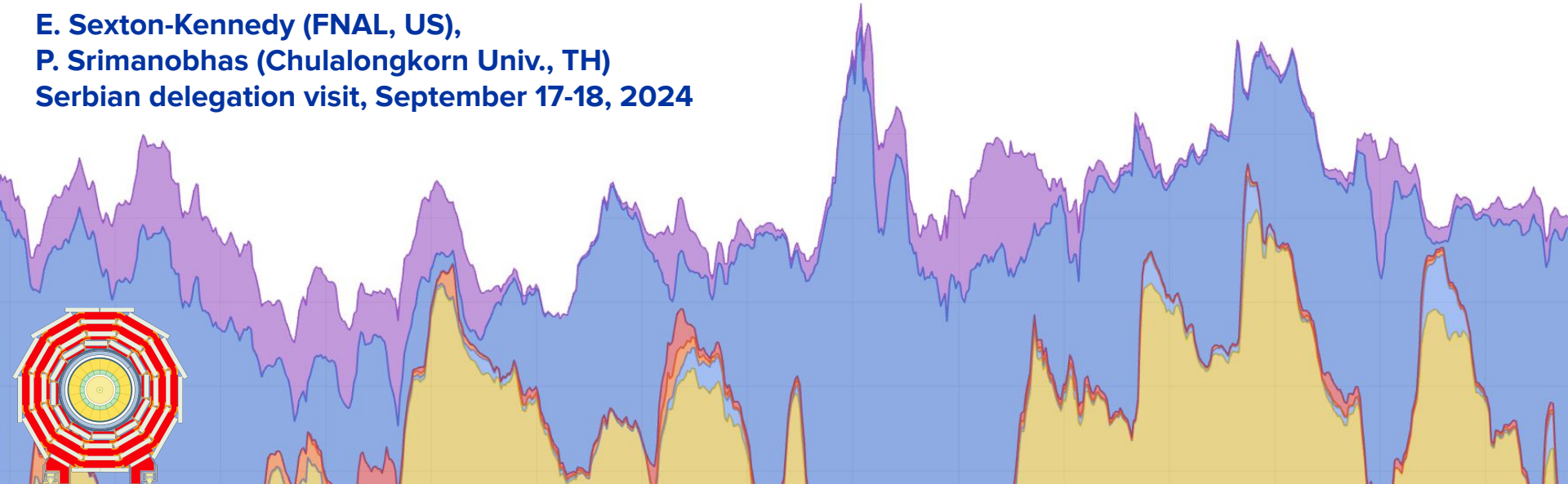
CMS use of WLCG resources

E. Sexton-Kennedy (FNAL, US),
P. Srimanobhas (Chulalongkorn Univ., TH)
Serbian delegation visit, September 17-18, 2024



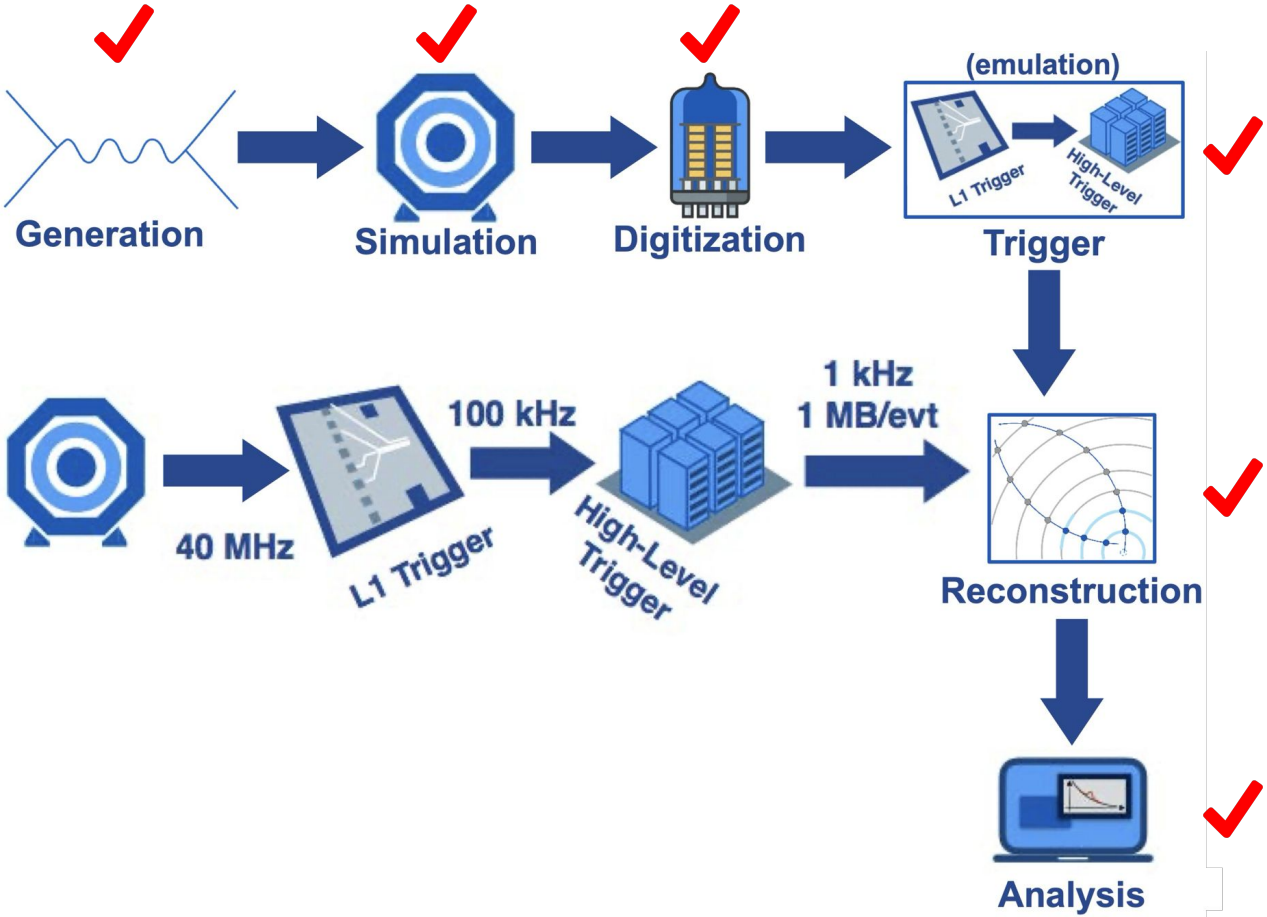
Outline of Topics

- **CMS Computing Infrastructure**
  - Data (re)processing
  - Resource utilization
  - Monte Carlo production
  - Disk and tape usage
- **HL-LHC Challenge for Offline Software & Computing**
  - Data transfer
  - Event processing
- **Site Setup and Personnel Training for Technology Implementation**
- **Conclusions**

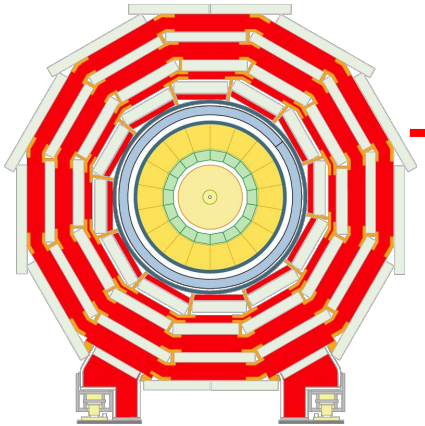
CMS Computing Infrastructure for Data (Re-)Processing

CMS uses Tier-1/Tier-2 (✓) resources for

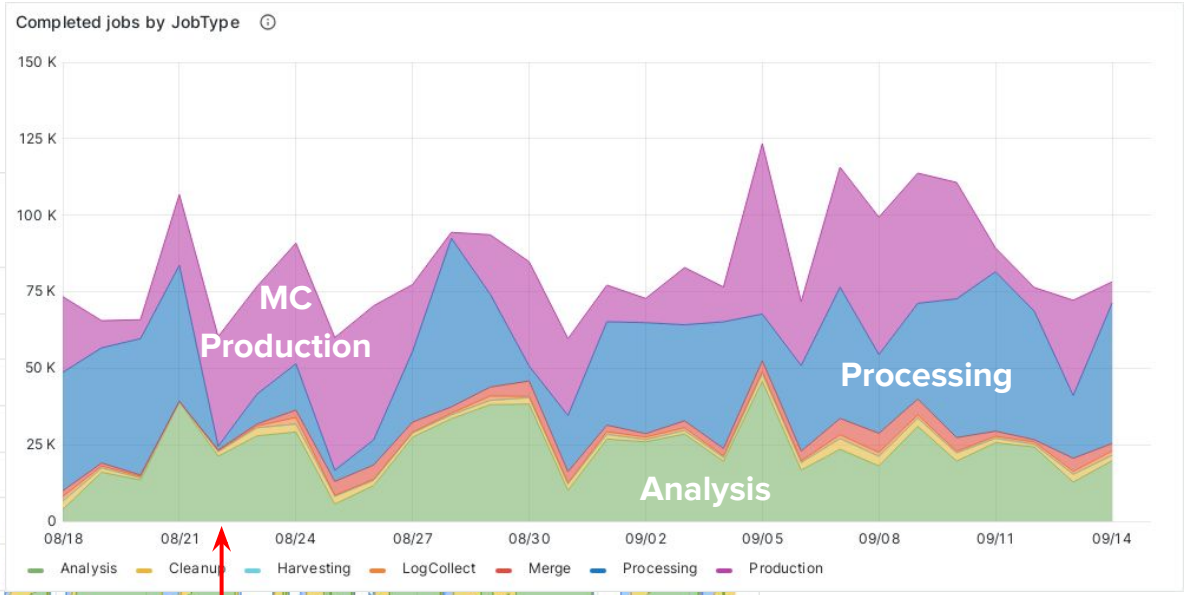
- **CPU**: Data “prompt” processing + reprocessing *(T1-only)*
- **CPU**: Monte Carlo production
- **Tape**: Archive of RAW data *(T1-only)*
- **Disk**: Store datasets used for analysis

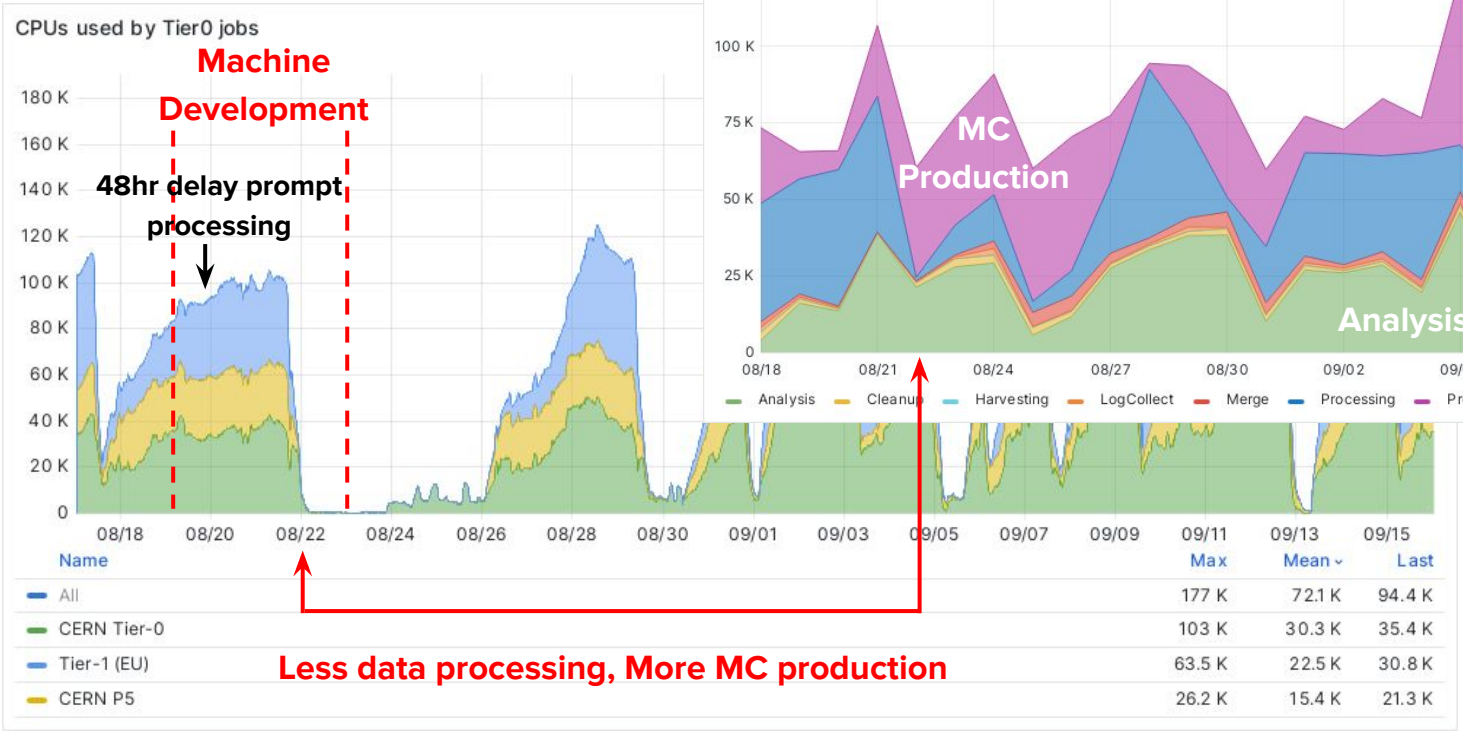


CMS Computing Infrastructure for Data (Re-)Processing



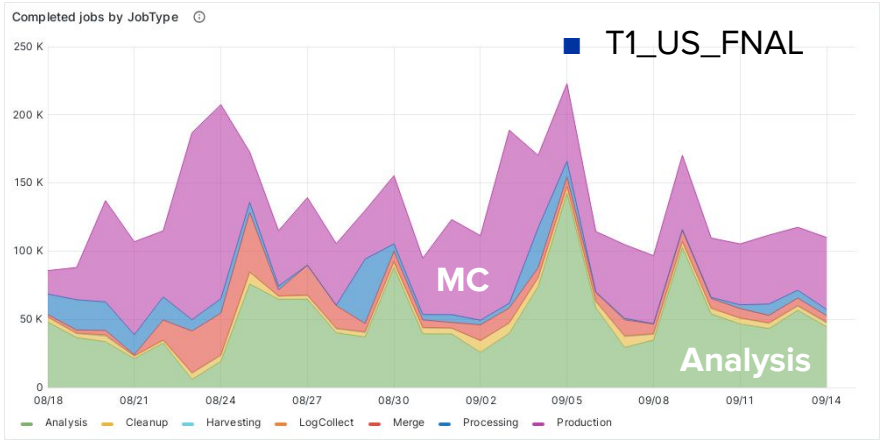
Tier-1(s): Resource Utilization

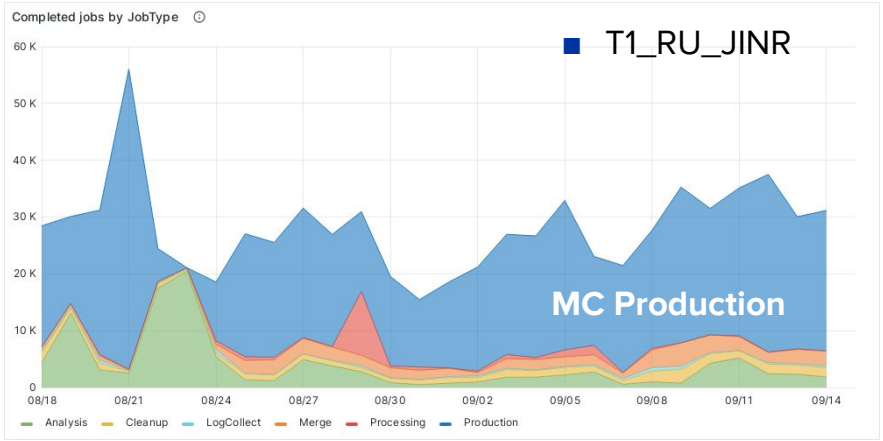


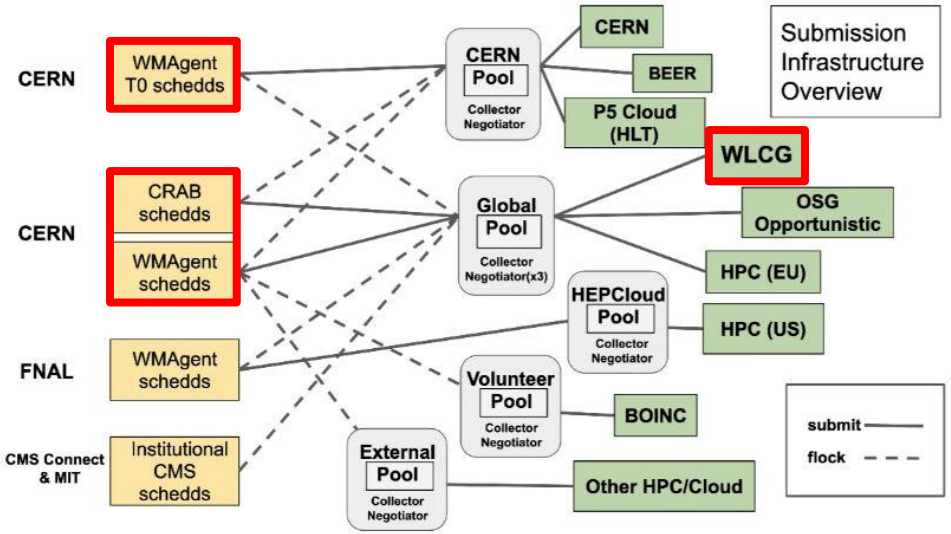


- T1_DE_KIT
- T1_ES_PIC
- T1_FR_CCIN2P3
- T1_IT_CNAF
- T1_UK_RAL

Tier-1(s): Resource Utilization

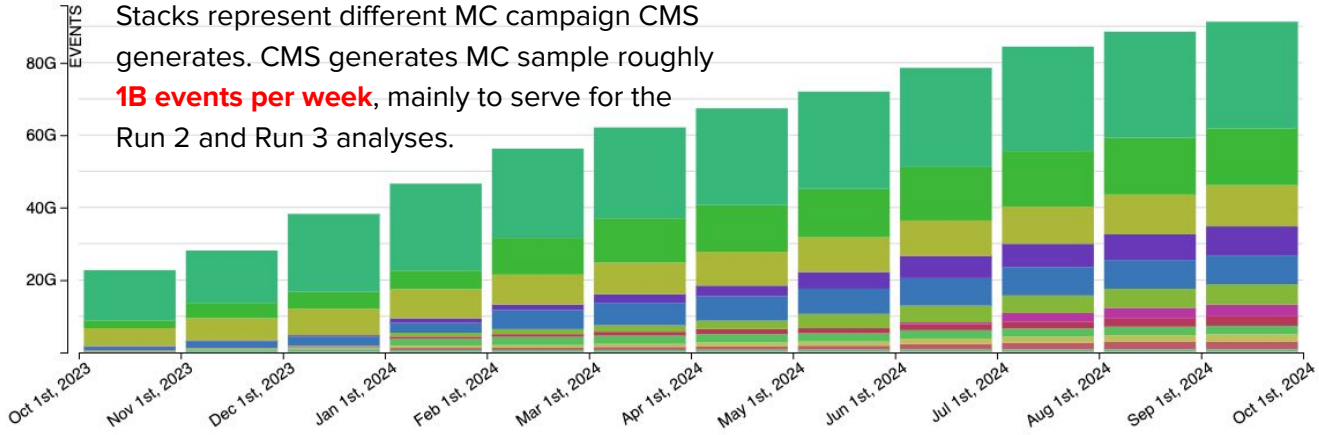






The CMS Submission Infrastructure includes distributed job schedulers (schedds) for managing Tier 0 operations, centralized production (WMAgent), and analysis submissions (CRAB). It handles data reconstruction, simulation, and analysis for CMS physics.

Monte Carlo Production

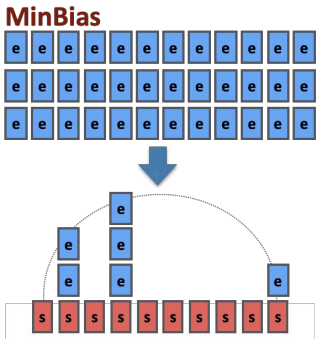


Stacks represent different MC campaign CMS generates. CMS generates MC sample roughly **1B events per week**, mainly to serve for the Run 2 and Run 3 analyses.

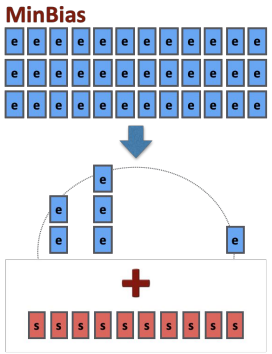
Monte Carlo Workflow

- Mostly start from scratch
- Read pileup to local storage or remote (mostly remote for premixing library from CERN or FNAL)
- Rerun data-reduction based AODDIM available at the site (disk or tape)

Classical mixing



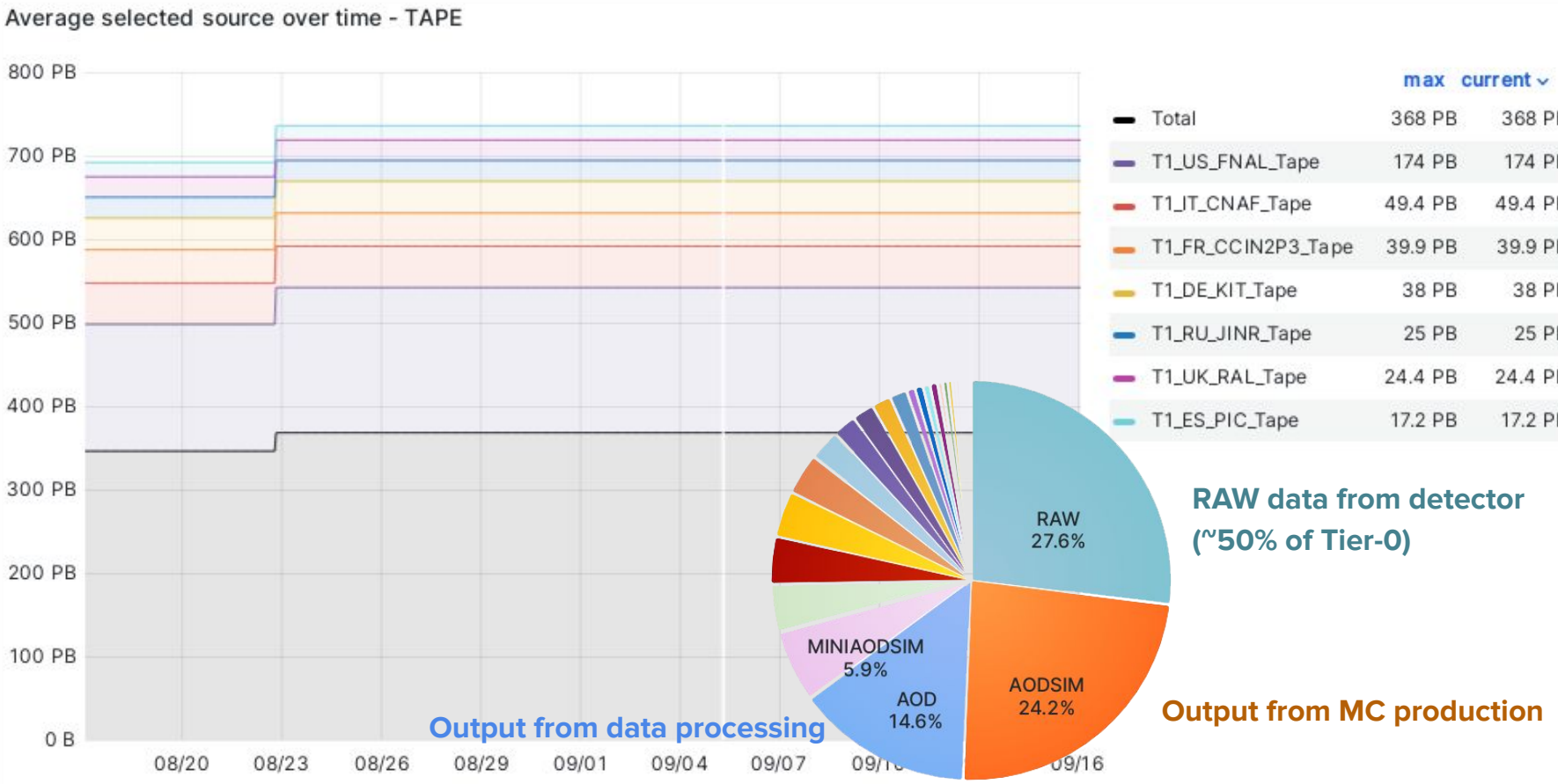
Premixing



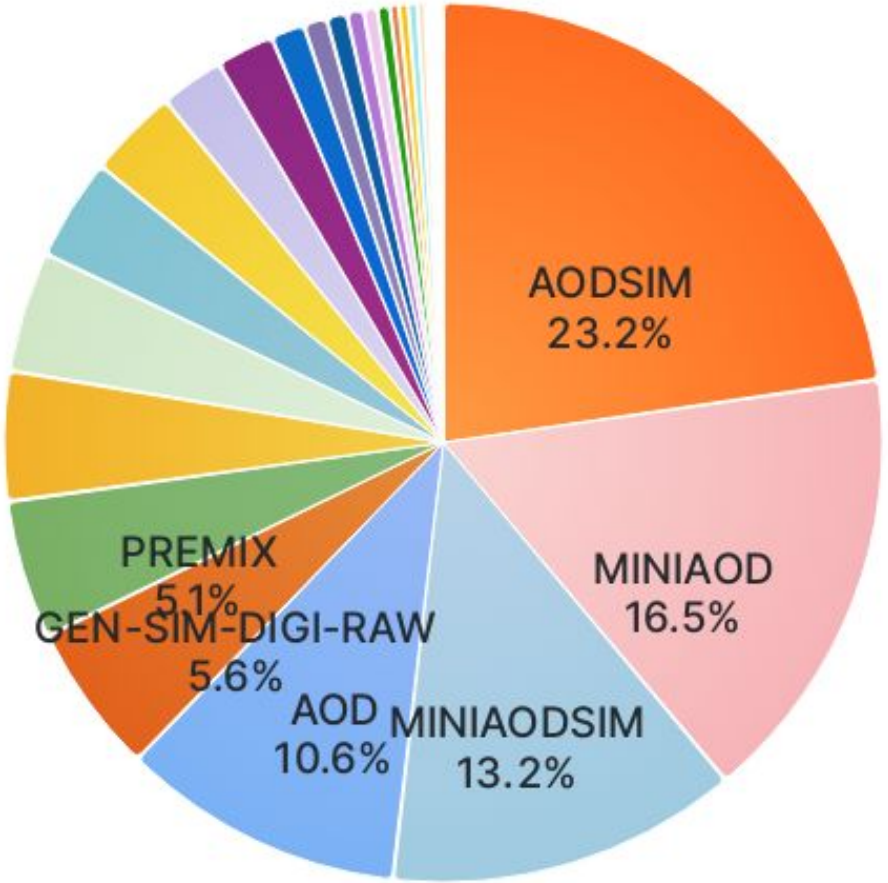
In LHC proton-proton collisions, the interactions are not one-to-one but many-to-many, with approximately 60 interactions for Run 3 and around 140 expected for Run 4. This phenomenon is referred to as pile-up (PU), and it must be accounted for in our simulations. Currently, there are two methods to handle PU:

- Classical mixing, which is I/O-intensive to read events from file simultaneously
- Premixing, where network performance plays a crucial role to read library from CERN

Tier-1 Tape Archive



Tier-1 Disk for Data-Reduction and Analysis

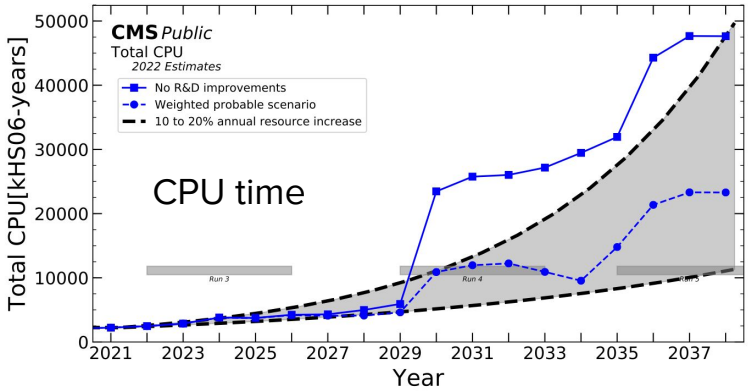


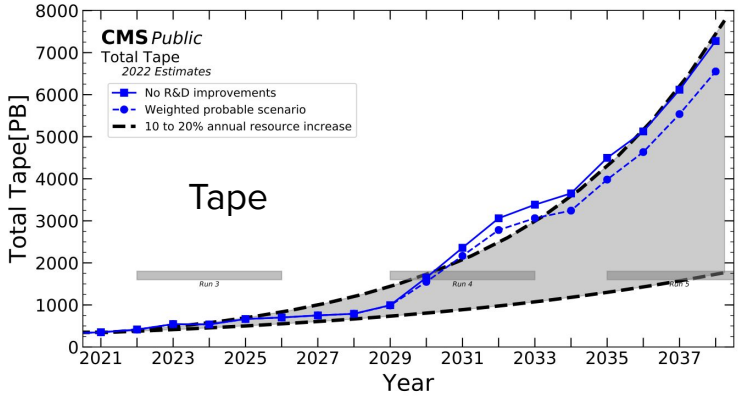
Tier-1 disk serves for many purposes:

- **AOD + AODSIM** (about one third of total disk): The main output from data reconstruction and MC simulations, used for data reduction, or analysis. Some datasets may be available on Tape only and recalled as needed.
  - CMS has data policy to manage datasets on disk
- **MINIAOD + MINIAODSIM** (about one third of total disk): datasets for analysis

The HL-LHC Challenge for Offline Software & Computing

CMS CPU time and tape time projected requirements (2022)





Run 4 of the LHC will begin in about 5 years:

- Increasing of instantaneous luminosity and simultaneous collisions per bunch crossing ⇨ higher trigger rate (events/second)
- More complex events ⇨ larger event size (MB/event)
- Upgraded detectors - more channels to read out ⇨ larger event size (MB/event)
- Event processing times will also increase due to more active events and complex detectors.

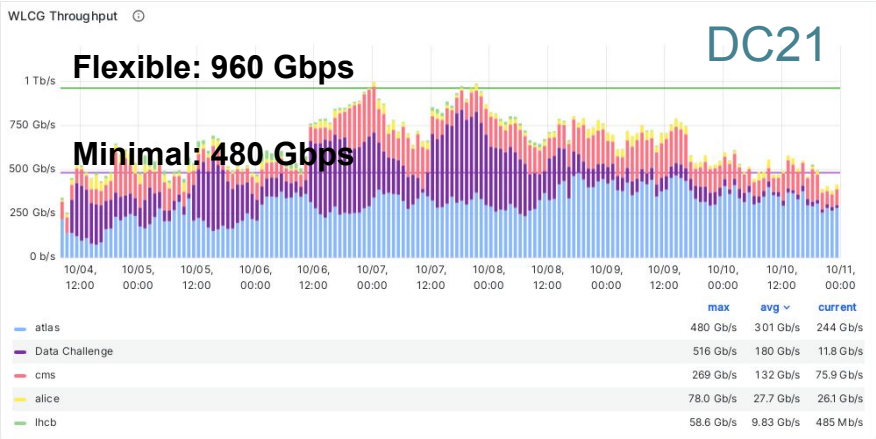
All of these drive order of magnitude increases in storage and computing needs. **Serbia Tier-1 will help!**

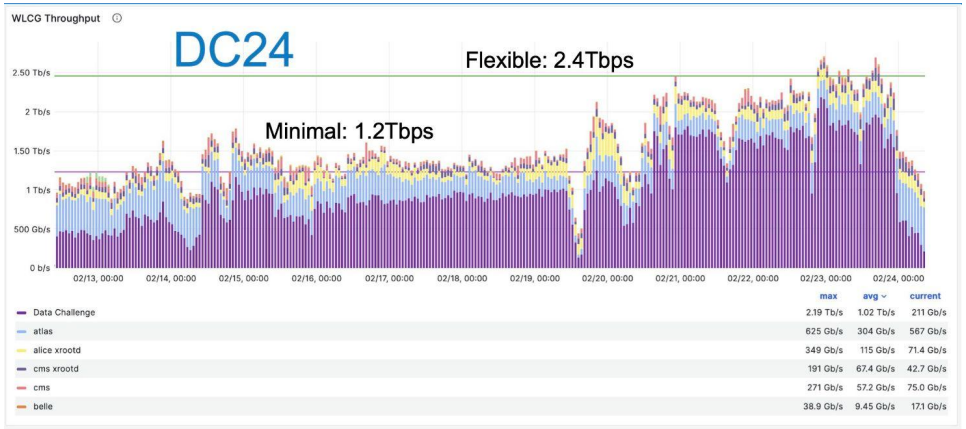
💥 **R&D efforts can drive down the resource (and timing) needs significantly**, as seen in the plots above!

HL-LHC Challenge: Data Transfer

With 2021 estimation on the network need for HL-LHC:

- CMS experiment is expected to generate ~350 PB of RAW data annually, and the expected RAW data export traffic is roughly 200 Gbps (to T1s in quasi-real time).
- An additional of 100 Gbps to account for other data formats.
- Network may need to accommodate 2x traffic for bursts and 2x overprovisioning.
- Estimated total network capacity required from CERN to CMS T1s by HL-LHC: 0.6-1.2 Tbps (CMS-Only)
- The Data Challenge (DC) has been organized to demonstrate network capability to transfer HL-LHC data with targeting 4.8 Tbps for 4 experiments



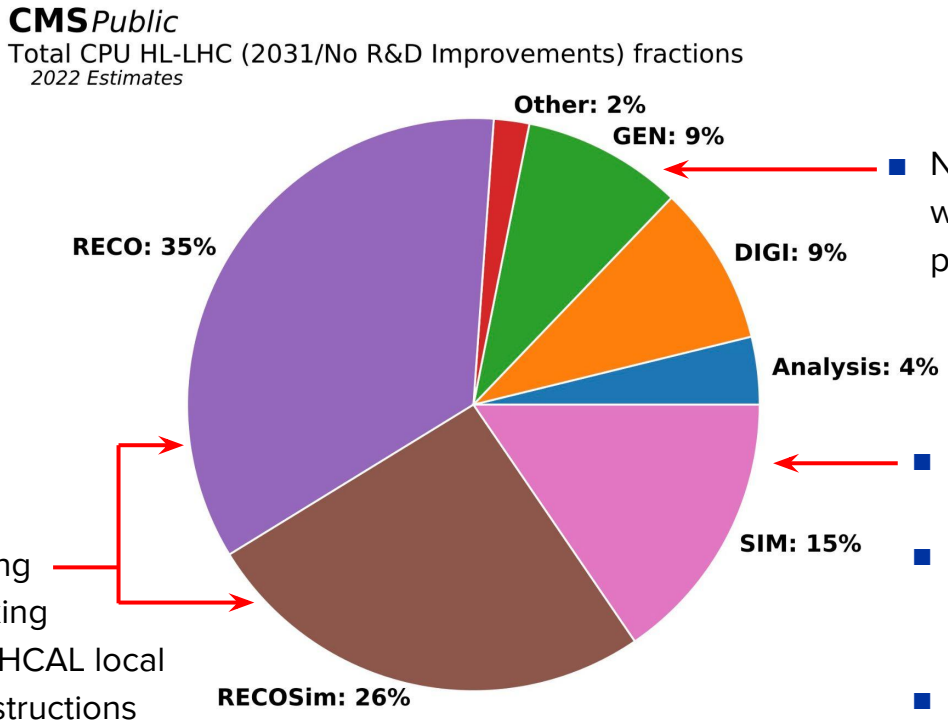


HL-LHC Challenge: Event Processing

Approximate breakdown of CPU time during a typical HL-LHC year. CMS is actively working on offloading various tasks to accelerators, such as GPUs.

Sites equipped with GPUs, either within the WLCG or through opportunistic resources, are encouraged to participate.

Significant progress has also been made in integrating heterogeneous resources into the CMS submission infrastructure. Prototype workflows have been successfully tested and validated on the system.



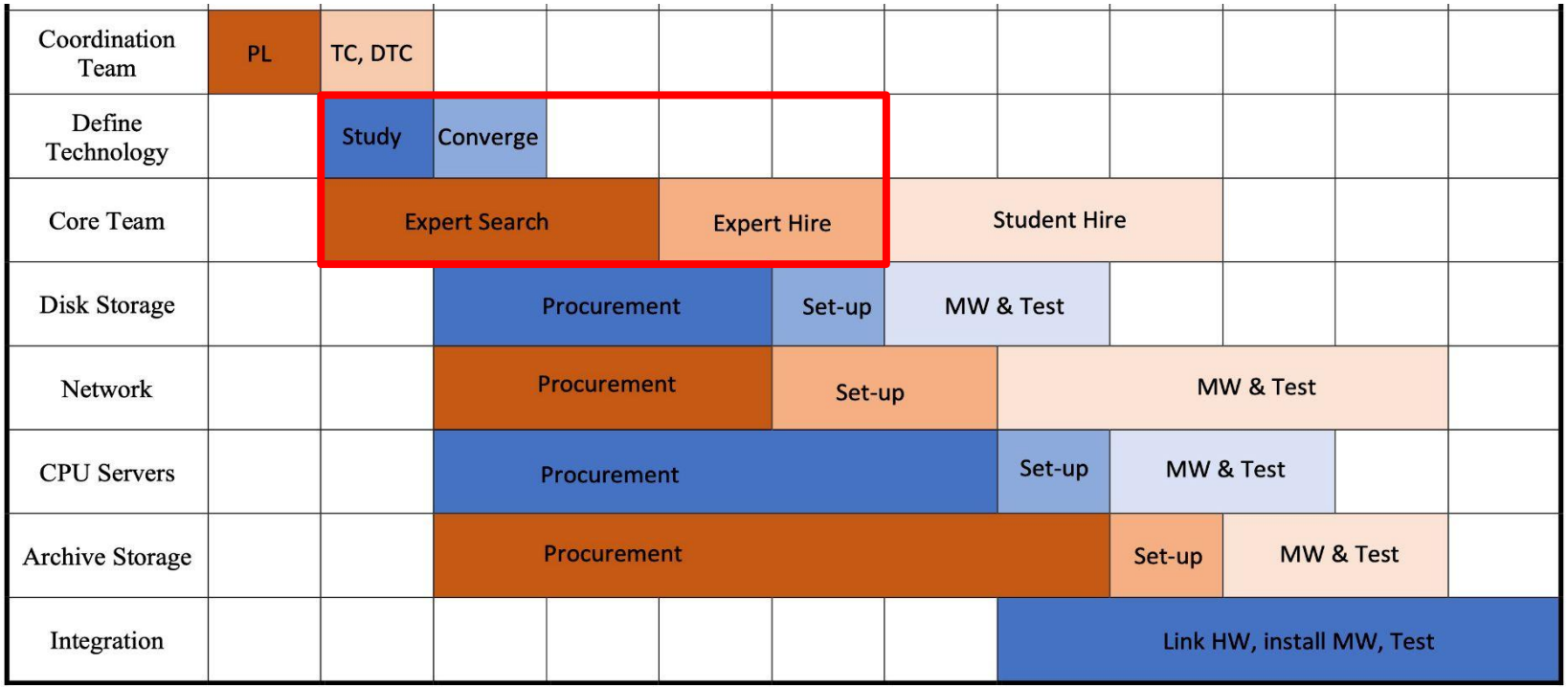
- New area to go with CPU+GPU in process calculation

- Full Simulation: AdePT / Celeritas
- Fast Simulation with ML refinement
- Flash Simulation using DL on end-to-end simulation

- Tracking
- Vertexing
- ECAL/HCAL local reconstructions
- Clustering & particle flow
- Electron seeding

Site Setup and Personnel Training for Technology Implementation

It is recommended to choose market-available hardware and implement your selected technology (e.g., Tape/Disk management) on top of it. If existing hardware is available, setting up a pilot site is advisable as part of expert training in all necessary areas, ensuring familiarity with the chosen technology.



Conclusions

- We would like to express our sincere gratitude to Serbia for considering the establishment of a Tier-1 computing infrastructure for CMS. CMS is fully committed to providing support in areas such as technical design, training, and operational guidance. We also extend our thanks to CERN IT and WLCG for their assistance in providing reviews and recommendations.

- It appears that you are opting for market-available hardware and planning to implement your chosen technology on top of it. We believe this is a solid approach. Once the technology is finalized, we would like to connect you with other sites using the same technology, allowing you to benefit from their experiences.

- In our view, one of the most critical early steps is to gain experience in site setup and operations. If some hardware is already available (even if it's not newly purchased), setting up a pilot site would be an excellent way to build practical experience.

- As we approach the HL-LHC era, where challenges abound in offline software and computing, this is a great opportunity to collaborate, participate, and grow together.