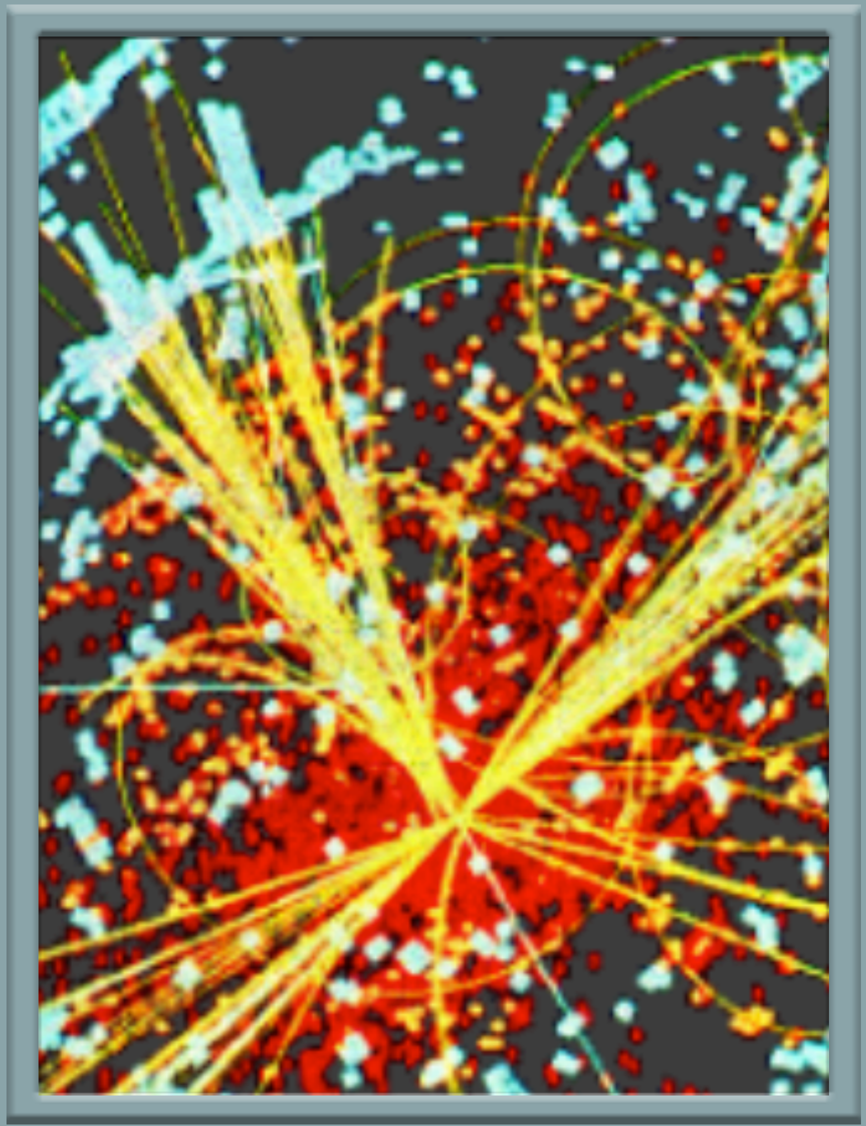
# Future challenges of HEP computing



Matthias Kasemann DESY
ESPPU Open Symposium
Granada, 15th May 2019

# Acknowledgement


- Material contributed my many:
  - L.Betev, I.Bird, T.Boccalli, C.Bozzi, V.Bric, S.Campana, J. Catmore, D.Costanzo, Y. Kato, M.Kluthe, M.Litmaath, A.McNab, B.Riedel, D.Schultz, F.Würthwein and more


- ESPP papers on computing:

[ID-5] "A European Data Science Institute for Fundamental Physics"
[ID-53] "HEP Computing Evolution" (WLCG Overview Board)
[ID-59] "Initial INFN input on the update of the European Strategy for Particle Physics: software and computing" (INFN)
[ID-79] "The Importance of Software and Computing to Particle Physics" (HEP Software Foundation)
[ID-114] "Monte Carlo event generators for high energy particle physics event simulation" (MCnetITN3 network)
[ID-126] "Deep Underground Neutrino Experiment (DUNE)" (also relevant for computing)
[ID-128] "Quantum Computing for High Energy Physics" (CERN Openlab)
[ID-150] "APS division of particle and fields white paper" (US input)
[ID-162] "The Importance of Research-Industry Collaborations on Emerging Technologies towards Exascale Computing" (CERN Openlab)

- All errors are mine.



matthias.Kasemann@desy.de

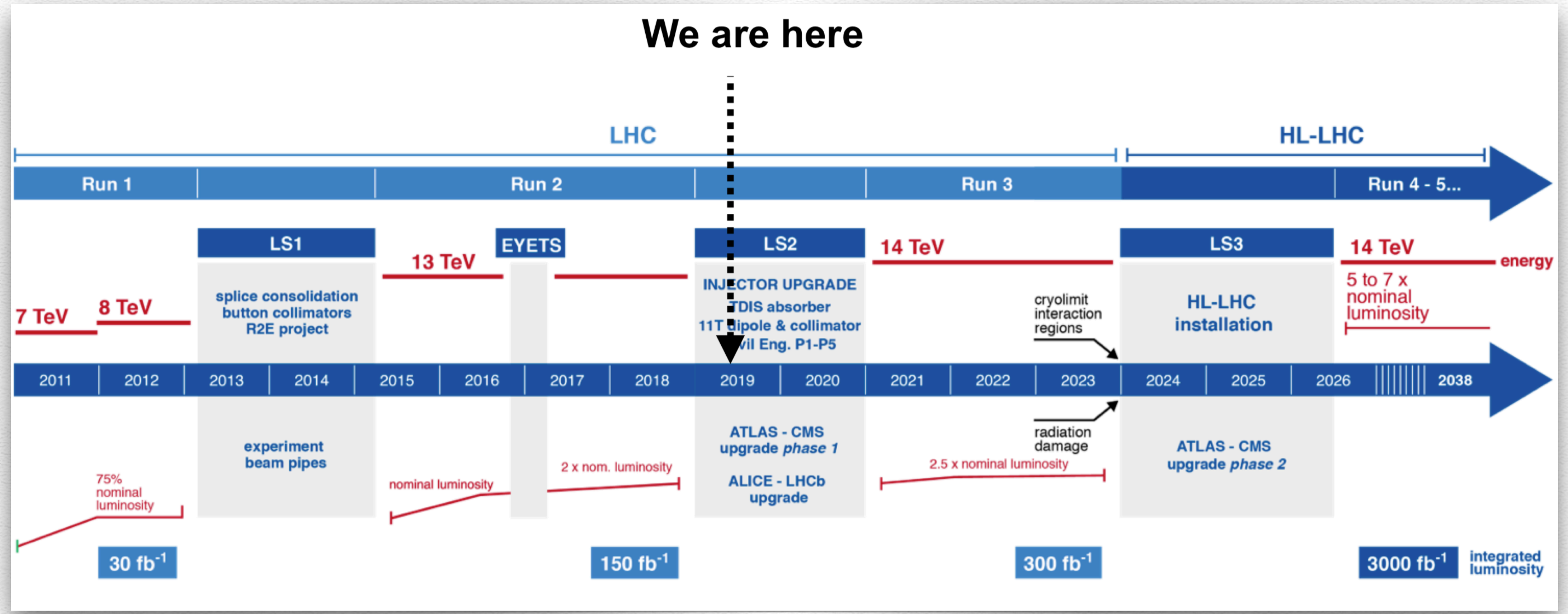
# Topics

- Where are we - where will we be
- Challenges - resources
  - LHC experiments: Run 3 and HL-LHC
  - Belle
  - Other HEP, Astro-Particle and Astronomy experiments
- Challenges - other

# LHC upgrade to High Luminosity

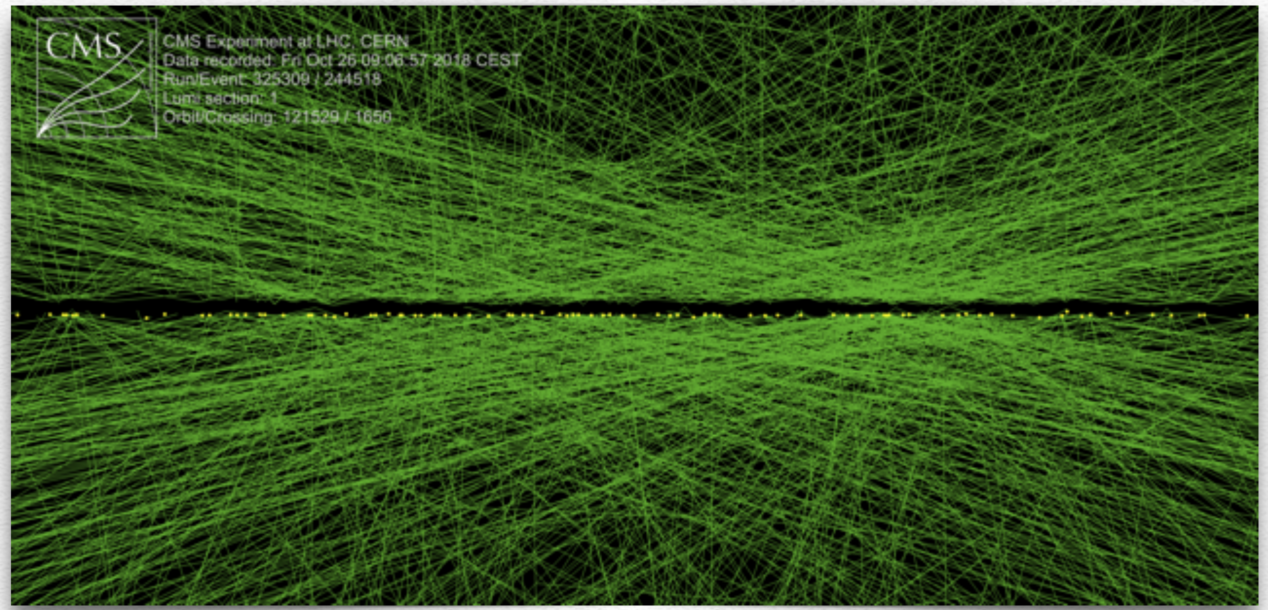
- The accelerator will be upgraded to provide ~3-4 times higher luminosity by 2026
  - Luminosity:
    Phase I: < 2.2 x $10^{34}$ $cm^{-2}s^{-1}$
    Phase II: (5)7.5 x $10^{34}$ $cm^{-2}s^{-1}$
  - Planned to deliver 3-4000 $fb^{-1}$ until 2037

| | LHC | HL-LHC |
|---|---|---|
| Pileup | ~60 | ~140-200 |
| Dataset | 300/fb | 3000-4000/fb |
| Instantaneous Lumi | ~2x$10^{34}$ | 5-7.5x$10^{34}$ |



# HL-LHC data taking

- 14 TeV center of mass energy
  - 6000 primary tracks per event
- Simultaneous events (Pileup) increases from ~60 to 140-200



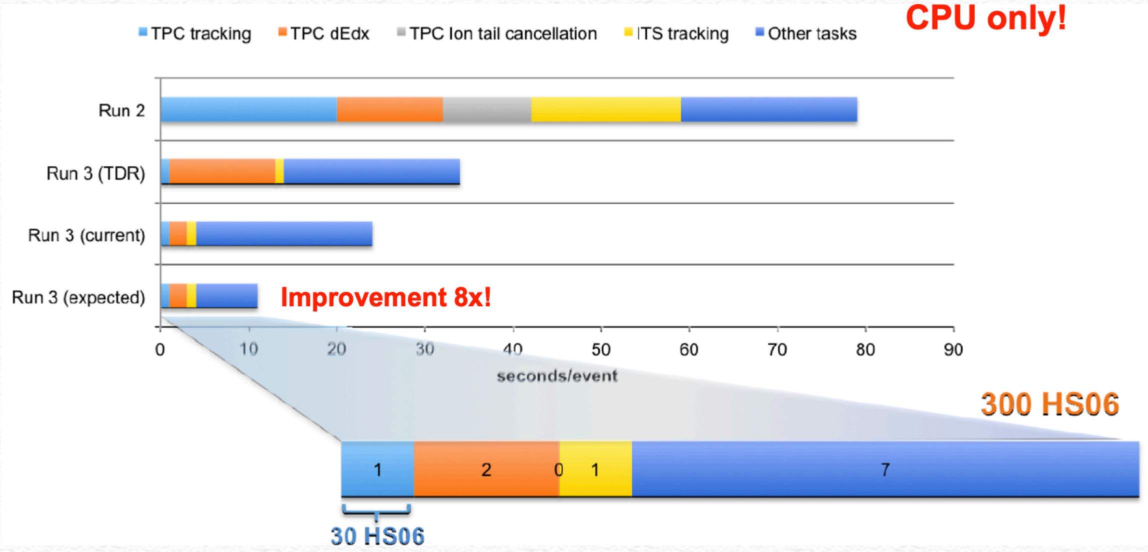

- Experiments will upgrade their detectors
  - To achieve similar performance for the new data taking conditions
  - To cope with increased trigger and data rates
  - To improve reconstruction, identification, and rejection of background
- Strategies:
  - Increased use of silicon sensors (radiation tolerant)
  - More granularity in silicon to deal with high pileup
  - Precision timing, resolution of 50 ps to separate collisions (space and time)
  - Faster processing of data in real time for trigger.

# Resources usage, estimates

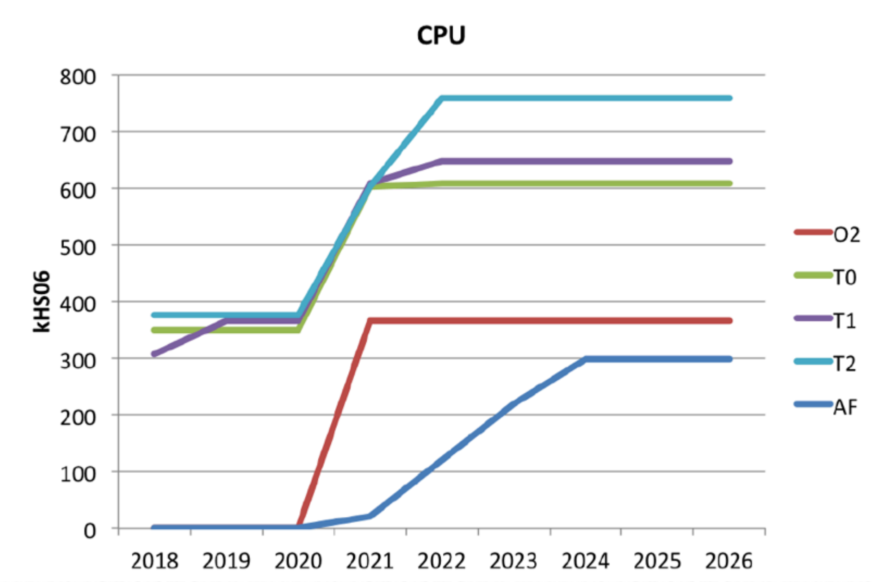
- WLCG in Run 2:
  - About 850k cores at 162 sites in 42 countries
- ATLAS, CMS in run 3:
  - evolution of Run 2,
  - Lumi-leveling will increase physics, and data volume
  - resources needs hopefully matched,
    including trigger farms, HPC computing and other
    opportunistic resources

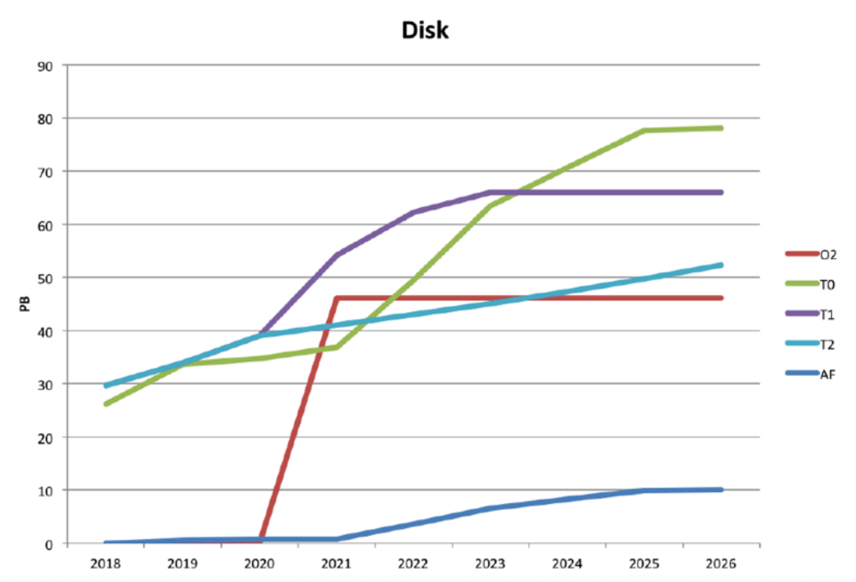
# Resources usage, estimates

- ALICE - Run 3
  - Substantial improvements achieved in reconstruction
    - Factor 8 compared to Run2



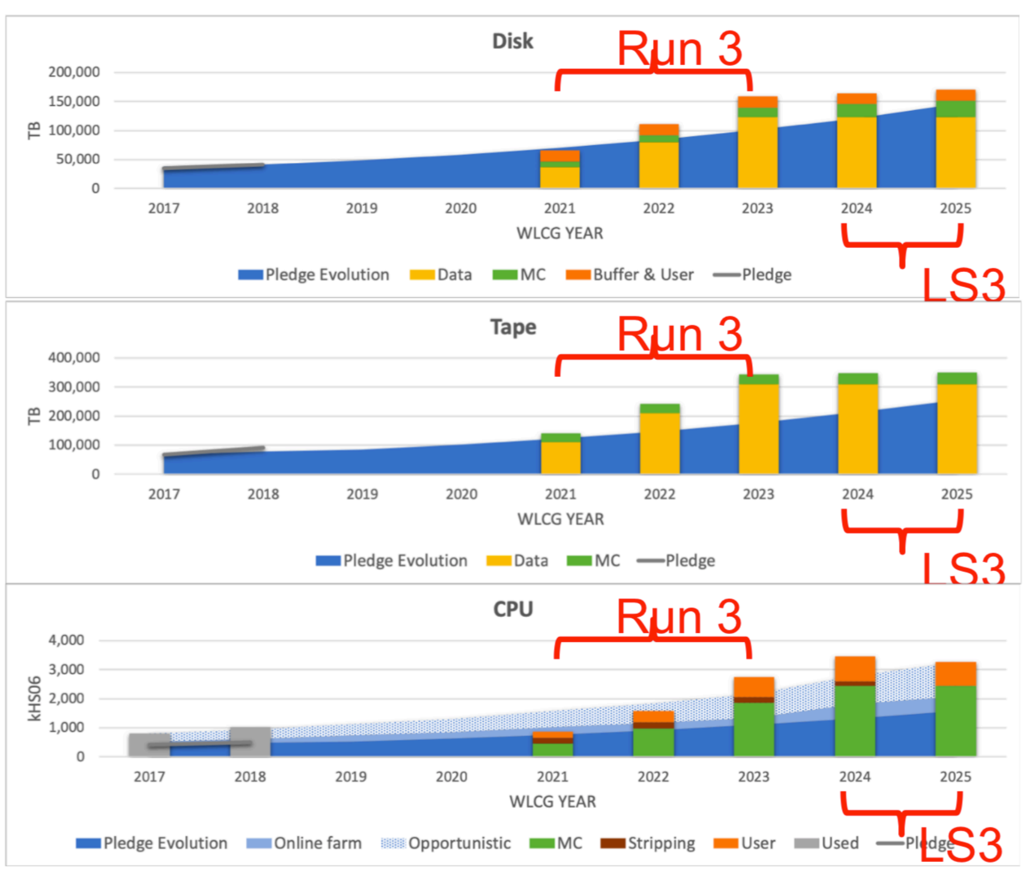

  - Resources needs are compatible with flat budget scenario



# Resources usage, estimates

- LHCb - Run 3
  - Order of magnitude more physics events from detector
  - Reconstruction done "online"
  - Offline dominated by simulation
  - WLCG resources needs almost compatible with flat budget scenario



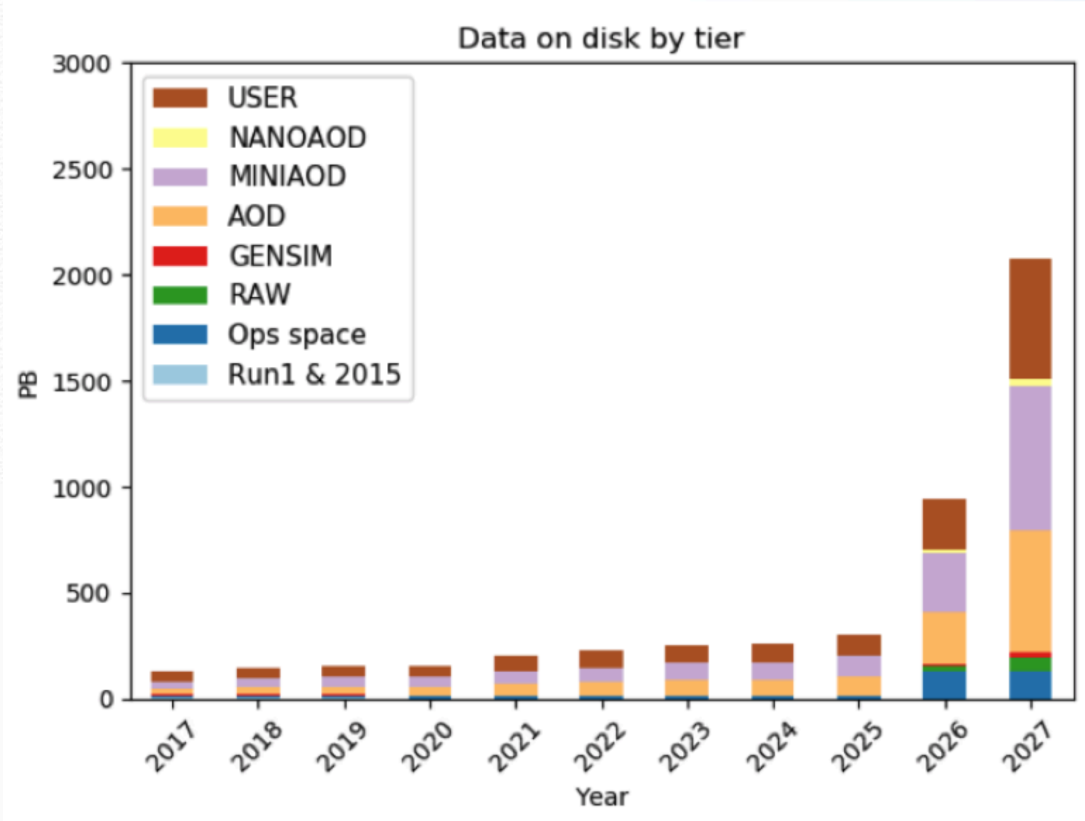

**"Making efficient use of HPC and of co-processor resources is challenging"**

- LHCb - upgraded detector for Run 4
  - accurate hit timing (~50-100 ps), high granularity, radiation hardness
  - and huge data rate

# HL-LHC - Run 4

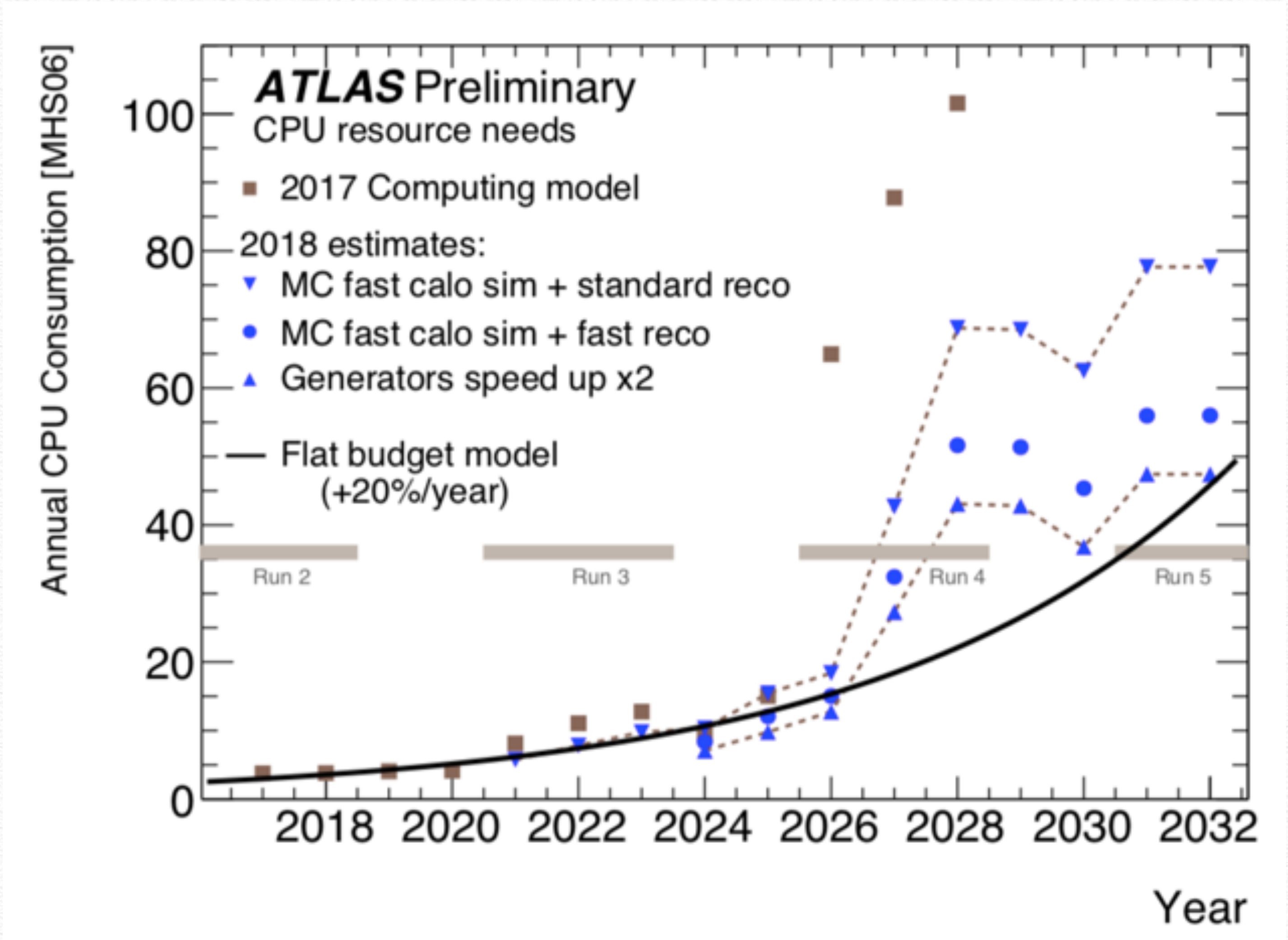
- ATLAS and CMS resources challenges
  - ~5 x Pile-up, ~7 x HLT rate,
    more readout channels,
    exponential CPU requirement for tracking
  - factors 50-100 expected on paper compared to Run 2

- Developments to reduce analysis data started
  - CMS estimates:
    - **CPU: 44 MHS06**
    - **Disk: 2.2 EB**
    - **Tape: 3 EB**
    - **(with respect to 2019 pledges, these are 22x, 13x and 15x)**



# HL-LHC Run 4

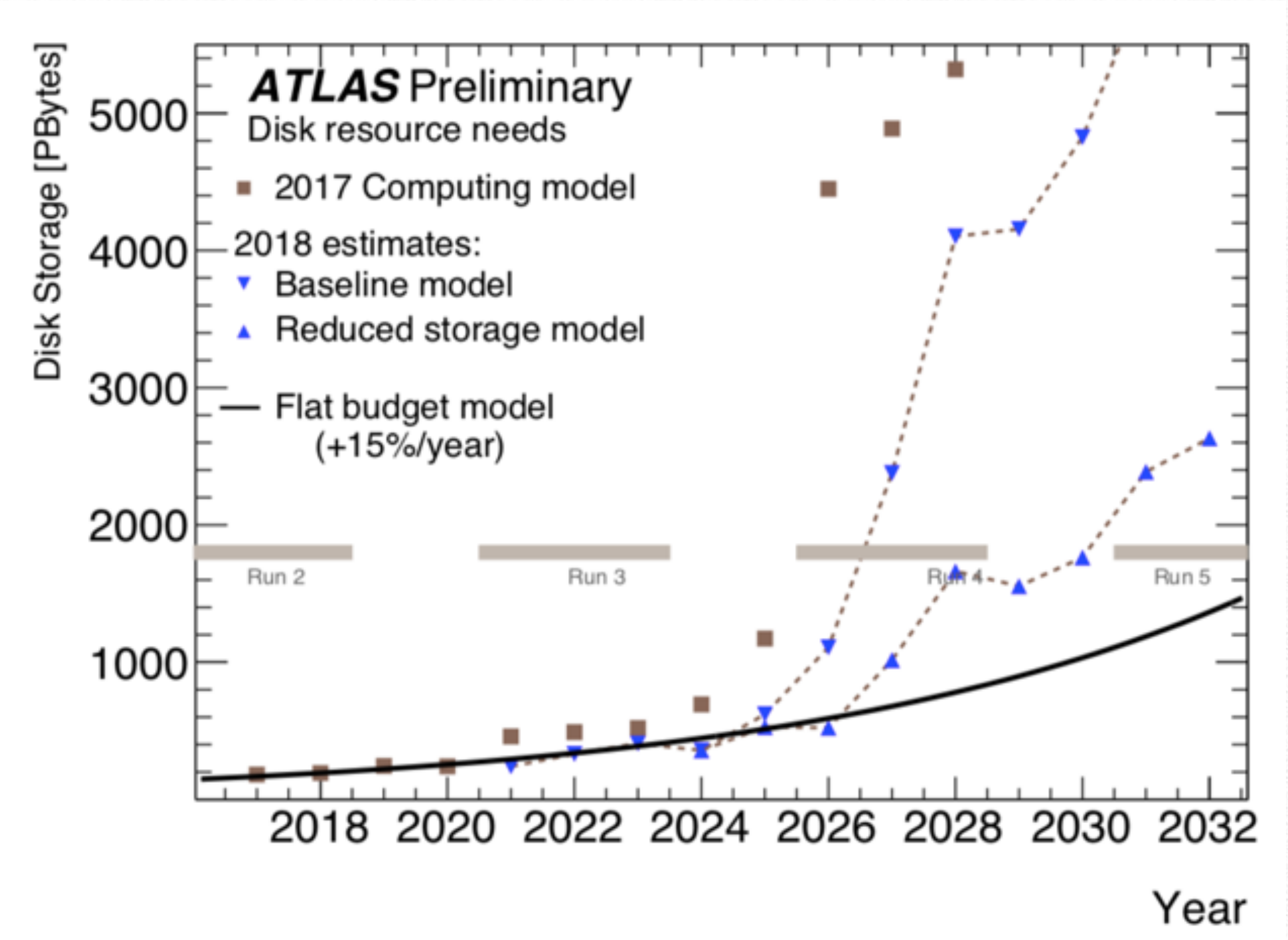
- CPU projections for Run 4




- Expected (based on experience) resources do not meet the requirements (factor up to 2)
  - "flat budget" assumption is identified risk

# HL-LHC Run 4

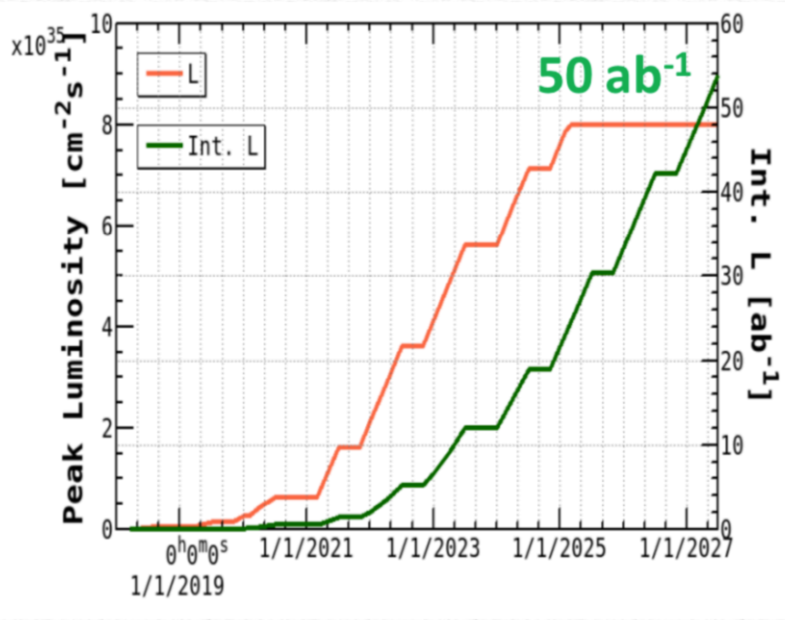
- Disk storage projections for Run 4




- Expected (based on experience) resources do not meet the requirements (factor up to 2)
  - "flat budget" assumption is identified risk

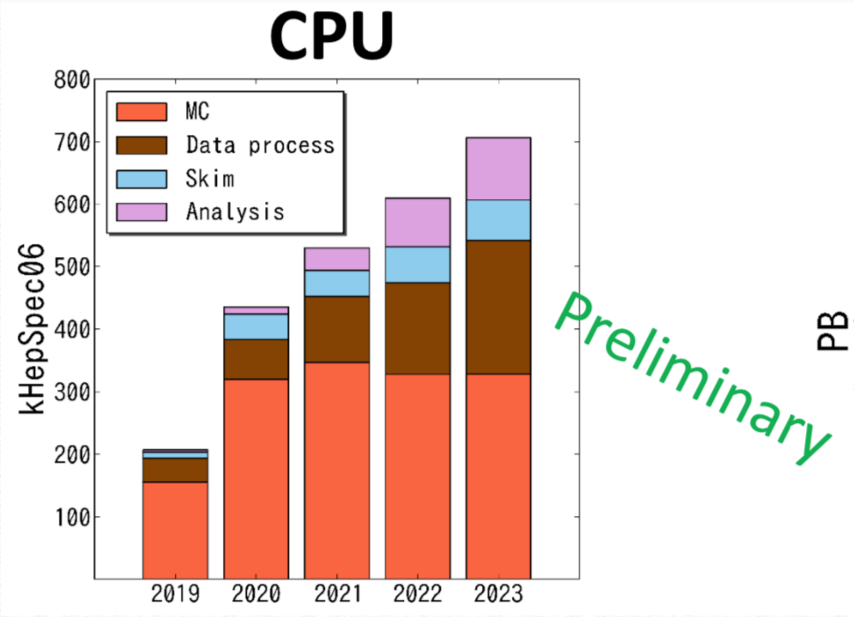
# Resources usage, estimates

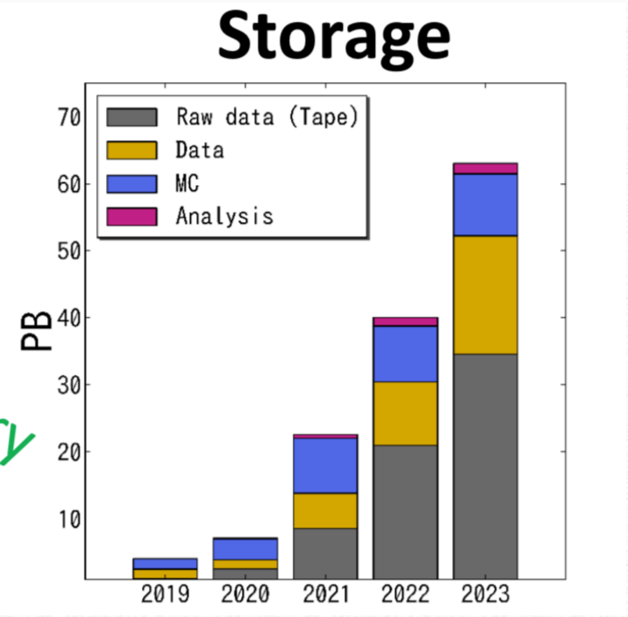
- Belle II
  - Plan to accumulate 50 $ab^{-1}$ (50x of Belle)
  - Collision data taking with full Belle II detector started




- Now: use about 20k job slots at 50 sites
- Resources needs (until 2023, 15 $ab^{-1}$)

(~similar to LHCb usage)

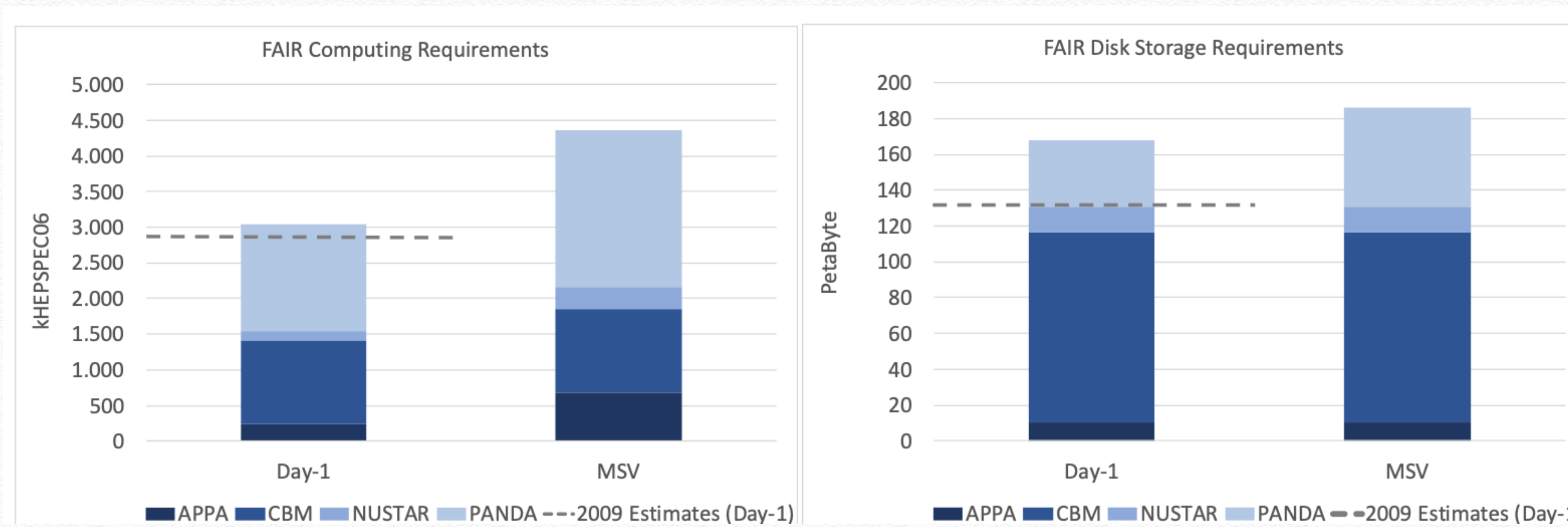


# Resources usage, estimates

- Dune
  - Foresees to produce ~70 PB/year in late 2020's

- FAIR Facility for Antiproton and Ion Research (FAIR) (under construction in Darmstadt, Germany)

- 


**Jürgen Eschke**
**GSI & FAIR GmbH**
**ESCAPE kick off meeting, Annecy, 07 Feb 2019**

# CPU and storage resources mix

- WLCG resources - typically grid computing centers
  - MoU signed with WLCG, annual pledging mechanism + review process allows planning for the experiments
- Other resources - large variety of architectures
  - HPC centers
  - Commercial resource providers ( like Clouds …)
  - Other opportunistic resources

  - All these common with various architectures, not necessarily WLCG like

# Heterogeneity challenge

- Variety of architectures and business models require large effort in software development
  - Adaptation of programs, libraries, workload and data management systems
  - All these have to be developed, commissioned and verified
- Needs a huge investment of skilled developers
  - Not currently available at the required level


matthias.Kasemann@desy.de

# Summary, Outlook

- Big challenges exist to meet the computing requirements for Run4
  - The current resources provisioning model is not sufficient
  - Collaboration with experiments beyond LHC beneficial
- Several other HEP/NP experiments require large amount of compute and storage resources
  - At smaller scale than ATLAS and CMS
- Technology, architecture changes
  - tracking is essential
  - $$$/resource unit is not a continuous and monotonic function, there will be surprises

# Summary, Outlook

- Strategies to meet future challenges are actively developed
  - In WCLG: resources provisions, coordination
    - Concepts for co-operation and coordination with other big data physics experiments is being developed
    - Data management, Middleware, networking
  - Experiments:  software development, data models
    - New tools and concepts in triggering
    - Data management, algorithms, software frameworks
    - Adaptation to use more heterogenous resources
- Needs a huge investment of skilled developers
  - Not currently available at the required level

# Backup

# Overview: Current View of HL-LHC

- Overview of "ultimate luminosity" scenario – 7.5x10³⁴, high availability: