End-user data analysis at the LHC

Nils Faltermann, Manuel Giffels, Günter Quast, Matthias Schnepf

Institute of Experimental Particle Physics (ETP)
Steinbuch Centre for Computing (SCC)

Karlsruhe Institute of Technology (KIT)

9. bwHPC-Symposium, Mannheim
October 23, 2023

Karlsruher Institut für Technologie

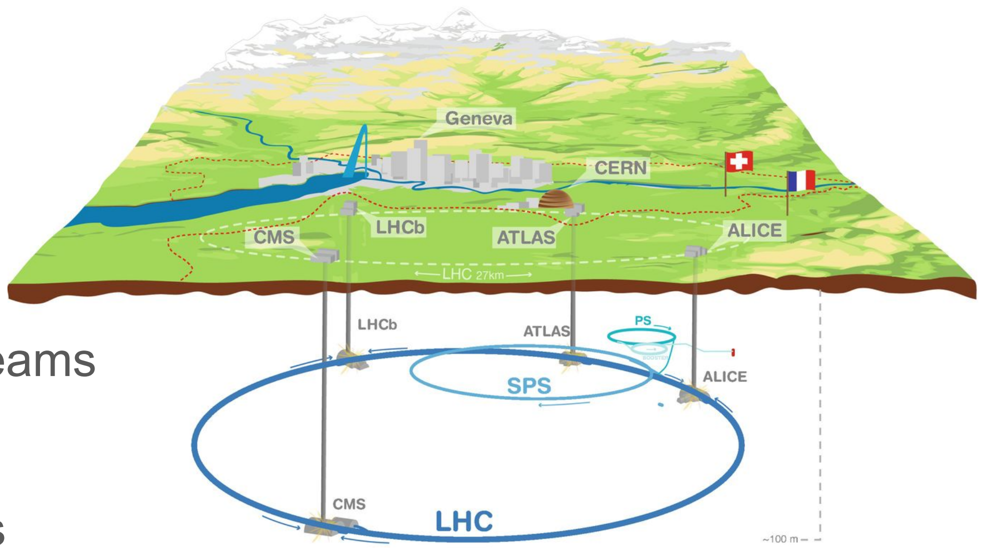
Institut für Experimentelle Teilchenphysik

bw|HPC

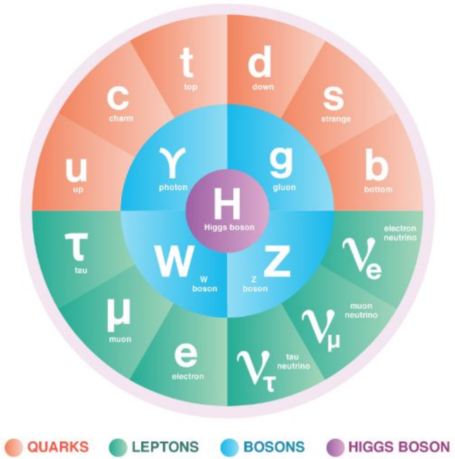
CMS Compact Muon Solenoid

Introduction

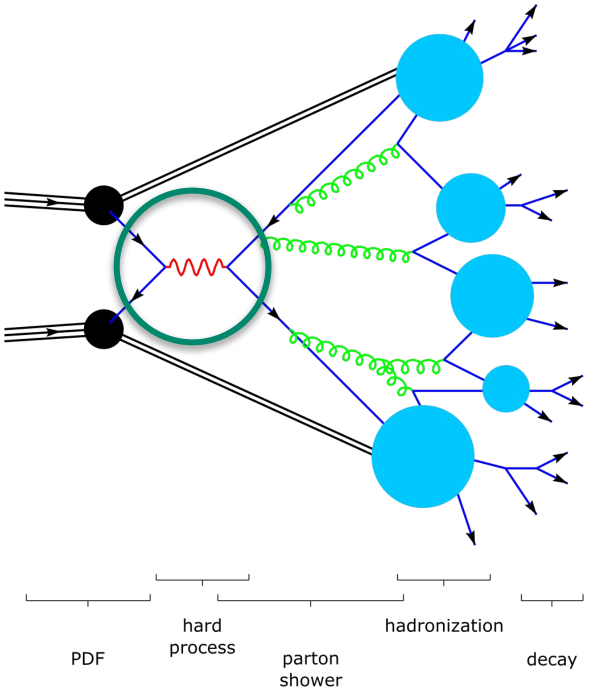
- **Large Hadron Collider (LHC)**, particle ring accelerator (27 km)
- Proton collisions via two particle beams
  - 2808 bunches, 10^{11} protons each
  - 40 MHz bunch crossing rate
- Collisions recorded by experiments



- Research goals:
  - Understanding subatomic nature
  - Verify current models and predictions
  - Search for unknown phenomena → dark matter

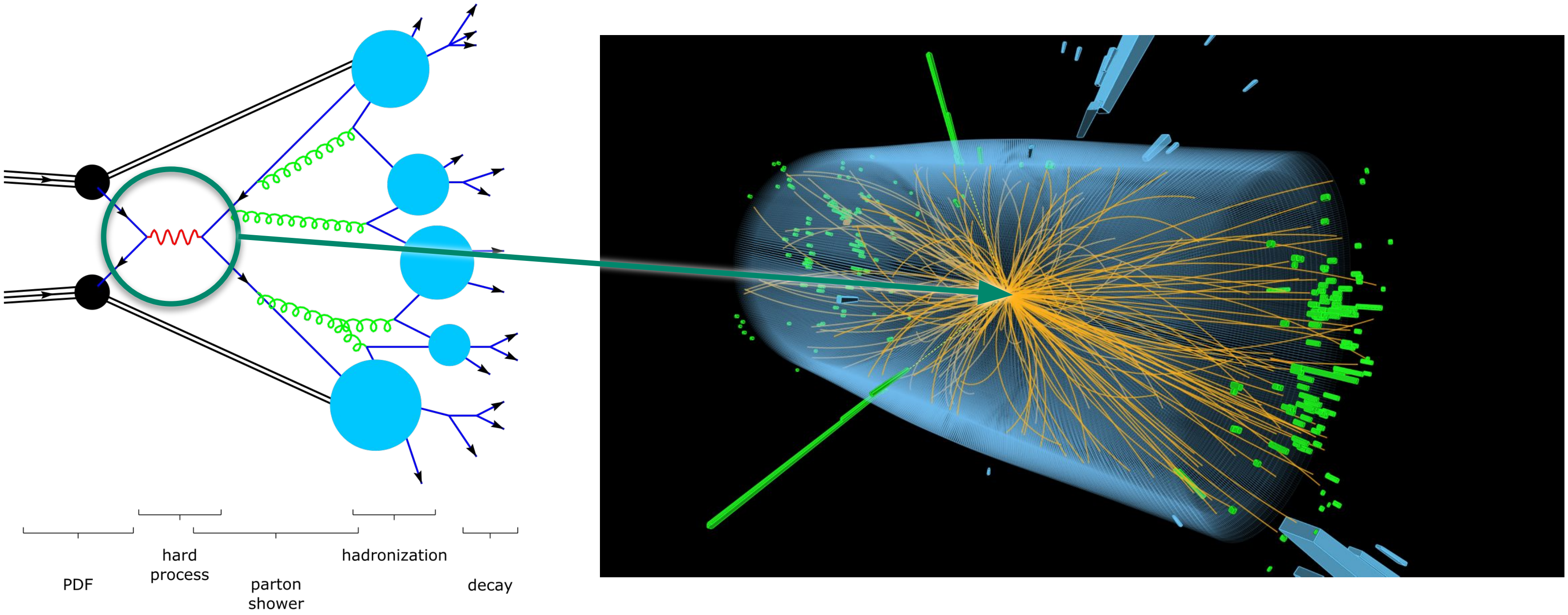


How we measure nature



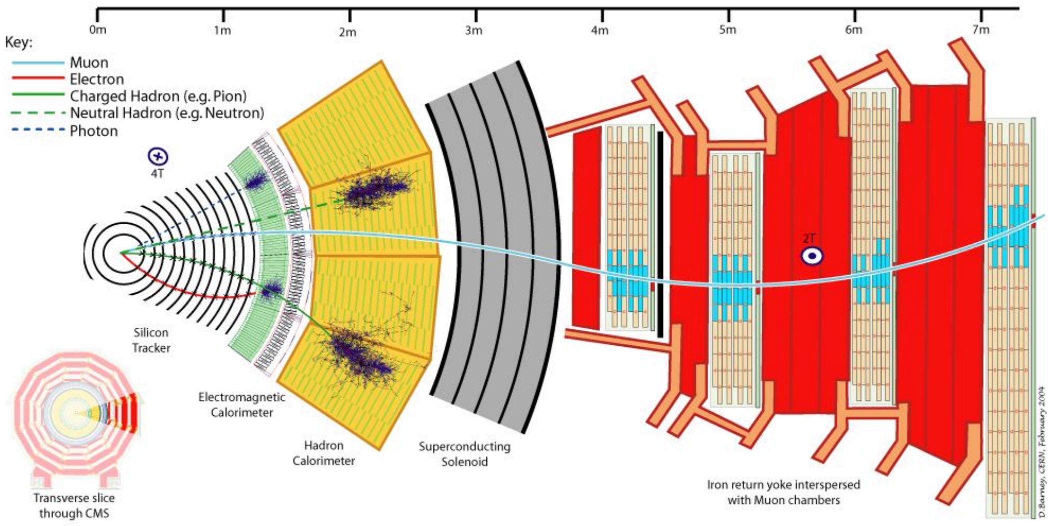
- Processes of interest rare, not directly accessible and diluted from background

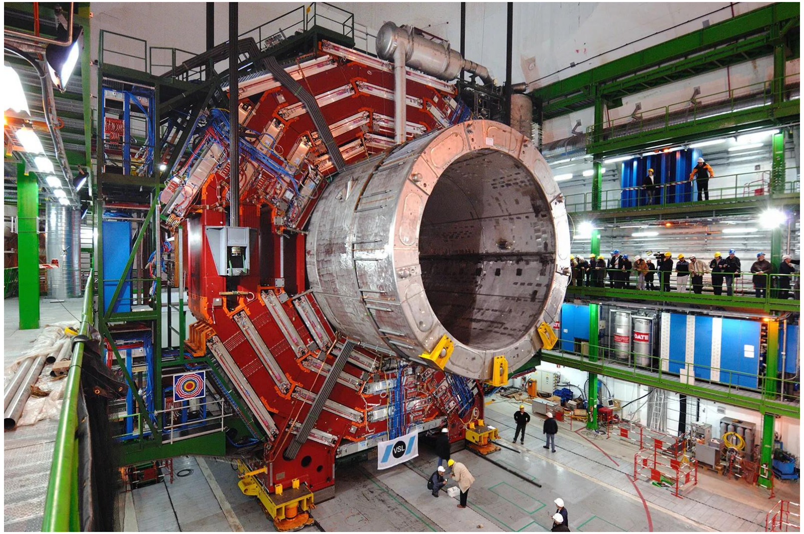
How we measure nature

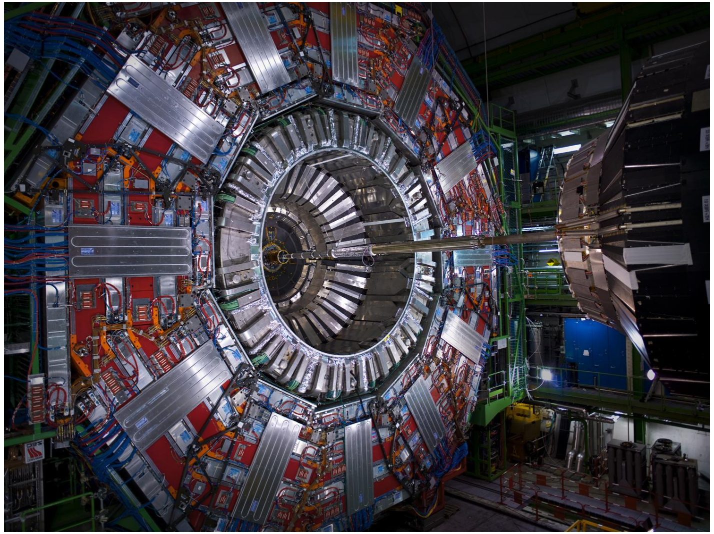


- Processes of interest rare, not directly accessible and diluted from background
- Infer underlying physics from complete detector information, “going back”

The CMS detector







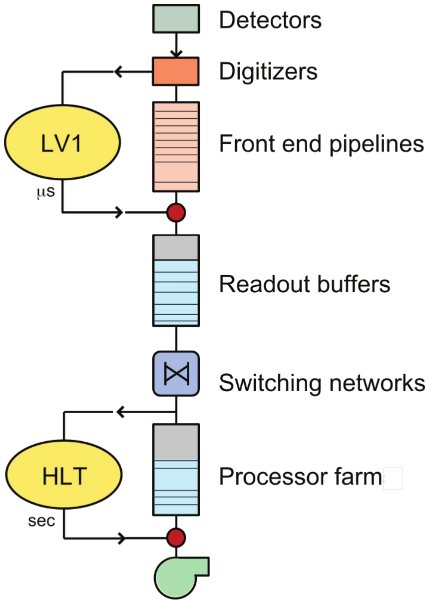
- Cylindrical general-purpose detector
  - Weight: 14,000 t
  - Dimensions: 15 m (diameter), 28.7 m (length)
  - Magnetic field: 3.8 T

Data acquisition

- 40 MHz rate + multiple collision per crossing → ~10^9 collisions/s (events)
- 55 million readout channels (~1 MB per event) in the CMS detector→ **~1 PB/s**

Data acquisition

- 40 MHz rate + multiple collision per crossing → ~10^9 collisions/s (events)
- 55 million readout channels (~1 MB per event) in the CMS detector→ **~1 PB/s**

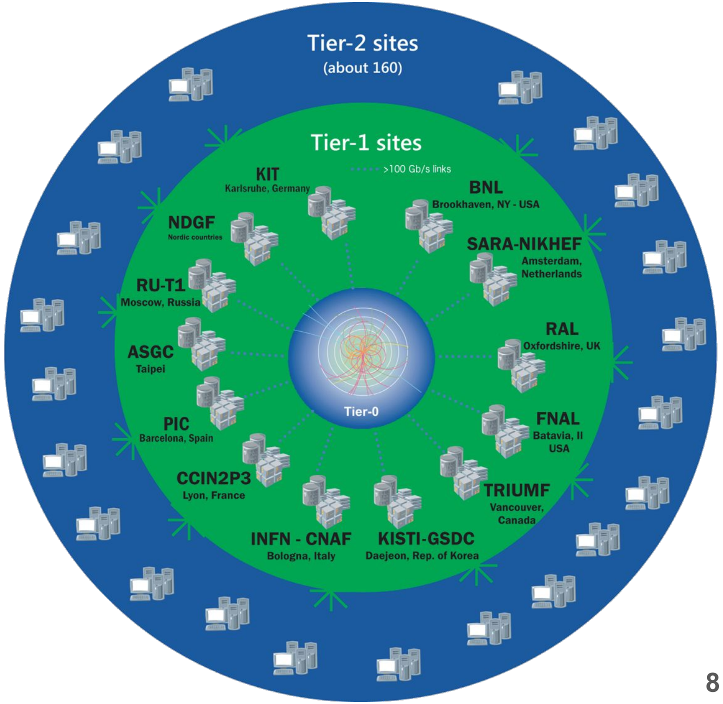


- Multi-level trigger system: store only interesting events
  - **Level 1 (L1) trigger:** custom electronics, coarse (sub)detector data, 40 MHz → 100 kHz
  - **High-Level trigger (HLT):** software on farm, partial reconstruction, 100 kHz → 2 kHz
- Assuming 200 days of data taking per year → **~20 PB/year**

Worldwide LHC Computing Grid (WLCG)

- Distributed high-throughput computing infrastructure, store and process data from experiments, tiered structure → decentralized and (mostly) homogeneous
  - **~170 sites across ~40 countries**
  - **~1M CPU cores**, disk and tape storage **>1 EB**
  - Private and overlay NREN networking, up to 400 Gb/s

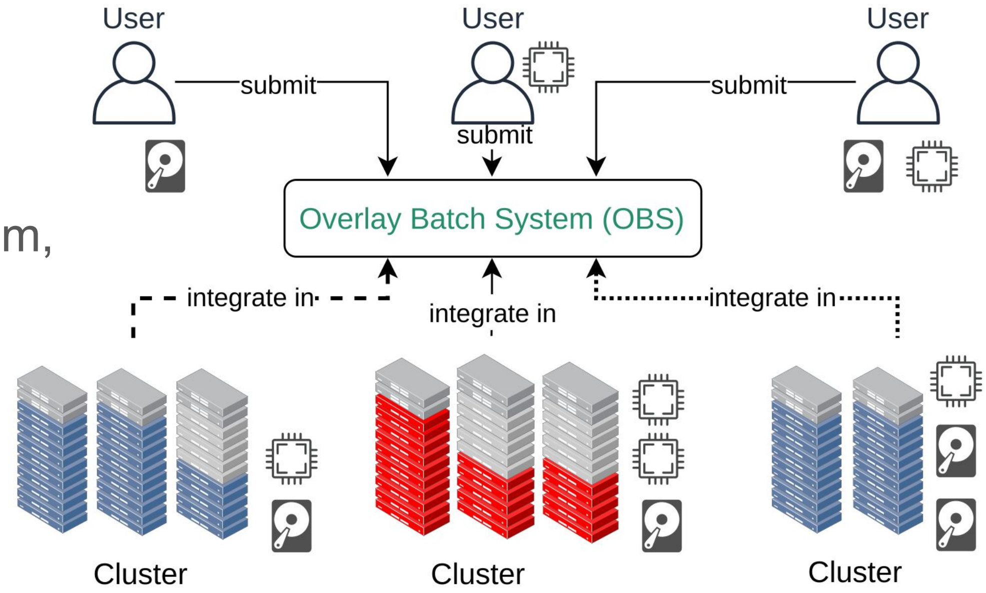
- Automated workflow management, ~100% utilization
- Only small fraction of custom user jobs



Data analysis

- Final analysis work done on so-called T3 sites → local institute clusters
- Broad and varying job requirements: cores, I/O, network

- OBS to integrate heterogeneous local and external resources
- User only interacts with one system, jobs automatically scheduled and distributed
- Never enough resources → **opportunistic resources**

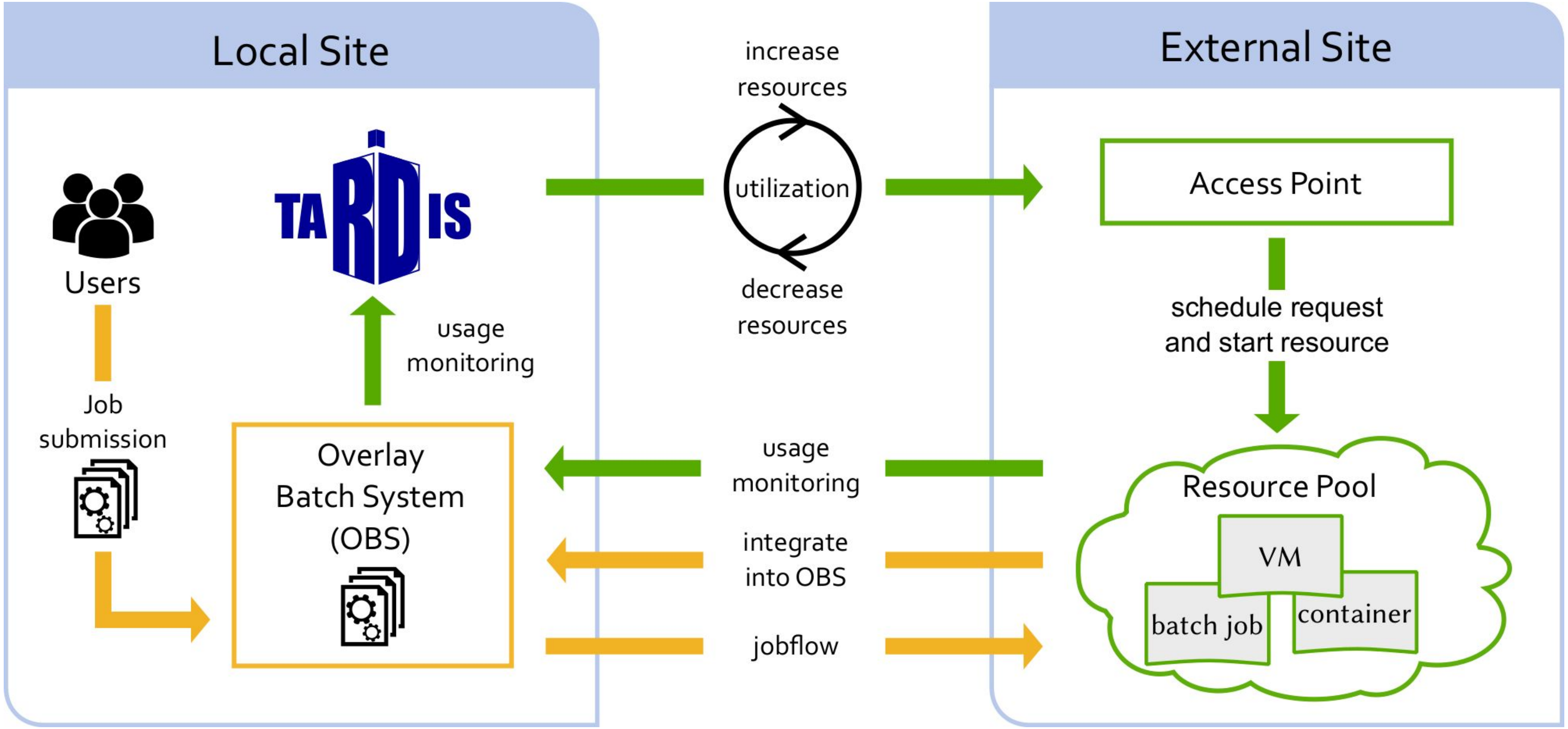


Requirements

- **Institute cluster:** ~80 researchers, not only CMS, general HEP
  - 1k cores and 2 PB storage, all static
  - Usually not enough, but sometimes no jobs on the cluster
    → economically not reasonable to extend, efficiency
  - Fluctuating demand ideal use case for opportunistic resources

- **Jobs:** mostly event-based tasks, trivial parallelization (except ML)
  - Environment usually fixed, usage of VMs and containers
  - (Pre-)compiled software and libraries, integrated to sandbox
    or streamed via CVMFS (cacheable)
  - Data shipped directly or via remote access (additional authentication)

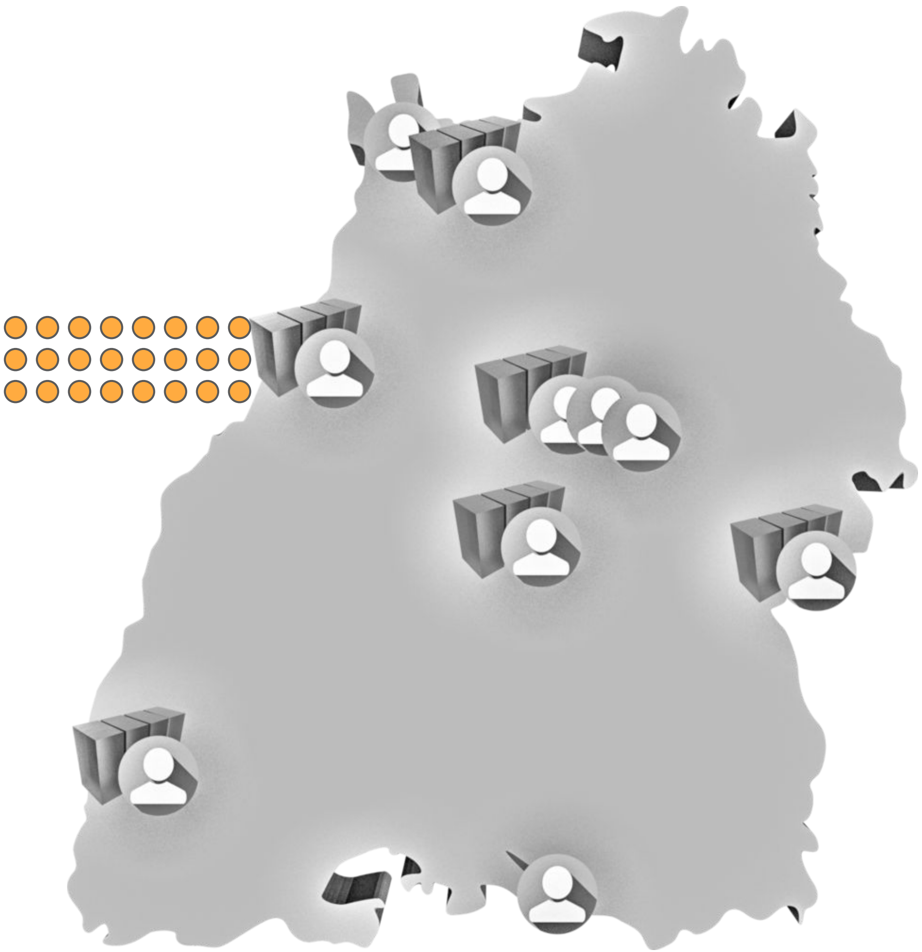
Integration of opportunistic resources



TARDIS

Example workflow

- Processing of simulation and experimental data for a physics analysis (10k cores, 10 TB data)

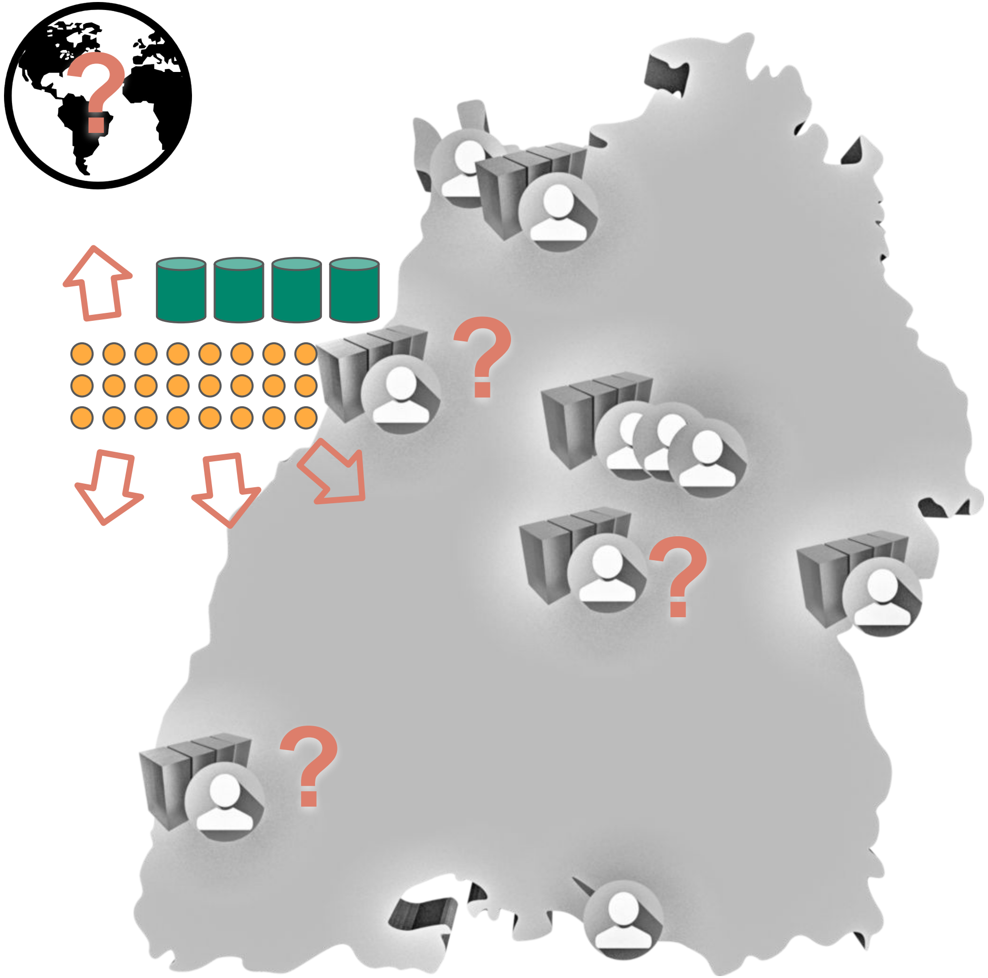


- User submits jobs to local batch system
- Specifies attributes: cores, memory, runtime, data locality

Example workflow

- Processing of simulation and experimental data for a physics analysis (10k cores, 10 TB data)

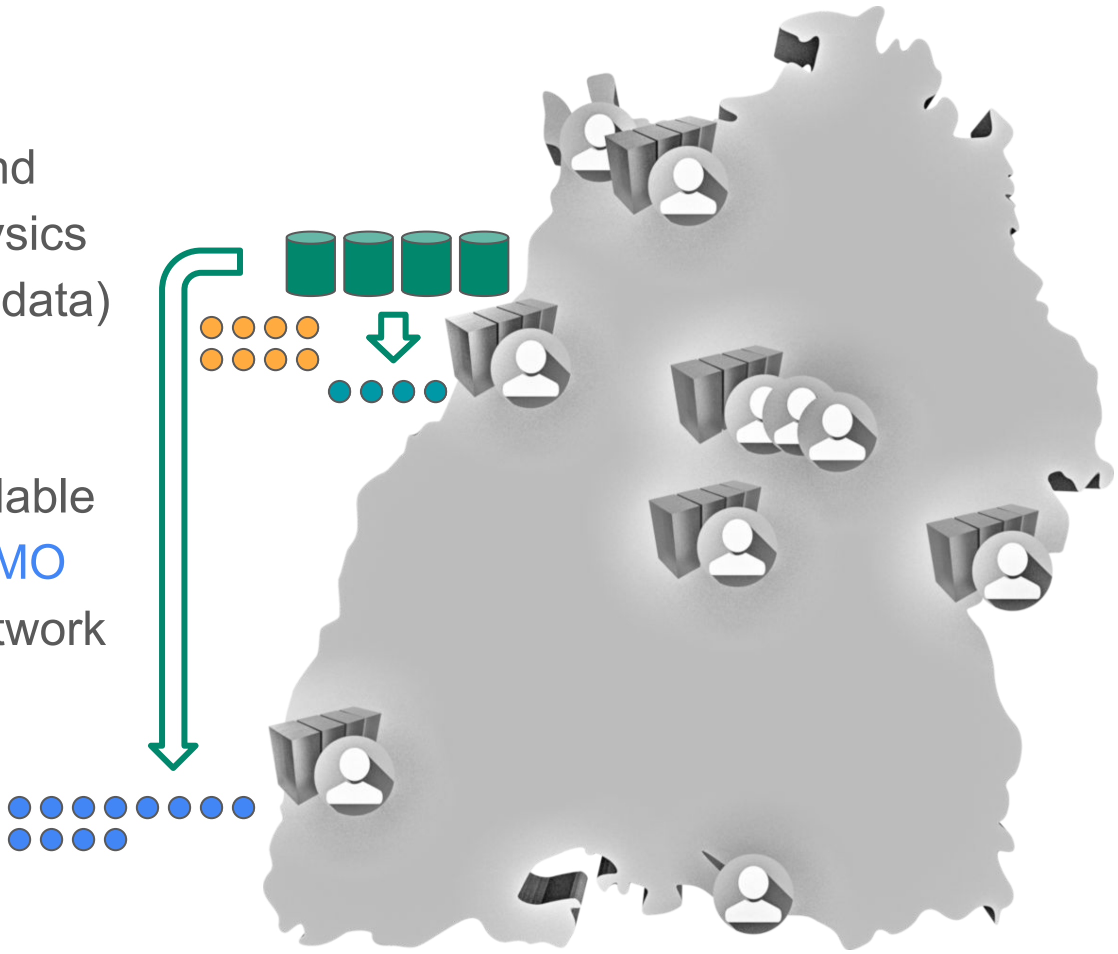
- Allocating matching resources
- Here: input data located at grid site, remote access possible



Example workflow

- Processing of simulation and experimental data for a physics analysis (10k cores, 10 TB data)

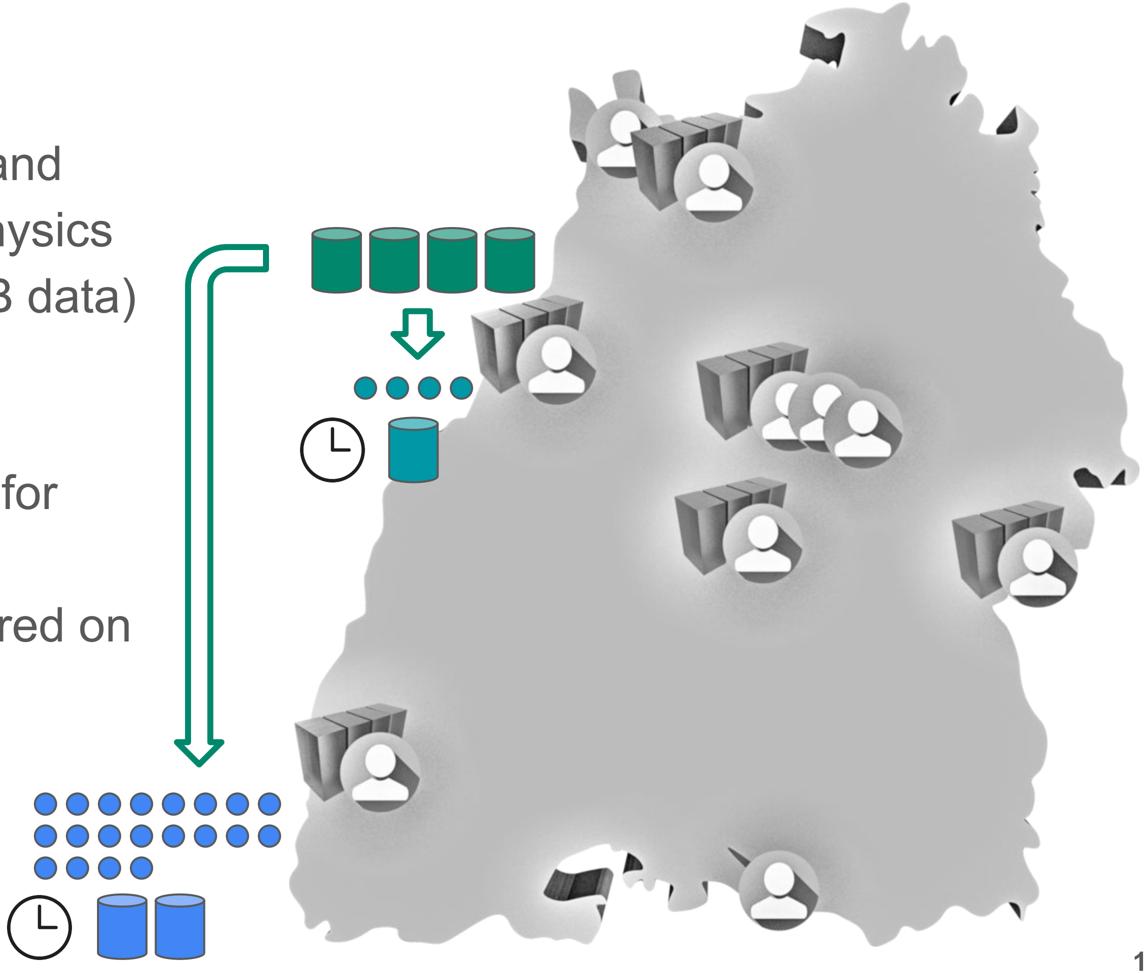
- Jobs are scheduled to available resources, e.g. local or NEMO
- Input data streamed via network (XRootD)



Example workflow

- Processing of simulation and experimental data for a physics analysis (10k cores, 10 TB data)

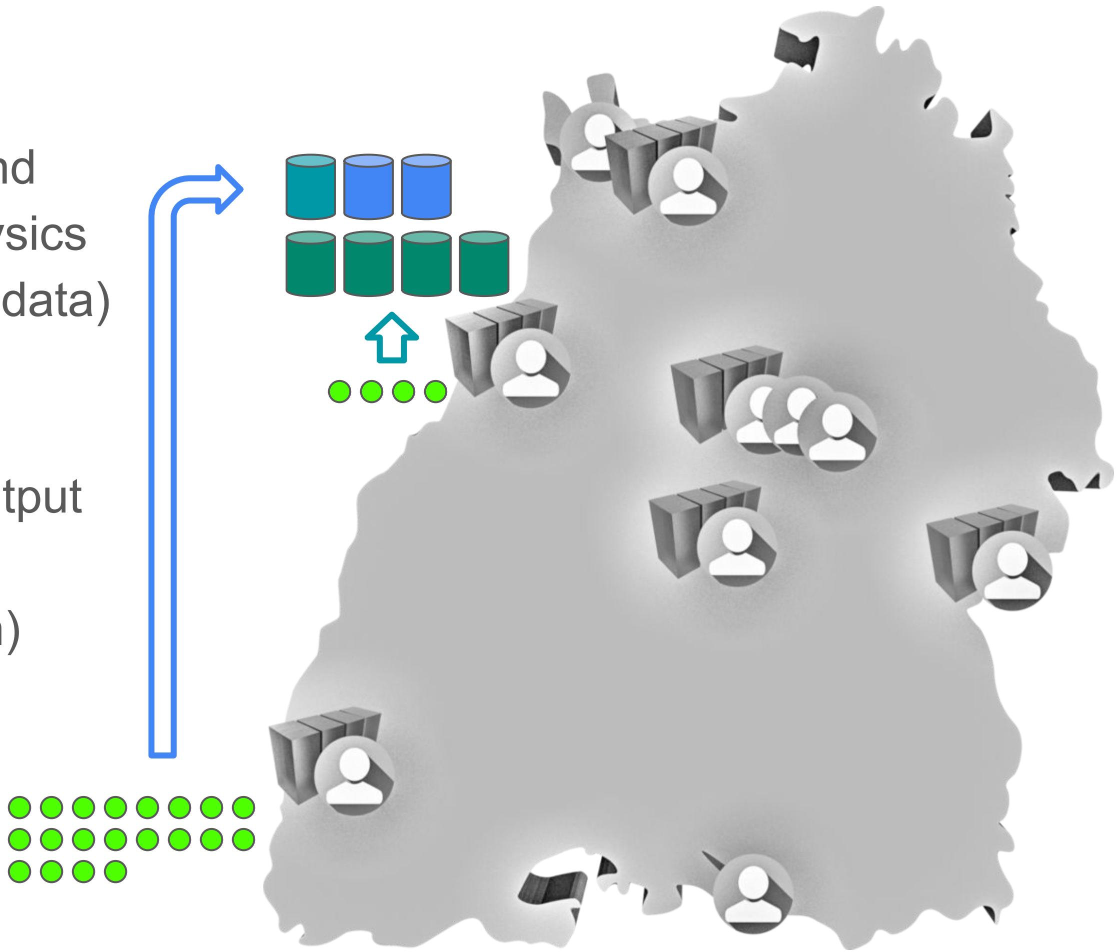
- More resources allocated for remaining jobs
- Immediate output files stored on WN scratch space



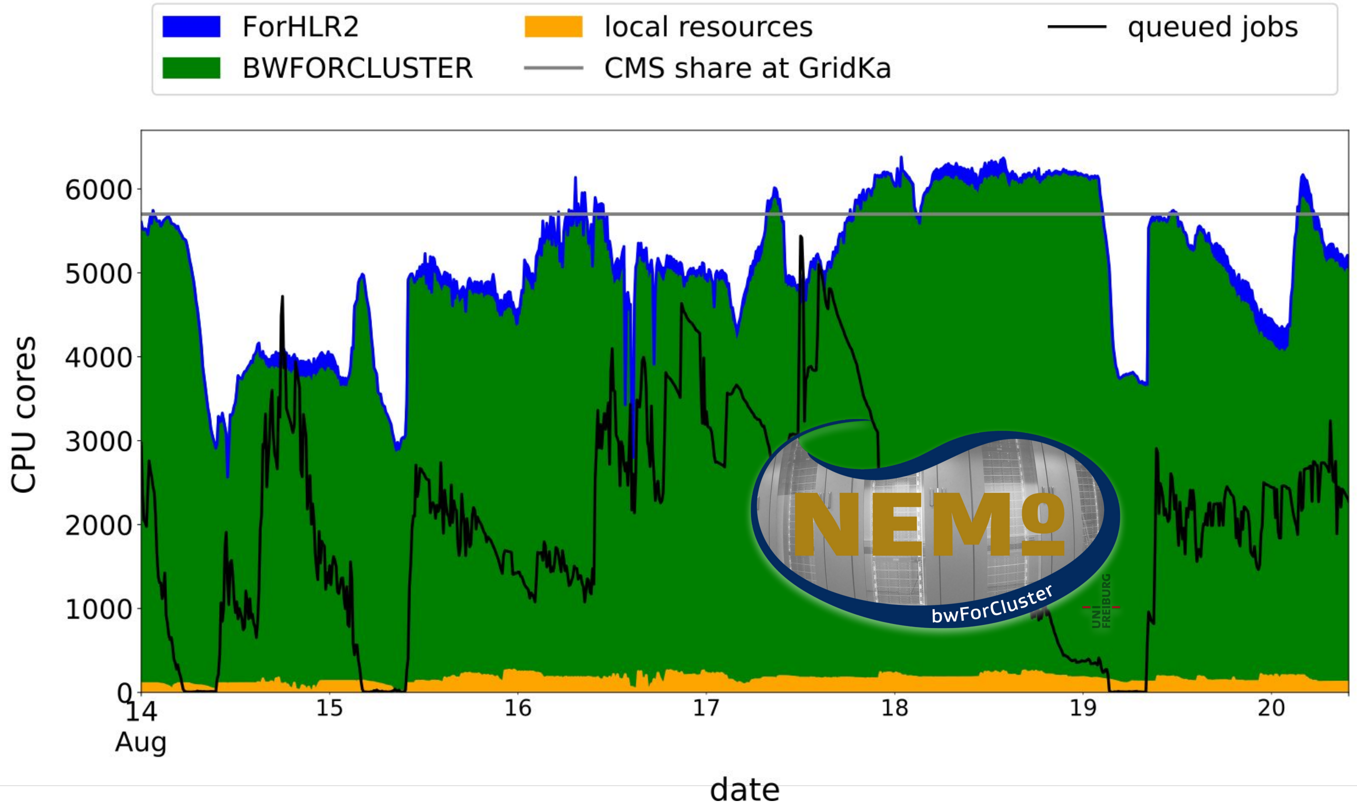
Example workflow

- Processing of simulation and experimental data for a physics analysis (10k cores, 10 TB data)

- After successful runtime output files are transferred to final storage (e.g. grid site again)



Example usage



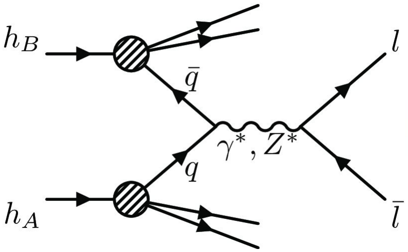
Scientific highlights

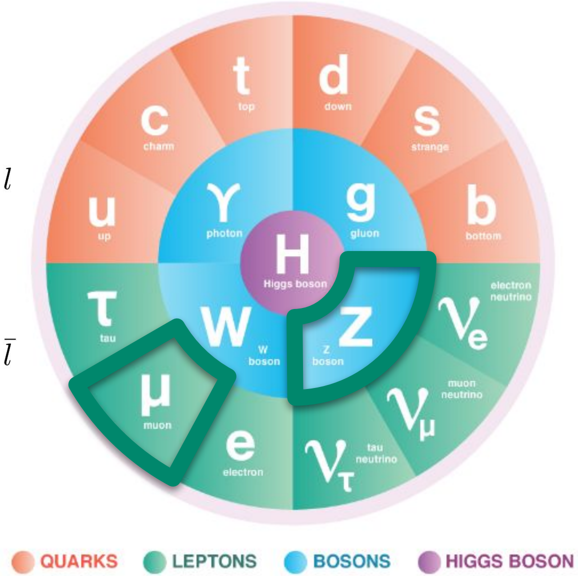
- Successful usage of opportunistic resources in physics analyses, multiple different sources over time:
  - **Research and education:** bwForCluster NEMO (Uni Freiburg), TOpAS (GridKa), GKS (GridKa), ForHLR II (KIT), bwUniCluster 2.0 (KIT)
  - **Commercial provider:** Open Telekom Cloud, Exoscale, 1&1
- Many published results utilize these resources (>50M CPU hours):

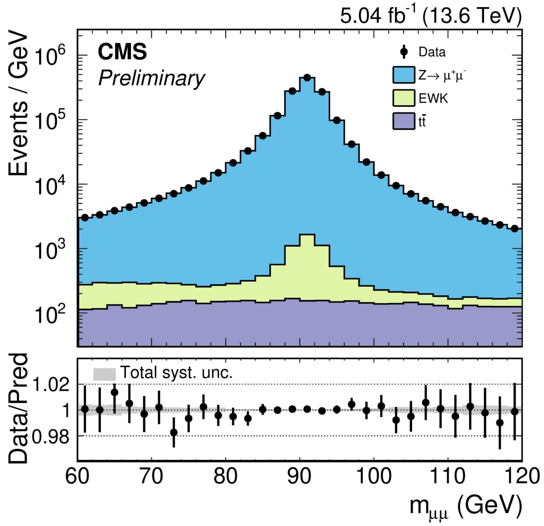
- **August 2023:** *Measurement of the inclusive cross section of Z boson production in pp collisions at $\sqrt{s}$ = 13.6 TeV*, CMS Collaboration, CMS-PAS-SMP-22-017, to be submitted to PLB
- **August 2023:** *Measurement of the ttH and tH production rates in the H → bb decay channel with 138 fb^{-1} of proton-proton collision data at $\sqrt{s}$ = 13 TeV*, CMS Collaboration, CMS-PAS-HIG-19-011, to be submitted to JHEP
- **July 2023:** *Searches for additional Higgs bosons and for vector leptoquarks in ττ final states in proton-proton collisions at $\sqrt{s}$ = 13 TeV,* CMS Collaboration, JHEP 07 (2023) 073
- >20 more journal publications in recent years + many to follow

Scientific highlights

- Measurement in early data of new LHC data taking run
- “Standard candle” → used to calibrate many other processes



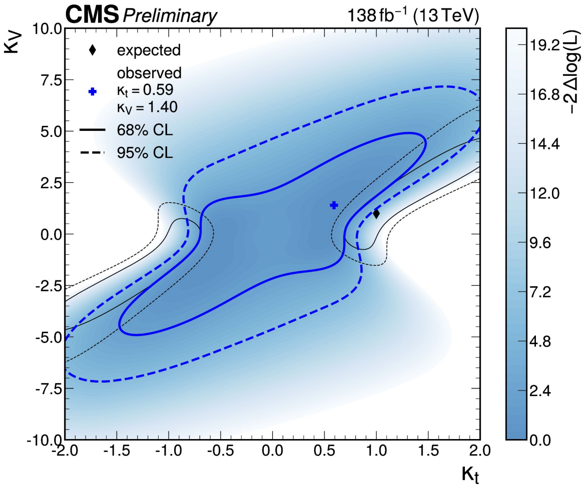


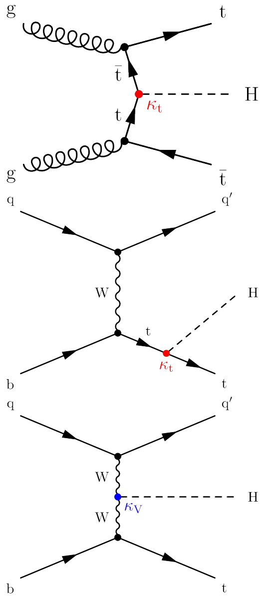


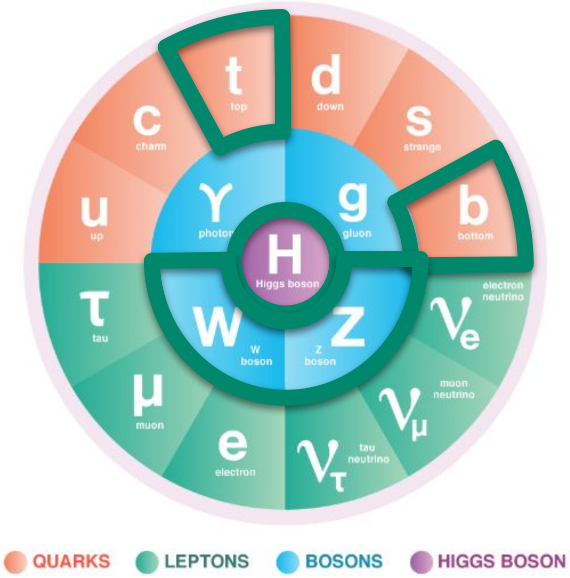
- **August 2023:** *Measurement of the inclusive cross section of Z boson production in pp collisions at √s = 13.6 TeV*, CMS Collaboration, CMS-PAS-SMP-22-017, to be submitted to PLB
- **August 2023:** *Measurement of the ttH and tH production rates in the H → bb decay channel with 138 fb^{-1} of proton-proton collision data at √s = 13 TeV*, CMS Collaboration, CMS-PAS-HIG-19-011, to be submitted to JHEP
- **July 2023:** *Searches for additional Higgs bosons and for vector leptoquarks in ττ final states in proton-proton collisions at √s = 13 TeV,* CMS Collaboration, JHEP 07 (2023) 073
- >20 more journal publications in recent years + many to follow

Scientific highlights

- Higgs boson couplings to top/bottom quarks + vector bosons
- Precise test of Standard Model
- Complex final state and analysis techniques



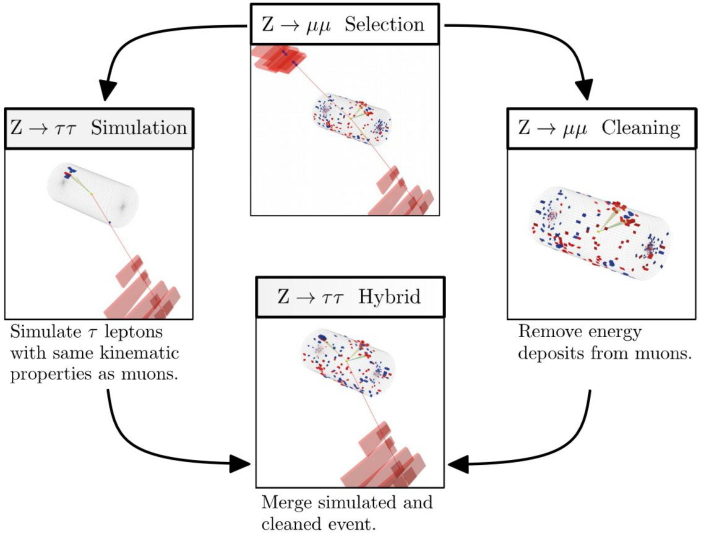


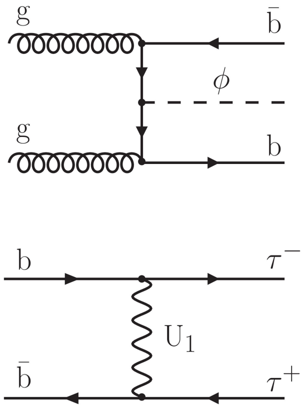


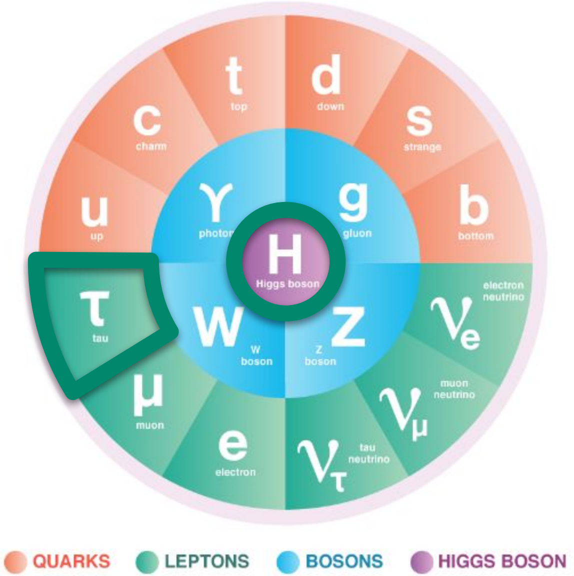
- **August 2023:** *Measurement of the inclusive cross section of Z boson production in pp collisions at √s = 13.6 TeV*, CMS Collaboration, CMS-PAS-SMP-22-017, to be submitted to PLB
- **August 2023:** *Measurement of the ttH and tH production rates in the H → bb decay channel with 138 fb^{-1} of proton-proton collision data at √s = 13 TeV*, CMS Collaboration, CMS-PAS-HIG-19-011, to be submitted to JHEP
- **July 2023:** *Searches for additional Higgs bosons and for vector leptoquarks in ττ final states in proton-proton collisions at √s = 13 TeV,* CMS Collaboration, JHEP 07 (2023) 073
- >20 more journal publications in recent years + many to follow

Scientific highlights

- Direct search for physics beyond the Standard Model
- Background estimation via τ-embedding technique







- **August 2023:** *Measurement of the inclusive cross section of Z boson production in pp collisions at √s = 13.6 TeV*, CMS Collaboration, CMS-PAS-SMP-22-017, to be submitted to PLB
- **August 2023:** *Measurement of the ttH and tH production rates in the H → bb decay channel with 138 fb^{-1} of proton-proton collision data at √s = 13 TeV*, CMS Collaboration, CMS-PAS-HIG-19-011, to be submitted to JHEP
- **July 2023:** *Searches for additional Higgs bosons and for vector leptoquarks in ττ final states in proton-proton collisions at √s = 13 TeV*, CMS Collaboration, JHEP 07 (2023) 073
- >20 more journal publications in recent years + many to follow

Summary

- Experiments in high energy physics typically require a lot of computing power
  - Large (static) resources available for fundamental processing, well-defined workflows
  - Additional resources needed for final analysis work, different requirements

- Opportunistic resources can help to provide short-term resources on demand
  - Integration not straightforward, especially if end-user access should be kept simple
  - Consider resource requirements, software availability, virtualization, containerization, …

- Successful integrated and used in the past
  - We will continue to explore these possibilities (e.g. GPU workflows on TOPAS)
  - With NEMO being on of the most used sites, we are looking forward to NEMO 2