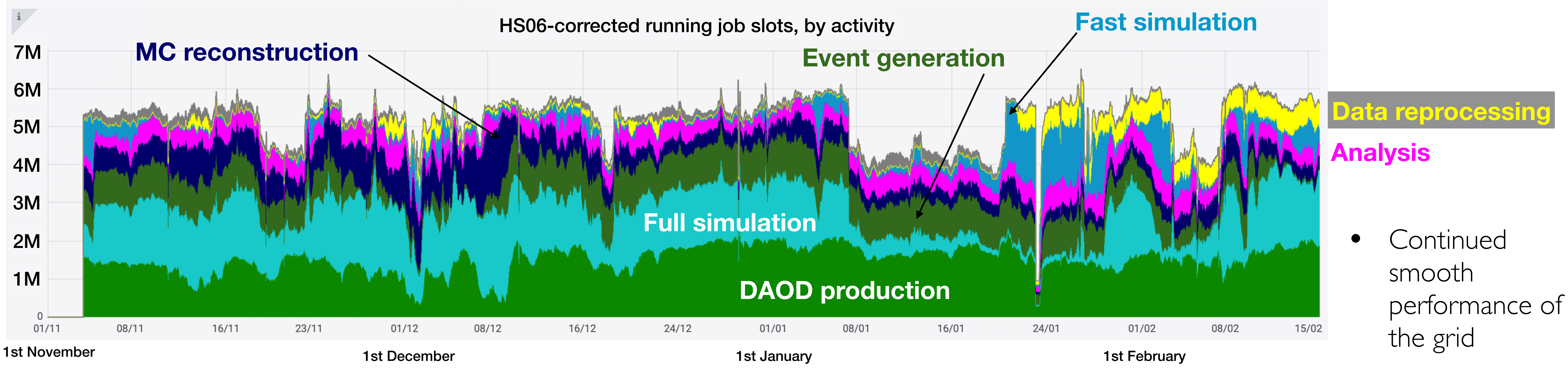
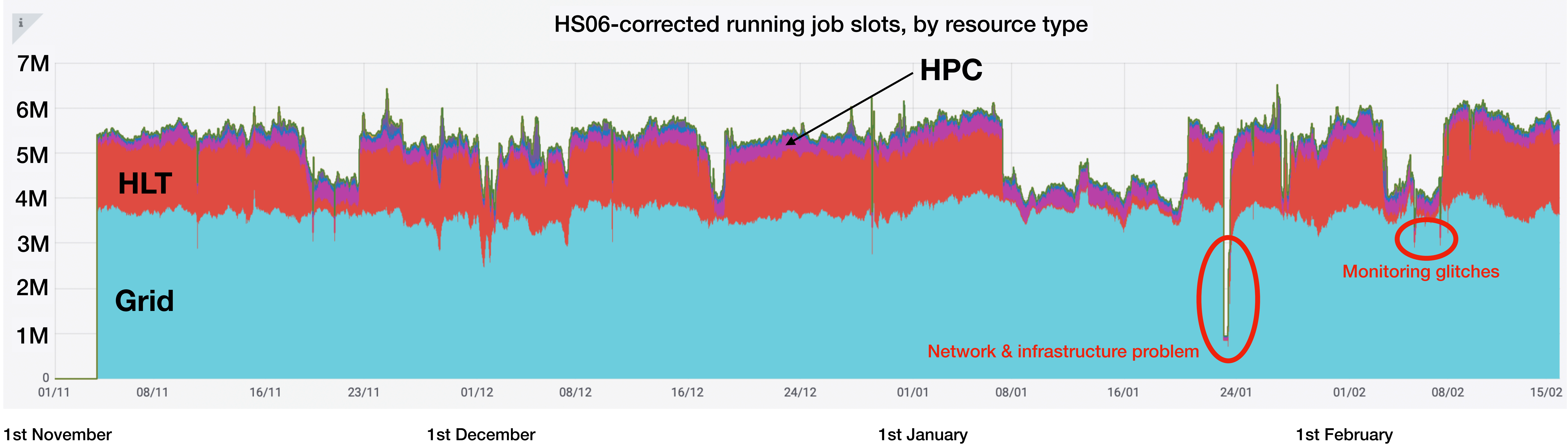
ATLAS computing and software update

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

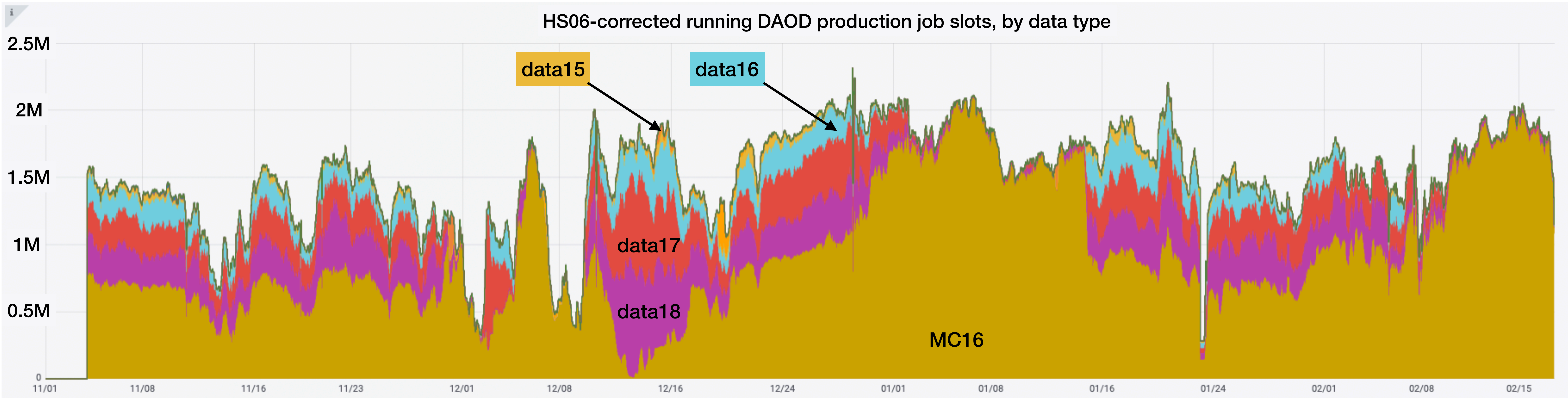
Production update





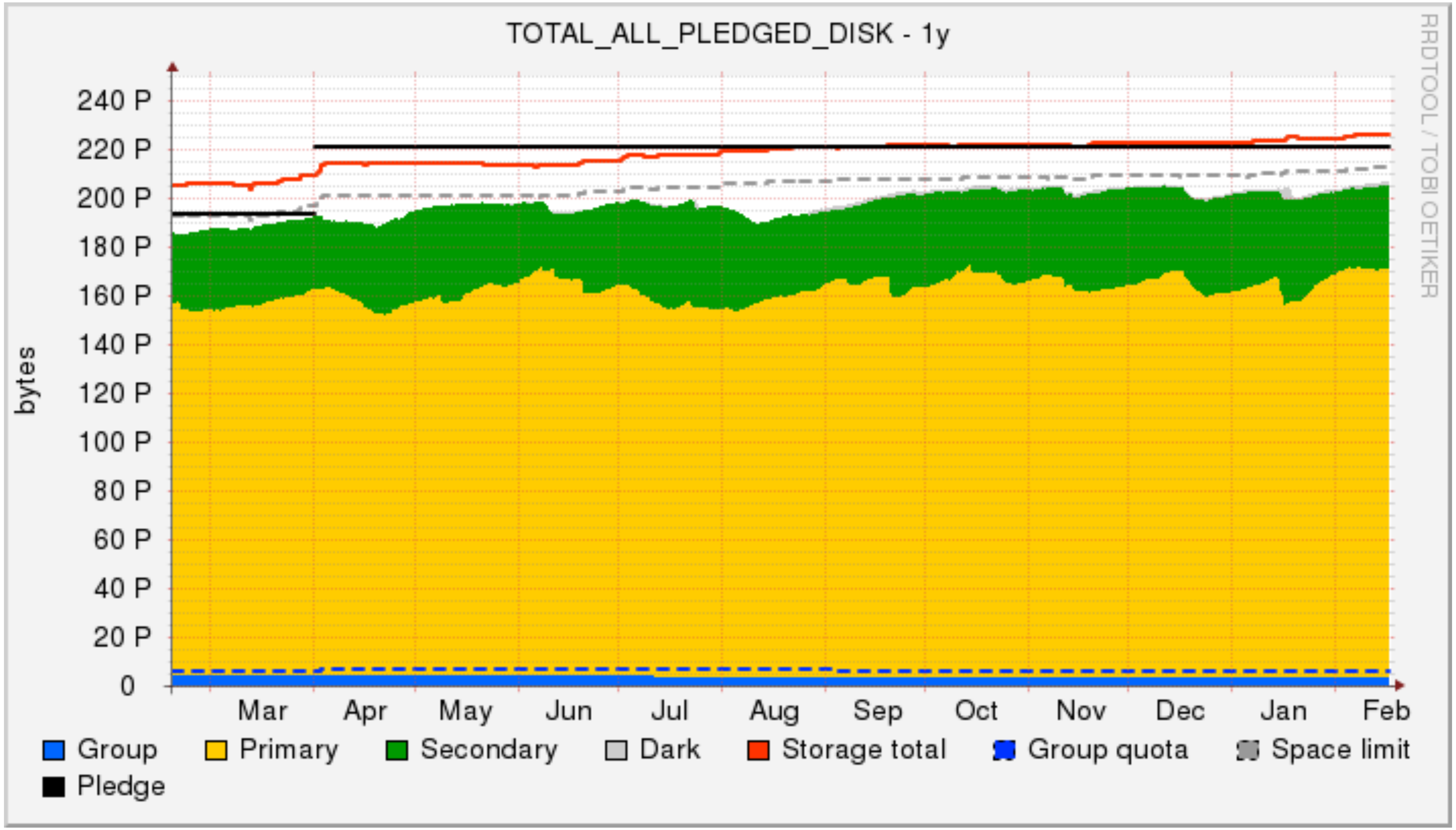
- Continued smooth performance of the grid
- MC and analysis format production dominates
- Grid substantially buttressed by HLT farm which continues to work smoothly

Production update: DAOD production



- DAOD production is mostly occupied by the larger search groups (SUSY and exotics), two measurement groups (top and Higgs) and two combined performance groups (flavour tagging and jets)
- Production of formats from MC dominates (the main target of the new analysis model - see later)

Storage update



- Disk space is becoming tight due to large volumes of DAOD production by physics and performance groups
- Enforcing a limit of 2 concurrent versions of DAODs
- The new analysis model will alleviate this situation once it is fully adopted

Production planning for 2020-2021

Sample	Activity	2020				2021			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Run 2 data	DAOD production for ongoing analyses								
	Production of DAOD_PHYS for transition to new model								
	Reprocessing in release 22 + DAOD(_PHYS)(LITE)								
Run 2 MC	New production for ongoing analysis								
	Production of DAOD_PHYS for transition to new model								
	Reprocessing in release 22 + DAOD(_PHYS)(LITE)								
Run 3 data	Reconstruction at Tier-0 + DAOD(_PHYS)(LITE)								
2021 MC	Generation/simulation								
	Reconstruction + DAOD(_PHYS)(LITE)								
	Reprocessing with improved conditions								
2022 MC	Generation/simulation								
Upgrade MC	Generation/simulation/reconstruction + DAOD(_PHYS)(LITE)								

Summary of C-RSG interactions

- 2021 requests were submitted to the C-RSG in August
  - Increases of 15% CPU, 20% disk and 10% tape requested w.r.t. 2020 (see next slide for full table)
- Face-to-face held in September, comments received in early October and presented to the RRB at the end of October
- April report to be submitted later this week: confirms the computing request made in August
  - New LHC schedule has no impact because the stable beam estimate has increased to compensate the shorter run

	Initial	**New**
Baseline luminosity	$17 fb^{-1}$	$20 fb^{-1}$
Tape contingency luminosity	$42 fb^{-1}$	None
Average pile-up	55	35
Stable beams (baseline)	20%	50%
Stable beams (contingency)	50%	None
Protons running time	3×10^6 seconds	3×10^6 seconds
Ions running time	1.2×10^6 seconds	1.2×10^6 seconds

Summary of C-RSG interactions

	2019 Agreed @ Oct 2018 RRB	2019 pledges	2020 Agreed @ April 2019 RRB	2021 Request @ Oct 2019 RRB	Balance 2021 wrt 2020 request
T0 CPU (kHS06)	496	496	496	550	11%
T1 CPU (kHS06)	1057	1084	1057	1230	16%
T2 CPU (kHS06)	1292	1293	1292	1500	16%
SUM CPU	2845	2873	2845	3280	15%
T0 DISK (PB)	27	*26*	27	30	11%
T1 DISK (PB)	88	94	88	107	21%
T2 DISK (PB)	108	101	108	132	21%
SUM DISK	223	221	223	269	21%
T0 TAPE (PB)	94	94	94	97	3%
T1 TAPE (PB)	221	217	221	249	13%
SUM TAPE	315	311	315	346	10%

Run 3 preparations: main objectives

- Multi-threaded software framework (AthenaMT) for simulation and reconstruction; more efficient memory utilisation and pre-requisite for accelerator offloading
- Require fast simulation to constitute 50% of all simulation at the start of run 3, rising to 75% by the end
- Use of overlay for pile-up simulation to avoid repeated re-digitisation
- New analysis model to improve disk space utilisation:
  - Most physics DAODs replaced with a single format (DAOD_PHYS, 50KB/event) that can be skimmed centrally when necessary. Uncalibrated; Contains all variables needed to extract instrumental systematic uncertainties
  - Pre-calibrated format (DAOD_PHYSLITE, 10KB/event) for fast analyses and (possibly) open data. Central values only. Candidate for main Run 4 format
  - Use of data carousel to reduce volume of AOD on disk

Status of the multi-threaded reconstruction

Milestone	Threads	Events	Description	Target	Status
I	1	O(10)	Technical operability of MT schedulers	May 2019	✓
II	1	O(10K)	MT schedulers with 1 thread give identical results to serial	July 2019 (achieved Sept)	✓
III	>1	O(100)	Multi-threaded operation	October 2019	✓
IV	>1	O(10K)	Physics validation & s/w performance with multiple threads	February 2020	✓
V	≥1	O(100M)	Feature freeze and run 2 reprocessing	October 2020	
VI	>1	O(1G)	Run 3 data and MC	2021	

Status of the multi-threaded software framework

AthenaMT reconstruction now running over 1000s of events with >6 concurrent threads. All systems including muons are now included

Initial assessment of the performance indicates that AthenaMT running with 8 threads uses at most 20% of the memory of an AthenaMP job running with 8 processes, for the same event throughput

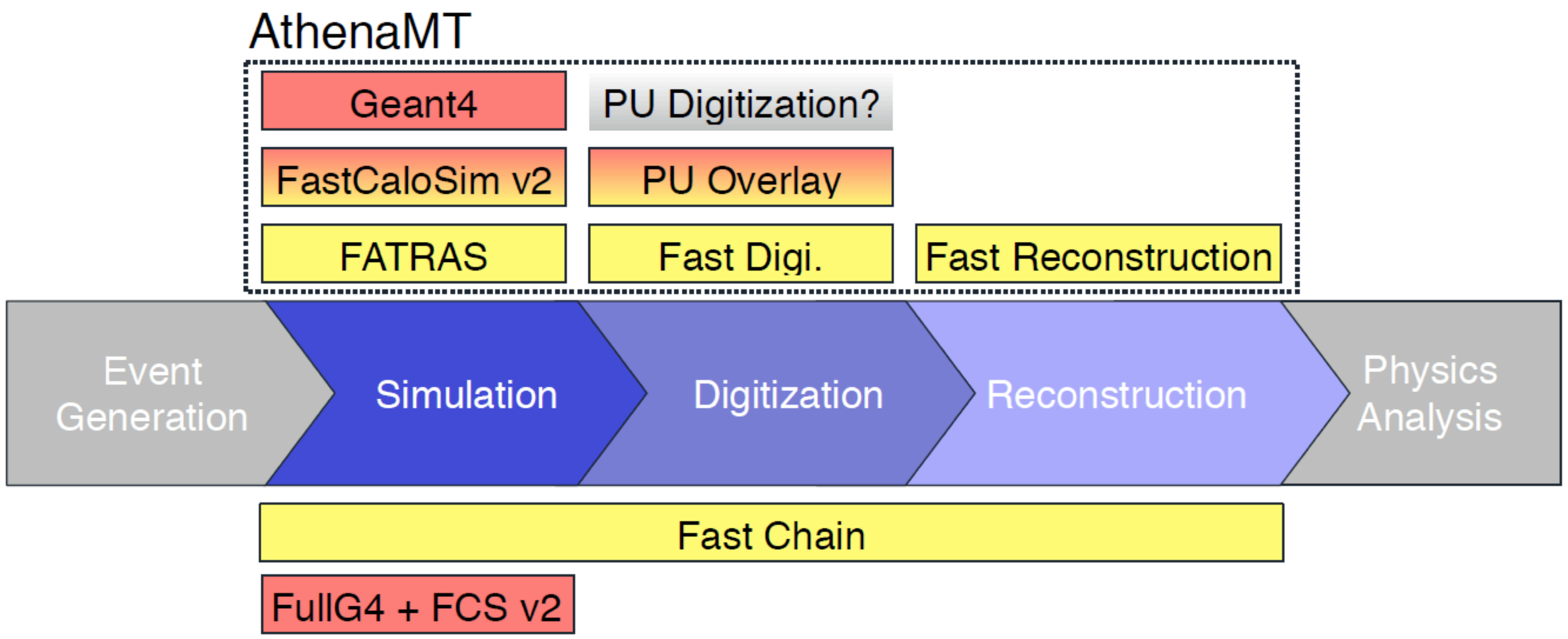
→ **AthenaMT is doing its job.**

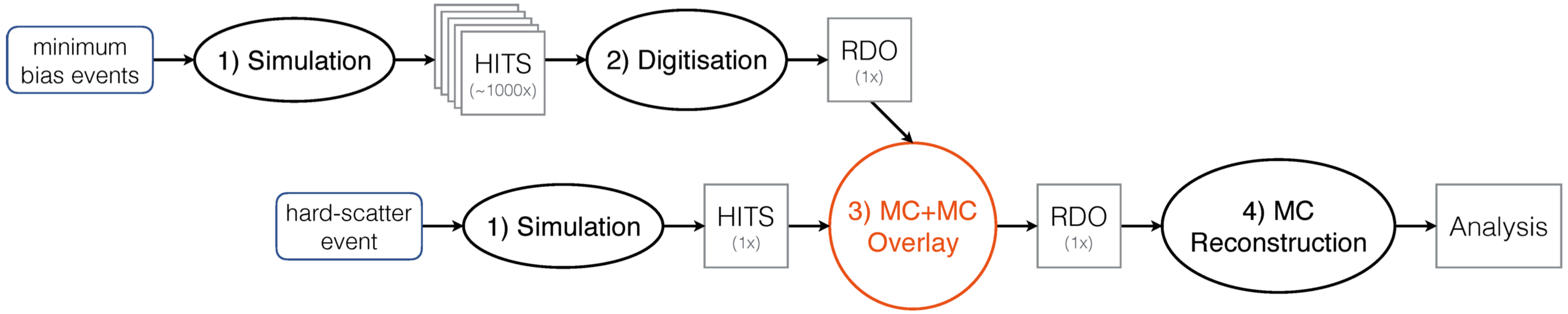
Extensive performance tuning, debugging and validation to follow in the next 5 months

Milestone	Threads	Events	Description	Target	Status
I	1	O(10)	Technical operability of MT schedulers	May 2019	✓
II	1	O(10K)	MT schedulers with 1 thread give identical results to serial	July 2019 (achieved Sept)	✓
III	>1	[illegible]	Multi-threaded operation	October 2019	✓
IV	>1	O(10K)	Physics validation & s/w performance with multiple threads	February 2020	✓
V	≥1	O(100M)	Feature freeze and run 2 reprocessing	October 2020	
VI	>1	O(1G)	Run 3 data and MC	2021	

Simulation in Run 3

- Main features of Run 3 simulation include
  - Upgrade of fast calorimeter simulation to FastCaloSimV2 → same computing performance, better physics performance. Starting to use it this year following improvements to the simulation of hadronic shower sub-structures for TeV scale jets.
  - New Geant4 version 10.6 and Geant4 optimisations and tuning (Neutron Russian roulette, EM range cuts, Birk's Law)
    - The optimisations demonstrate a CPU improvement at the 10-20% level and are planned to go into production in early 2020
  - Multithreading using AthenaMT
    - Geant4v10 already working: work ongoing for fast simulation; validation also needed
  - Use of overlay for pile-up simulation to avoid repeated re-digitisation
  - Continue development of FastChain for deployment in Run 4

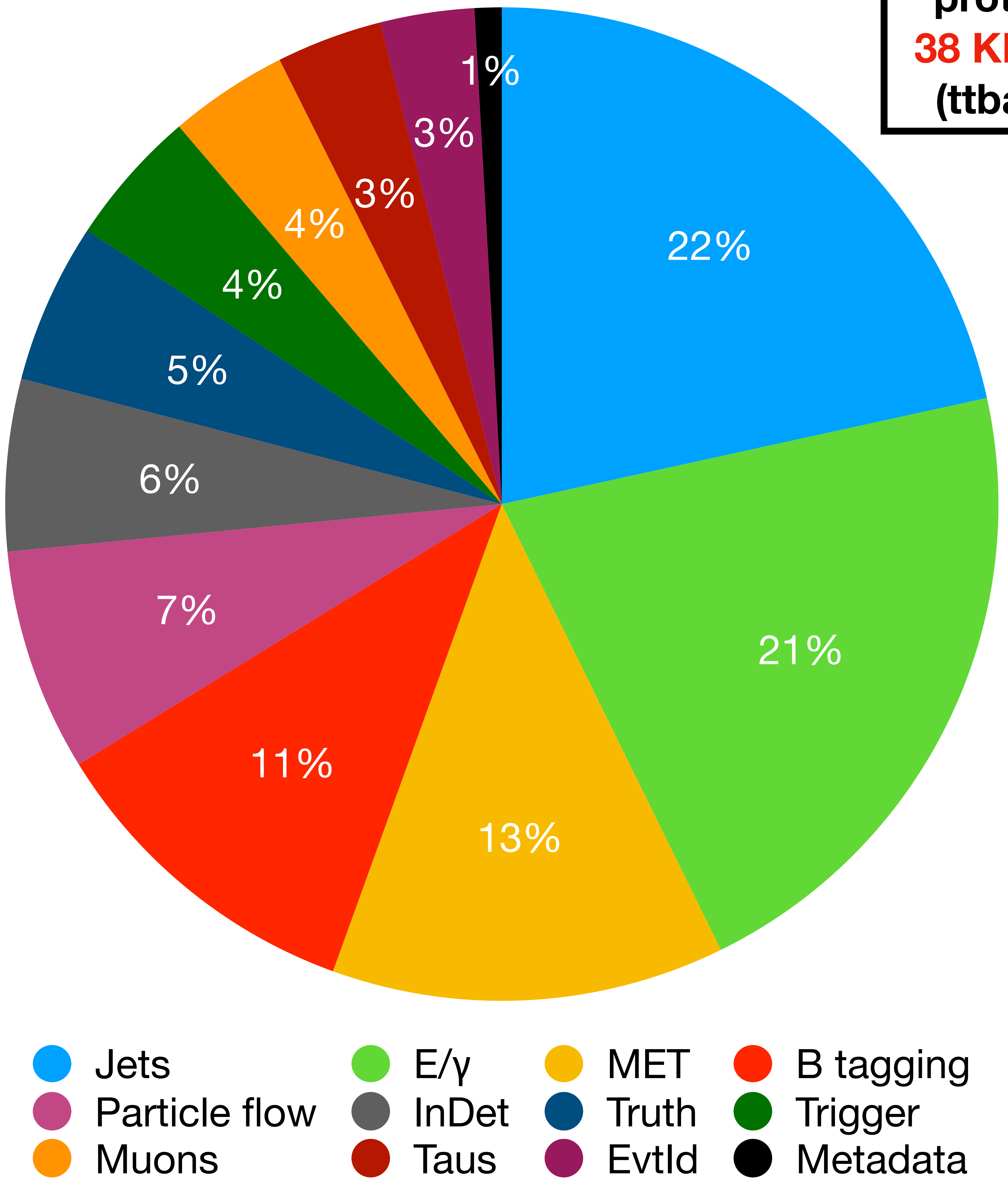




Status of the new analysis model

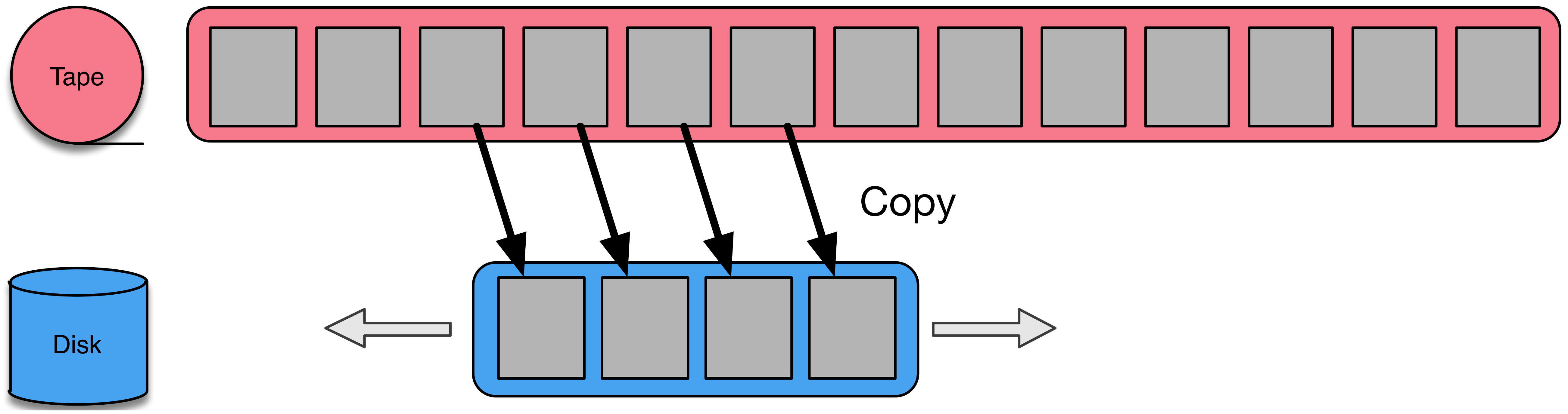
- DAOD_PHYS production under way to phase in the new model
  - Mechanism for skimming DAOD_PHYS to be tested in the coming weeks
  - New physics analyses will be asked to migrate to the new format once it is available
- Working on how to obtain systematics with DAOD_PHYSLITE → good progress made with SUSY analyses
  - Investigating using DAOD_PHYSLITE directly in the Python data-science ecosystem and via new access methods (e.g. for Run 4)
- Code for applying lossy compression ready and will be introduced progressively over the course of 2020, where full float precision exceeds the instrumental precision

DAOD_PHYS prototype:
38 KB/event
(ttbar MC)



Data carousel status

- Data carousel: on demand reading from tape without pre-staging
  - Uses a rolling disk buffer whose size can be tuned to suit available resources and production requirements
  - If used for DAOD production, allows us to reduce the amount of AODs on disk → significant savings



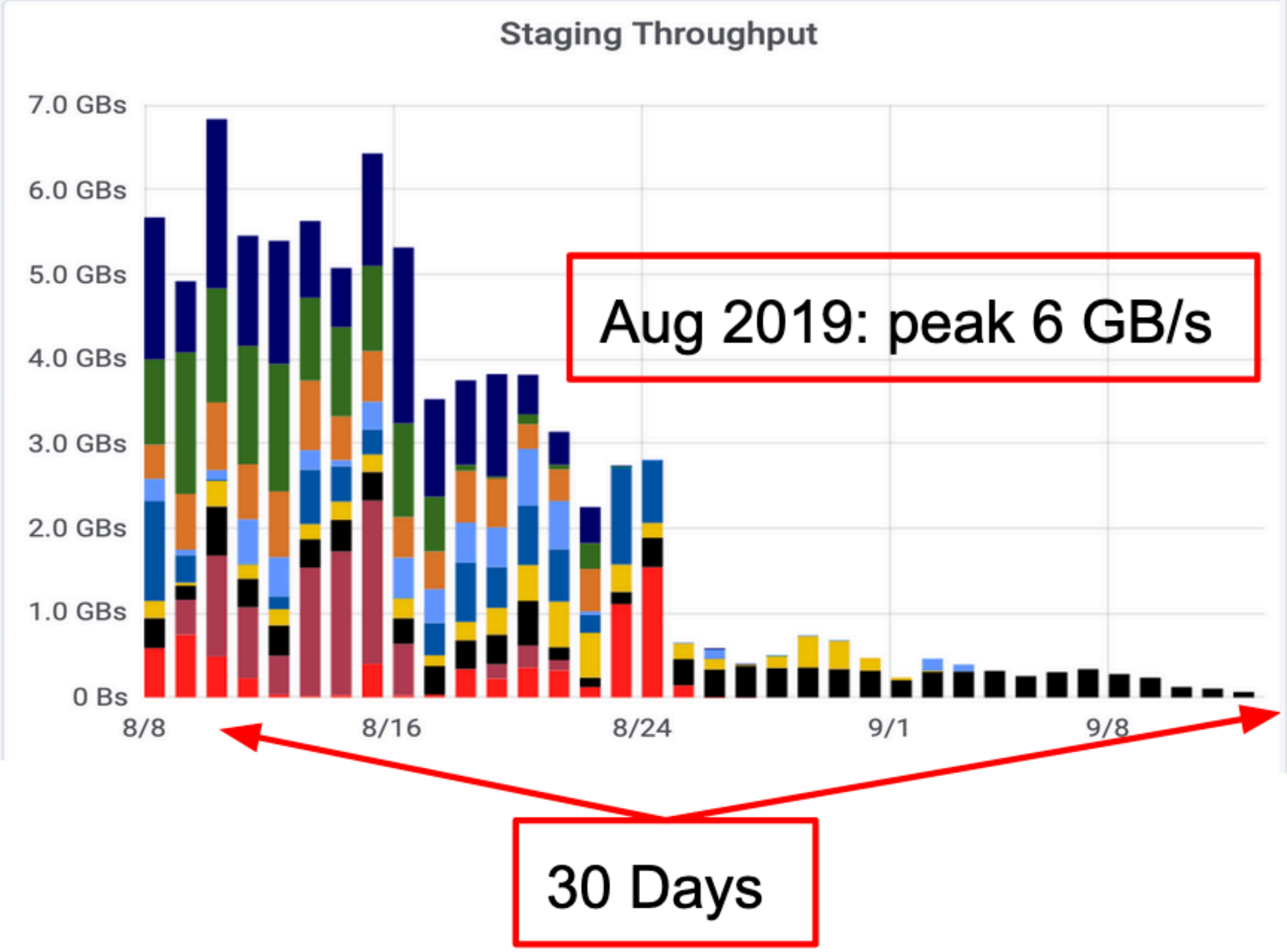
- Key to success: rate at which data can be staged to disk at the Tier 1 sites
- ATLAS has invested several FTEs for more than a year to improve the usage of tapes: discovered lot of room for improvements (without major hardware investments)
- With the introduction of the data carousel the FTS service will have an even more central role so it is essential that it remains solid and reliable, and flexible enough to integrate new workflows

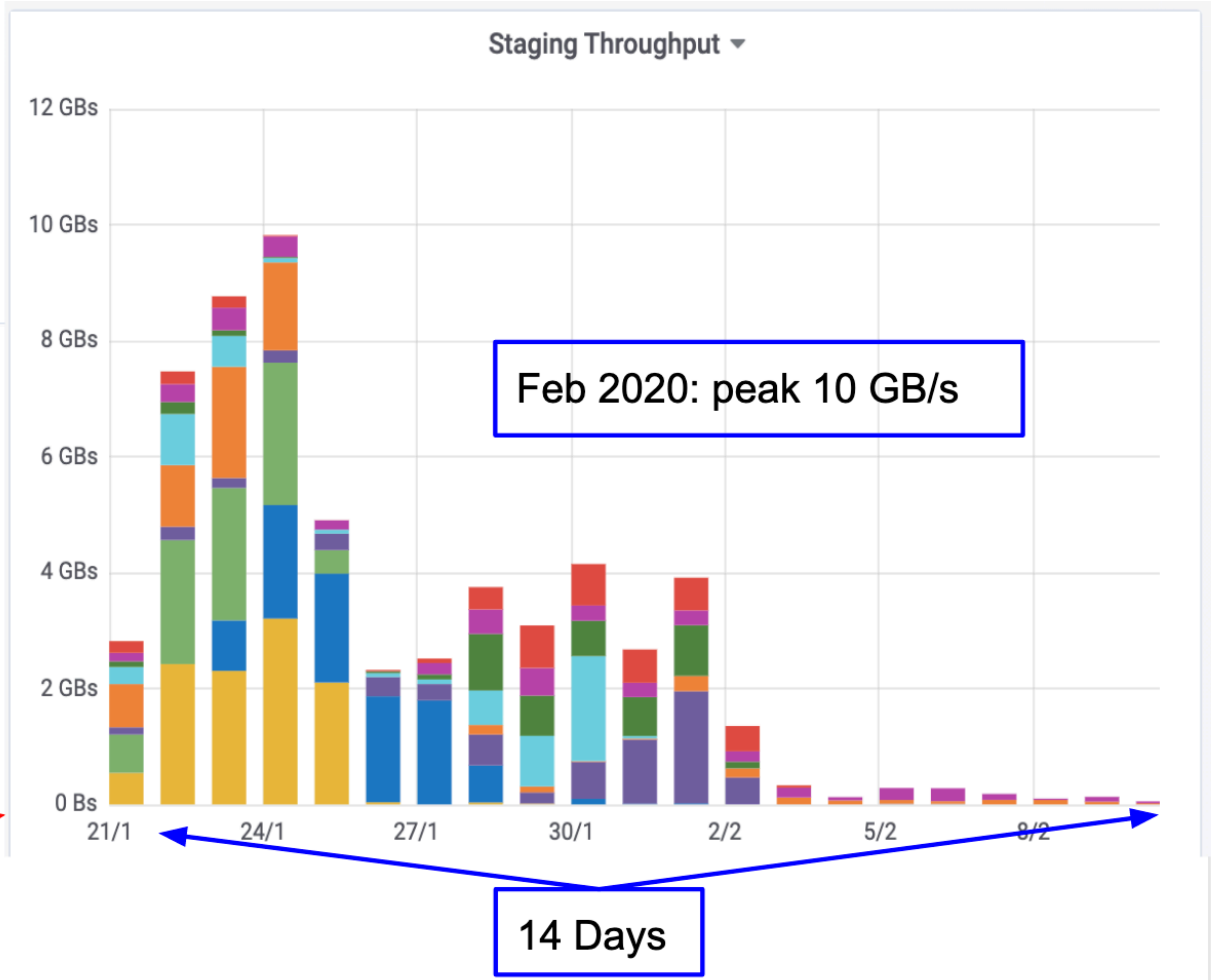
Data carousel status

- From August to February: new "site staging profile" extension

- Still lot of work to do, e.g. overall Tape recall efficiency is on the low end:

  - Push forward smart writing: crucial to engage dCache team





Run 4 preparations

- Our **Conceptual Design Report** is now at an advanced stage, will be ready in time for the May LHCC review
    - Main sections: core software, detector description, event generators, simulation, reconstruction, visualisation, analysis model, role of machine learning, CERN infrastructure, evolution of distributed computing, evolution of databases, physics impacts, resource requirements projections, milestones
- Heterogenous computing
    - Evaluation of Intel's new OneAPI toolkits ongoing
    - Ongoing work to equip the Athena framework to offload work (of any kind) to an accelerator - handles for building CUDA code as part of the larger Athena build now in place
- Disk space can be overlooked as a problem for HL-LHC - important approaches for ATLAS include:
    - Data carousel and very small analysis formats (DAOD_PHYSLITE) [both for Run 3]
- Migration new Conditions data model CREST

Summary

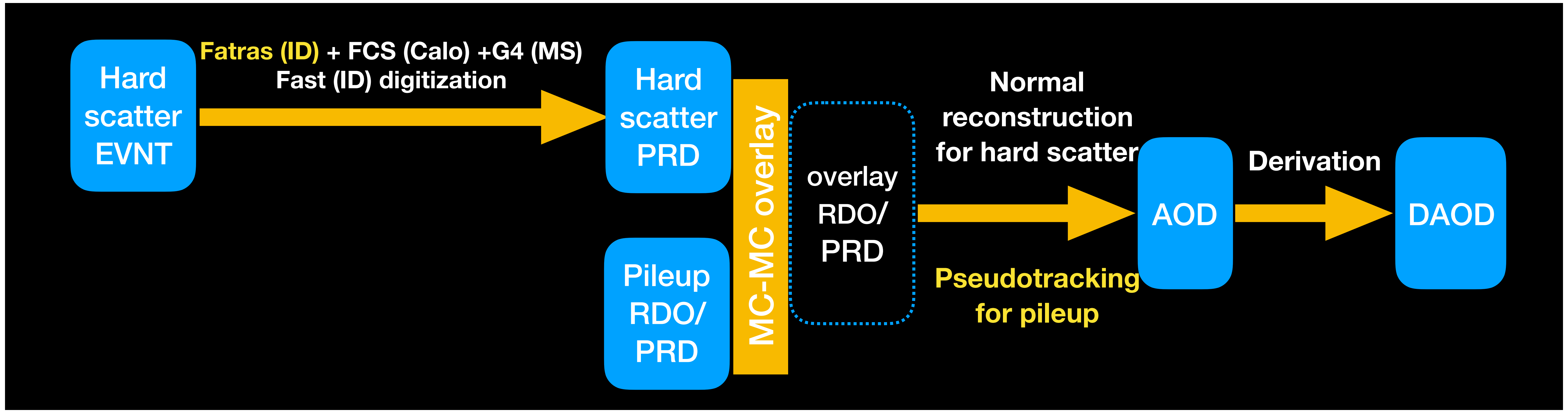
- ATLAS continues to make efficient use of its computing resources
- Resources for 2021 confirmed and involve increases of 15% CPU, 20% disk and 10% tape w.r.t. 2020
- Significant progress towards all of ATLAS' main goals for Run 3
  - New analysis model with data carousel
  - Multi threaded software framework
  - Heavier use of fast simulation and pile-up overlay
- The Run 4 CDR is in an advanced state and will be ready in time for the May review
  - Preparations are intensifying across a range of activities including tracking, simulation, heterogeneous computing

Extra slides

Size of DAODs on disk

Simulation in Run 4

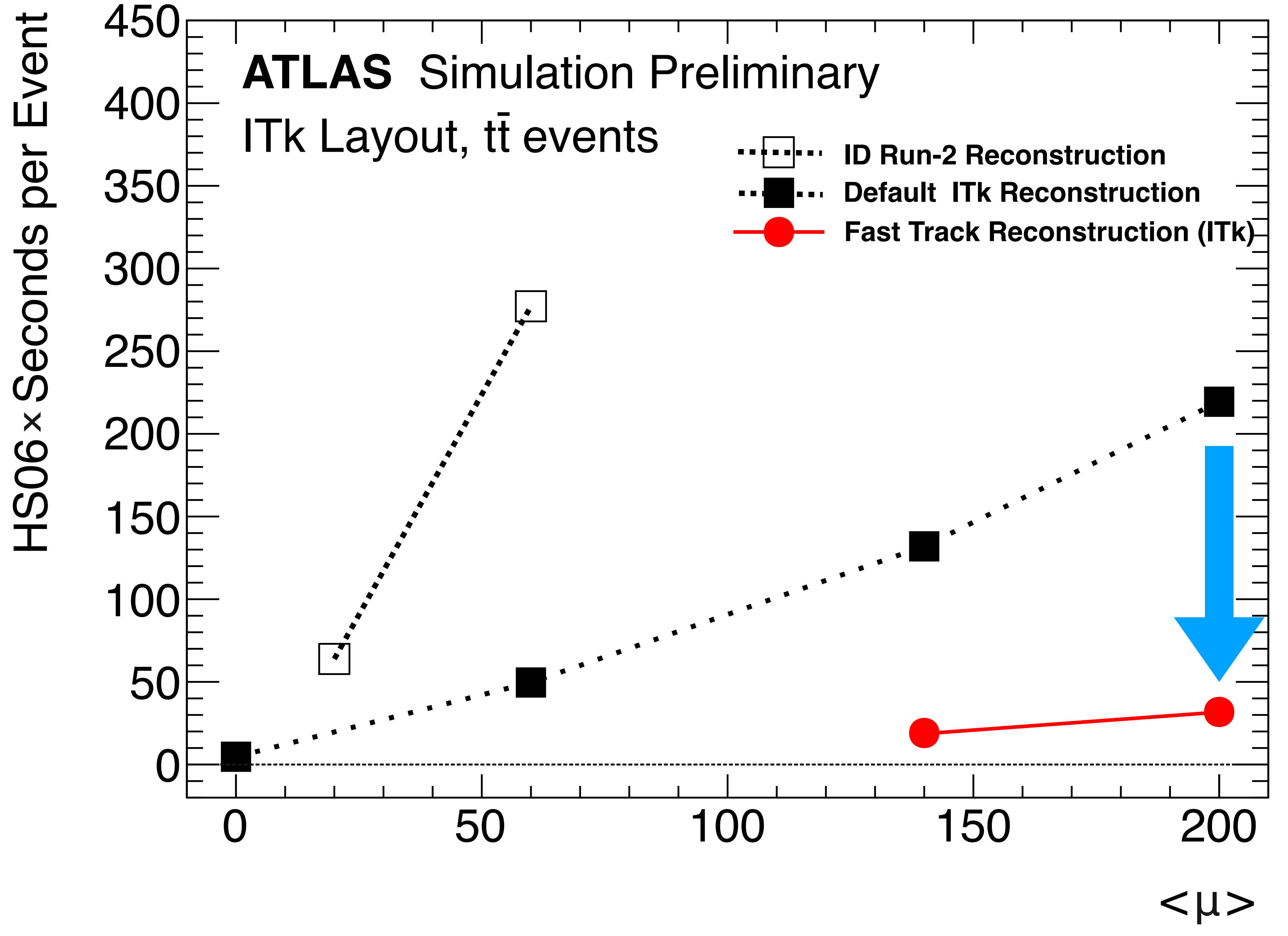
- Main project: *fast chain*
  - Addresses CPU and disk by significantly speeding up all parts of the chain, such that fewer intermediate data formats needs to be stored
  - FastCaloSim v2 is used for the calorimeter simulation
  - Pileup simulation not critical (events are re-used across datasets with MC-MC overlay)
  - Hard scatter and pileup treated independently
  - Fast simulation, digitization, reconstruction independent, can be deployed separately
  - Validation under way for the various parts of the chain: ambition to use fast chain already in Run 3 for signal MC

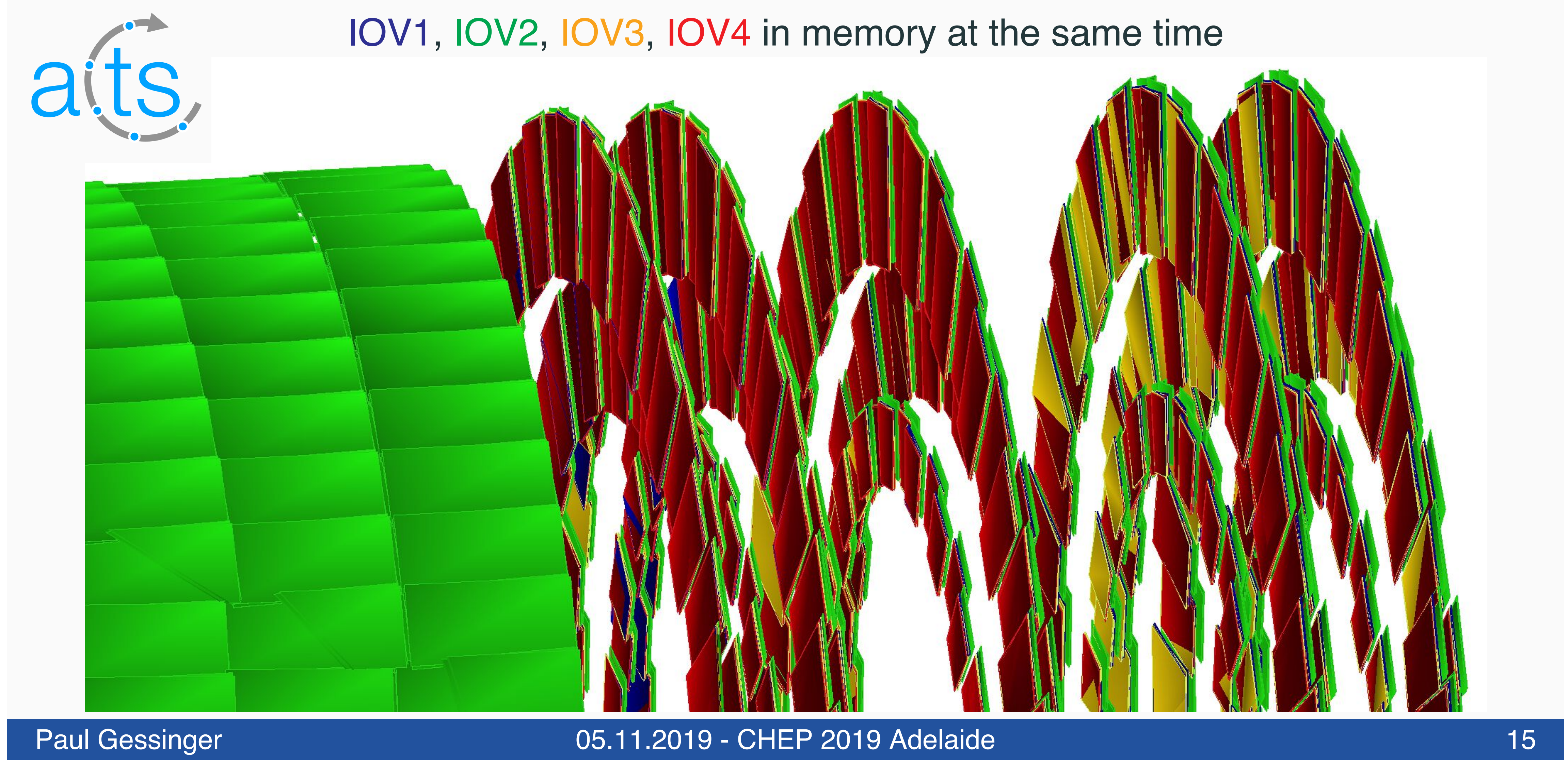


- Making small physics compromises can lead to significant speed-up. Example: Neutron Russian Roulette. Requires continuing and intensifying the existing good interactions with the G4 team
- Minor computing efficiencies can also have a big impact. Example: single static G4 library
- Heterogeneous computing: simulation is clearly *the* application that could make the biggest impact
- Intensive work on simulation with deep learning (GANs, VAEs) but overall impact on software and physics performance not yet clear. Can be plugged into fast chain along with the other components.

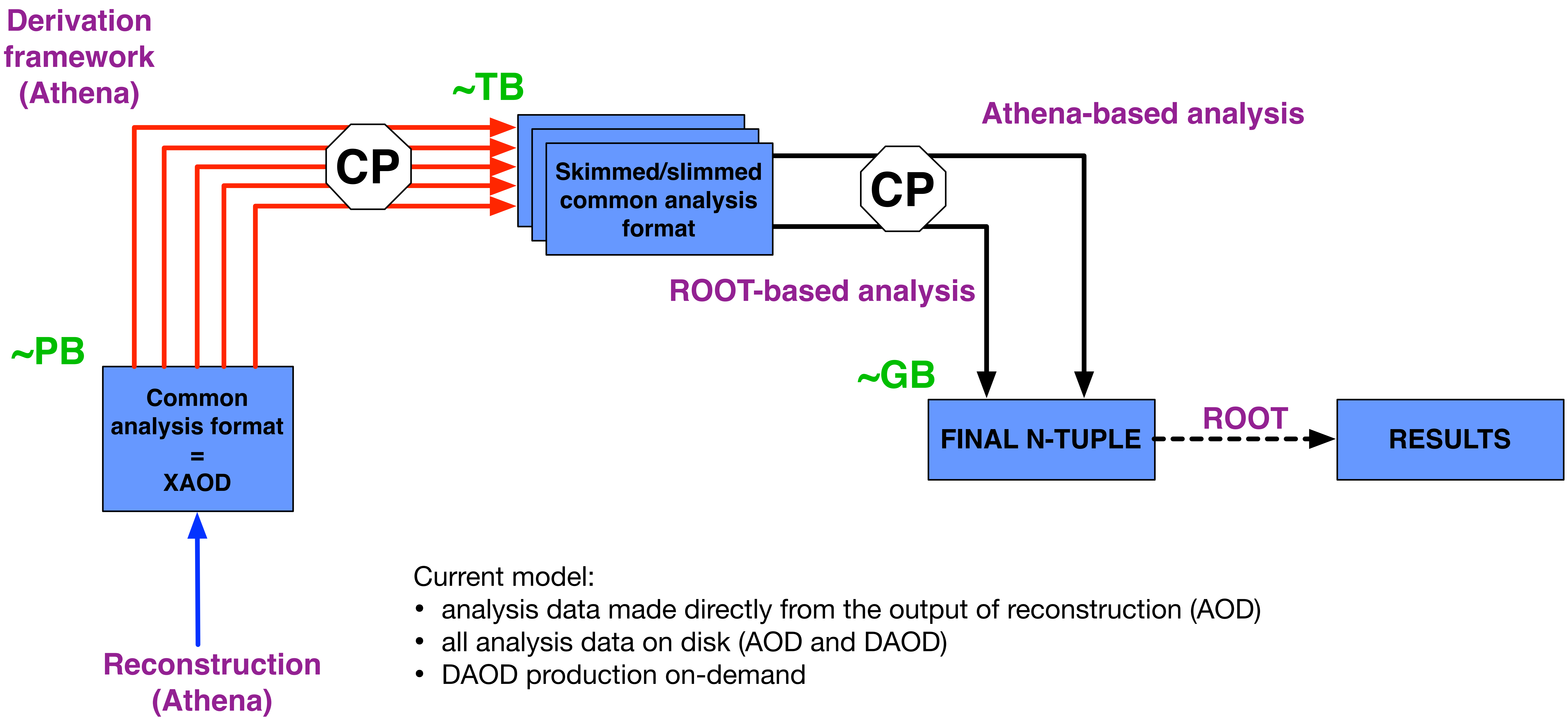
Reconstruction in Run 4

- Main work related to inner detector track reconstruction, the main consumer of CPU:
  - Optimising, revising and adapting the existing reconstruction workflows to HL-LHC conditions and detector layouts
    - Some of this work will be applicable to Run 3 with a potentially large CPU saving → under intensive study by the inner tracking performance group
  - ACTS: significant re-design and simplification of the core tracking software and EDM, with concurrency built in from the start (talk by Paul Gessinger @ CHEP)
    - Contains a fast-tracking module for fast chain (acts-fatras)
    - Some components will be in place for Run 3
- Contribution of heterogeneous computing to reconstruction not clear; to be investigated in detail once the core software is able to support it





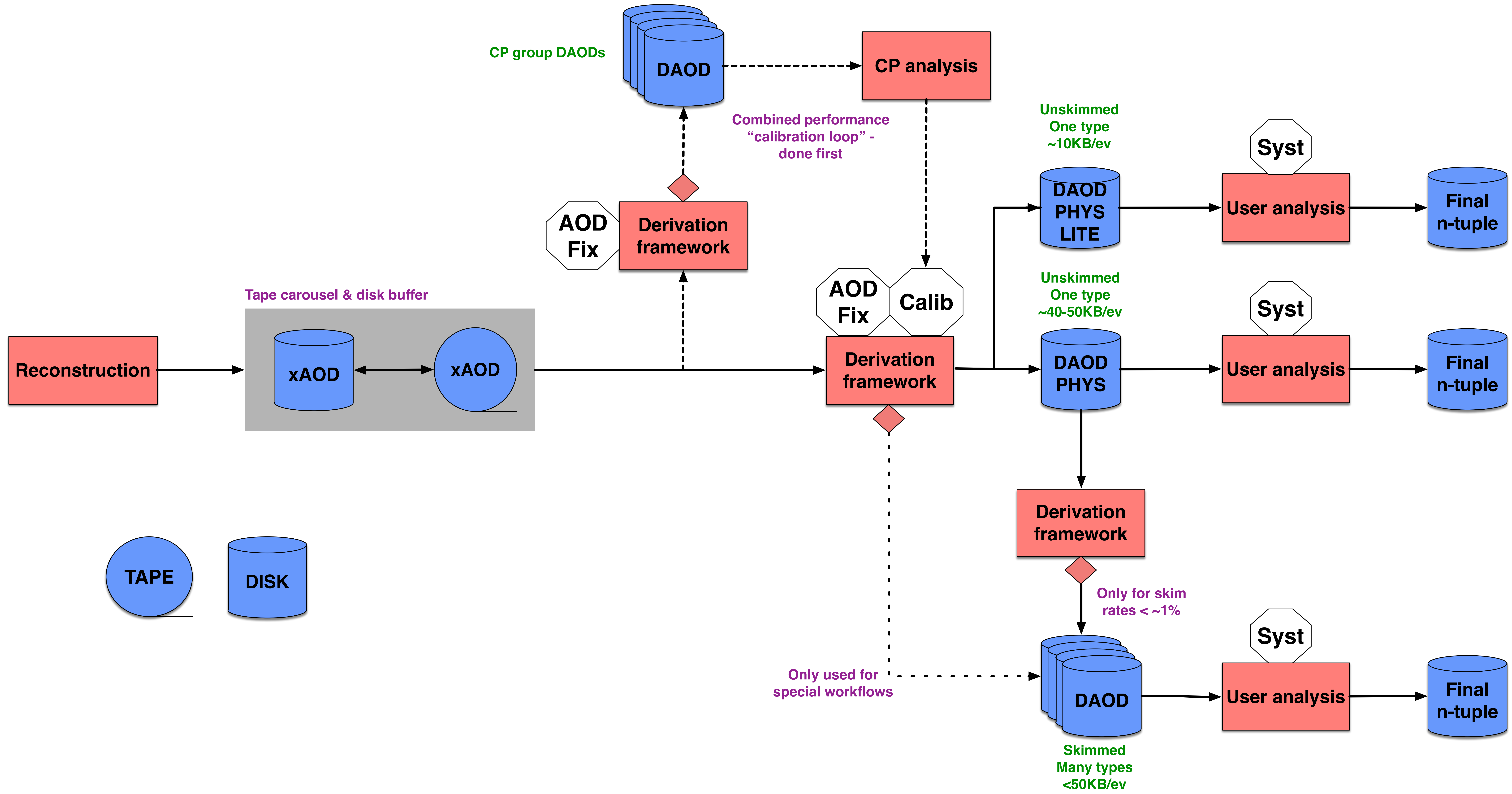
Run 2 analysis model



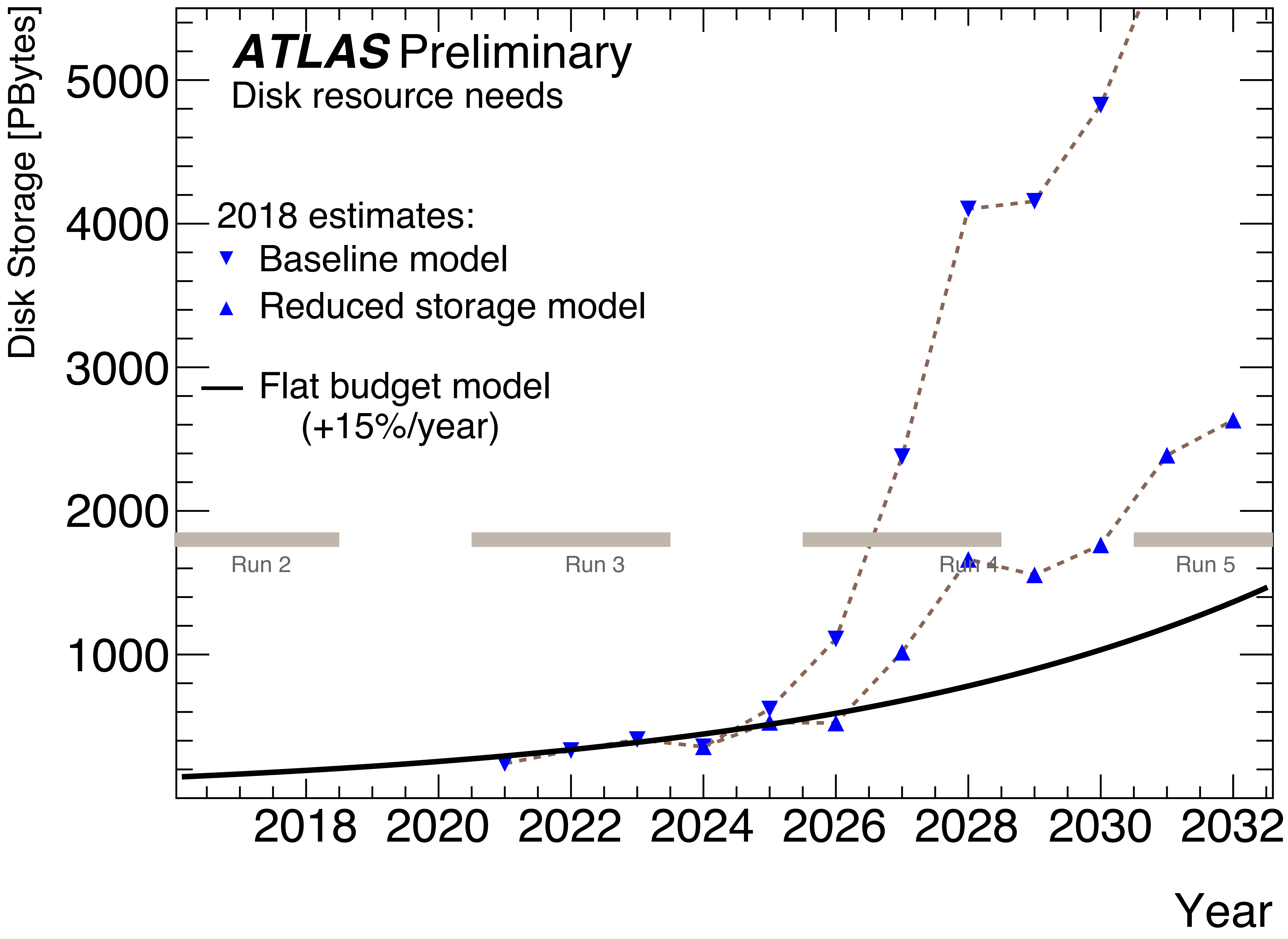
Current model:

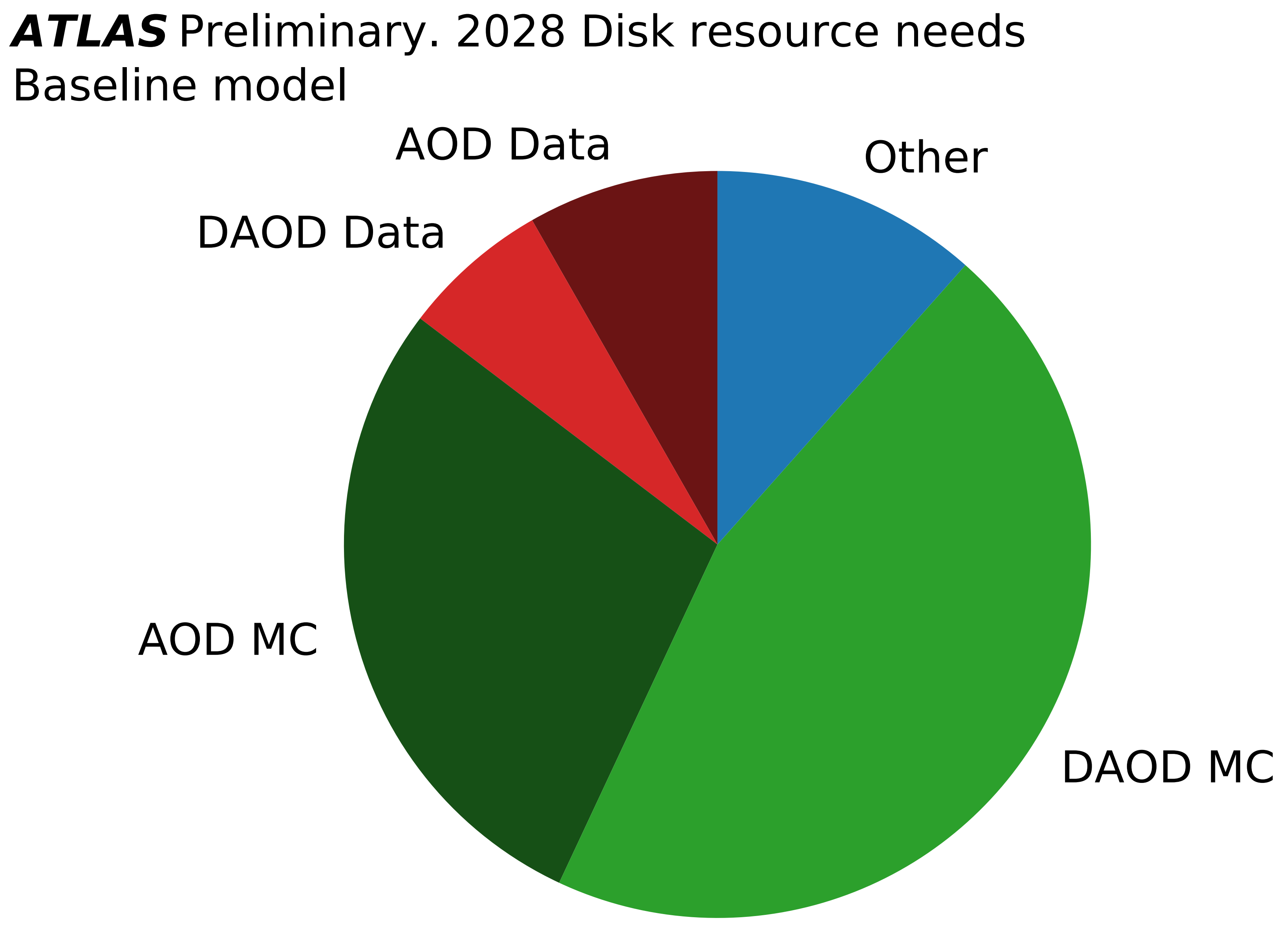
- analysis data made directly from the output of reconstruction (AOD)
- all analysis data on disk (AOD and DAOD)
- DAOD production on-demand

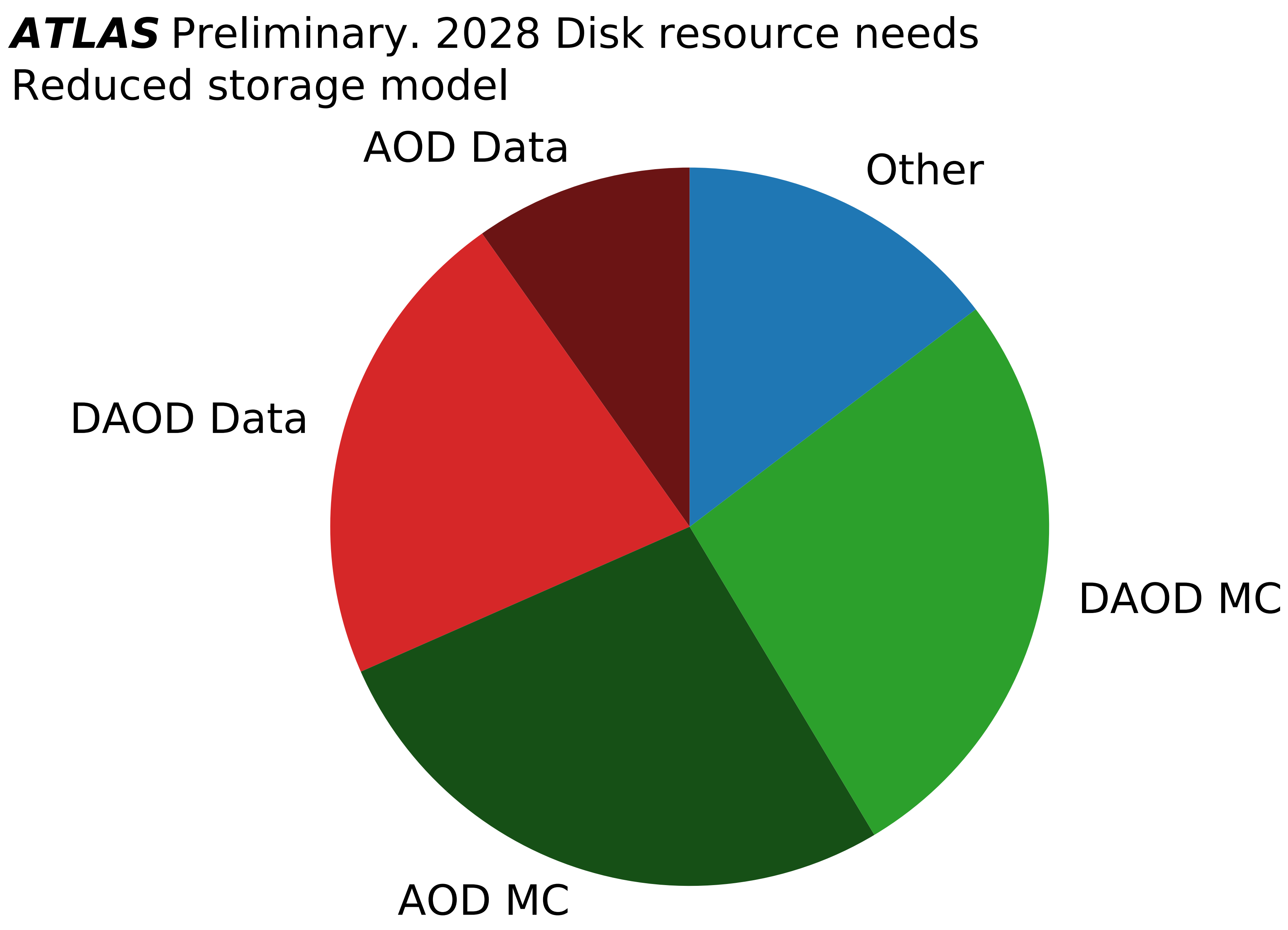
Run 3 analysis model



Run 4 extrapolations: storage







Run 4 extrapolations: CPU