The future of computing at the LHC

20 Jahre GridKa

Colloquium KIT Karlsruhe

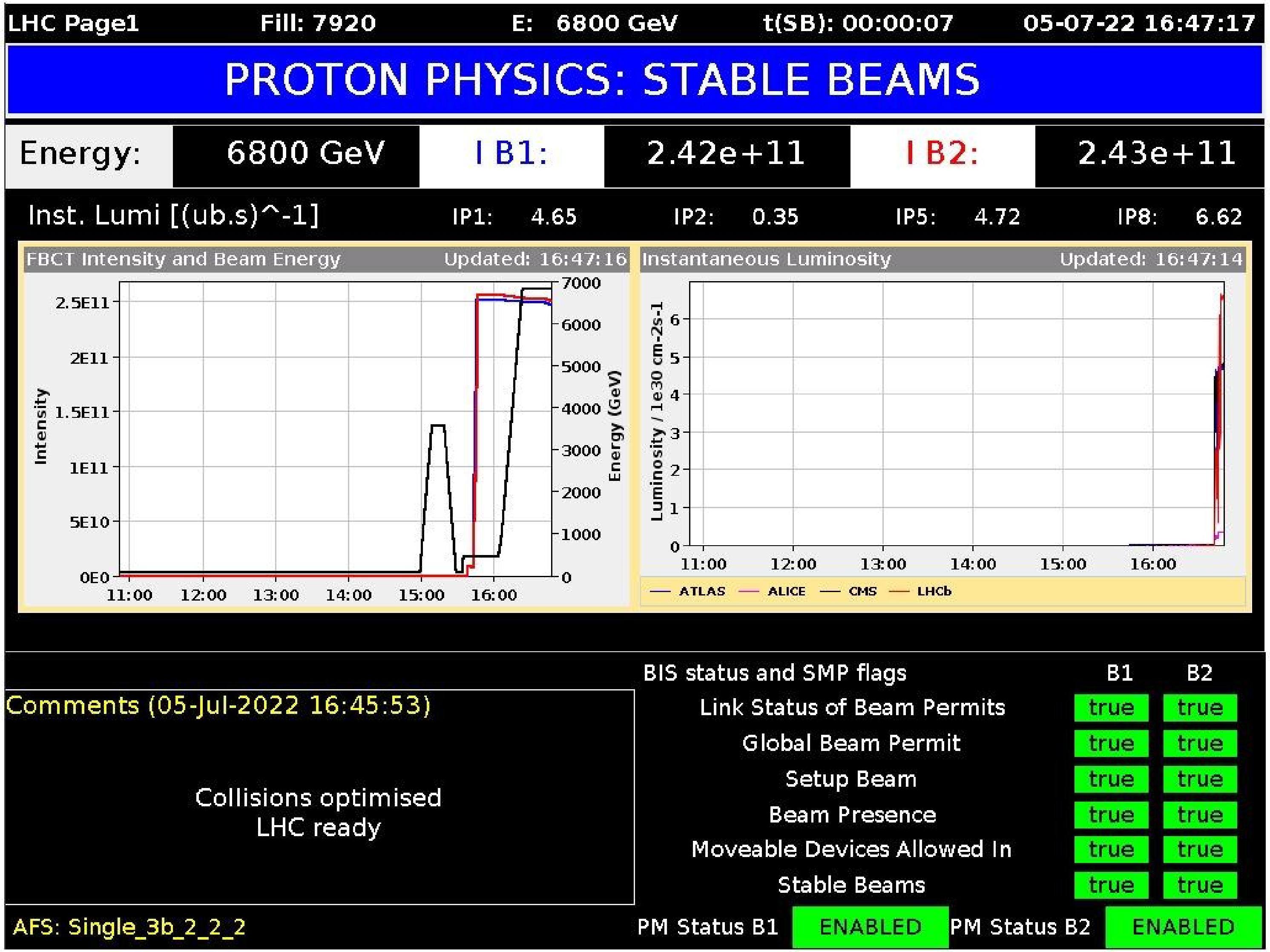
Joachim Mnich

July 22nd, 2022

The good news: the future has just started!

First Stable Beams at the record energy of 13.6 TeV – 5th July

July 5th, 2022: Start of LHC Run 3



New beam record energy: 6.8 TeV

Careful start:

- 3 on 3 proton bunches, i.e. only approx. 0.1% of the nominal luminosity
- Today already at 600 on 600 bunches
- And on the way to 2748 on 2748

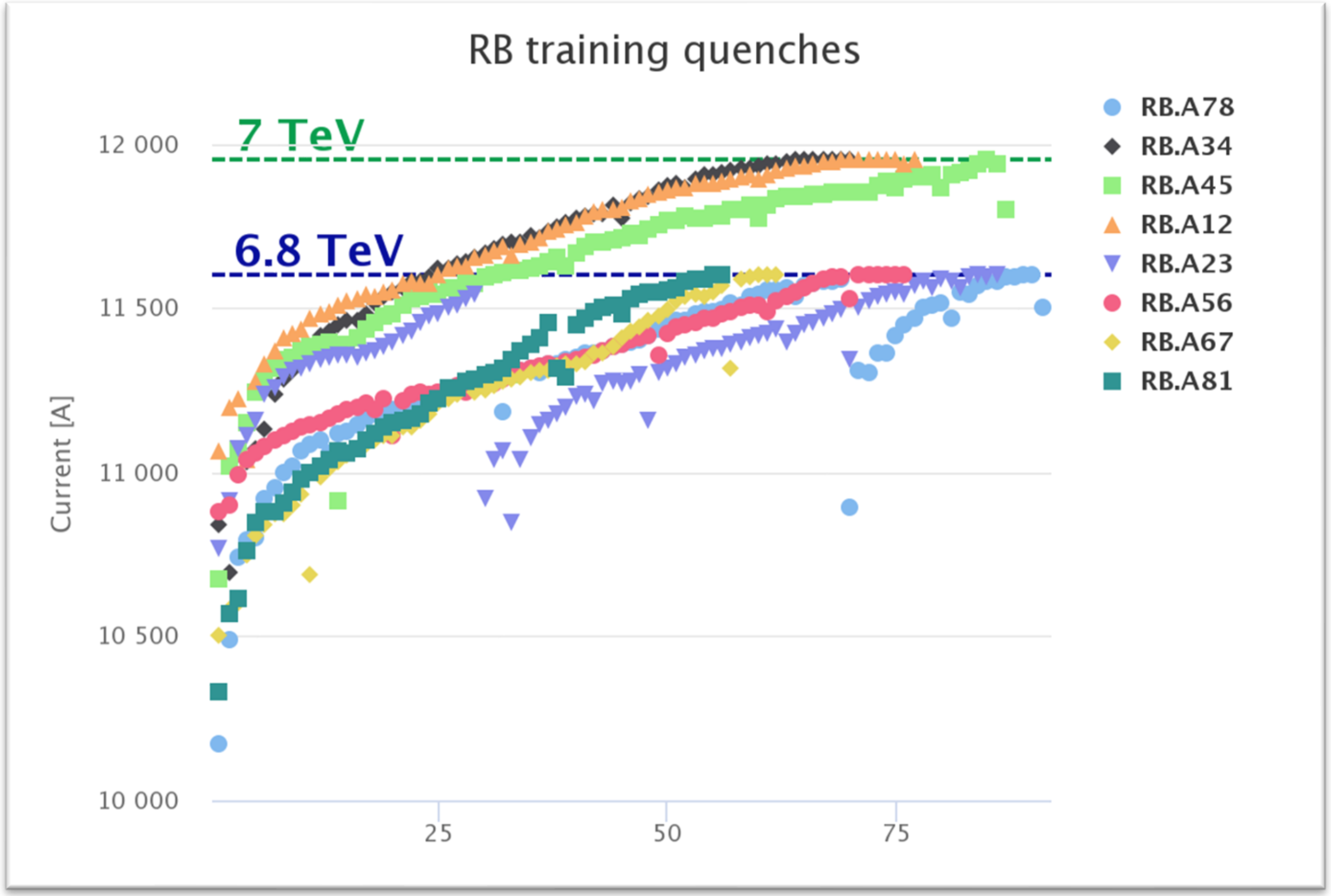
LHC – Dipole Training to 6.8 TeV Completed April 11th

During the main dipole training campaign spanning 2021 and 2022 over 600 primary training quenches

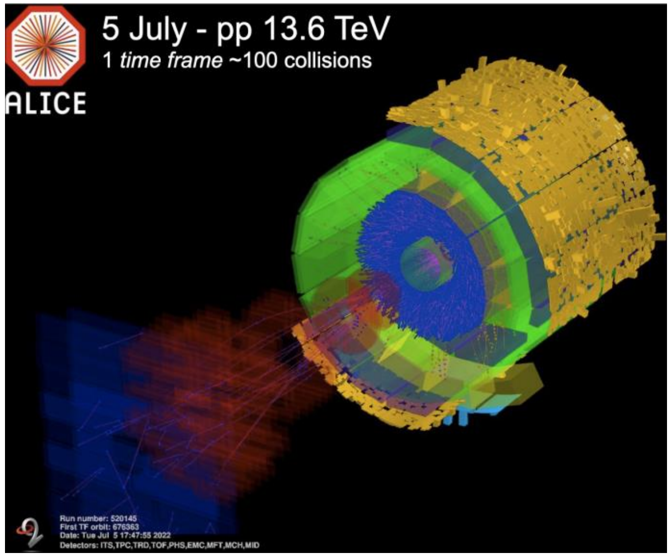
- 3 sectors trained to 7 TeV
- 2 sectors (78 & 23) were trained twice

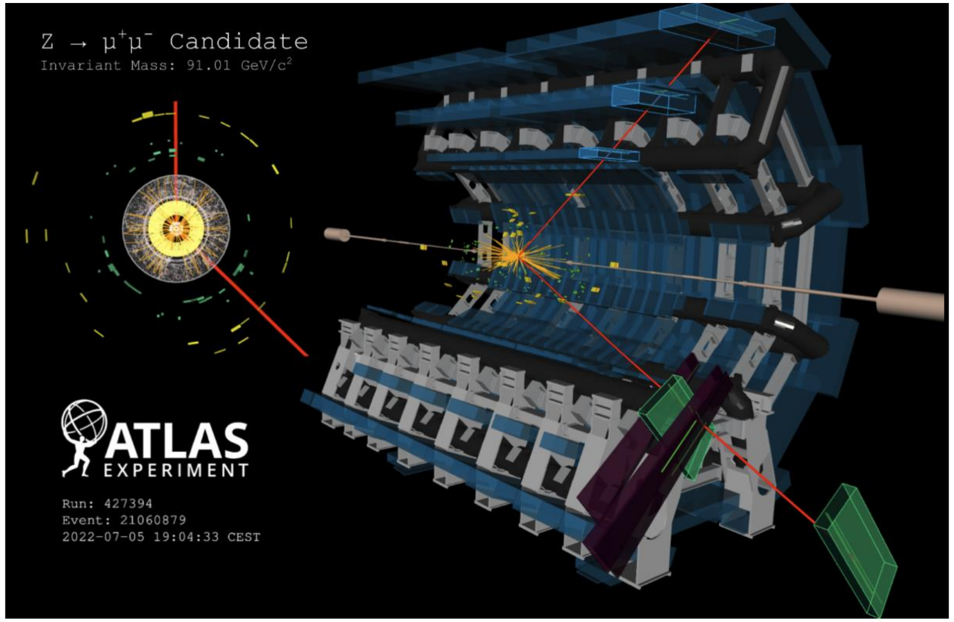
Very long training and powering test campaign

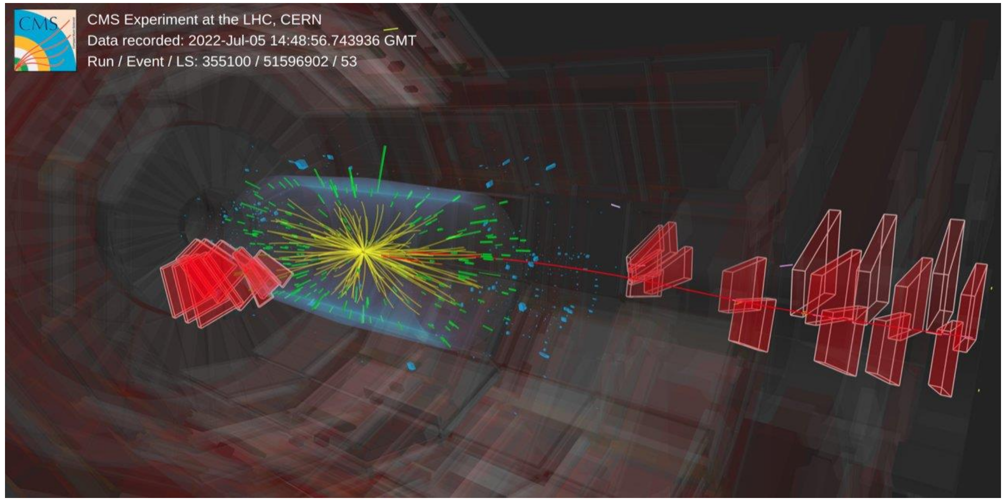
- risks have to be well managed throughout

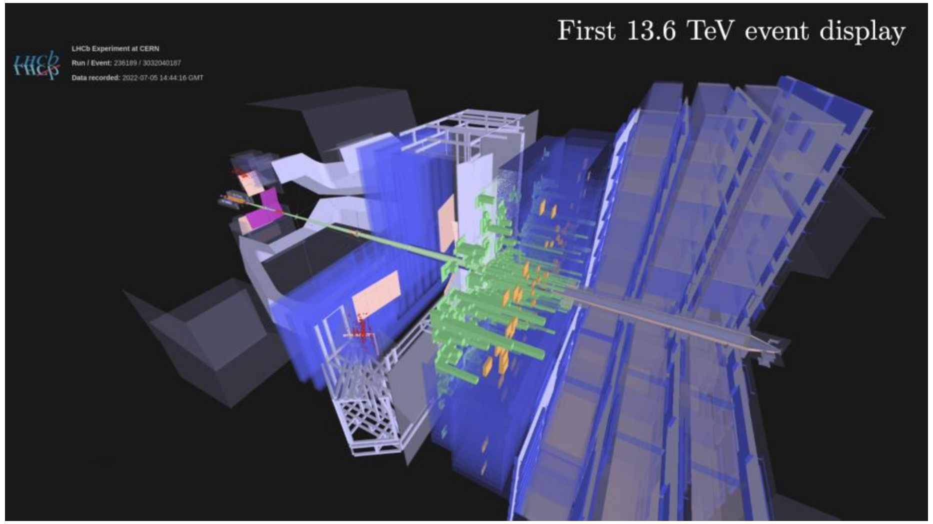


Experiments: Phase I Upgrades Completed & Ready for Run 3





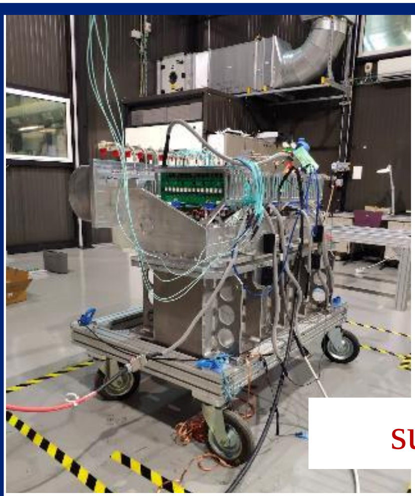




LHCb: VELO A-side Installation

leaving Liverpool
April 27th

Arrived at CERN
April 29th



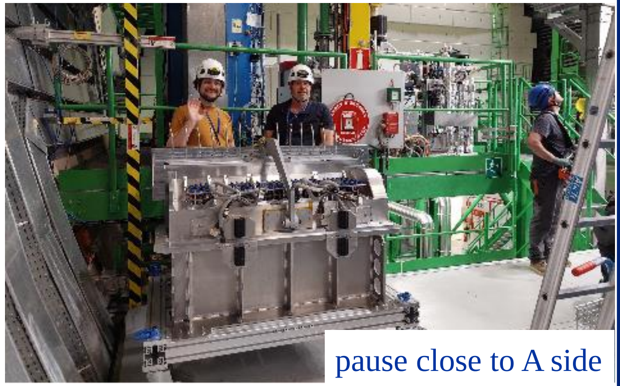
Connectivity tests & leak tests

Grey Room preparation

surface test **April 30th –May 9th**

installation preparation **May 10th**

lowered down shaft



pause close to A side

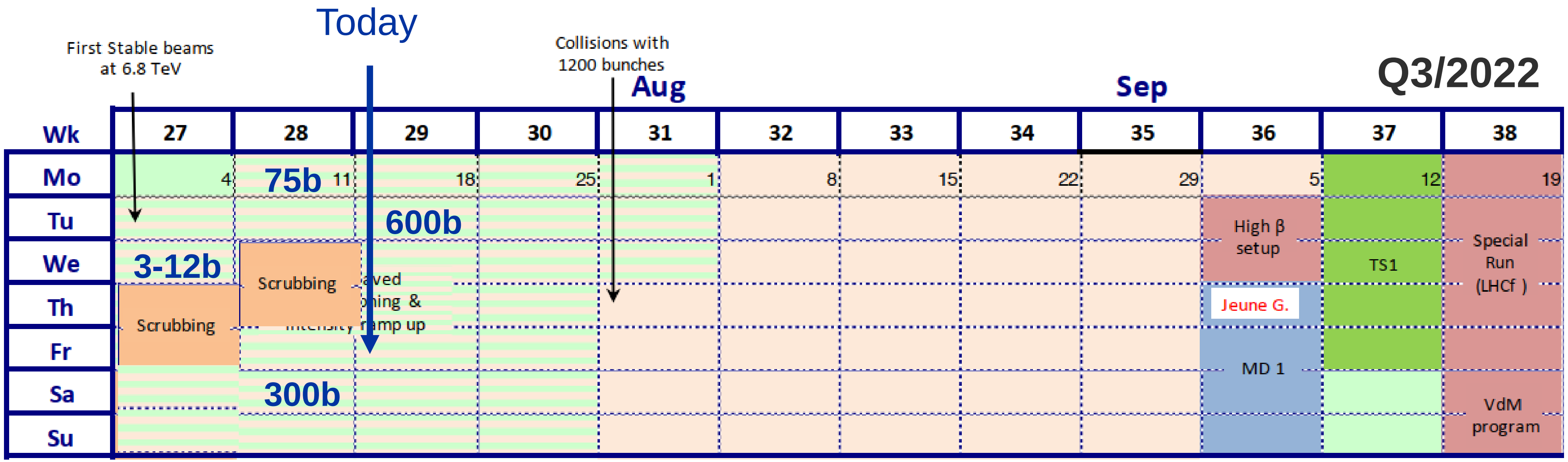
installation **May 11th**

May 12th – 16th
electrical connectivity
fibres connectivity
cooling commissioning
…

LHC Plans 2022

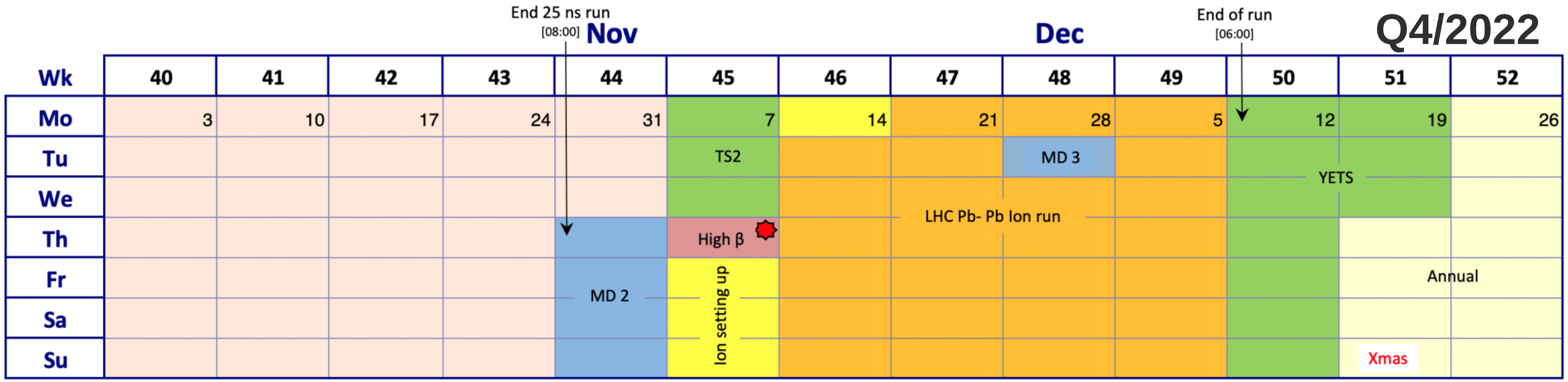
Up to 600 bunches
by Tuesday 19th July

End July 1200 bunches thereafter up to 2748 bunches in quanta of ~300 bunches - dictated by machine protection considerations



End 25 ns proton run:
08:00 3rd November

End of Run:
06:00 12th December



Expectations for Run 3

Indicative luminosity expectations for Run 3 (2022-2025):

Mode	ATLAS/CMS	LHCb	ALICE
p-p	**250/fb**	25 - 30/fb (~50/fb by LS4)	200/pb
Pb-Pb	7/nb (13/nb by LS4)	1/nb (2/nb by LS4)	7/nb (13/nb by LS4)
p-Pb	0.5/pb (~1/pb by LS4)	0.1/pb (~0.2/pb by LS4)	0.25/pb (~0.5/pb by LS4)
O-O	0.5/nb	0.5/nb	0.5/nb
p-O	LHCf 1.5/nb	2/nb	

Note: the integrated pp-luminosity delivered so far to ATLAS and CMS is about 190 fb^{-1}
Run 3 will more than double the integrated pp luminosity!

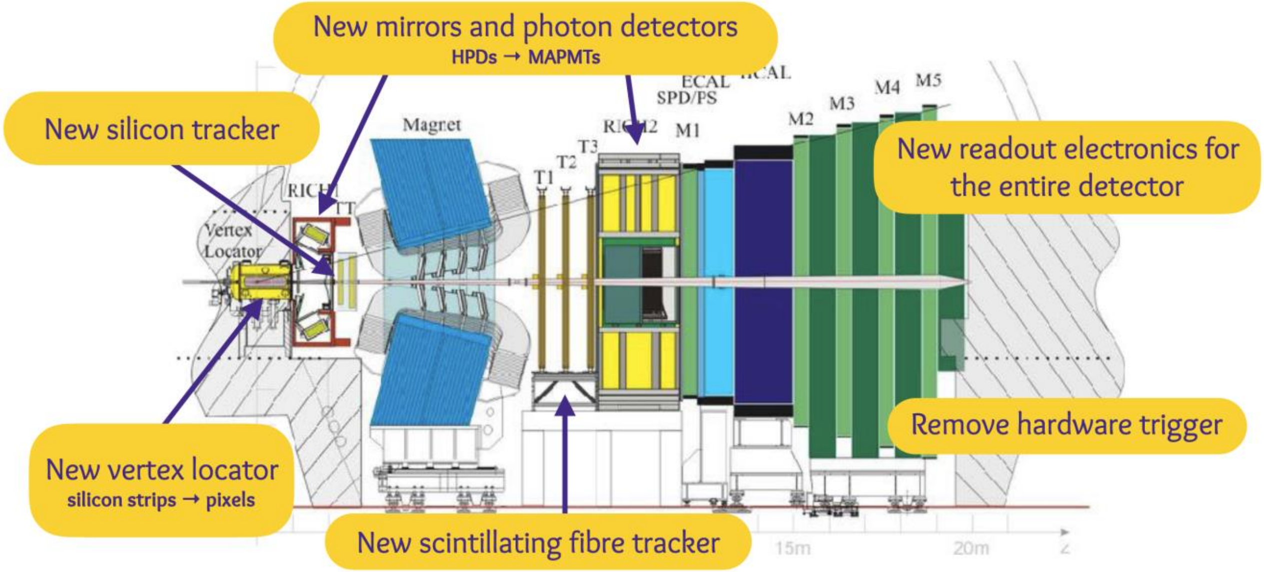
Run 3: Higher Data Rates for LHCb

LHCb upgrade:

- Raising the instantaneous luminosity by a factor five to $2 \times 10^{33} cm^{-2} s^{-1}$
- Implementing a full software trigger to overcome limitation of L0 hardware trigger

Full software trigger at 30 MHz inelastic collision rate

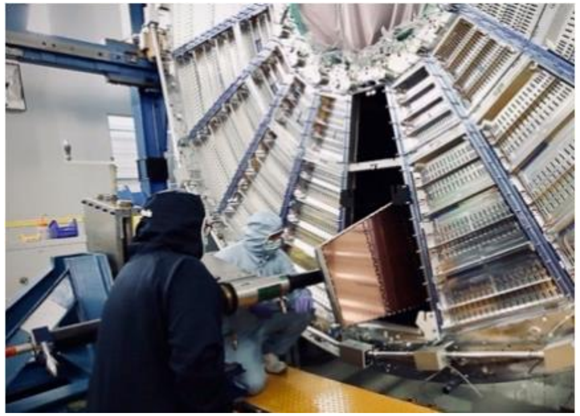
- Factor 2 increase in trigger selection efficiency
- At least 10x more recorded events
- Selective persistency: write out only the "interesting" part of the event

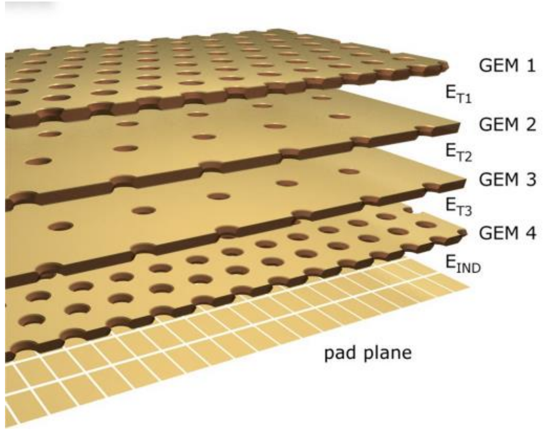


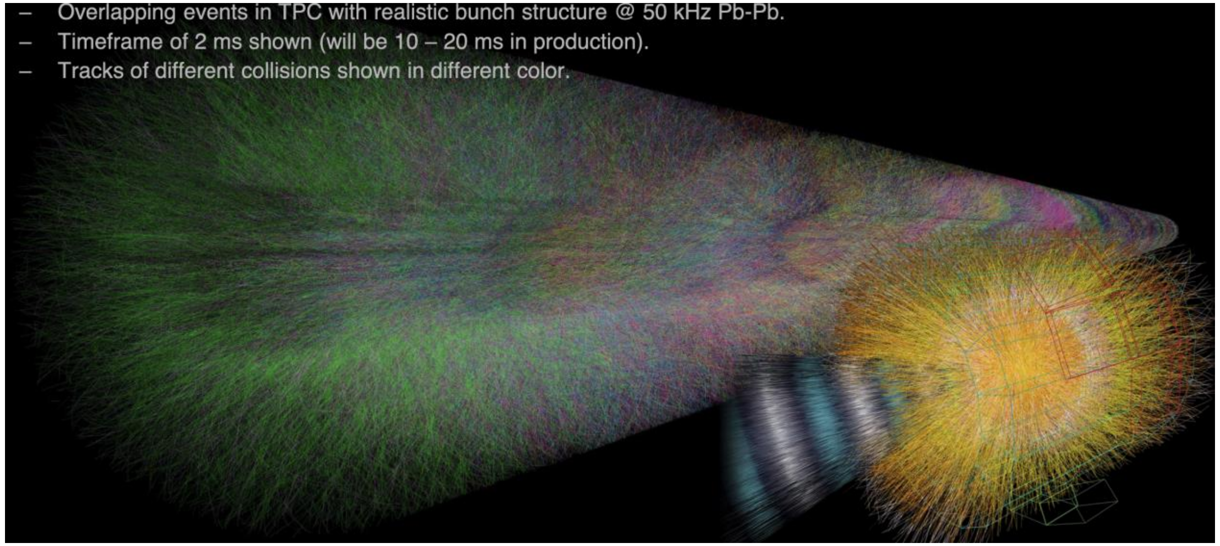
Run 3: Higher Data Rates for ALICE

ALICE upgrade

- New inner tracking system and especially GEM readout for TPC: move from gated/triggered operation to continuous readout (min Bias)
- At the same time, increase from 1 kHz to 50 kHz Pb-Pb interaction rate
- Collect 13 nb^{-1} of Pb-Pb collisions at 5 TeV in a minimum bias (MB) mode (compared ~0.1 nb^{-1} MB and ~1 nb^{-1} triggered during Run 1 and Run 2)
- 100x more recorded MB events wrt Run 1 and Run 2
  - Mainly addressed via an aggressive compression based on online reconstruction (including GPUs)
  - Analysis needs to gain factor 20 (disk and CPU) through smarter strategy and algorithms maintaining (or better improving) the physics performance



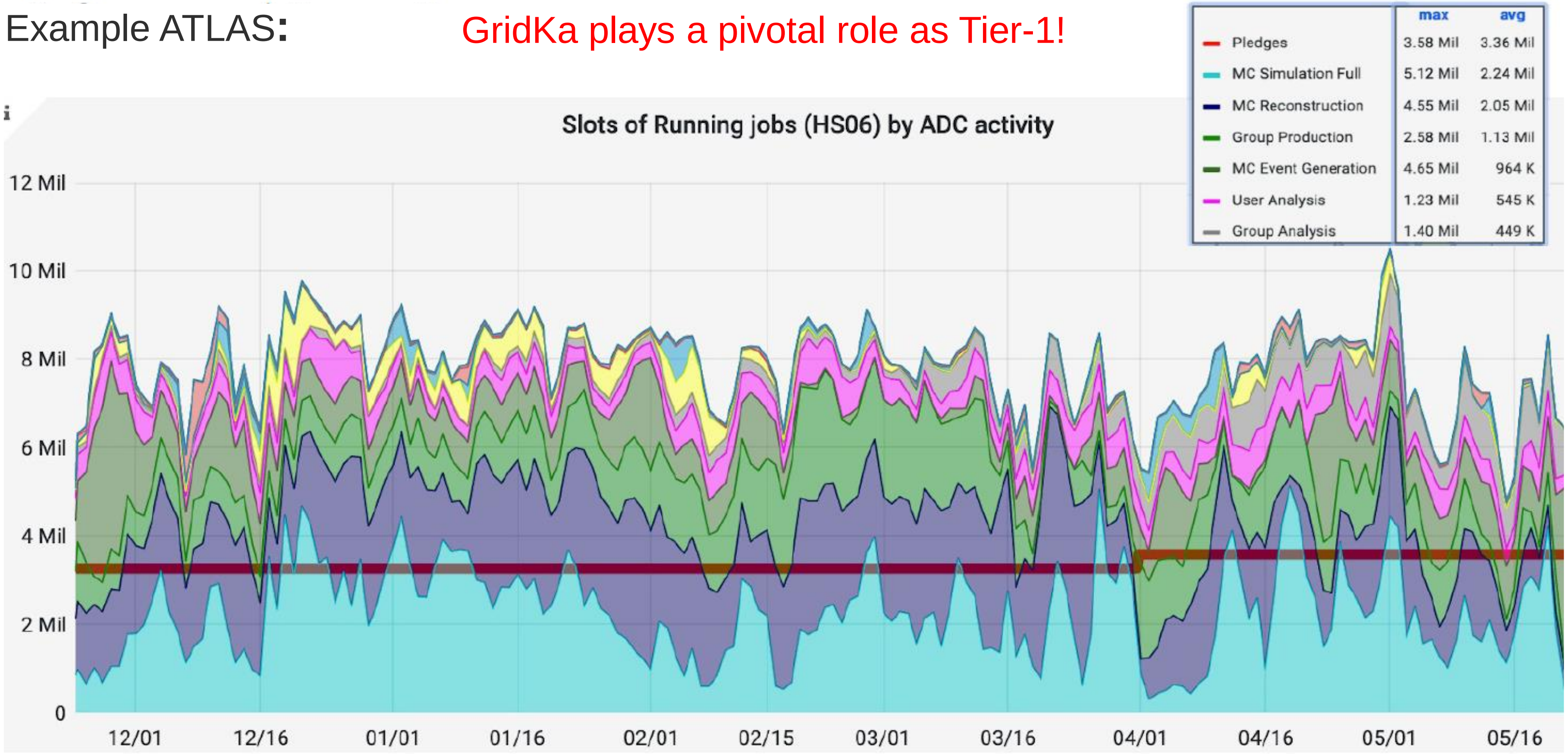




Status Computing at the start of Run 3

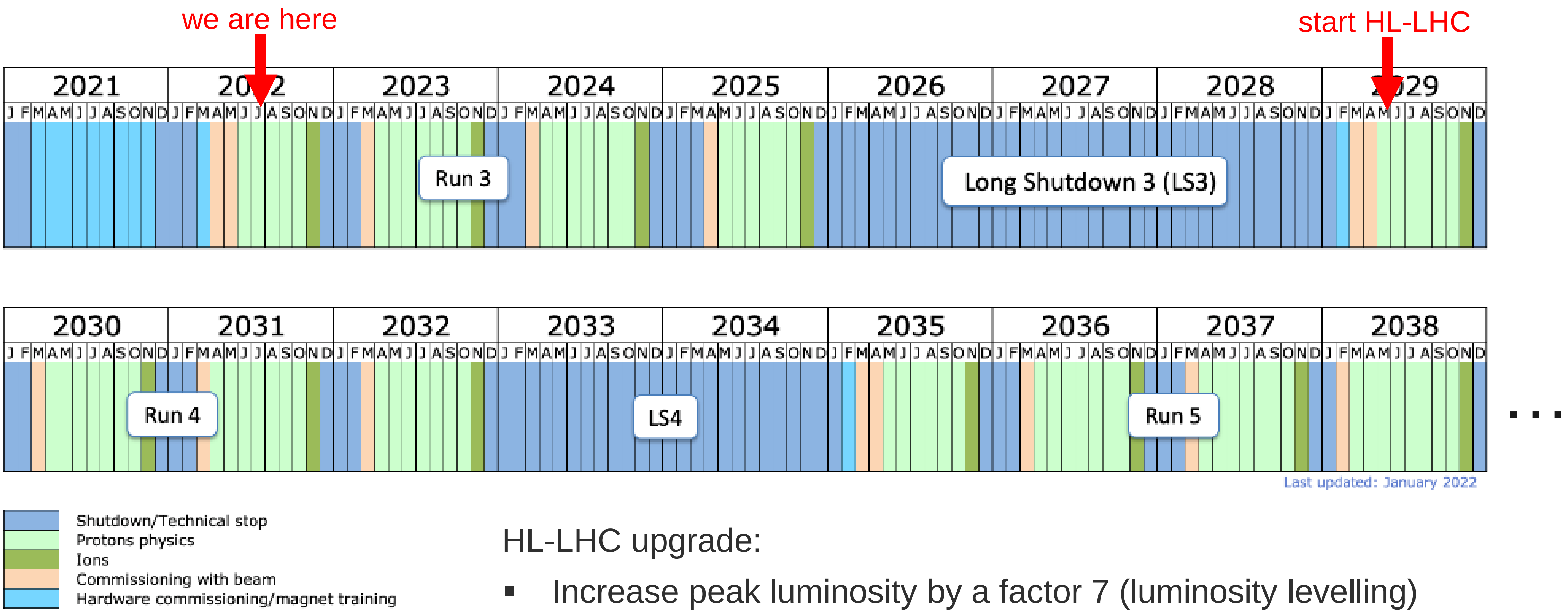
Example ATLAS:

GridKa plays a pivotal role as Tier-1!



Distributed computing infrastructure continues to perform extremely well

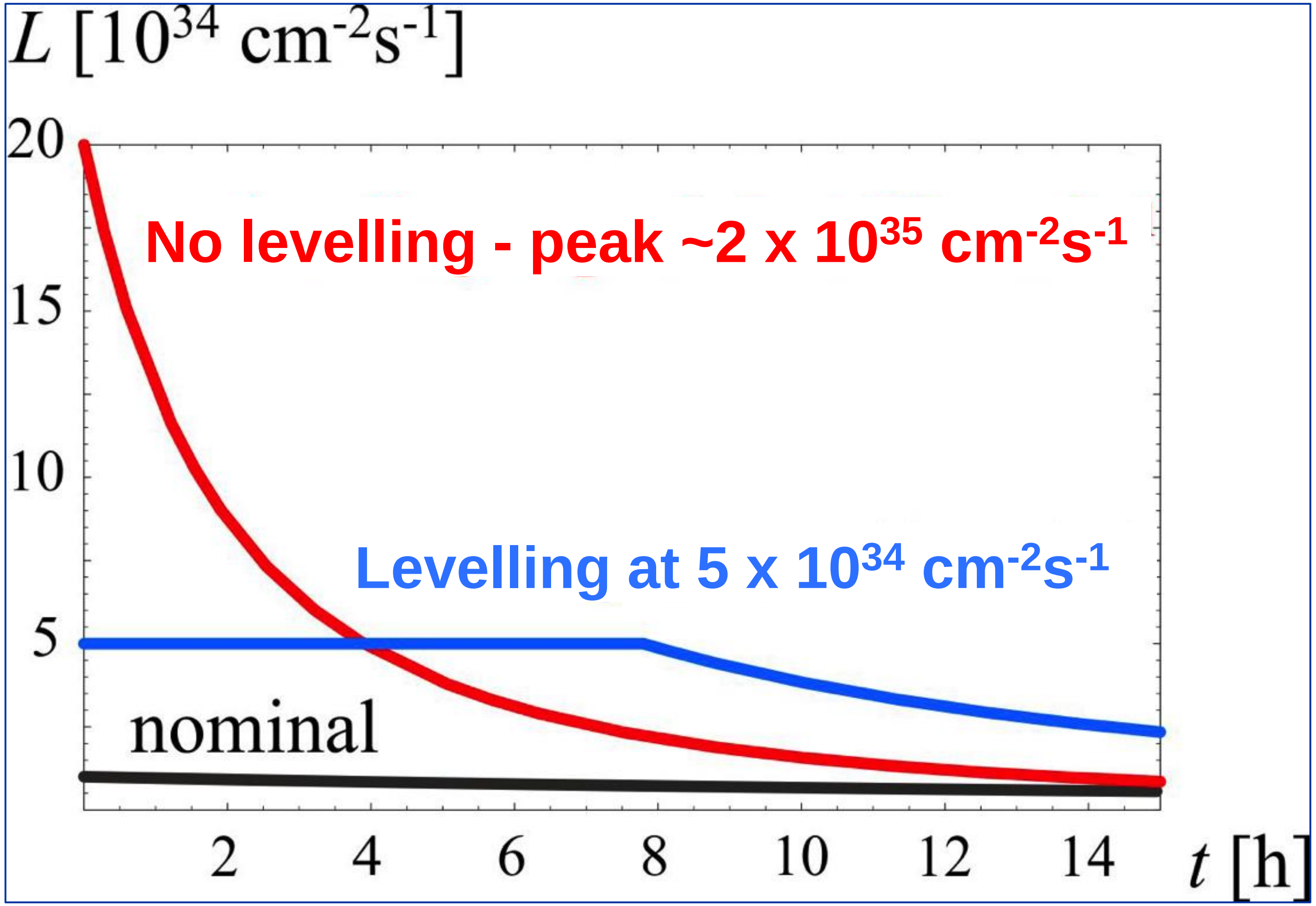
From LHC to High-Luminosity LHC (HL-LHC)



HL-LHC upgrade:

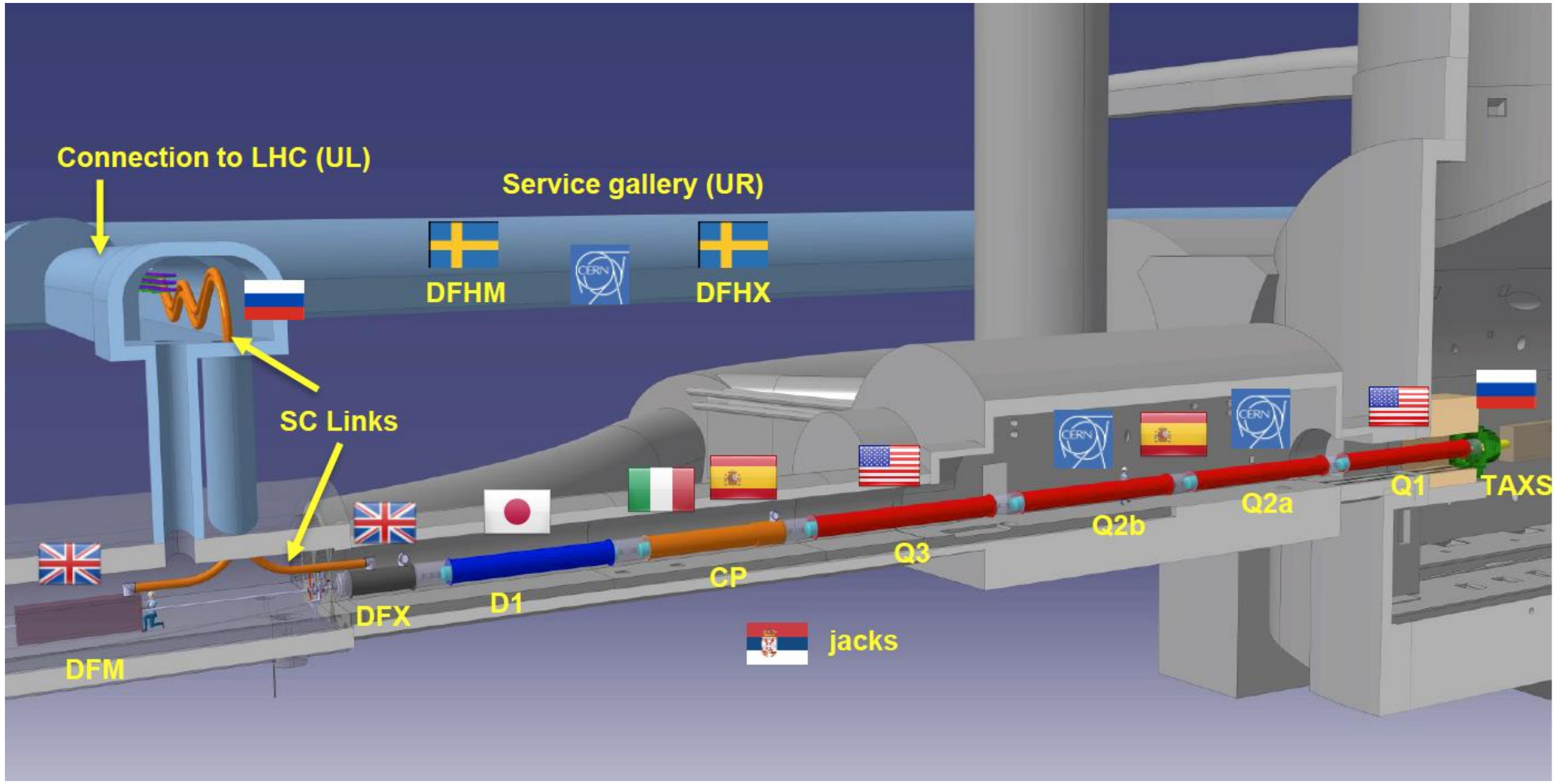
- Increase peak luminosity by a factor 7 (luminosity levelling)
- By approx. 2040 the integrated pp luminosity will have increased by a factor 20 wrt. today

HL-LHC Operational scenario



Nominal levelling time ~7 hours

Status HL-LHC

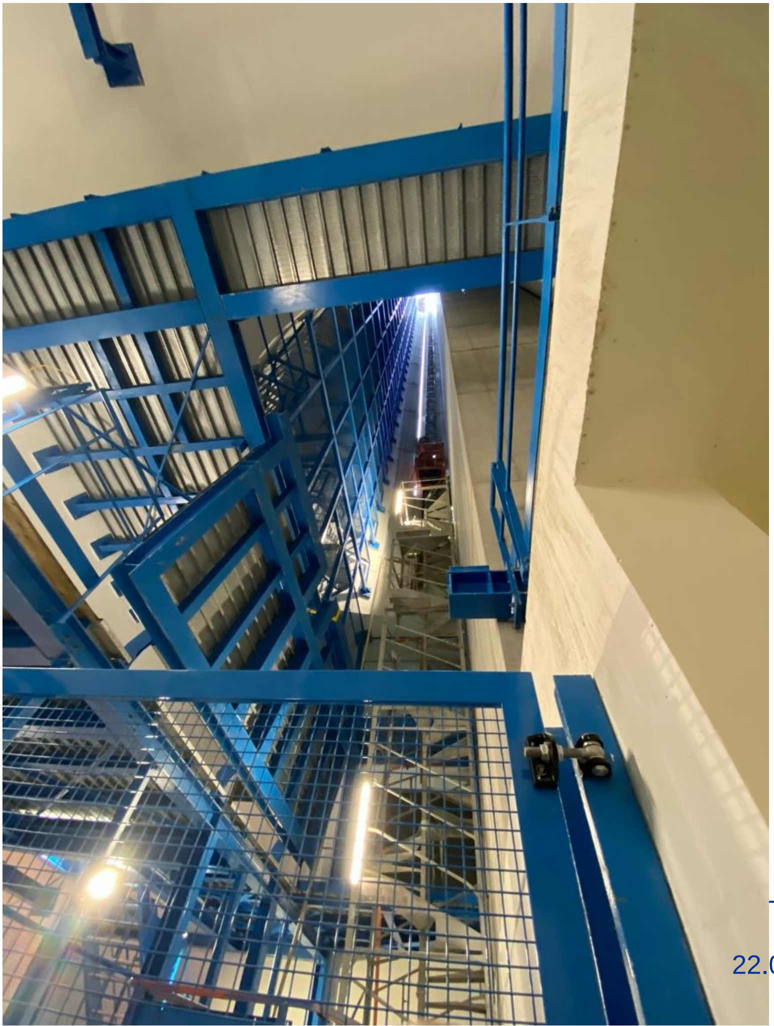


Accelerator:

- Increase luminosity by improving the final focus system
- Good progress in 2021/22 on all fronts despite COVID-19
- Completion of prototypes and start of series production for many components
- 2022 is a crucial year:
  - Production phase, resolution of important technical issues to be confirmed
  - Procurement (& in-kind contributions) in very challenging global market conditions
  - Plan B for in-house production of Russian in-kind contributions being fully developed

Point 1 surface buildings

Underground works at P5



US57 cavern

UR55 gallery

PM57 shaft

Progress SC Magnets

June 2022: MQXFBP2 (Q2) and corrector MCBXFBP1 on the alignment bench (second full-size prototype of the MQXFB magnet family)

MgB_2 flexible prototype electrical transfer line (SC Link) for the powering of the HL-LHC triplets

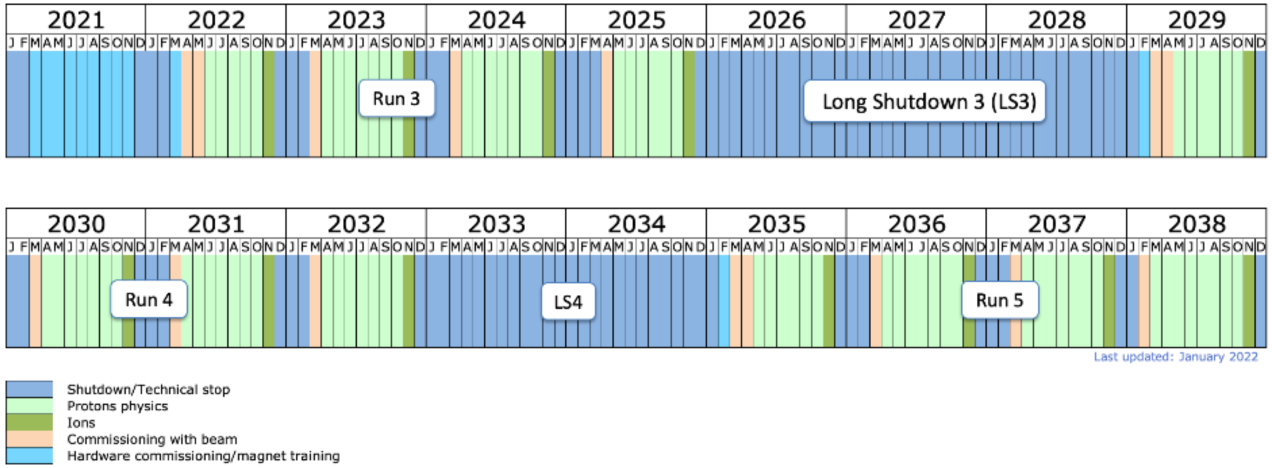
ATLAS & CMS Phase II Upgrades: Adapting to new Schedule

Adapting to the new HL-LHC schedule (re-baselining)

Brings all projects back on schedule

- i.e. no negative floats
- but ATLAS ITk and CMS HGCal remain on critical path

Additional person power at laboratories and universities is needed to mitigate schedule risks

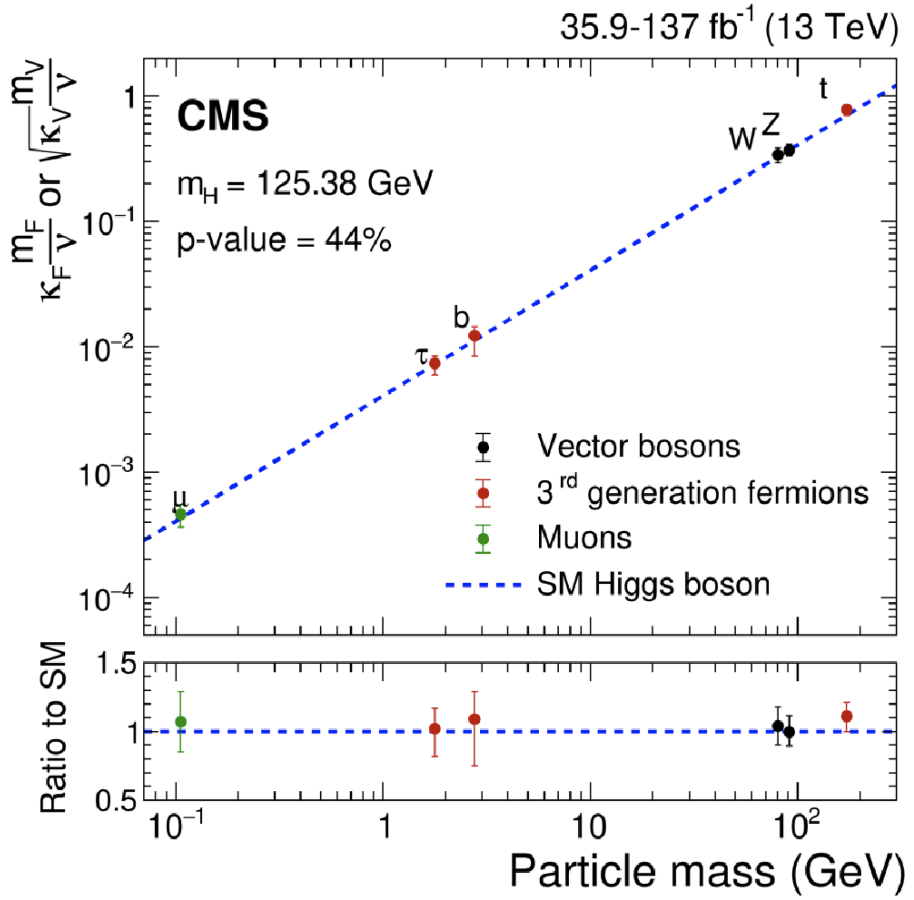


Challenges:

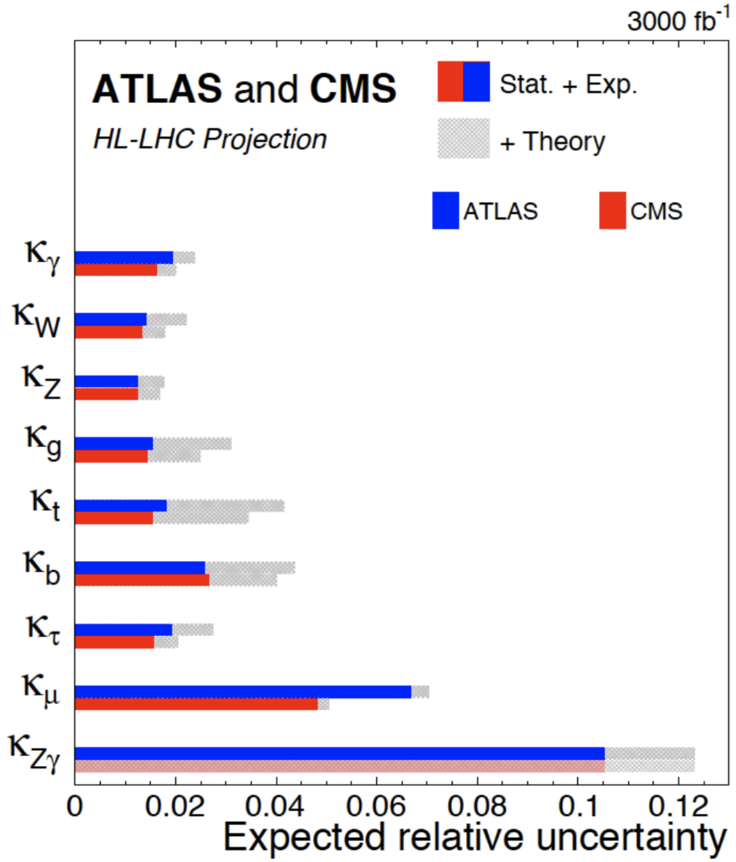
- Price increase: inflation, energy, material cost, specifications, electronics crisis, supply-chain, new green energy – pandemic
- Procurement: electronics, FPGAs, ASICs – also dicing, packaging, testing, hybridization
- Impact of Russian invasion of Ukraine is being studied

Physics goals HL-LHC: Higgs coupling as an example

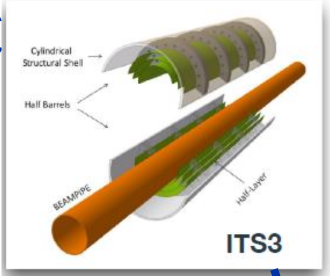
Today:

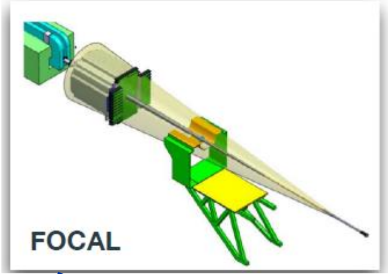


HL-LHC expectation:



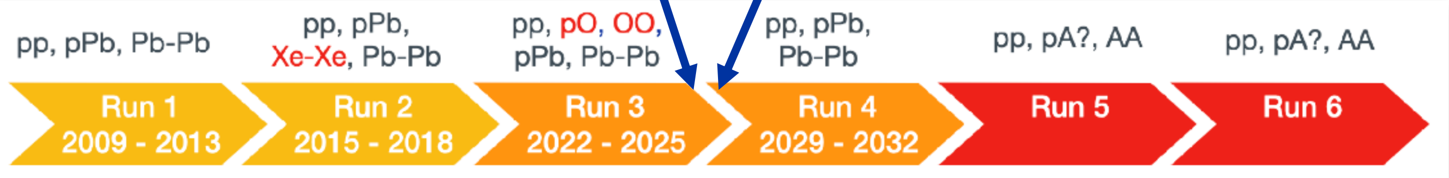
ALICE: Future Plans for HL-LHC





ITS3:

- Replace ITS2 barrel innermost 3 layers
- Reduced inner radius (22 mm → 18 mm)
- TDR in preparation

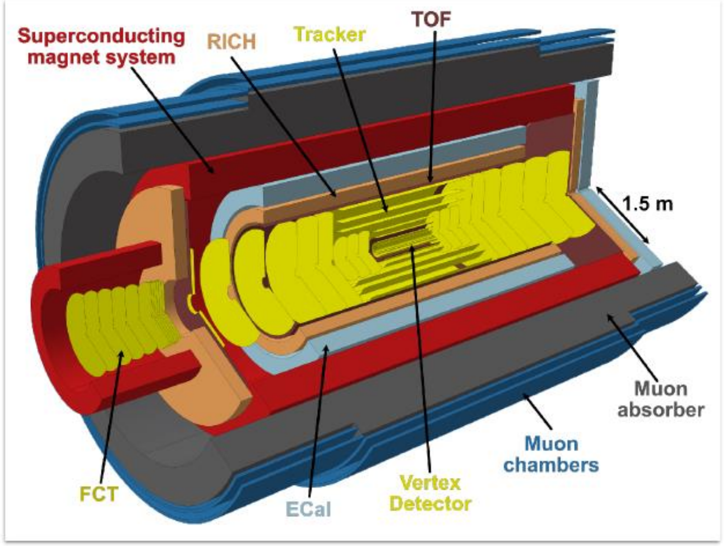


FOCAL:

- Physics: saturation & shadowing at low x
- TDR in preparation

ALICE 3 (for installation in LS4):

- Compact low-mass all-Si tracker, excellent vertex reconstruction and PID
- Letter of Intent reviewed by LHCC in March
- First discussion with funding agencies took place June 27th



[CERN-LHCC-2022-009]

LHCb: Future Plans for HL-LHC

LHCb Upgrade II (for installation in LS4):

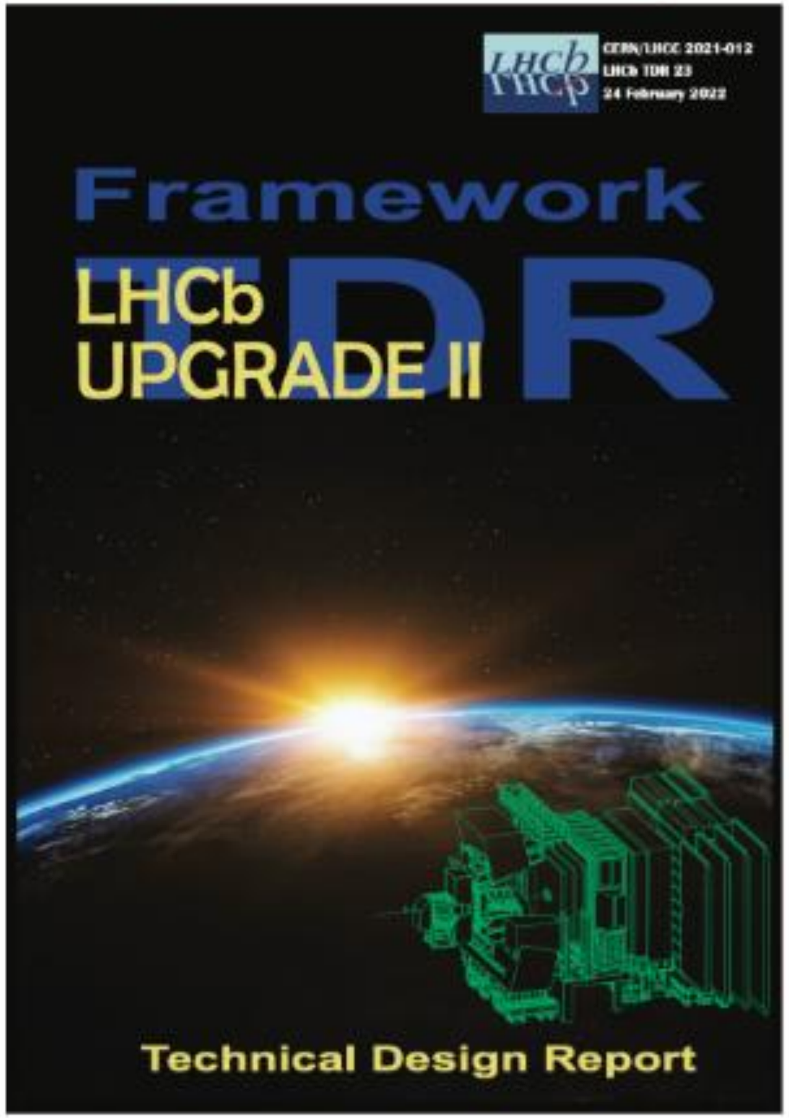
- Fully exploit the HL-LHC for flavour physics

$L_{peak} = 1.5\times10^{34}\ cm^{-2}\ s^{-1}$

L_{int} ~ 300 fb^{-1} during Run 5 & 6

- Targeting same detector performance as in Run 3, but with pile-up ~40!
- New detector technologies (e.g. precision timing, low-cost monolithic pixels)

- Framework TDR reviewed by LHCC in March
- First discussion with funding agencies June 24th



A factor 20 in luminosity and a diverse physics program need to be reflected in computing. This poses exciting challenges!

European Strategy Update 2020

European Strategy Update

2020 Strategy Statements

4. Other essential scientific activities for particle physics

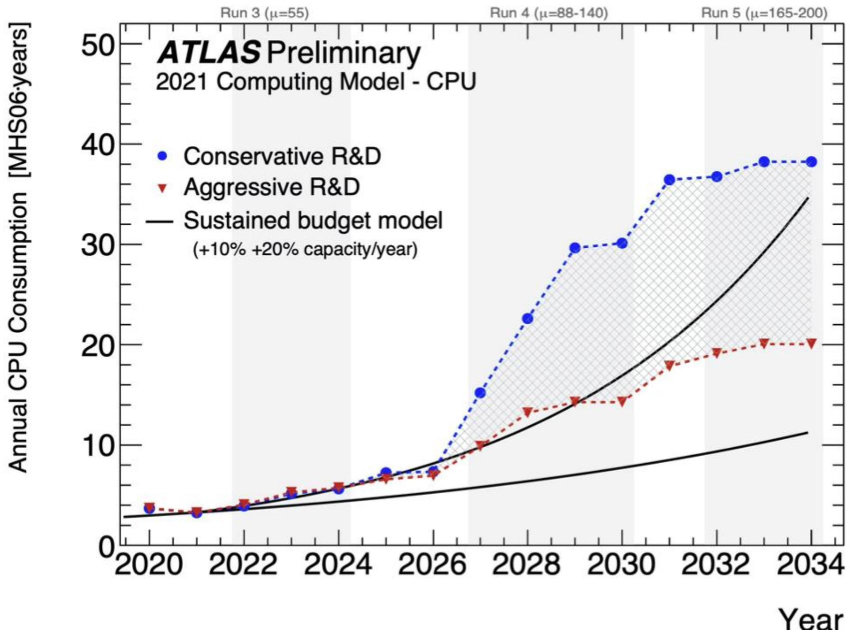
Computing and software infrastructure

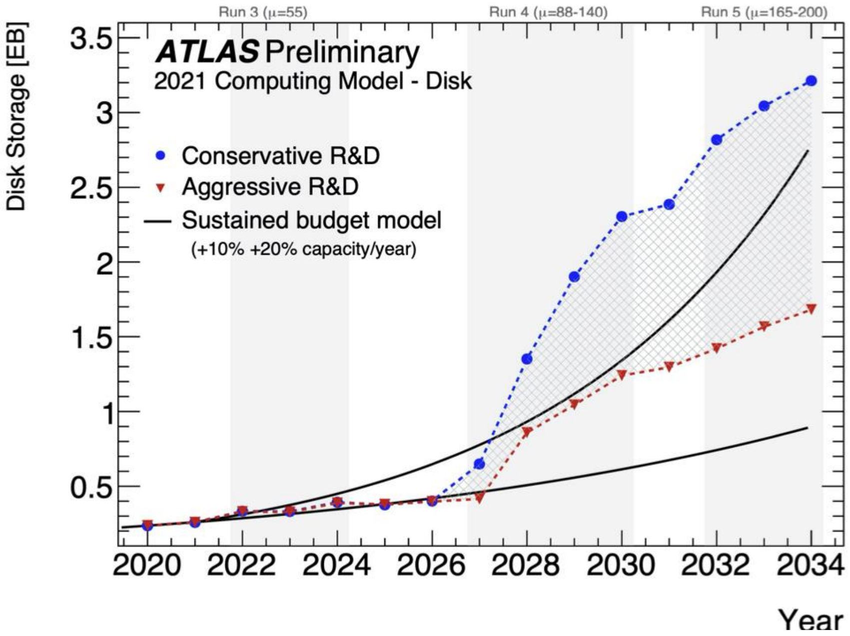
- There is a need for strong community-wide coordination for computing and software R&D activities, and for the development of common coordinating structures that will promote coherence in these activities, long-term planning and effective means of exploiting synergies with other disciplines and industry
- A significant role for artificial intelligence is emerging in detector design, detector operation, online data processing and data analysis
- Computing and software are profound R&D topics in their own right and are essential to sustain and enhance particle physics research capabilities
- More experts need to be trained to address the essential needs, especially with the increased data volume and complexity in the upcoming HL-LHC era, and will also help in experiments in adjacent fields.

d) Large-scale data-intensive software and computing infrastructures are an essential ingredient to particle physics research programmes. The community faces major challenges in this area, notably with a view to the HL-LHC. As a result, the software and computing models used in particle physics research must evolve to meet the future needs of the field. *The community must vigorously pursue common, coordinated R&D efforts in collaboration with other fields of science and industry to develop software and computing infrastructures that exploit recent advances in information technology and data science. Further development of internal policies on open data and data preservation should be encouraged, and an adequate level of resources invested in their implementation.*

19/06/2020 CERN Council Open Session 24

ATLAS and CMS needs for HL-LHC





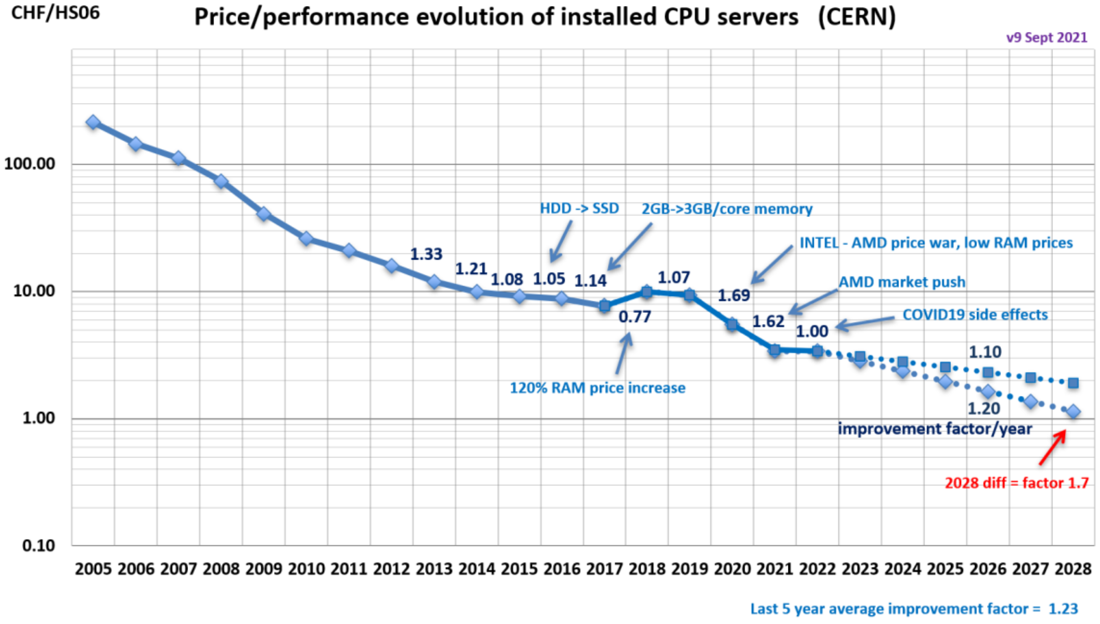
These are the current estimates from ATLAS. Similar plots exist for CMS (will be public soon)

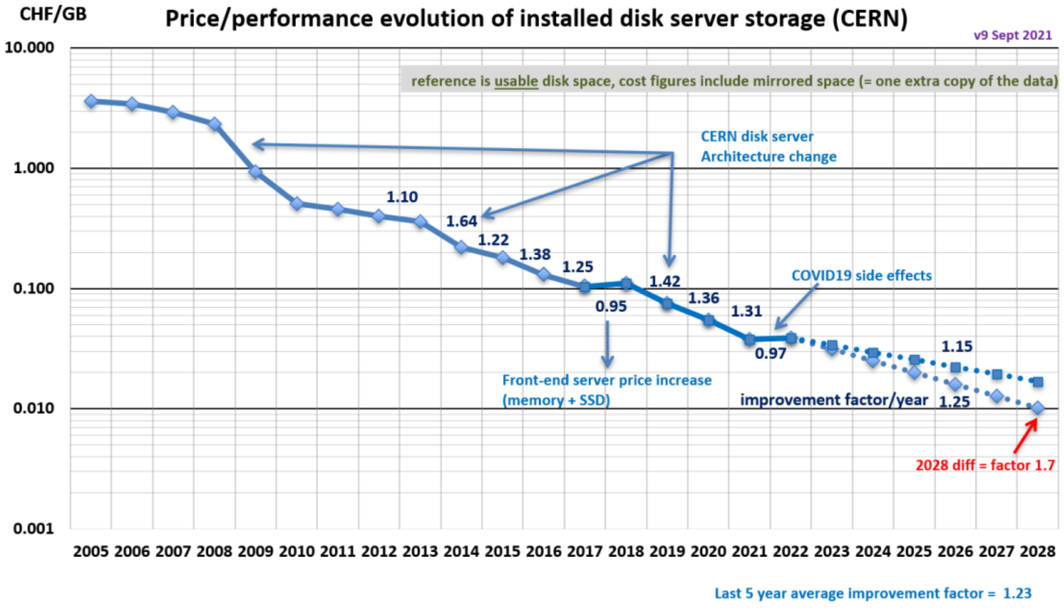
- The gap between available and needed resources is filling up, assuming the main R&D activities are successful
- Investing in further (identified) R&D activities would fill this gap further. Needs more effort
- There are still large uncertainties

Hardware cost and market trends

Hardware cost is more and more dominated by market trends rather than technology

The assumption of +20%/year in storage and CPU for the same budget is not necessarily holding anymore and there are large fluctuations due to the Covid crisis and the Russian invasion of Ukraine (see plots below and notice the log scale) → these do not only affect the price/performance ratio, but also delivery times!





Opportunistic resources

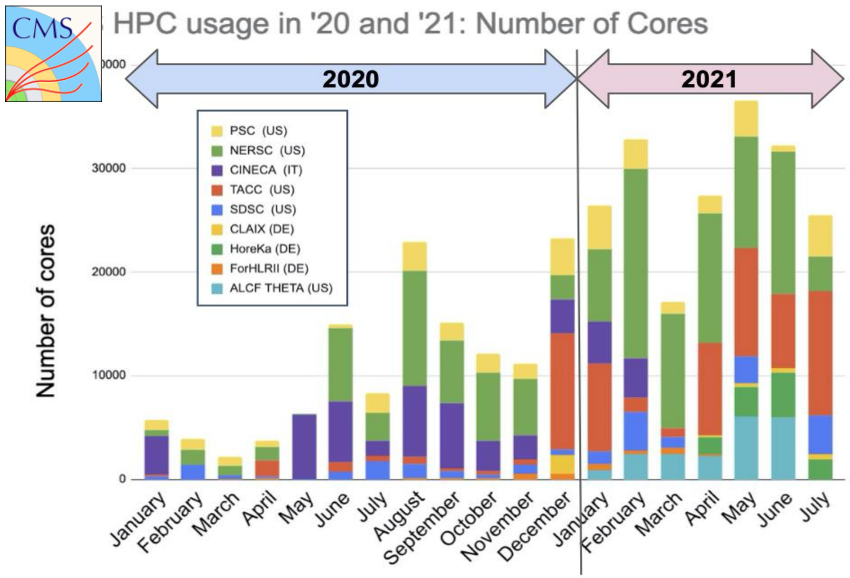
A mitigation for the gap in resources comes from opportunistic CPUs. HPC centers offer a unique opportunity

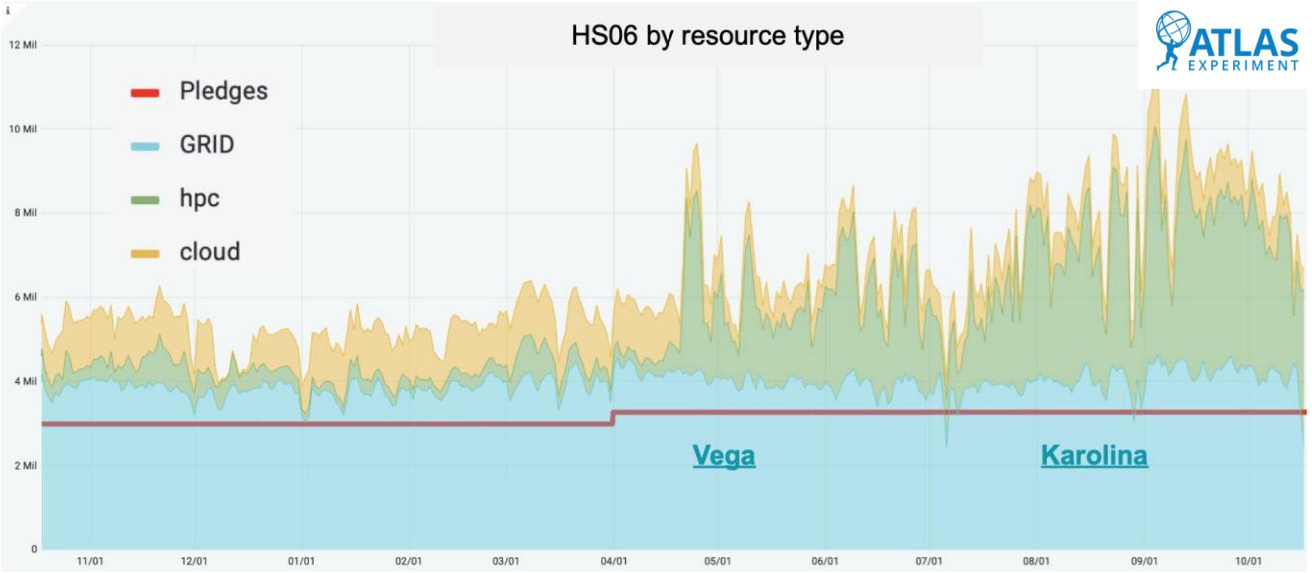
Accessing and using resources at HPC centers comes with different challenges:

- Diversity in access and usage policies, edge services, system architectures
- Heterogeneous computing architectures: non x86 CPUs and GPUs

Software portability and the success in integrating accelerators will play an important role

Need a more strategic approach for the use of HPCs





Heterogeneous Architectures

Heterogeneous architectures: complementing CPU capacity with accelerators (e.g. GPUs)

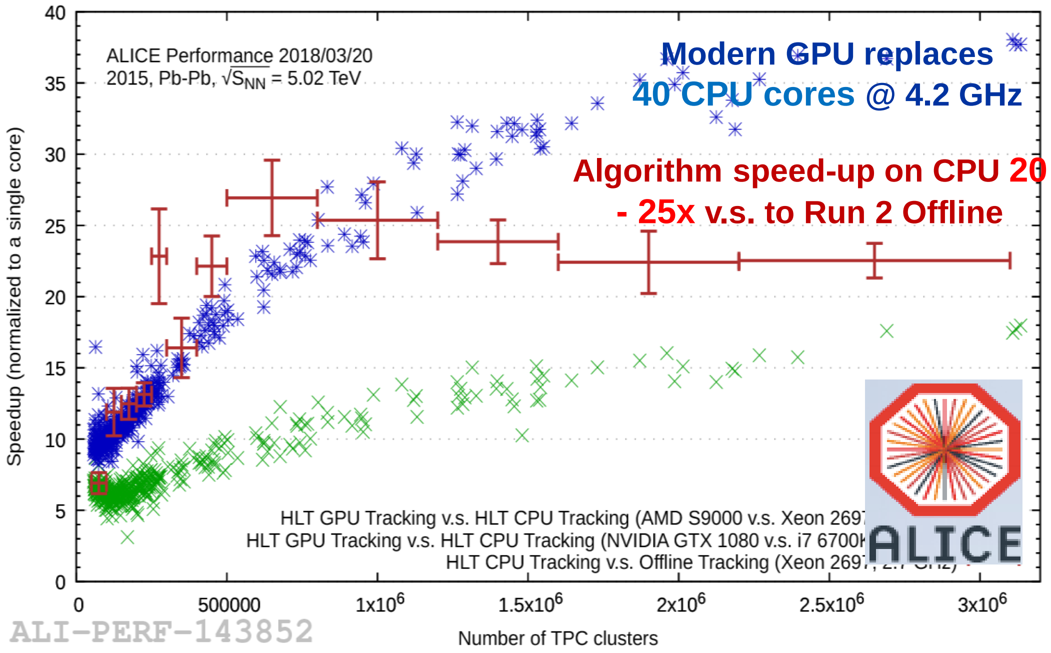
Playing a fundamental role in Run 3 already, in most online systems. Non exhaustive examples:

LHCb: exploitation of heterogeneous architectures:

- for partial reconstruction in Run 3
- for full reconstruction in Run 4?
- for fast reconstruction in Run 4 or Run 5 ?

Ongoing work in all experiments

Alice O2: Speed up from GPU usage + from algorithmic improvements + tuning on CPUs



CERN will continue its computing infrastructure investments

Prévessin Computing Centre (PCC)

2021

- July: Contract, but detailed design already started in April 2021
- September: Submission of planning permit request

2022

- January: Construction works started
- April: First Stone Ceremony held on the 22nd
- End 2022: Civil engineering and structural work to be completed
- December: Tender for first installation of servers to be adjudicated at the FC

2023

- Summer: Delivery and installation in PCC planned (very dependent on actual delivery delays at the time)
- 3rd quarter: Data Centre ready for commissioning
- End 2023: Inauguration Ceremony foreseen

Green computing

CERN annual electricity consumption for ≈ 1.25 TWh

- Mainly for accelerators, lot of effort to modernize and reduce consumption
- But bomputing plays important role in the overall effort of CERN to reduce energy consumption
  - Heat recovery from PCC for CERN Prévessin site
  - Heat recovery from LHC point 8:

LHCb online farm:

- 25 GWh/y to ZAC Ferney-Voltaire
- Heating for ≈ 8000 people

Sustainability through collaboration

The Belle-2 and DUNE HEP experiments leverage the same infrastructure as WLCG (and they are "associate members")

DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT

Belle II

Evolution of Scientific Computing in the next decade: HEP and beyond

WLCG Overview Board
17th December 2018

Contact: Ian Bird (Ian.Bird@cern.ch), Simone Campana (Simone.Campana@cern.ch)

Astroparticle Physics European Consortium (APPEC)

APPEC

APPEC Contribution to the European Particle Physics Strategy

December 17, 2018

Editorial Board:
S. Katsanevas, A. Masiero, T. Montaruli, J. de Kleuver, A. Haungs

Contact Person:
T. Montaruli (APPEC Chair from Jan. 1, 2019)
Email: teresa.montaruli@unige.ch

Website : http://www.appec.org

European Science Clusters

The Science Clusters are EU project launched in 2019 to link the world-class Research Infrastructures with the European Open Science Cloud (and influence its evolution)

We are part of the ESCAPE cluster (HENP, Astronomy, Astro-Particle)

The European Science Clusters produced a joint Position Statement on expectations and long-term commitment (role) in open science

Many opportunities to potentially leverage:

- Influence the direction of EOSC, EuroHPC
- Create synergies with other sciences in other clusters
- A step forward in our vision to share technologies and services with other experiments and sciences, with the same needs, and in many cases sharing the same infrastructure, for long term sustainability

ENVRI FAIR

EOSC-Life

ESCAPE

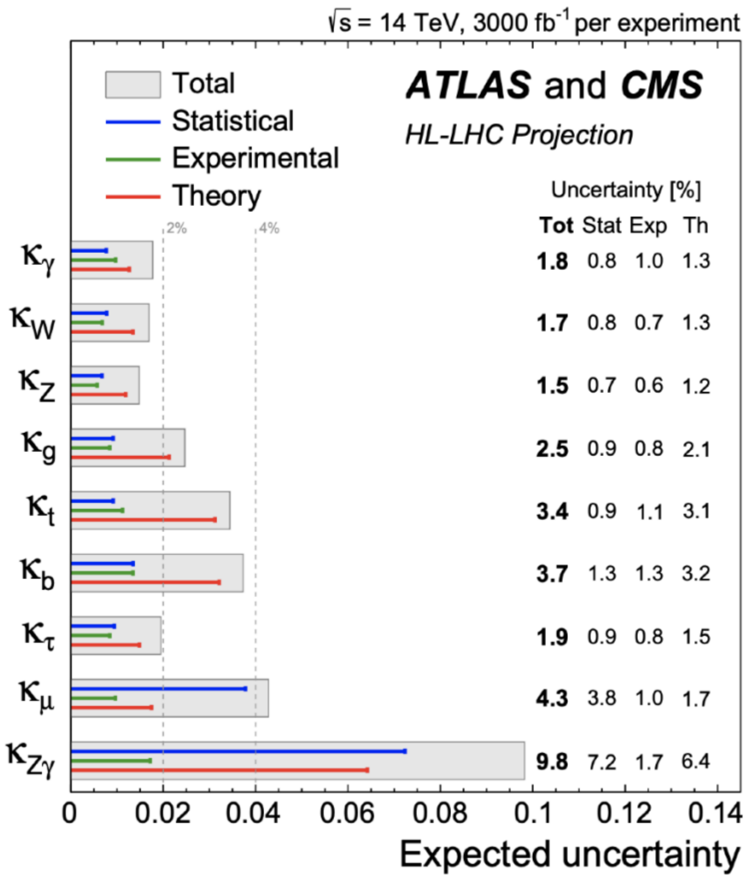
panosc

SSHOC

Software Challenges: Event Generation

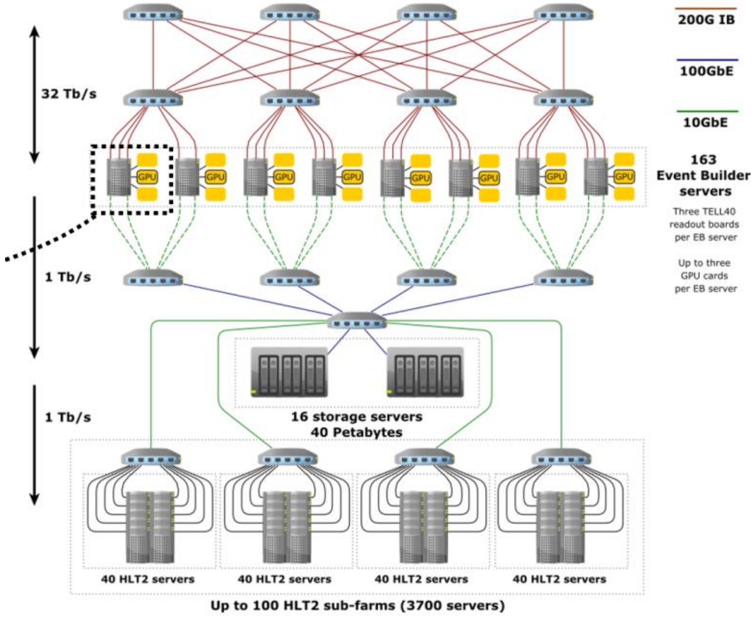
- **High luminosity LHC will require a very good control over uncertainties**
  - Theory errors are very significant
  - Need to move from LO calculations to NLO and even NNLO
  - These calculations are inherently more demanding
    - And introduce problems of negative weight events from MC@NLO matching
- **Generally there is a mismatch between the incentives for the theory community and the needs of the experimental community**
  - Work in the HSF and SWIFT-HEP is helping to bridge this gp

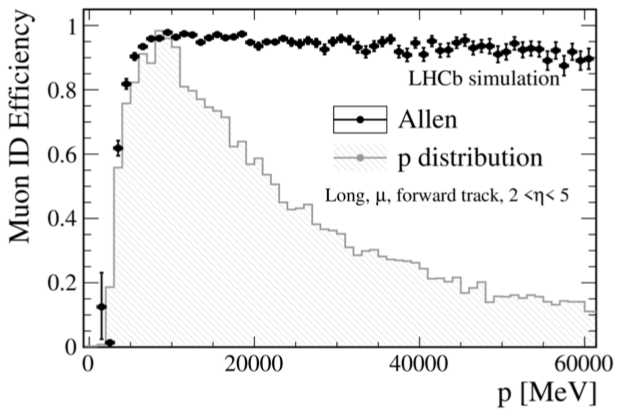
Higgs couplings at HL-LHC:
Expected uncertainties in κ formulation



Reconstruction

- **New detectors are optimised for high-rate data acquisition**
  - Physics needs require high quality selection at close to beam-crossing rates – software triggers
  - This is drinking from the fire-hose at 30 MHz (LHCb have MHz rates of charm and beauty)
- **Break with the past and rewrite software targeting GPU architectures**
  - Pioneered by ALICE in Run 2, now revamped for Run 3; being introduced by CMS too (Patatrack project)
  - LHCb have a new implementation of their HLT1 running on *GPU*, the Allen framework
    - GPUs integrated into event builder nodes, up to 3 GPUs per server
  - Lessons learned: keep data model simple, bulk data, be asynchronous, minimise data transfers
- **Allows for real-time analysis**
  - Keep partial event information for many more events recorded





Machine Learning and Other Technologies

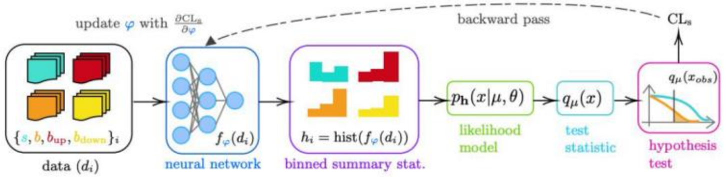
- **Machine Learning**
  - Used in HEP for a long time
  - Discrimination, classification, anomaly detection are all in use in HEP
  - Centre of cutting edge software outside HEP, developed and supported by industry
    - Our models are generally very simple, cf. GPT-3 (autoregressive language generator) that has 175 billion parameters

Table 1 | Effect of machine learning on the discovery and study of the Higgs boson

Analysis	Years of data collection	Sensitivity without machine learning	Sensitivity with machine learning	Ratio of P values	Additional data required
CMS[24] $H \to \gamma\gamma$	2011–2012	2.2σ, $P = 0.014$	2.7σ, $P = 0.0035$	4.0	51%
ATLAS[43] $H \to \tau^+\tau^-$	2011–2012	2.5σ, $P = 0.0062$	3.4σ, $P = 0.00034$	18	85%
ATLAS[99] $VH \to bb$	2011–2012	1.9σ, $P = 0.029$	2.5σ, $P = 0.0062$	4.7	73%
ATLAS[41] $VH \to bb$	2015–2016	2.8σ, $P = 0.0026$	3.0σ, $P = 0.00135$	1.9	15%
CMS[100] $VH \to bb$	2011–2012	1.4σ, $P = 0.081$	2.1σ, $P = 0.018$	4.5	125%

Radovic et al., [10.1038/s41586-018-0361-2]

- **Auto-differentiation**
  - Neural networks (amongst other things) are differentiable
  - Can evaluate the relationship between underlying calculations and observables in both directions
    - Requires the development of *surrogate models* for non-differentiable parts
  - But with this could do an end-to-end optimisation of everything from analysis to experiment design (MODE talk, Pietro Vischia)



Auto-differentiation adds a backwards pass to optimise an analysis, Simpson and Heinrich [2203.05570]

- **Quantum Computing**
  - Definitely a very hot topic
  - Remains to be seen if this will really be a revolution for software (and on what timescale)
    - Doing theory calculations directly on a quantum computer would be really exciting!

CERN QUANTUM TECHNOLOGY INITIATIVE

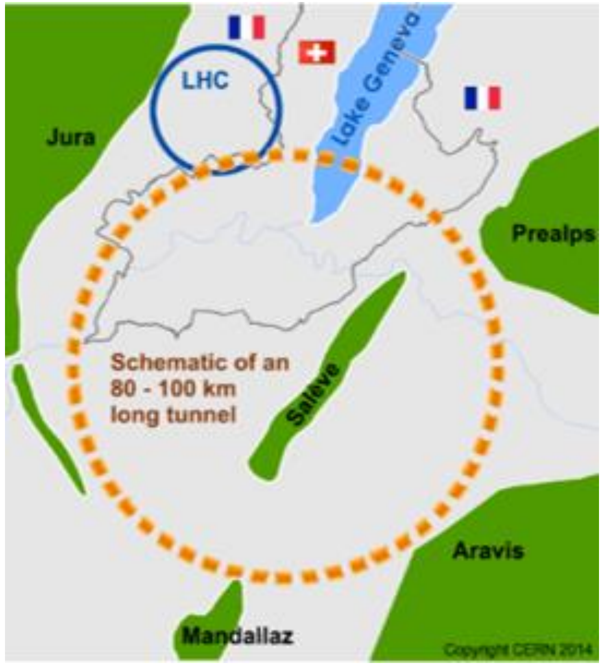
Outlook beyond HL-LHC: Future Circular Collider (FCC)

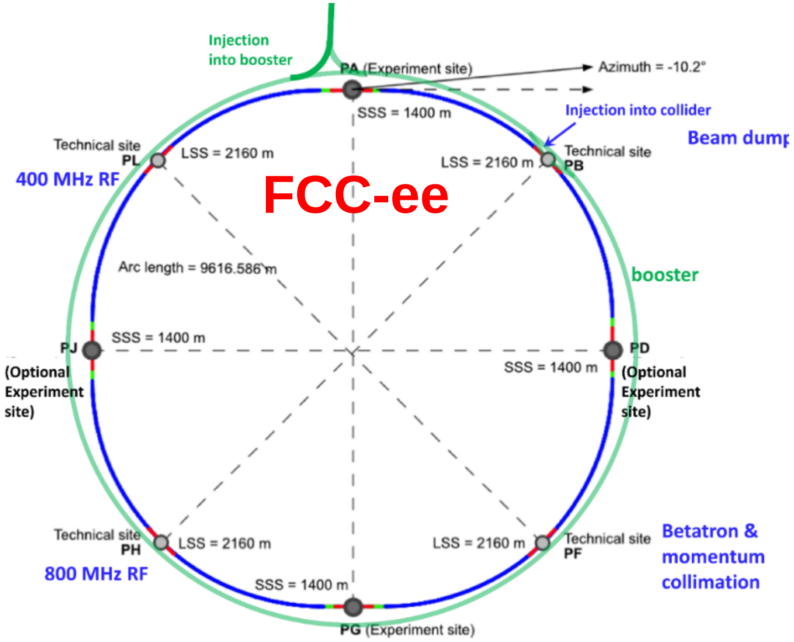
The FCC integrated program inspired by successful LEP – LHC programs at CERN

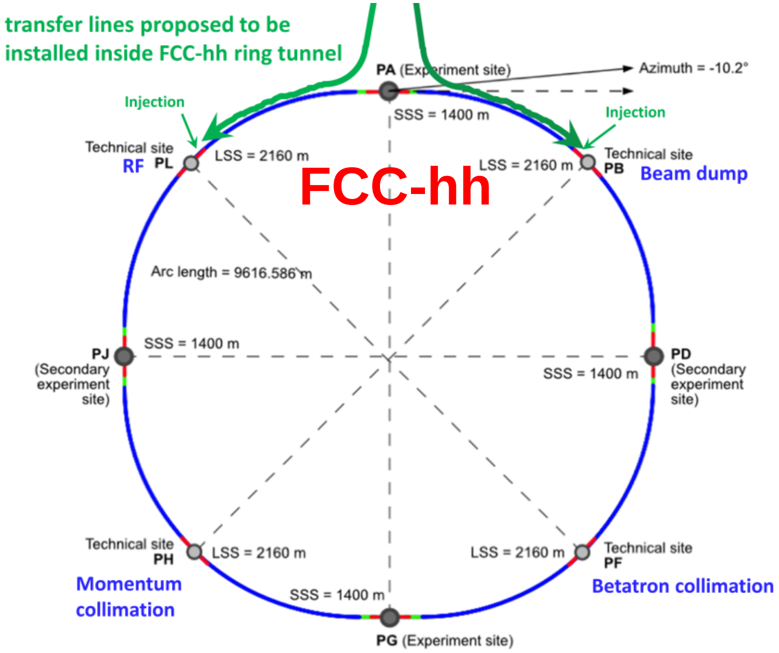
M. Benedikt
April 2022

comprehensive long-term program maximizing physics opportunities

- **stage 1: FCC-ee (Z, W, H, $t\bar{t}$) as Higgs factory, electroweak & top factory at highest luminosities**
- **stage 2: FCC-hh (~100 TeV) as natural continuation at energy frontier, with ion and eh options**
- complementary physics
- common civil engineering and technical infrastructures, building on and reusing CERN's existing infrastructure
- FCC integrated project allows seamless continuation of HEP after completion of the HL-LHC program







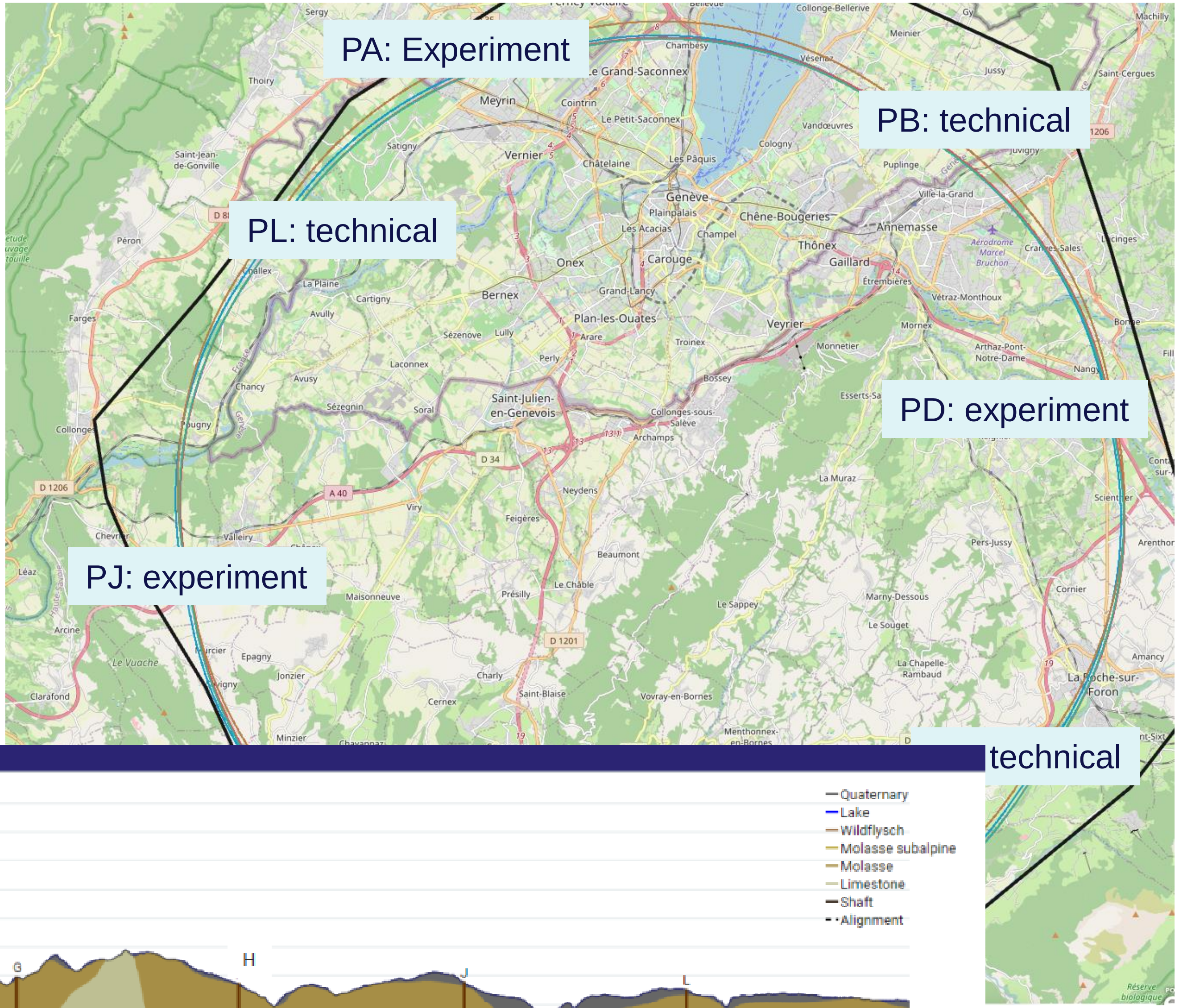
2020 - 2040 | **2045 - 2060** | **2065 - 2090**

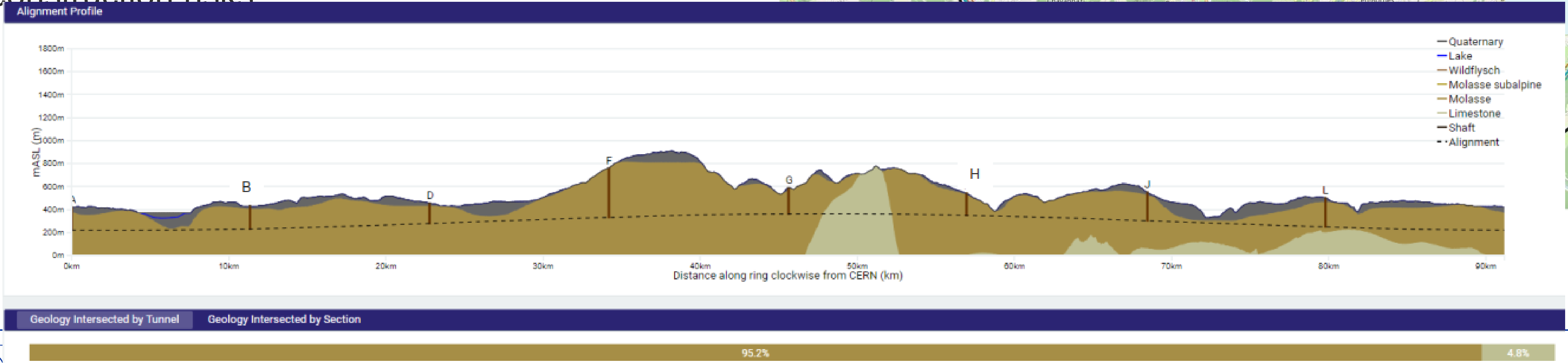
FCC Baseline Design

- 8 surface sites
- C = 91.2 km
- 4-fold symmetry and 4-fold superperiodicity
  - FCC-ee 2 or 4 IPs
  - FCC-hh 4 IPs

Present implementation variant was established considering:

- Geological 3D model and tunnelling risks
- 95% in molasse geology for minimising tunnel construction risks

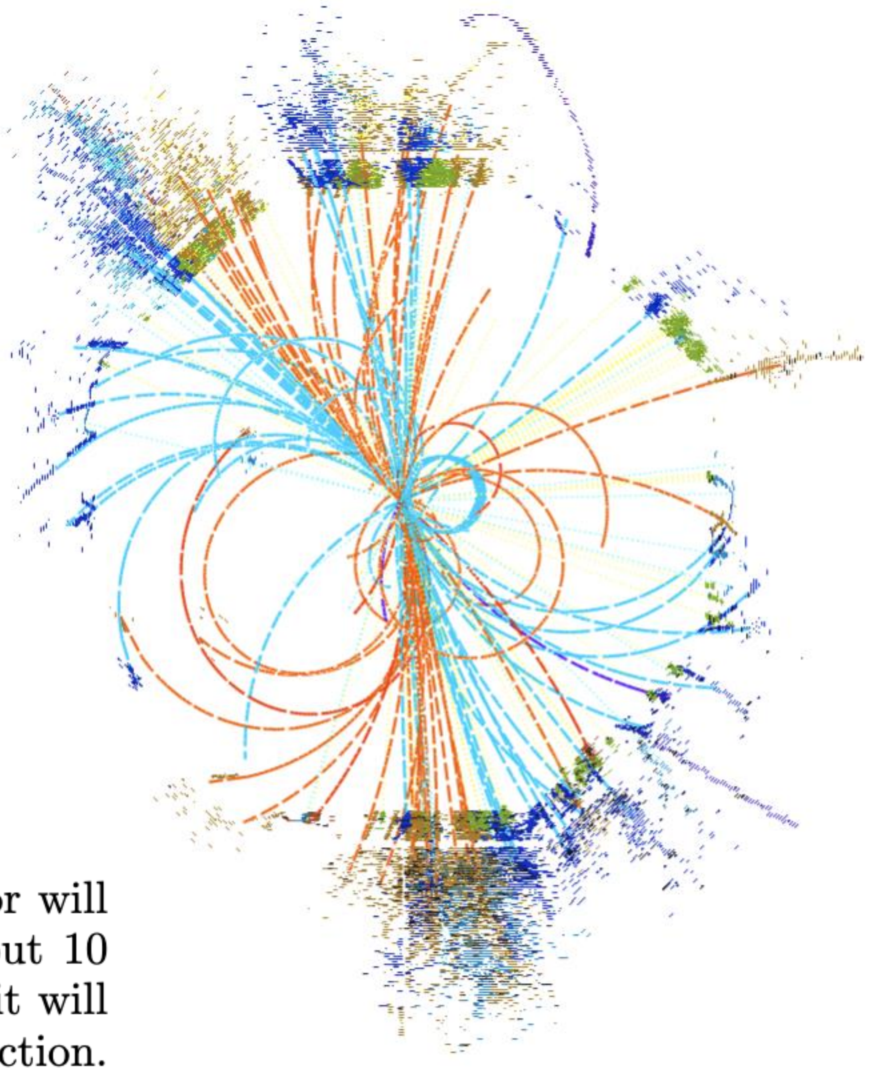




Computing for the FCC

→ **Software tools for simulation and reconstruction are already being developed**

→ **CPU and storage have not been assessed yet. It is difficult to project how the computing world will look in 25 years from now…**



FCC-ee: CERN-ACC-2018-0057

FCC-hh: CERN-ACC-2018-0058

The increased granularity and acceptance of the FCC-hh reference detector will result in about 250 TByte/s of data for calorimetry and muons systems, about 10 times more than in the ATLAS and CMS Phase-II scenarios. It is likely that it will be possible to cope with these amounts of data by the time of detector construction.

Summary

- LHC Run 3 successfully started
  - Doubling of pp luminosity until end 2025
- All experiments are taking data at 13.6 TeV
- Computing ready for Run 3
- HL-LHC and Phase II detectors upgrades proceeding
  - Big challenge, not only for computing

Congratulations to GridKa and looking forward to continuing the successful collaboration!