
- ATLAS and CMS will face several challenges wrt Run3 (now):  $\sim 3x$  instantaneous luminosity, 3-4x pileup ( $\mu \sim 60 \rightarrow 200$ )



# HL-LHC Software and Computing Challenges

- Computing Resource Estimates show need for major R&D (or budgetary) effort to achieve HL-LHC physics potential.
  - We have defined *Conservative R&D* and *Aggressive R&D* scenarios
    - Conservative should be achieved with today's effort and people
    - Aggressive requires more people, more R&D projects bringing resource savings
  - The black lines indicate the “flat budget” of 10% (lower line) and 20% (upper line).



**Investing in brainpower is paramount!**

# Tackling the HL-LHC Computing Challenges

- Compute..
  - Optimize the usage of all the available resources, including new HW technologies and facilities
  - Make physics choices that save computation without sacrificing the physics reach
- Storage
  - Optimize data format
  - Evolve the computing models to make use of different (cheaper) storage technologies
- Sustainability
  - Environmental sustainability (Use - not waste!)
  - New technologies!
  - Boomers, GenX/Y/Z: we are all different, different times different challenges!





# Time to act: NOW

Experiments defined their roadmaps toward HL-LHC:

[ATLAS S&C HL-LHC Roadmap](#) ; [CMS Phase 2 Computing Model](#)

- Based on the HL-LHC schedule, it is clear that R&D projects need to start now
  - from R&D to demonstrators, with measurable impact
  - We see already lot of engagement!
- Integration and validation will require lot of time
  - Late arriving R&D is risky



# Some caveats: lots of “business as usual”

We have to keep on running the experiments while we are planning for major upgrades

- “we are building a new ATLAS while we are running ATLAS” (Karl Jakobs, ex ATLAS spokesperson)
- Failing is not an option

Lot of efforts that need to go into non-R&D work (or at the boundaries)

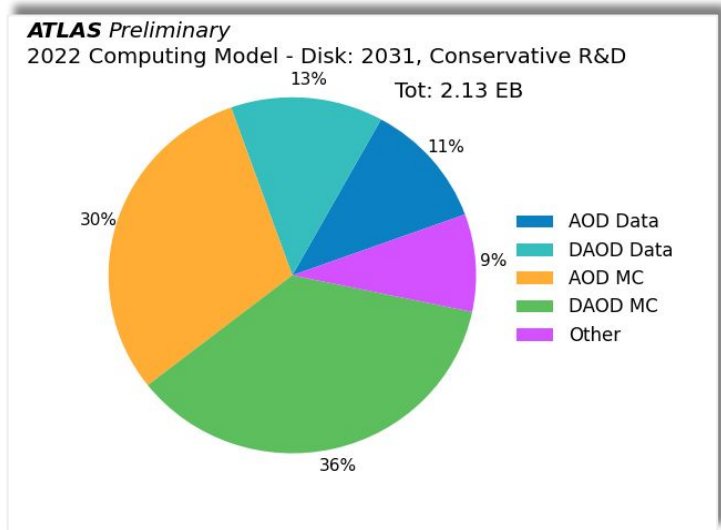
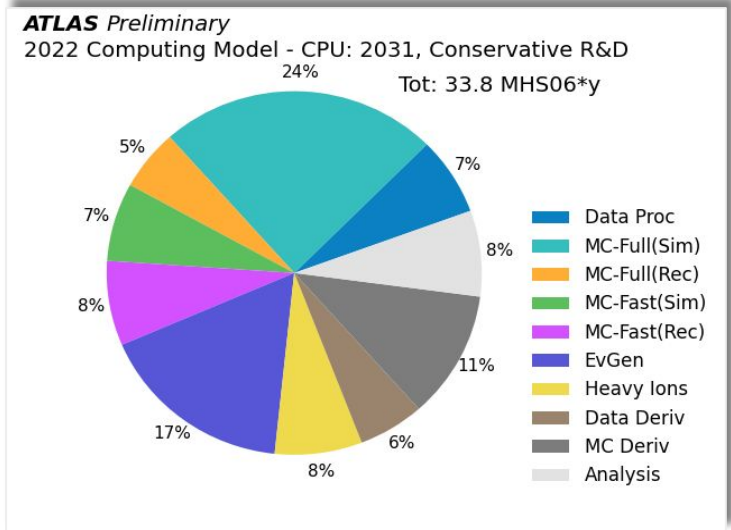
- Maintaining our current software
- Updates to the distributed computing infrastructure
- Upgrade geometry, digitization, re-tuning simulation...
- SW performance improvements

R&D projects are on top of this:

- balance (between R&D and “business as usual”) is key
- **We need a strong focus on “impact”**



# HL-LHC - Resource Estimates for 2031



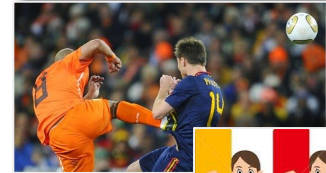
- There is no “silver bullet”:
  - it’s not that we do have 1 (one) single problem and if we solve that “we are done”.
- we need to work on each and every part of our system
- The opportunities are “everywhere”: we need to work hard to scavenge a few % here and there
  - An “evergreen” motivational speech, [Any Given Sunday](#) : true we do not play \*against\* anyone, but we still want to bring out the best from ourselves - while having fun!



# Tackling the challenges



- There is nothing bad in being “two” in tackling difficult challenges!
  - Important we work together and we organize ourselves
  - We don’t compete
- There are no shortcuts and not so many “low hanging fruits”
  - It requires disciplines from all of us
  - Others might suffer from our “being late”



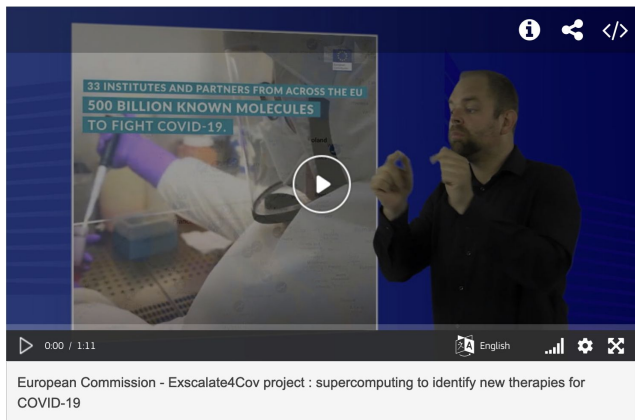
- Expectations management is paramount
  - We do not dictate! We do bring people onboard, we (try to) motivate them (and people motivate us!)
  - Crucial to clearly define our commitments taking into account all the “other” stuff we have to do
  - It is critical that we are able to estimate *by when* “that something” will be done.



# Tackling the challenges: infrastructure evolution

- HPC and Clouds (and Industrial partners)
  - More and more often we are seeing them discussed together
  - Not really much overlap, except that they “both” are not standard Grid sites.
- They really deserve two different “set of slides” → next

High performance computing refers to computing systems with extremely high computational power that are able to solve hugely complex and demanding problems.

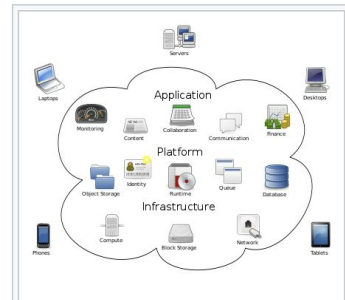


In the digital decade, high performance computing (HPC) is at the core of major advances and innovation, and a strategic resource for Europe's future.

**Cloud computing**<sup>[1]</sup> is the on-demand availability of **computer system resources**, especially data storage (**cloud storage**) and **computing power**, without direct active management by the user.<sup>[2]</sup> Large clouds often have functions **distributed** over multiple locations, each of which is a **data center**. Cloud computing relies on sharing of resources to achieve coherence and typically uses a "pay as you go" model, which can help in reducing **capital expenses** but may also lead to unexpected **operating expenses** for users.<sup>[3]</sup>

## Contents [hide]

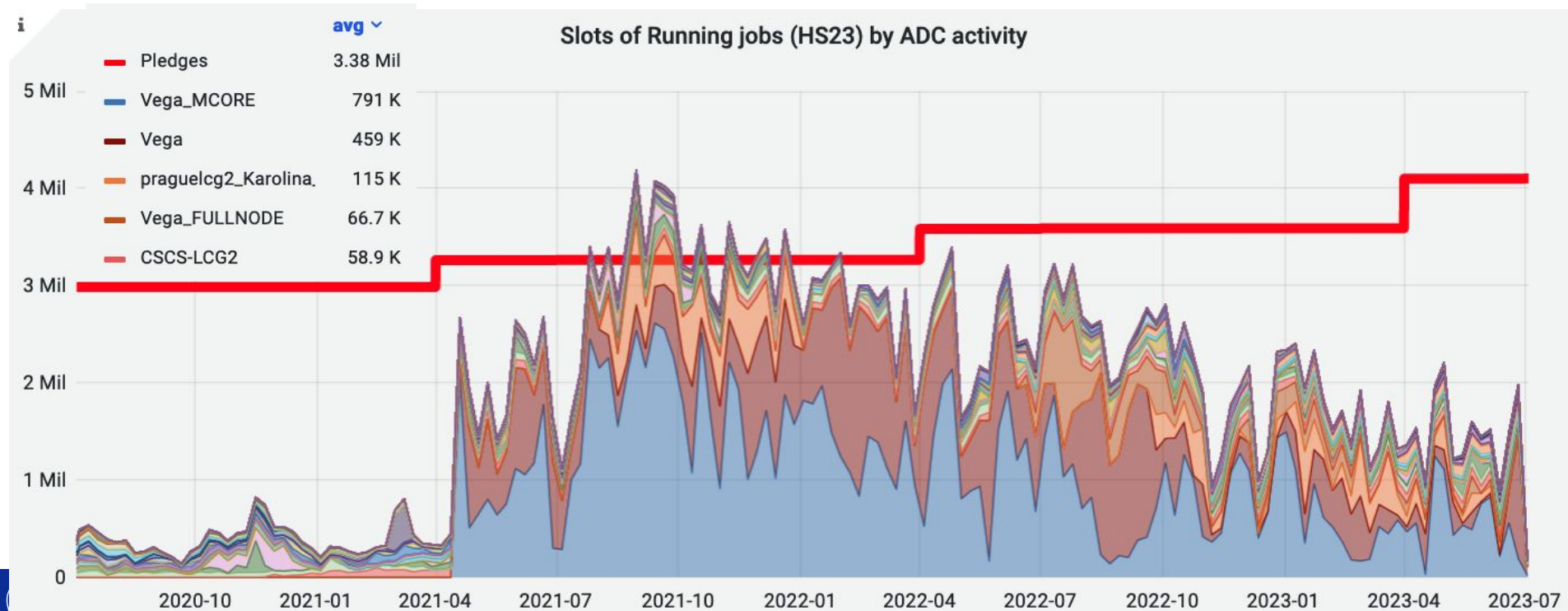
- 1 Value proposition
- 2 Market
- 3 History
  - 3.1 Early history
  - 3.2 2000s
  - 3.3 2010s



Cloud computing metaphor: the group of networked elements providing services need not be individually addressed or managed by users; instead, the entire provider-managed suite of hardware and software can be thought of as an amorphous cloud.

# HPC and ATLAS - last 3 years - HPC only

- A few [HPCs](#) make a big difference
  - But still it's important to highlight that many HPCs are integrated!
  - Looking forward to Leonardo(n4) and Perlmutter(n8)



# A strategy towards opportunistic HPC

- 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/with internal good support ones gives ample CPU.
  - All the “easy” HPCs transparently integrated, running all workflows, are much welcome!
- 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 big HPCs is ongoing
  - E.g. [Perlmutter](#), [Fugaku\(2\)](#)/[CSCS](#), [Toubkal](#) ...
- The details can make the difference:
  - HPC are not easy, need expertise on I/O, shared file systems, scheduling, AAI...
- Without forgetting:
  - The risk that the opportunistic CPU resources will disappear at any time is very real
  - The main physics program of the experiment should be supported through pledged resources
  - A constant yearly increment for both CPU and Disk is reasonable to minimize the risks that, if these resources disappear, the computing centers providing pledged resources won't be able to provide what is needed

# R&D Projects with Industry

- The experiments and the labs are engaging with industrial partners on R&D projects, e.g.:
  - [SEAL](#), a decentralized cloud storage start-up
  - [Google](#) Cloud and [Amazon](#) Web Services
- Opportunities to evaluate technologies and to define a resilient long term strategy
  - Use of cloud compute for *all* workflows, and a comparison to traditional resources of the total cost of ownership (TCO)
    - Google already used for interactive analysis
  - Evaluation of test technologies (e.g. TPU, ARM) using cloud resources without requiring large-scale purchases
    - AWS extensively used in setup and validation of ARM Athena nightlies builds
  - Evaluation of remote cold storage as a complement / supplement to tape systems
- These projects also present great connections with strong partners
  - Amazon and Google are major players that many groups already collaborate with
  - Being able to speak the same language and learn how they operate is invaluable



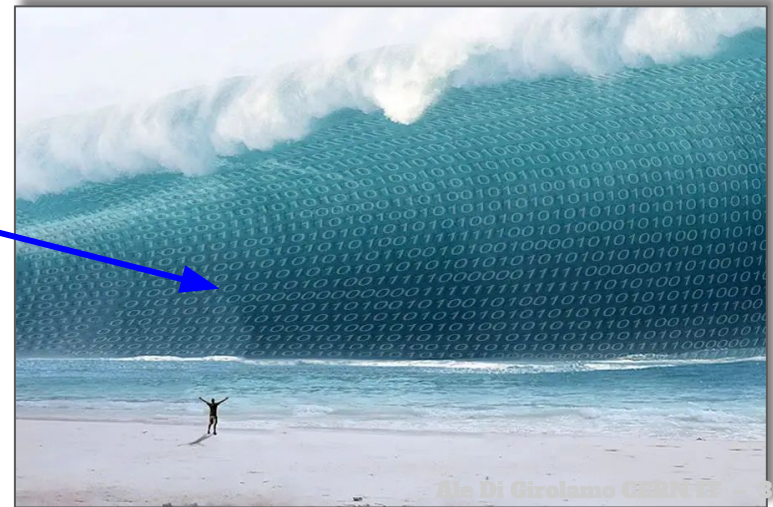
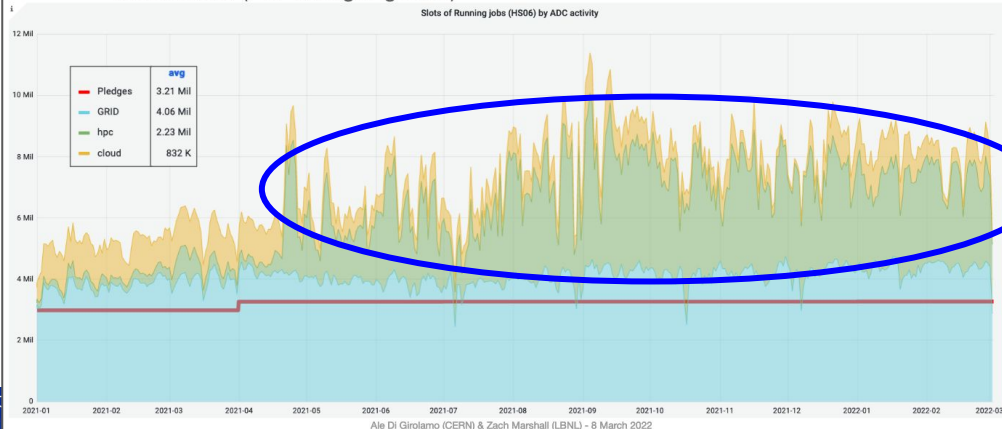
# Opportunistic CPU but NO opportunistic Disk

- Resources extremely useful for the physics community
  - We have developed a detailed programme for secondary MC simulation samples
  - E.g.: large number of diphoton background events simulated → significant reduction of spurious signal systematic uncertainty → enabled a low diphoton-mass analysis; V+jets w both new Sherpa and MadGraph → primary setup for upcoming V+jets modelling pub, used in Z+jets measurements paper
- Still, they produce a huge challenge for our storage: we need to store all these data!

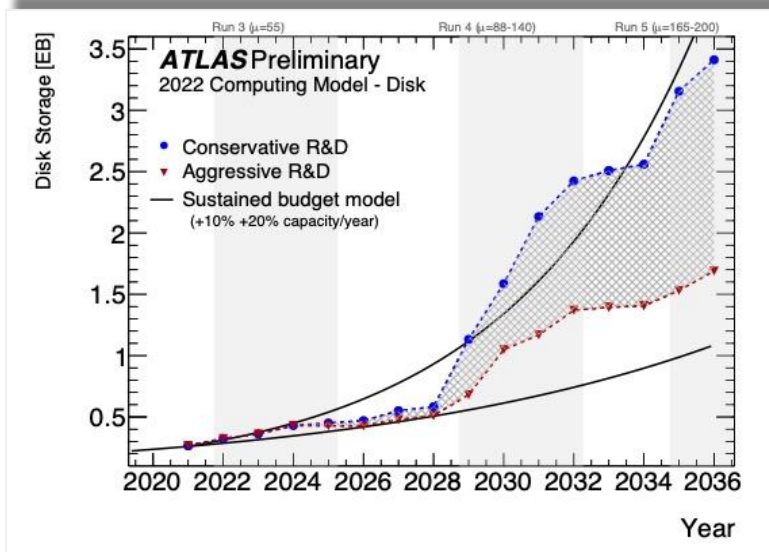
## Resource Usage - Compute - from Jan 2021



- Excellent performance of our distributed computing infrastructure
  - Opportunistic resources (HPCs and HLT farm) are still providing very useful and sizeable contribution. (HLT farm going soon)



# Storage challenges



- Our data sample is going to be 10x bigger!
  - today trigger rate is  $\sim 4$ kHz, will be 10kHz in HL-LHC,
  - RAW size from 1.5MB/event to  $\sim 4$ MB in HL-LHC
  - MC simulated events from 30B/year to 150B/year
- We are working on several R&D (and lots of “business as usual”) to make sure we make the best use of our Disk and Tape:

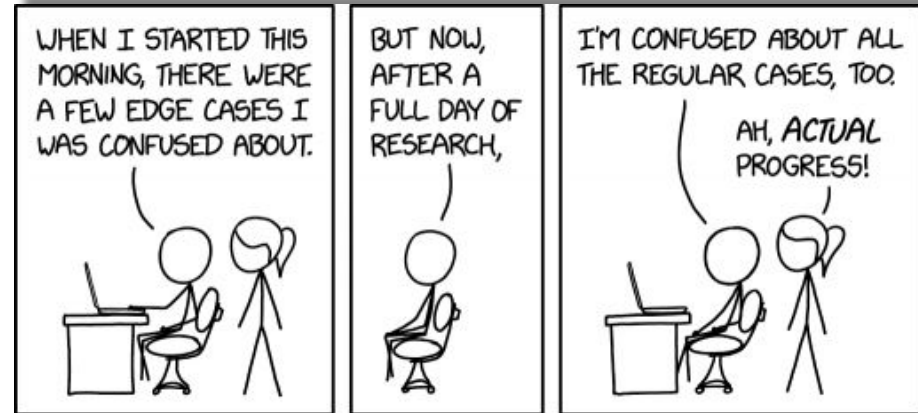
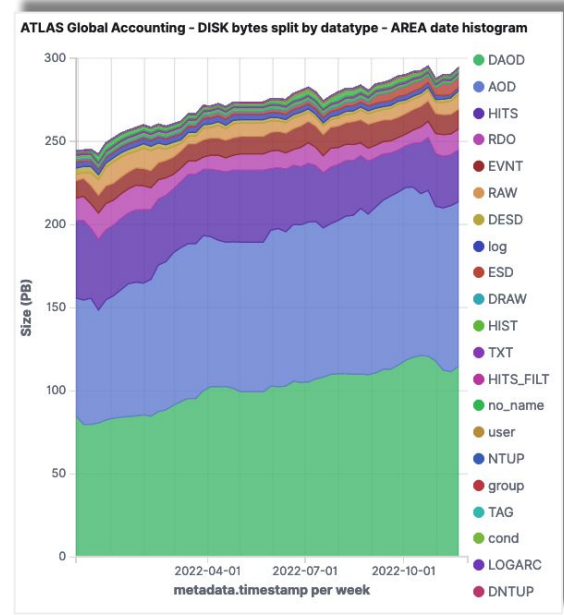
- New analysis models: smaller data types
- Data Carousel: trading Tape (cheaper, but slower) for Disk
- Constant review of the content and the objects stored in each datatypes

# Storage challenges: more

- Lot of analytics:
  - Too much, never enough - and always complicated to keep all the chain working! - help welcome!
  - understand what we use, how often, how many replicas ...

Some more blue-sky R&D are fascinating, e.g.:

- fastchain: trading CPU for Disk
  - A fast end-2-end simulation, we could choose to save less (or zero) intermediate data - potentially important impact on disk and tape savings
- More aggressive deletion - and recreation on demand
  - Again trading CPU for Disk, one of the complexity is on the timely reproduction (with proper metadata)



Data Scientists:

**WANTED**



# Tackling challenges: considerations

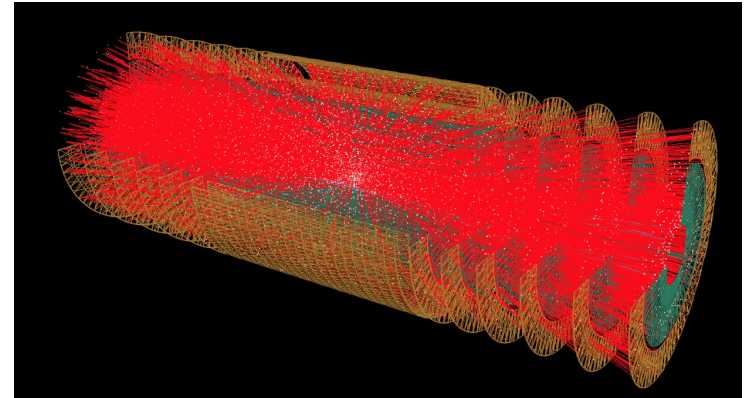
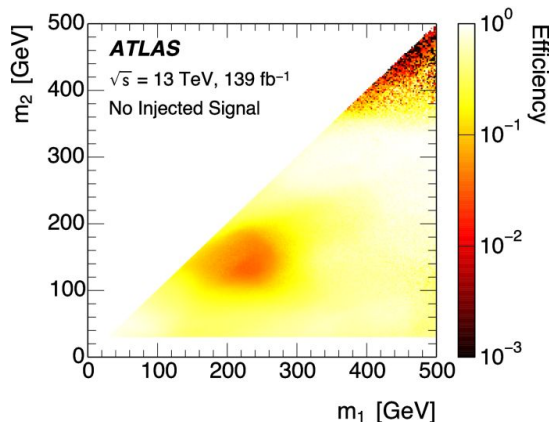
- HPCs are very useful to mitigate the risk of being short in CPU resource
  - With some caveats - as we discussed
- Clouds and in general project with industry help us in improving our frameworks, make it more sustainable and “community standard”
  - Some interesting “opportunistic storage” R&D
  - And TCO evaluation is important
- Many more R&D to mitigate the HL-LHC not-enough-CPU risks:
  - We are following with great interest the Geant4 collaboration (+Adept and Celeritas) and Event Generator groups’ R&D efforts
- Many R&D to mitigate the storage challenges’ risks:
  - Our systems are complex: need lots of brains and analytics expertise to take the right informed decisions
- Note that not all our R&D will reduce our resource consumption!
  - We want to do \*great\* physics, some of the R&D will help us on this too
  - And some others will push us to rewrite our (reco and tracking) code, to be able to be accelerators (e.g. GPU) friendly
    - → independently on the usage or not of accelerators, we will for sure gain having better written code: sustainability and capability to engage more people

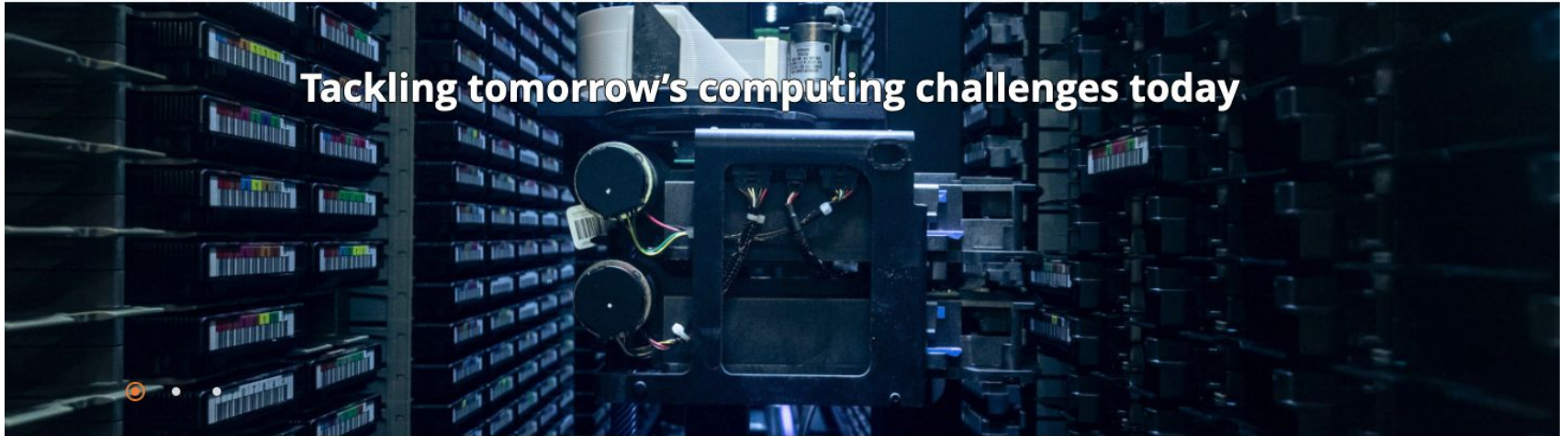




# ML, AI, and more

- Machine Learning (and AI) is pervasive in the experiments
  - For trigger, reconstruction (e.g. pixel clusters, electrons and hadrons), analysis (e.g. signal/background discrimination), fast detector simulation (e.g. generative networks for calorimeter simulation)
  - Not only inside the experiments, but also work in collaboration with external machine learning experts (e.g. TrackML challenge, Higgs ML challenge)
- AI for computing
- Quantum Computing and Algorithms





Tackling tomorrow's computing challenges today



# .... **The Role of Computing in Society**

... in the scientific community actually... but also in the society at large.

**What do YOU think?**  
**What's YOUR forecast?**

<https://www.menti.com/al99q1kj4iqi> or  
<https://www.menti.com/33023009>



# The road ahead: “Welcome”!

- Exciting Challenges:
  - The experiments at CERN are facing interesting, difficult, but solvable software and computing challenges for the HL-LHC.
- Critical Need for Talents:
  - One of the biggest challenges not mentioned here is engaging, supporting and retaining skilled developers – YOU!!
- Unified Goals:
  - Focusing our efforts on common shared objectives is paramount.
- Impactful Collaboration:
  - The way in which we work can make the difference between success and failure!

Time to act is *now!*

- Ad Augusta per angusta!

**We are part of a team. We are all unique. We can make a difference.**





# Thank you

**... a special thanks for the many very useful discussions and providing material for this talk, to: Antonella Del Rosso, Zach Marshall, David South, Concezio Bozzi, Stefano Piano, Fernando Barreiro Megino, Maria Girone and many more!**

WHEN A USER TAKES A PHOTO,  
THE APP SHOULD CHECK WHETHER  
THEY'RE IN A NATIONAL PARK...

SURE, EASY GIS LOOKUP.  
GIMME A FEW HOURS.

... AND CHECK WHETHER  
THE PHOTO IS OF A BIRD.

I'LL NEED A RESEARCH  
TEAM AND FIVE YEARS.



IN CS, IT CAN BE HARD TO EXPLAIN  
THE DIFFERENCE BETWEEN THE EASY  
AND THE VIRTUALLY IMPOSSIBLE.



Success is no accident.  
It is hard work, perseverance,  
learning, studying, *sacrifice*,  
and MOST of all,  
*love of what you are doing.*

-Pele

# Plan B

