# Status of ATLAS software and computing

James Catmore (Oslo/CERN), Alessandro Di Girolamo (CERN)

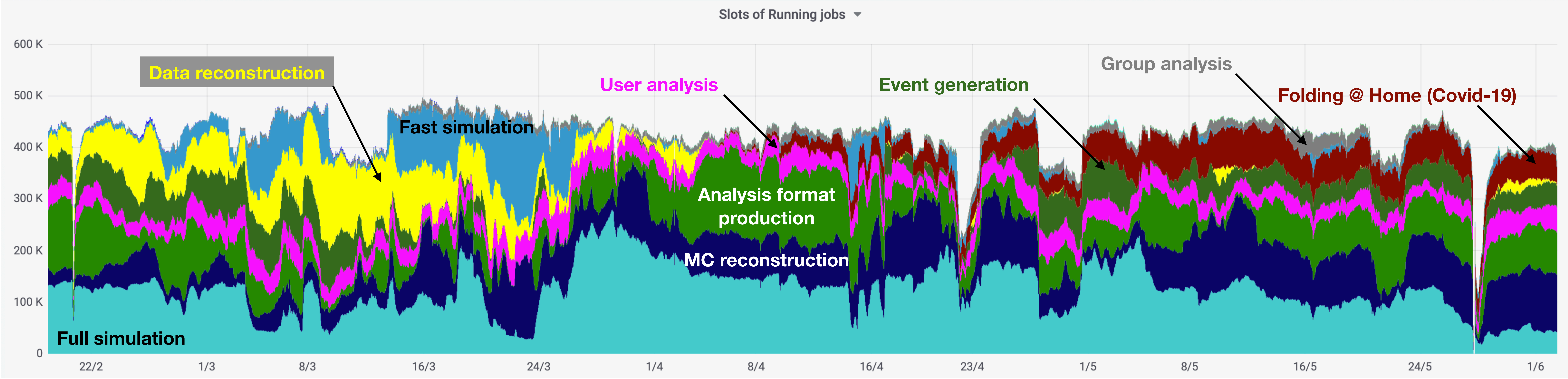
# Overview

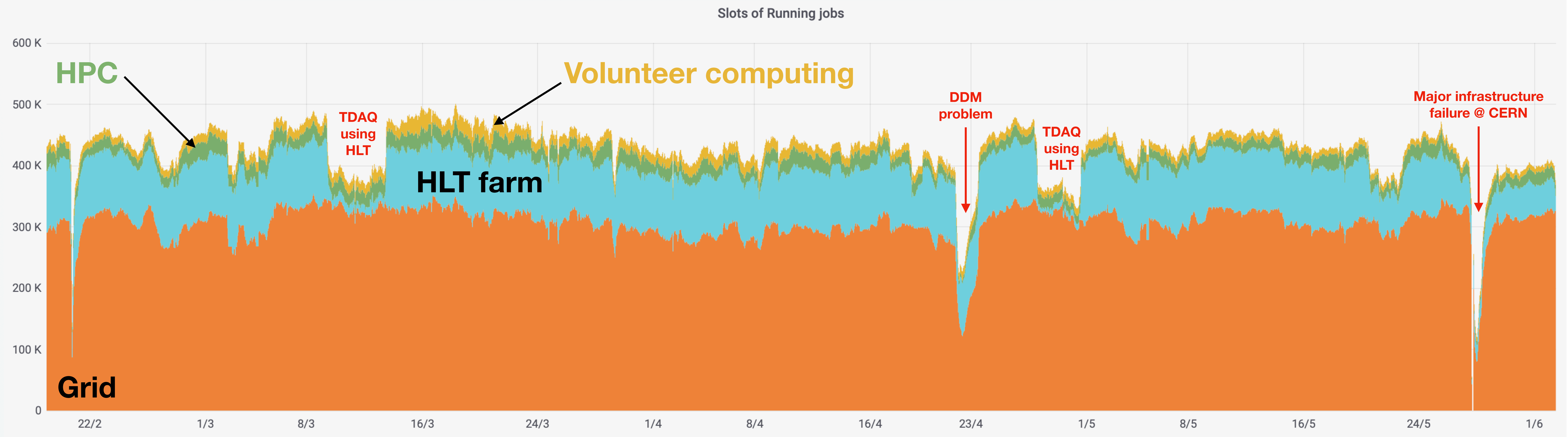


- A review of HL-LHC software/computing was held on May 19th; in this session we focus entirely on preparations for Run 3
- In this session
  - Distributed computing usage since the last LHCC week
  - Impact of Covid-19 on computing and contribution to CERN Against Covid-19
  - Preparation for Run 3
    - Distributed computing
    - Software and simulation
    - Databases
    - Analysis model

# Distributed computing usage since last LHCC: compute






- WLCG continues to run at full strength, mainly running simulation, analysis format production and user analysis
- Significantly buttressed by the HLT farm for the majority of the time
- Slight reduction in demand for computing resources due to cancelled conferences, delays in preparation of certain Sherpa samples
- New activity: Folding@Home jobs to support the CERN Against Covid-19 initiative (see later)
- HPC: high performance computing infrastructure that can be accessed without special treatment

# Distributed computing usage since last LHCC: special HPC

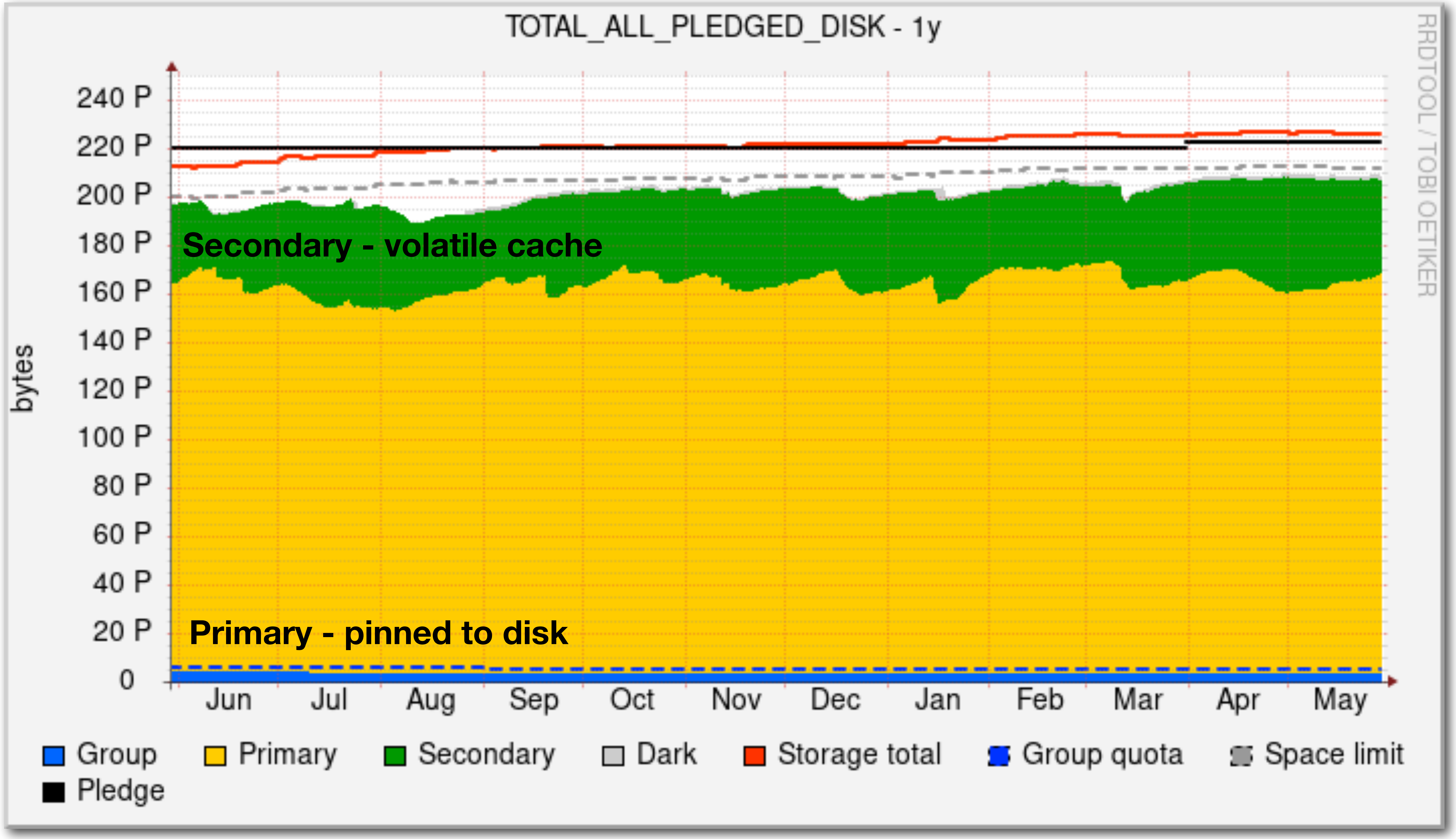


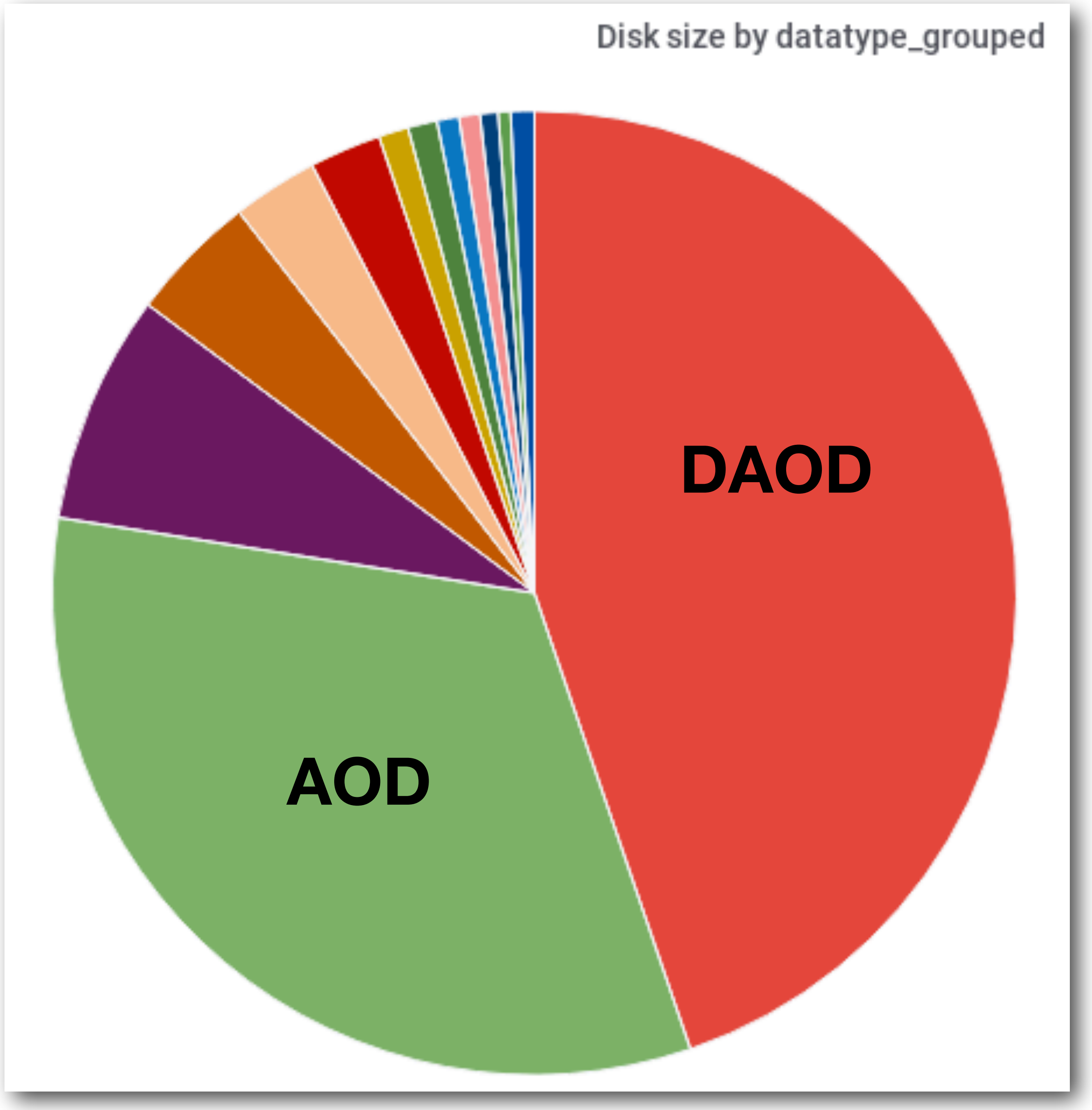

Significant use of "Special HPC" resources: very large "leadership class" machines in the United States that need special treatment in addition to the usual ProdSys machinery

During this period almost all resources of this type came from Cori @ NERSC and was used for Geant4 simulation

# Distributed computing usage since last LHCC: storage





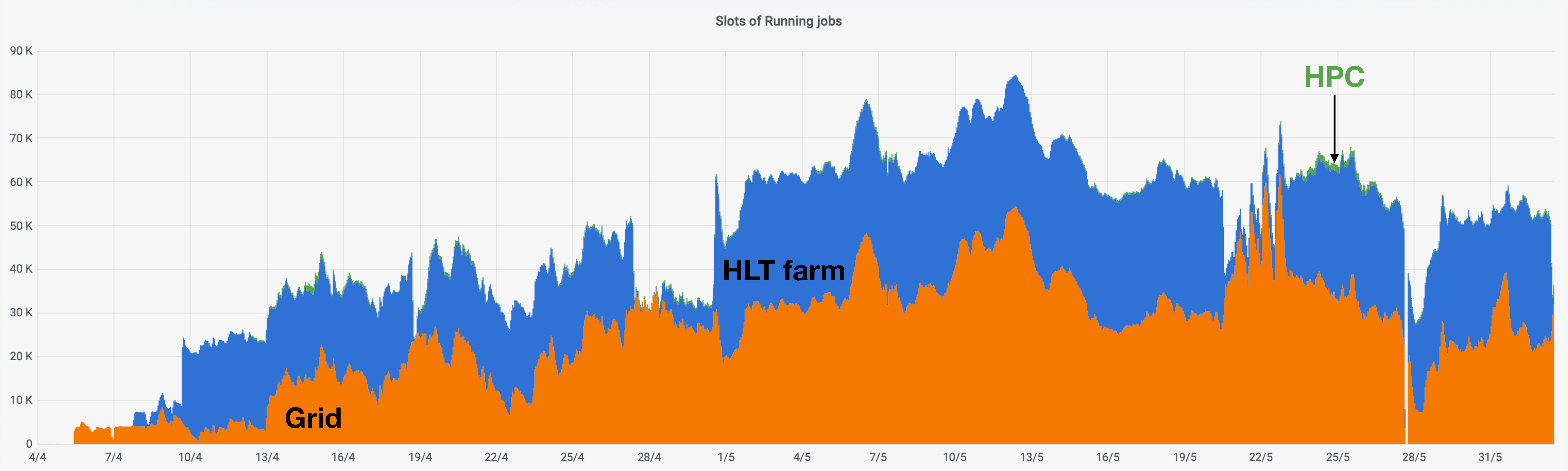

| | total | percentage |
|---|---|---|
| DAOD | 94.9 PB | 48.3% |
| AOD | 62.0 PB | 31.5% |
| HITS | 15.6 PB | 7.9% |
| EVNT | 8.55 PB | 4.3% |
| log | 3.54 PB | 1.8% |
| RAW | 3.11 PB | 1.6% |

- Storage continues to be heavily occupied by post-reconstruction analysis formats (DAODs and AODs)
- Active disk space management (deleting obsolete or unused data) has kept the situation under control and there is no current risk of the disks filling up

# Computing contribution to CERN Against Covid-19



- Since the beginning of April ATLAS has contributed WLCG and HLT resources to the **CERN Against Covid-19** initiative
- As agreed by the initiative steering group, ran Folding @ Home jobs: simulations of internal molecular dynamics of proteins, primarily for identifying potential pathways for medication
- Configured to run computations relevant to SARS-COV-2
- Full flexibility in the system to throttle F@H jobs when more capacity is needed for physics (physics always given priority)
- Targeting approximately 10% of resources, agreed in advance with the ATLAS International Computing Board (funding agency computing representatives)
- Several new techniques exercised during this exercise: containers, direct access to GPUs




- As the CERN Against Covid-19 initiative moves to longer term collaborations, the F@H activity will be wound down

# Impact of Covid-19 on computing operations



- ATLAS distributed computing operations **continue to work as before**
- Response and reaction time of sites may be affected in the future
  - a few sites starting to report **hardware issues** which cannot be promptly fixed
  - some important **upgrades have been postponed** (would require physical access)
- Disruptive tests of infrastructure in these weeks (e.g. stress tests) avoided
- Although the daily business continues, **months of teleworking does impact the mood and motivation** of people
  - The lack of spontaneous or ad-hoc face-to-face discussions for new and innovative ideas is noticeable
  - New people find it harder to integrate without any chance to meet their colleagues
  - "*Creativity comes from spontaneous meetings, from random discussions. You run into someone, you ask what they're doing, you say, 'Wow,' and soon you're cooking up all sorts of ideas.*" - Steve Jobs

# Distributed computing

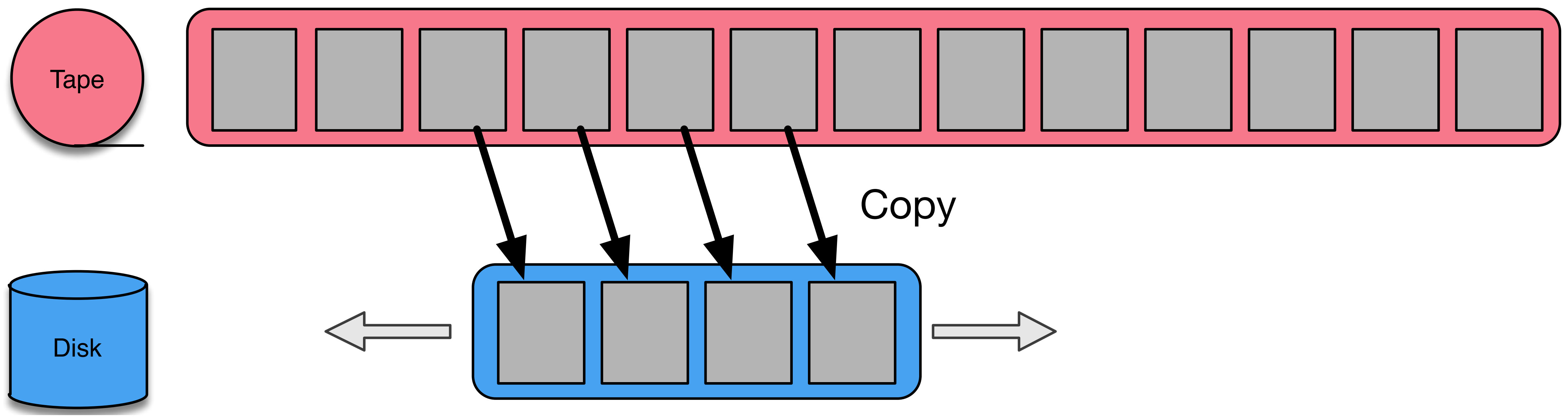


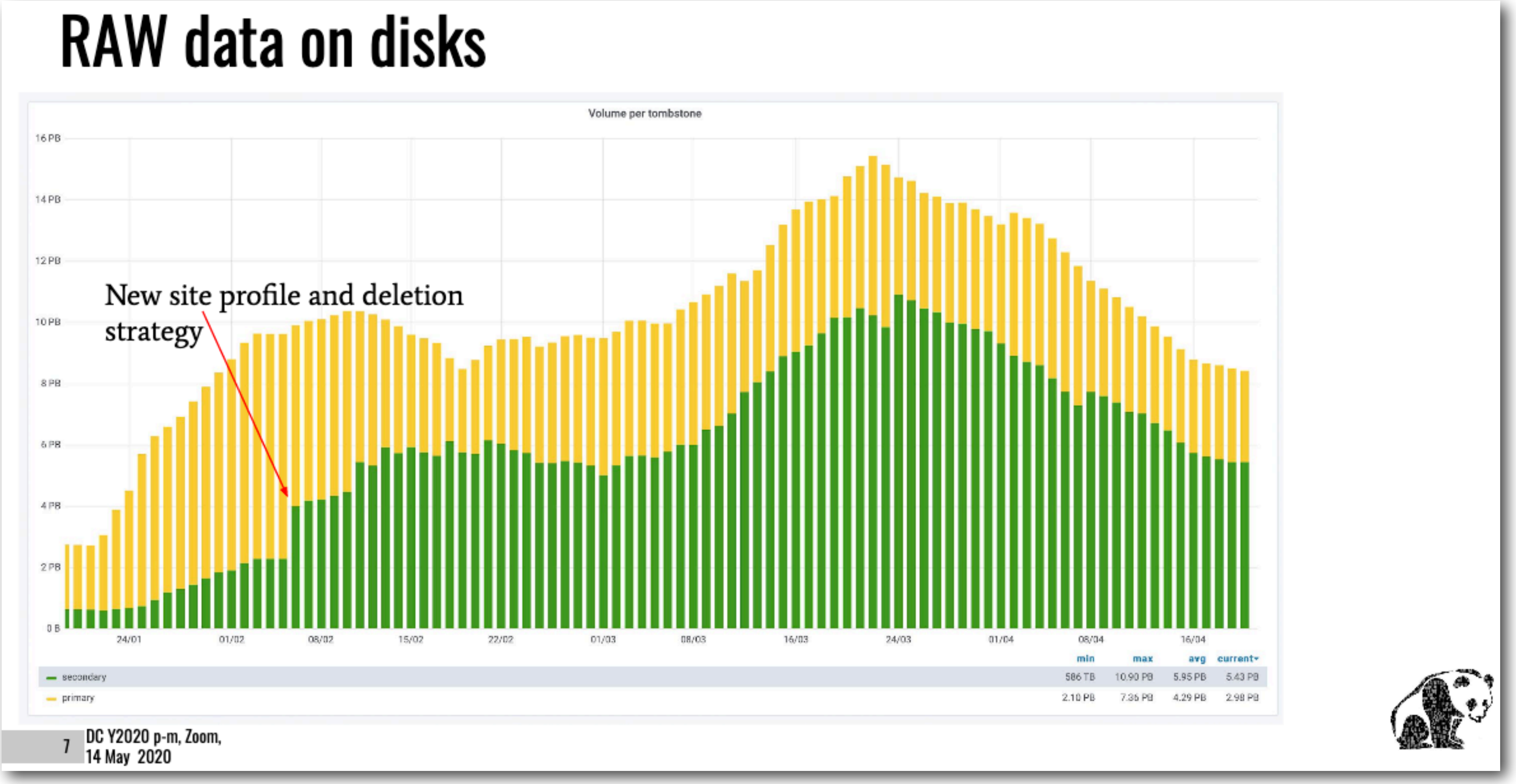
- Unified (Computing) Queues
  - Motivation: **optimization of site resources** by enriching job dispatch information with detailed job requirements. Allows flexibility to scale up analysis resources if needed. Ongoing: to be finished by end of summer
- Data Carousel (see next slide)
- Third-party copy (**TPC**): continuing the long process of moving to non-gridFTP third-party-copy for data movements between grid sites within the WLCG DOMA project.
- Use of **standalone software containers**
  - Important part of the strategy to simplify user analysis and embrace new ideas and technology (run on HPC)
  - Images building and distribution are the biggest challenge for user analysis
- New **CERN Tape Archival** (CTA) integration: ATLAS is pioneering the CTA integration
  - Several tests performed, reaching ~5GB/s read and (separately) 5GB/s in write.
  - Integration with ATLAS Point1 TDAQ ongoing, handshake still to be addressed.

# Data carousel



- Activity (started in 2018) progressing well
- Reprocessing of complete run 2 (2015-2018) data
- 18 PB of RAW data from 10 sites ; started on 19 of January and finished on 4 of April.
- data18: 2PB rolling buffer on disk, optimized (next staging requests launched only when previous recall were almost all processed) to almost 1PB for data17-15
- Periods with almost half of the Grid (150k-200k slots) running reprocessing
- Now beginning tests with DAOD production



# Software



- Main change for Run 3: **multi-threaded framework** AthenaMT
  - Motivations: **memory footprint reduction** and support for heterogeneous computing environments
- Most areas of the **reconstruction** are now fully thread safe, allowing a bulk processing test involving millions of events to be run on the grid with multiple threads, allowing the performance to be measured and rare crashes identified
- Major work remaining
  - Combining reconstructed **muons** with the inner detector, or extrapolating muons, leads to irreproducible differences. Under investigation.
  - **In-file metadata** (luminosity information, cut-flow etc) relies on thread-unsafe infrastructure. Migration being carried out by a short-term task force
    - Associated issue: simultaneous writing of multiple files not yet possible. Being addressed by the same task force
  - Finalisation of interval-of-validity transitions and alignment in MT
  - Detailed **profiling** of the computing performance and optimisation
  - Full understanding of the **physics performance** of the new release
- Simultaneous overhauling of the Python side (job configuration): new modular configurations mechanism; Python 3 migration

# Simulation



- Fast simulation
  - New version of the **fast calorimetry simulation** (FCSv2) with same computing performance but much better physics performance in final validation, to be ready for Run 3 simulation
  - Version with a **generative-adversarial network** (GAN) model for hadronic physics also in validation
- Multi-threading
  - Geant4v10 already MT-compatible; simulation framework now MT-compatible as well
  - Work ongoing to run fast simulation under MT conditions
- Full simulation (Geant4)
  - **Plan to introduce version 10.6:** hadronic physics has changed (pions have a ~10% response change in the 10 GeV momentum range) → retuning of Birk's constant required to compensate. Major ongoing task.
  - Work on further tweaks to **improve speed without impacting physics**: updated calorimeter geometry; static library builds; photon Russian roulette; new stepper algorithms → each could yield speed-ups of O(1-10%)
  - EM range cuts and neutron Russian roulette already included

# Databases



- **"COOLR" developments**
  - ATLAS Run 3: continue to use LCG Conditions DB for storage but is developing a RESTful service "COOLR" to access the data (bypassing LCG Cool API)
- **Frontier Launchpad and Squid:** deployment and monitoring improvements
  - Automating dissemination of software and configuration updates via AGIS and/or CRIC
  - Expanding monitoring of Launchpads and Squids and their failovers not only shows dysfunctional sites but also failing sites.
  - A number of 'stress tests' are ongoing to better understand and address bottlenecks which need to be identified before the real stress of Run 3 demands.
- **Frontier analytics:** Elasticsearch and Kibana
  - Collection of Frontier Launchpad logs into Elastic Search has been a key to understanding problematic tasks and requests as well as assessing throughput.
- **Metadata catalog evolution**
  - Dedicated metadata catalogs continue evolution to meet growing use cases and in anticipation of changes in data taking and dataset processing, and increased event counts of Run 3.

# Physics validation

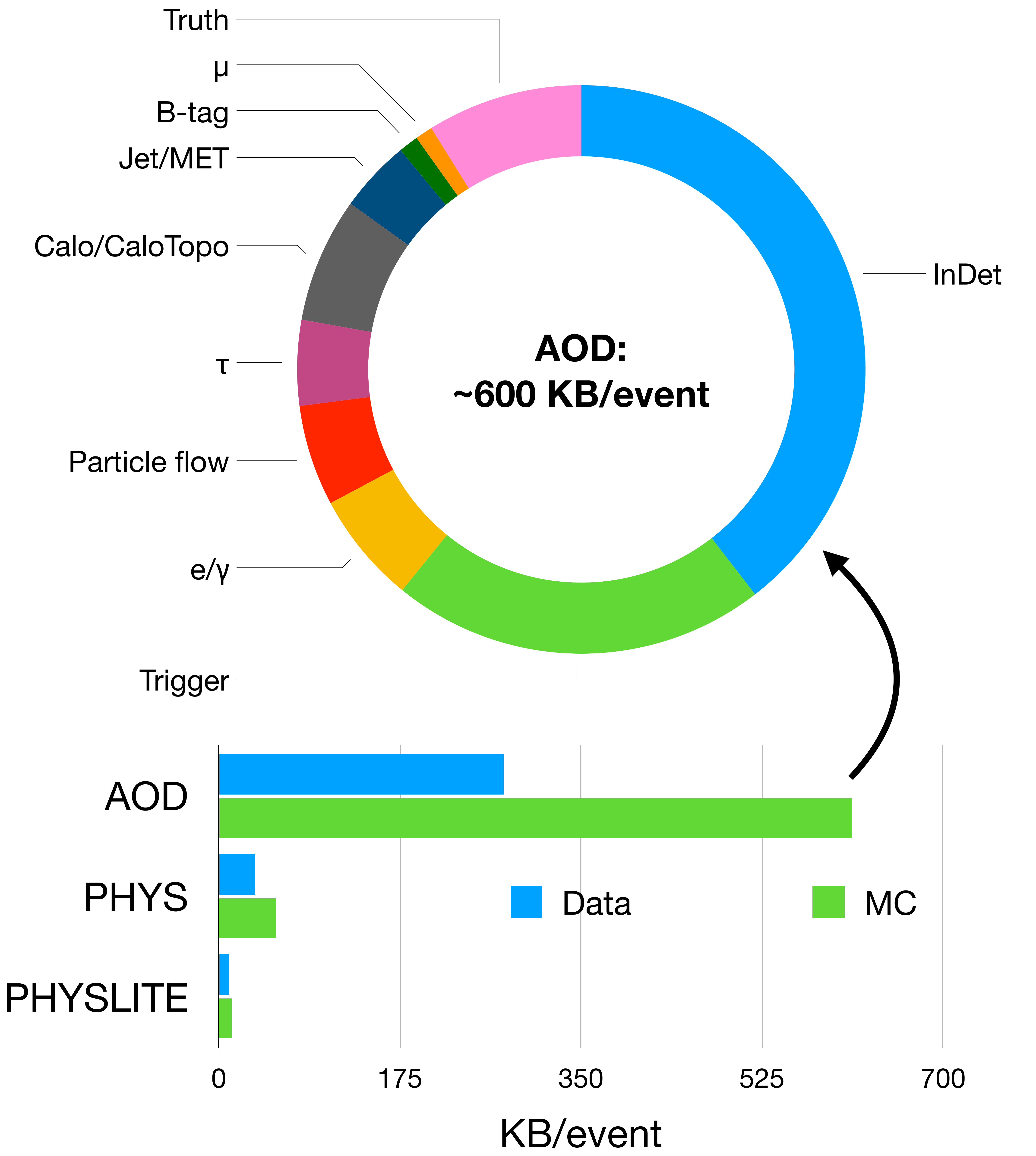


- Physics validation of release 22 is required for
  - Validating changes related to multi-threading and checking that multi-threaded jobs produce the same output and multi-process and serial jobs in the same release
  - Validating changes made to the reconstruction software in release 22 versus release 21
- Several bulk validations are planned before fall. Each validation is separated by ~2-3 months.
  - Validation is performed on Run-2 data and simulated samples
  - Had first bulk validation last February and second validation is on-going
  - Identified a few crash sources that could only be caught with large scale tests
  - Understanding some changes in performance
- Plan to include physics analyses into validation procedure for next bulk validation

# Analysis model



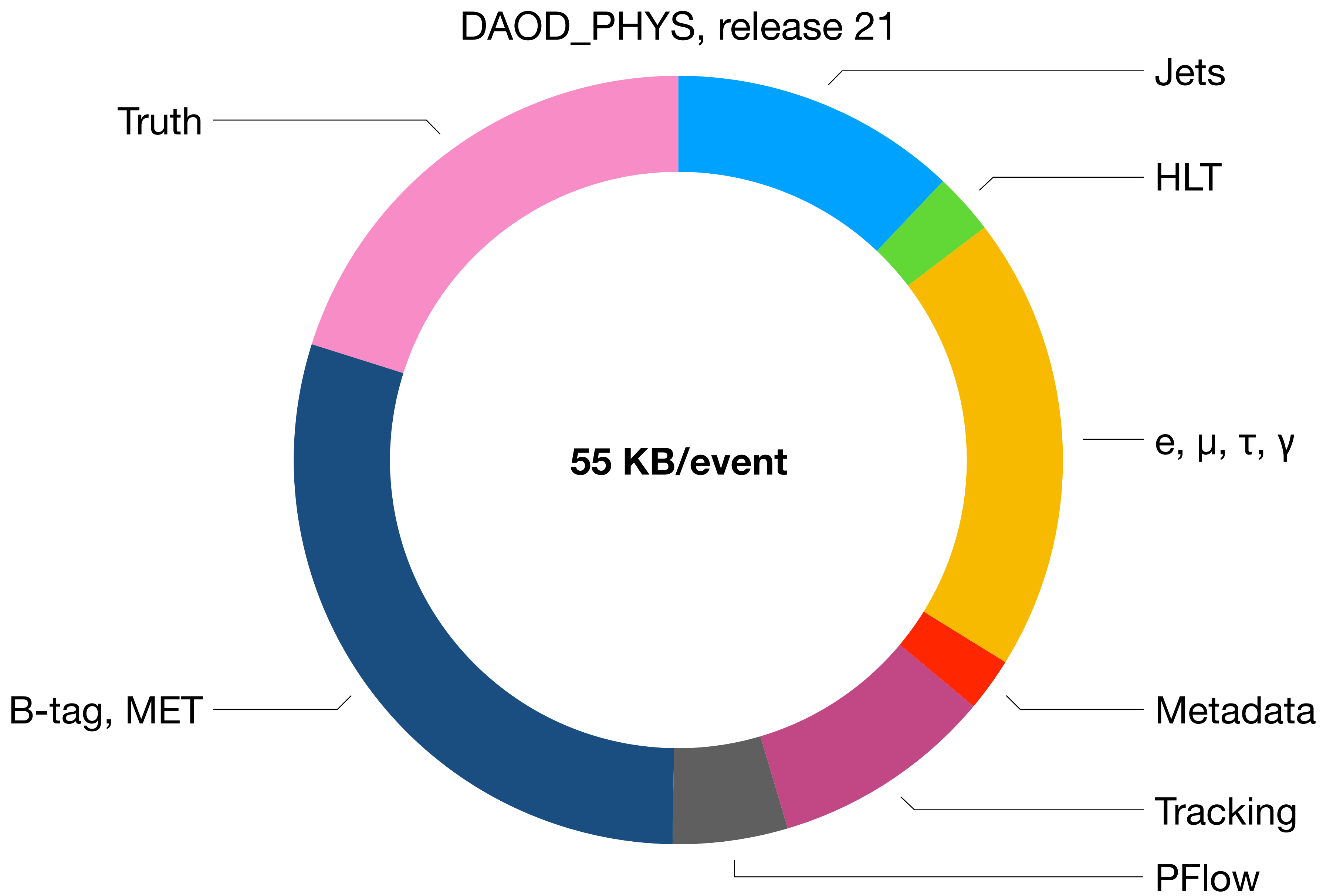
- Analysis data formats constitute the largest part of the data stored on disk by ATLAS
- Primarily a result of the analysis model, in which each physics analysis team defines their own centrally produced analysis format according to a common template ➞ 84 such formats in current use
  - Heavy overlap between them, especially in MC
- Run 3 will see the introduction of a new analysis model, based around a single analysis format (DAOD_PHYS) which contains all variables required to calibrate and apply systematics ➞ much less overlap
  - Central skimming of this format envisaged for analyses with low acceptance
- Run 4 will see a transition to an even smaller format (DAOD_PHYSLITE) which will contain only calibrated quantities
  - This will be commissioned in Run 3

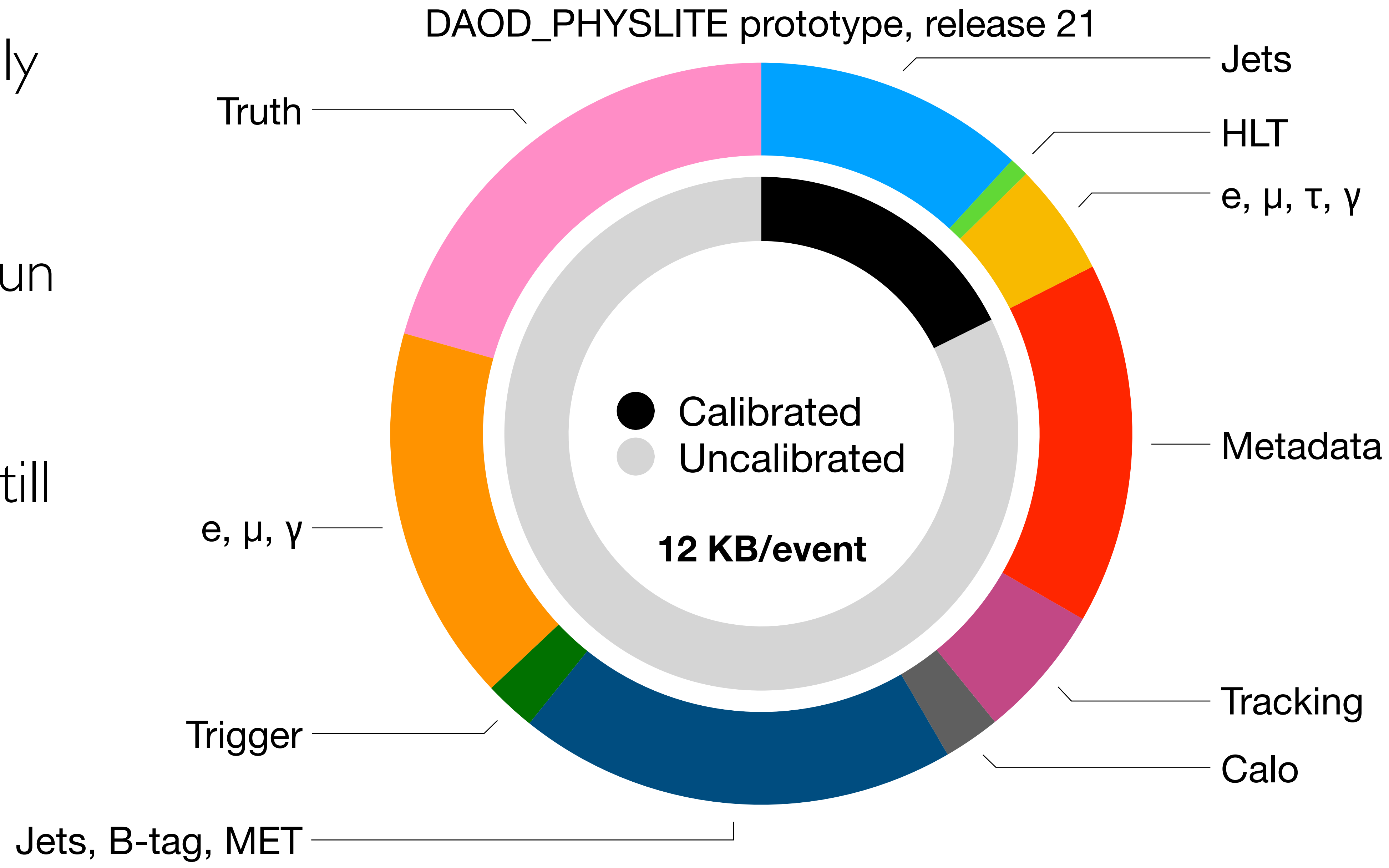


# Analysis model



- DAOD_PHYS is being produced in release 21, analyses are beginning their transition
  - Some DAOD_PHYS production using the data carousel
  - Mechanism for skimming DAOD_PHYS ready, being tested in production
- DAOD_PHYSLITE: only variables needed directly in analysis; fully calibrated; central values only → means of assessing systematics will need to be established
  - This is the major challenge to the success of this format and is the focus of current work of the Analysis Model Group
- Migration of the analysis software stack to release 22 (especially the derivation framework needed to produce DAOD_PHYS) is under way: tedious but not technically challenging
  - DAOD production (like event generation) is not memory-limited so no need to run it in MT
- Not yet clear how many physics analysis will be unable to use DAOD_PHYS, but we still expect <20%
  - Combined performance groups may need continue to use old-style DAODs in some cases



# Conclusions



- The pandemic has not unduly impacted computing/software operationally but months of teleworking are starting to have an impact on morale and motivation
- Nevertheless ATLAS continues to make good progress in its preparations for Run 3
  - Multi-threaded software
  - Better simulation
  - More efficient distributed computing workflows
  - New analysis model

# Additional material

# Run 3 analysis model