Derived Physics Data Production in ATLAS: Experience with Run 1 and Looking Ahead

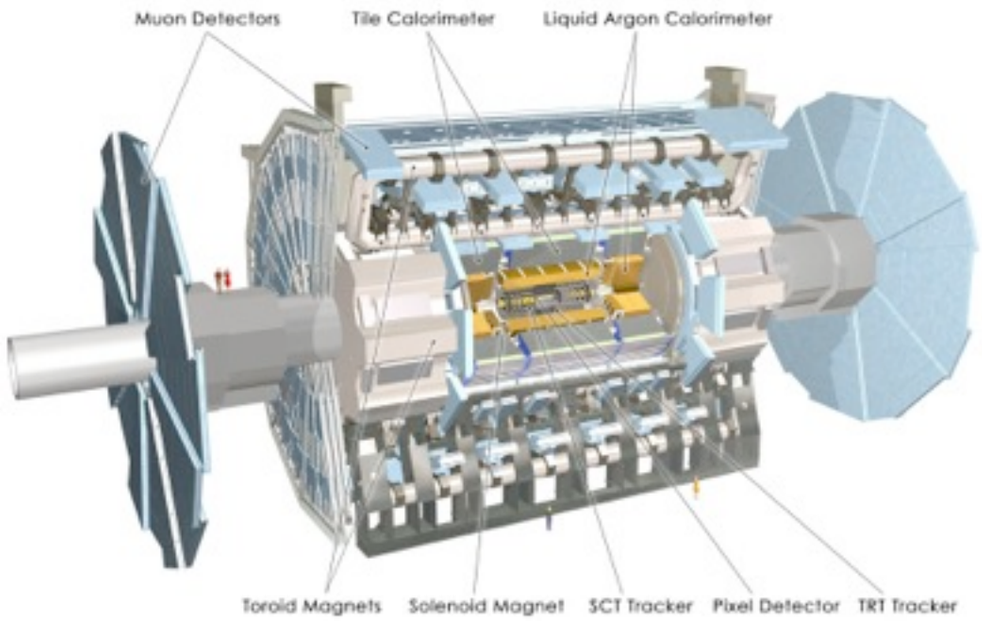
Paul Laycock*
on behalf of the ATLAS Collaboration

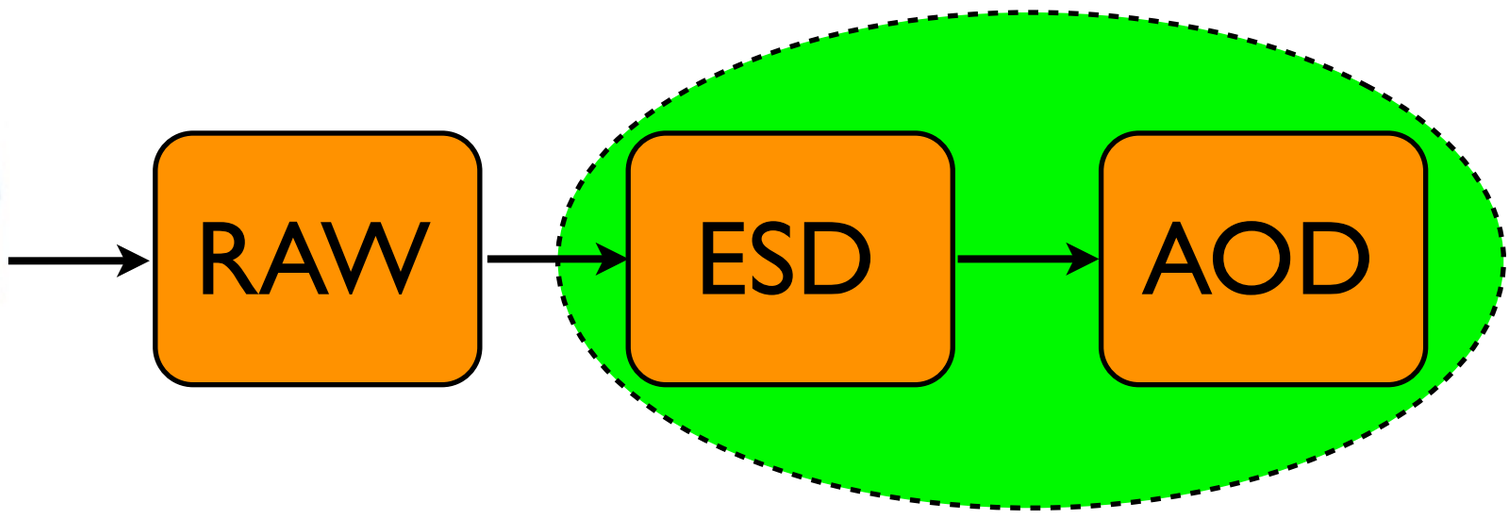
UNIVERSITY OF LIVERPOOL

CHEP, October 17th, 2013

**with N Ozturk, M Beckingham, R Henderson and L Zhou*

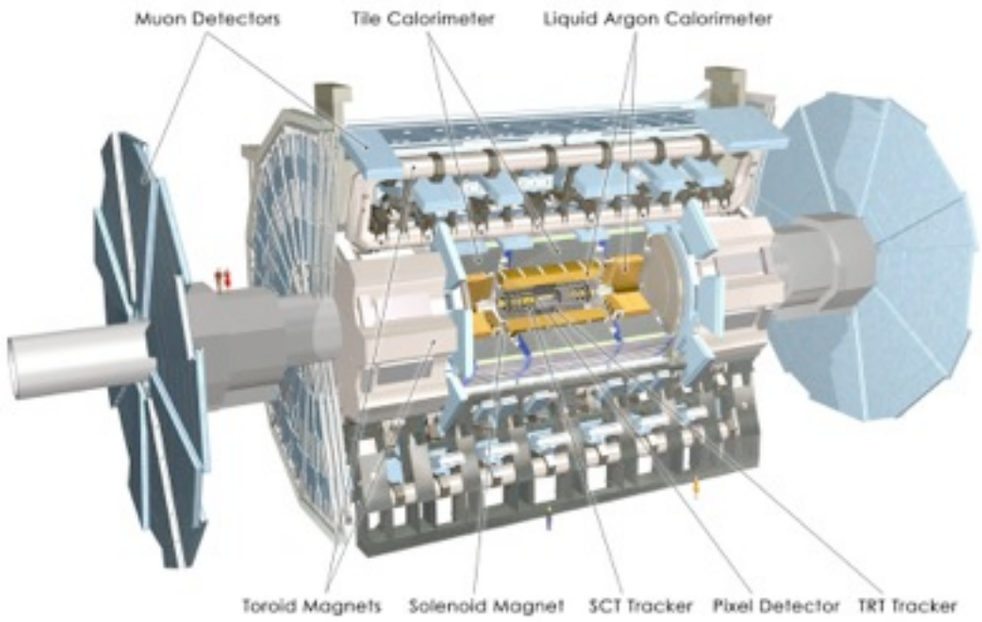
From ATLAS to the AOD

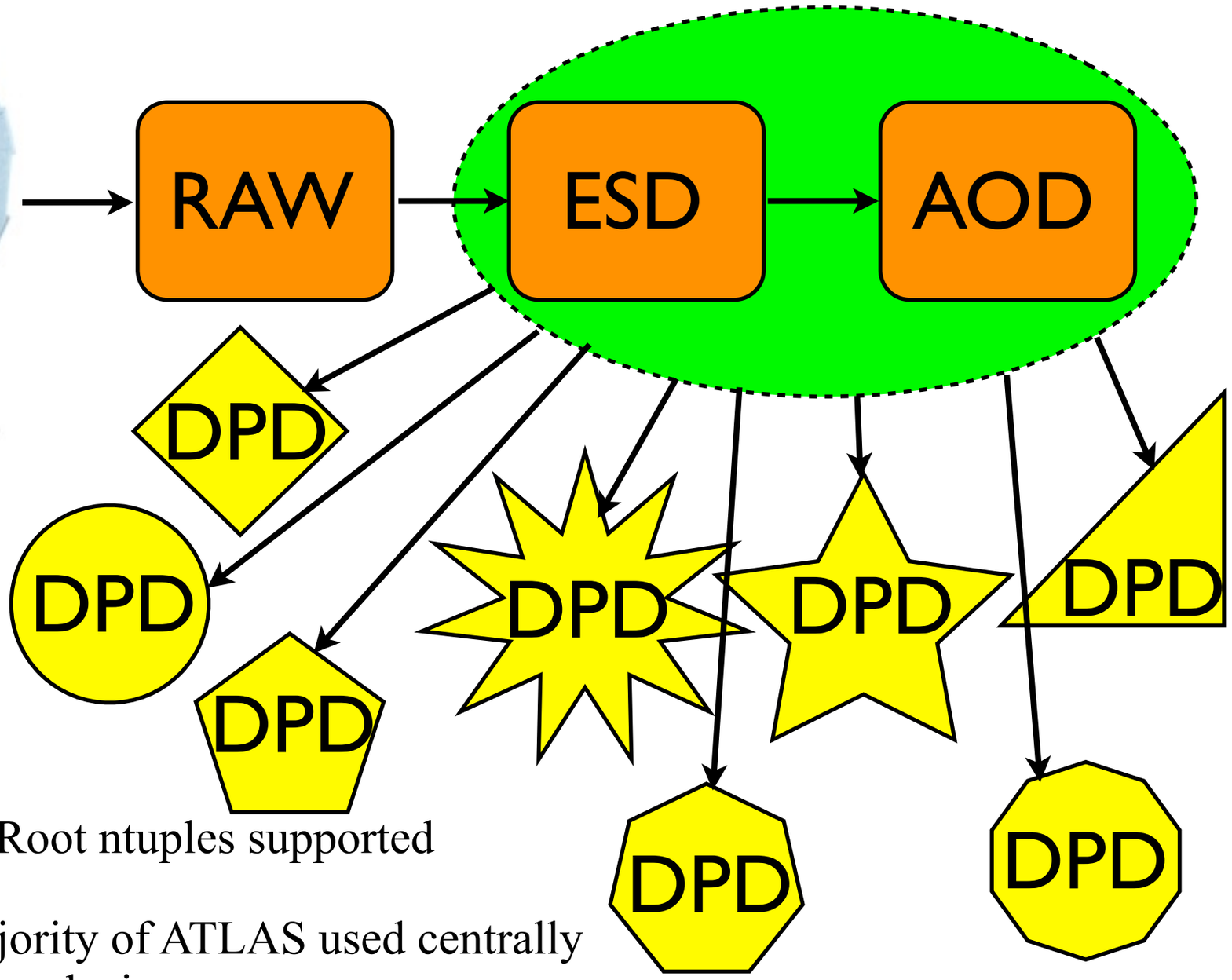




- The RAW detector data were promptly reconstructed at the Tier0 into ESD (Event Summary Data) then AOD (Analysis Object Data) using the Athena framework, based on Gaudi
- ESD became a temporary format, deleted after a finite amount of time
- AOD was originally the go-to analysis format for the end user, requiring analysis in athena
- Tools provided to create derived formats from AOD (or ESD) with the main intention of being able to remove events and remove objects (or whole containers of objects)
  - Derived Physics Data or DPDs
- These end-user formats were organised at the physics and combined performance (CP) group level and then later centrally managed by “group production”

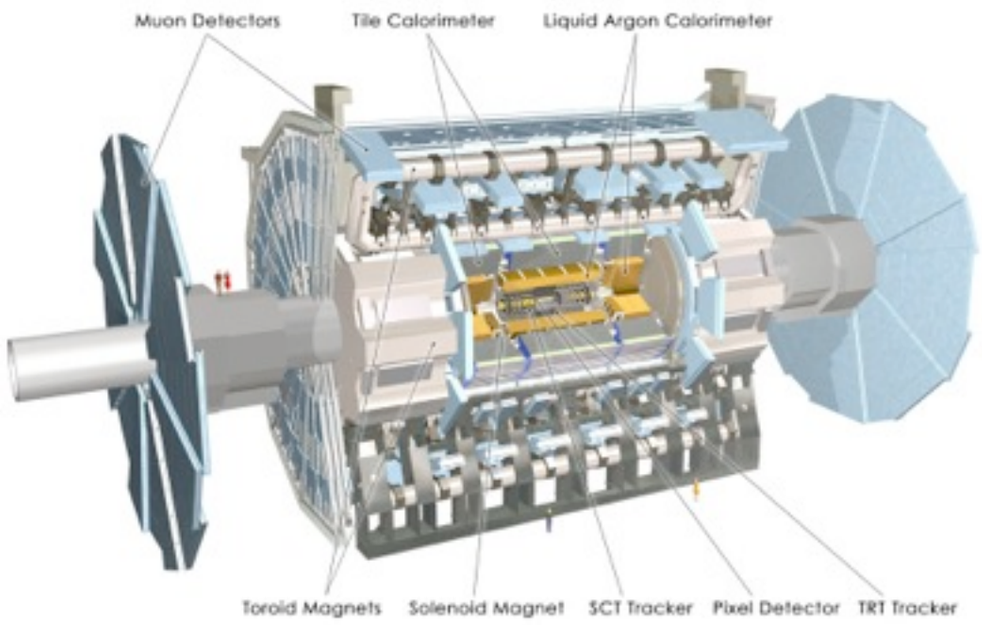
Derived Physics Data (DPD) production

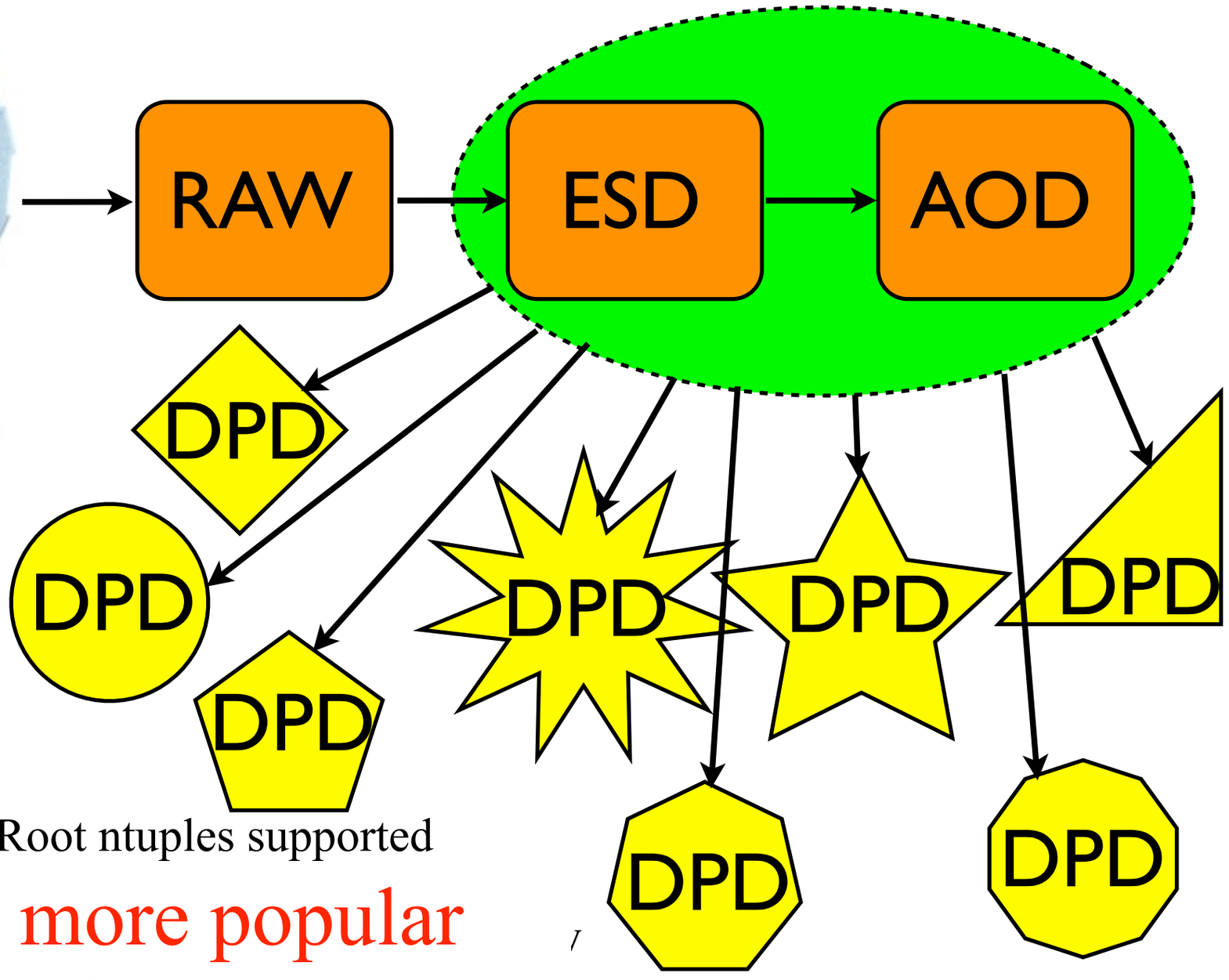




- Both athena-readable DPDs and Root ntuples supported
- By the end of Run 1, the vast majority of ATLAS used centrally produced DPDs to perform their analysis
- Each group's DPD had evolved independently, while using the same software such that common improvements could be incorporated

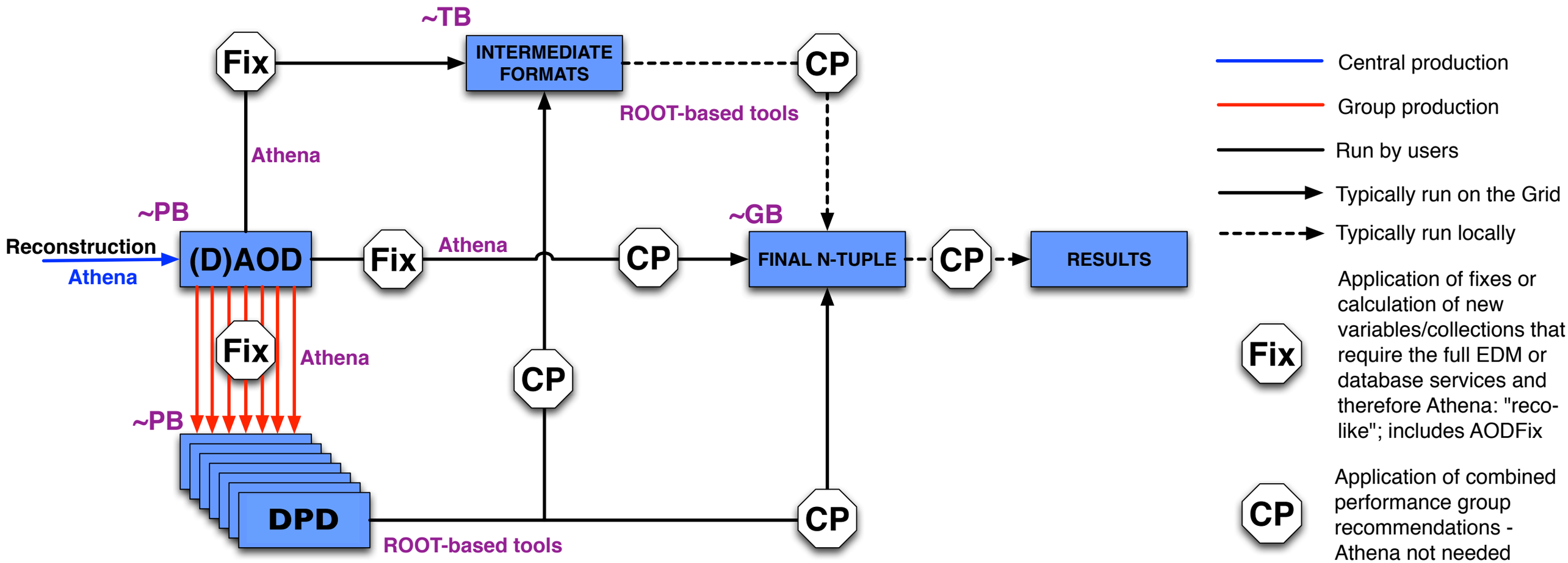
Derived Physics Data (DPD) production





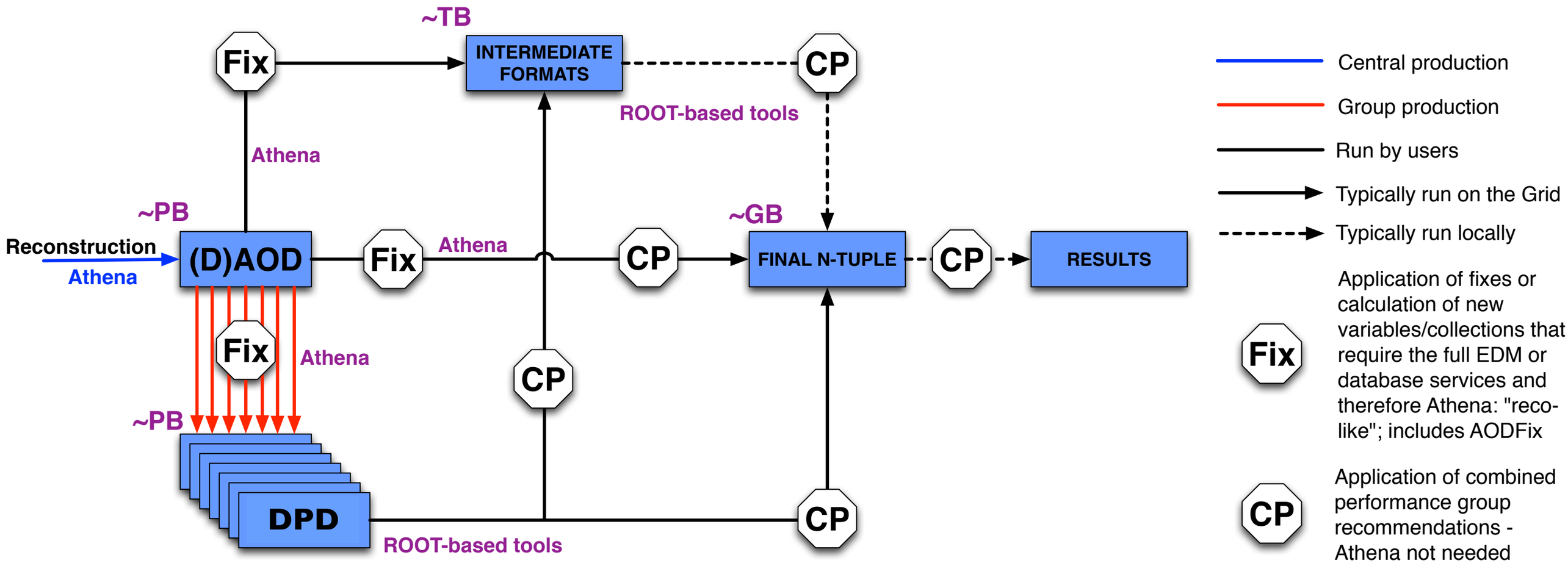
- Both athena-readable DPDs and Root ntuples supported
- Root format far more popular than athena format
- Each group's DPD had evolved independently, while using the same software such that common improvements could be incorporated

The run 1 ATLAS analysis model



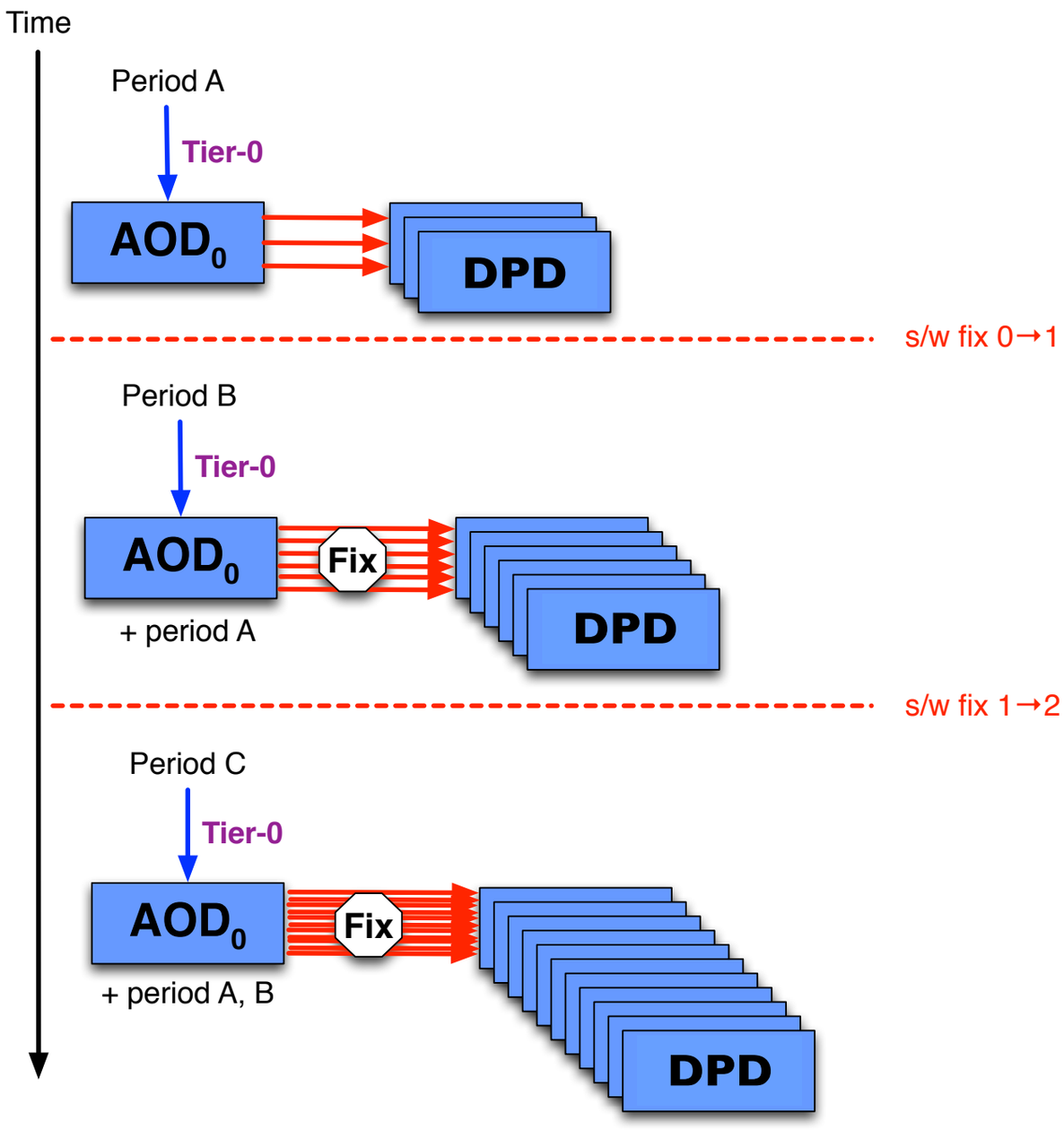
- Towards the end of 2012, many of the group's Root DPDs had become very large, in order to incorporate all of their group's requests for variables and variants of objects
  - They were so large that in several cases they approached the size of the AOD
- These PB-size formats would often be reduced by individual users down to smaller, intermediate formats of order a TB in size

The run 1 ATLAS analysis model



- The content of the DPDs was the “best knowledge” available at the time of production and in many respects superior to the AOD, as the AOD s/w was frozen before running
  - Fixes applied in a controlled way in group production
- Model of maximum flexibility was very successful
  - High publication rate same as CMS, around 270 papers from each experiment
- However, resource usage was not optimal...

The run 1 analysis model vs time



- The AOD production software was frozen before running
- Software fixes built up which could be applied to AOD, e.g. in DPD production
- Over time, the amount of software fixes became non-negligible in terms of CPU
- These fixes were being repeated for every variant of DPD

DPD production and disk

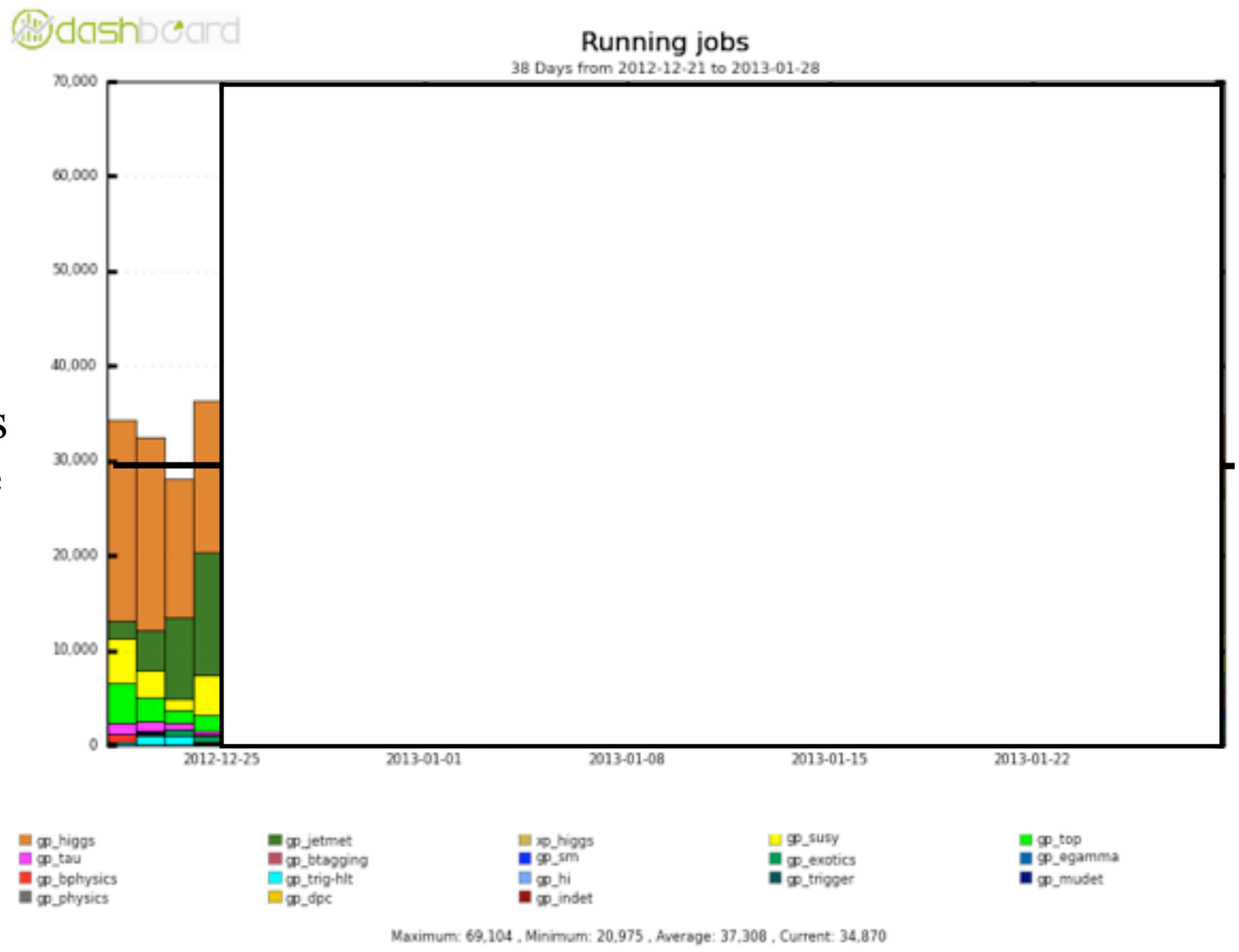
- DPD production was intended to reduce AOD to something smaller
  - Simple reduction operations assumed, but in reality reconstruction-type operations were adding large extra object collections fulfilling everyone's wishlist
- Much value was added to the DPD analysis format, but there was much repetition
  - Despite the similarities between much of the content, the independent evolution made comparisons and the prospect of merging surprisingly difficult
- The thinking was, different groups had different requirements, so different formats needed
- The consequence was that DPDs took up a lot of disk space

DPD production and CPU

- DPD production was also supposed to be CPU-light, but some reconstruction operations were very heavy
  - e.g. adding new jet collections, different optimal MET calculations...
- The combined CPU time / event for all DPDs was similar to a full RAW to AOD reconstruction job, around 30 seconds / event
- Usually DPD production ran both synchronously and asynchronously to AOD production
  - Synchronous production tended to be more conservative (faster turnaround)
  - Asynchronous production was staggered between groups
- No computing resource problem seen... until **all** data and MC AODs were reprocessed and consequently **all** groups wanted to remake their DPDs at the same time
- The massive production was prepared, but if all group DPD production jobs were produced at the same time, the large CPU time would require a lot of resources...

DPD production and resources at peak times

- Following a successful AOD reprocessing campaign, the DPD-making software struggled to converge on a software version, as all groups suddenly tried to update at the same time



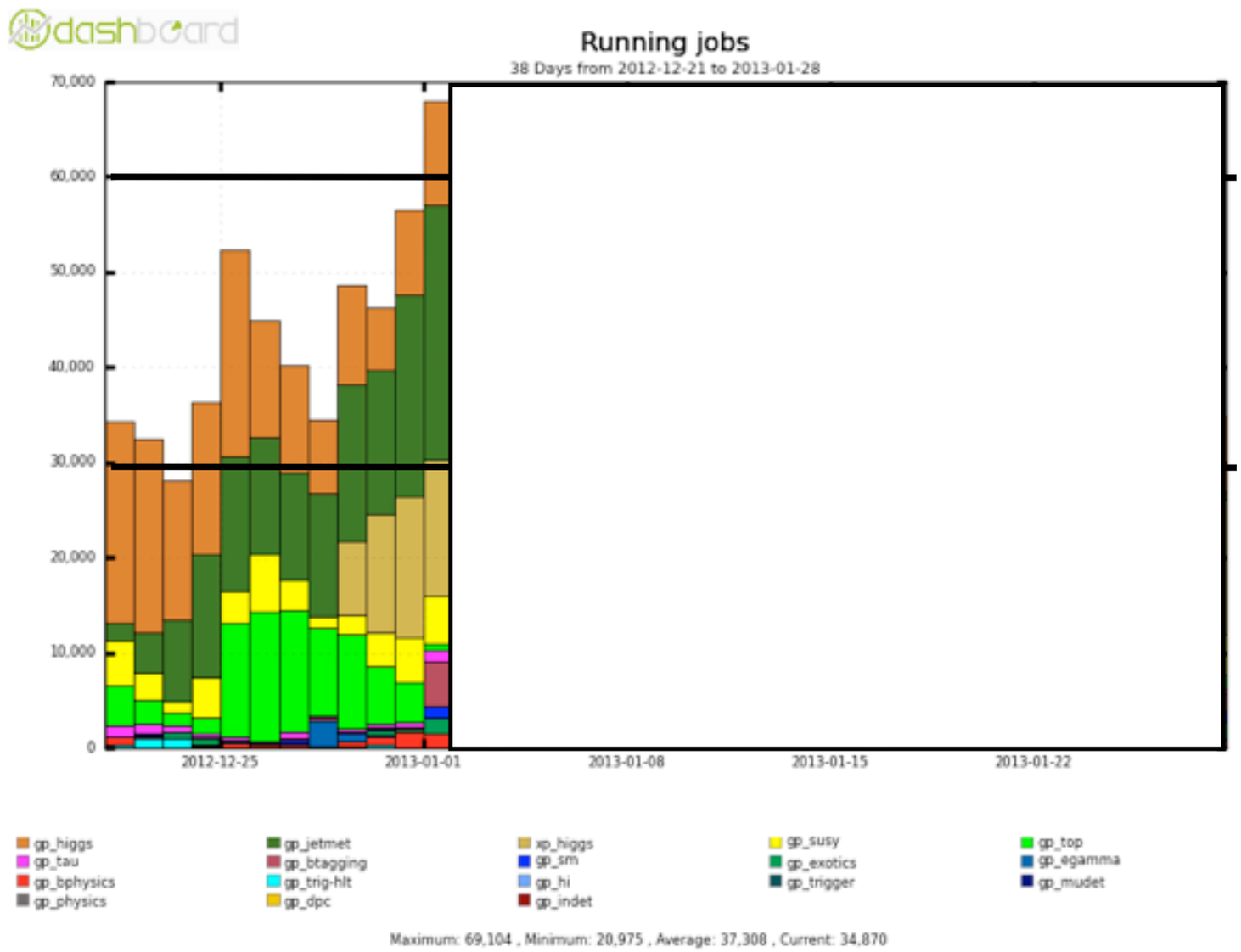
- Around 50 DPDs were made

- Jobs began to be launched on the 21st December, but the number and complexity was overwhelming, despite the fact that the production team had automated the submission procedure and had several months of experience

DPD production and resources at peak times

This didn't look like saturating of it's own accord

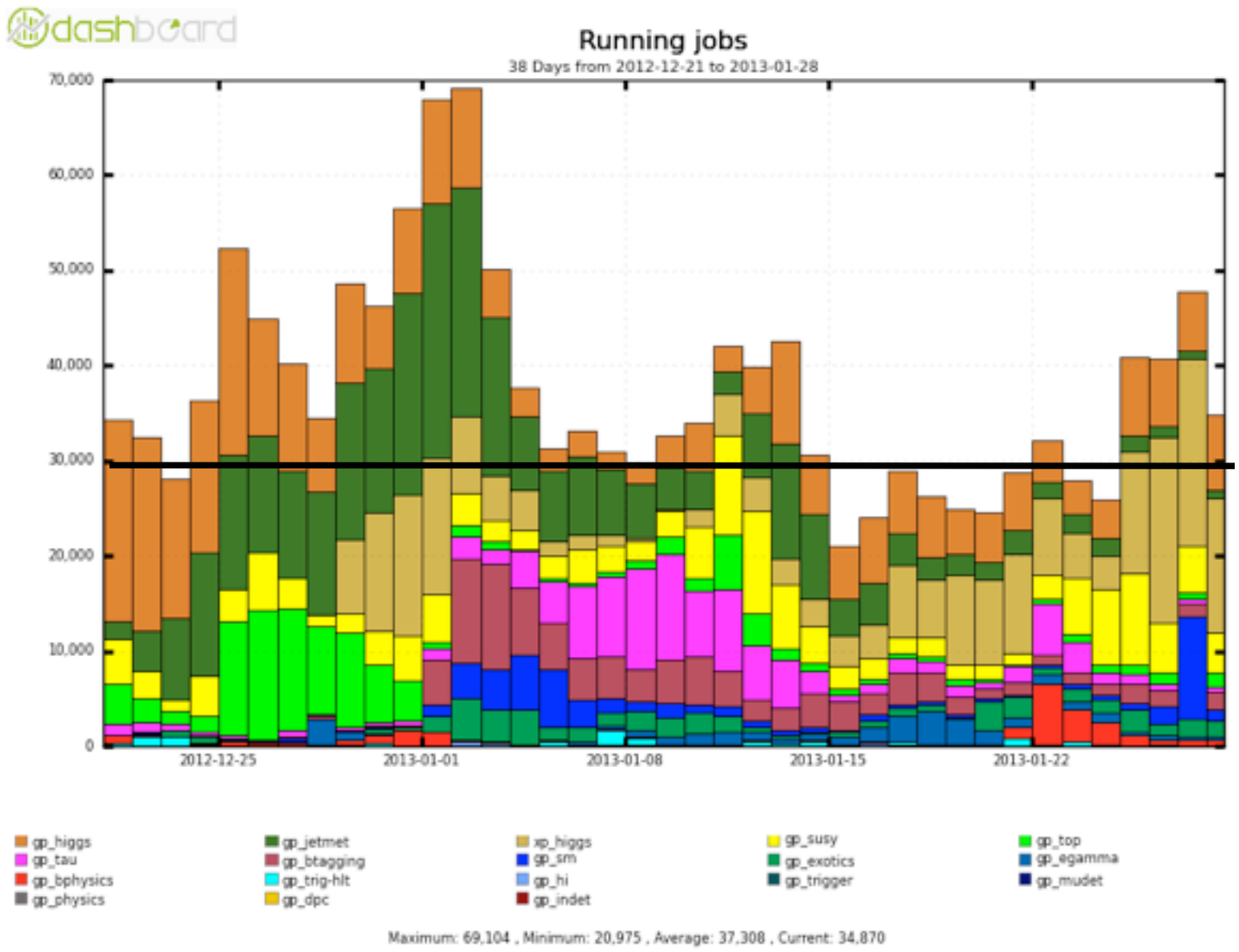
DPD production resource allowance



- Jobs continued to go into the production system, until at some point a decision had to be made as this had started to severely affect MC production, eating up resources

DPD production and resources at peak times

DPD production resource allowance

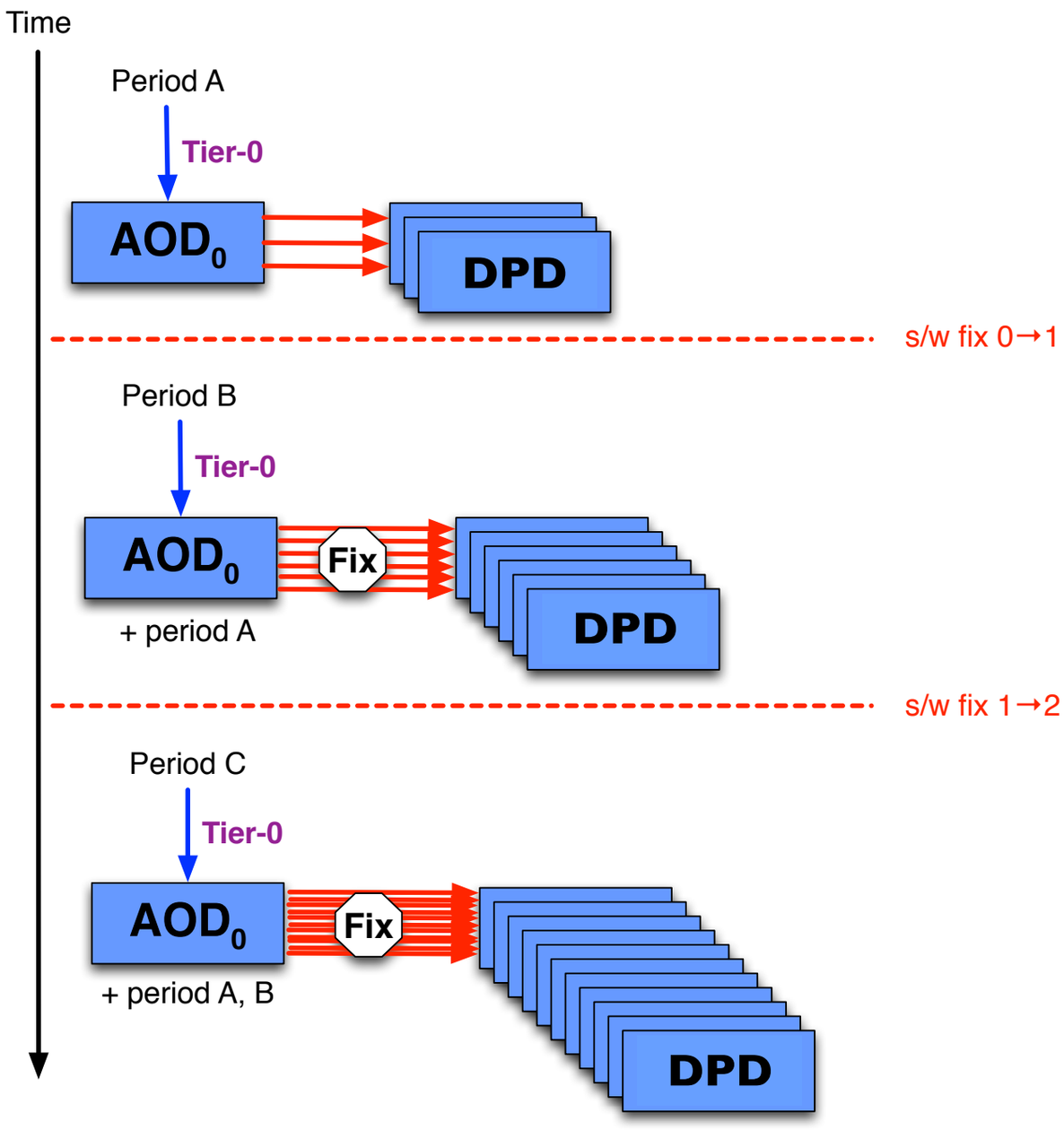


- The resources were now set and the DPD production campaign took as long as it had to
- The question was, could we do better?

The Common DPD

- The simplest solution one could think of with the minimum disruption to analysis was a unification (or at least partial unification) of the large Root-based DPDs
    - The Common DPD prototyping started
- After only a couple of months it was clear that this was actually possible
- More than that, we discovered that practically all of the AOD to Root-based DPDs could be accommodated by the Common DPD
    - The exception was the b-tagging group which needed a lot of detailed information
- After some work on consolidation on some of the larger object collections (tracks and clusters), as well as adding the final run 1 wishlist of variables that groups had planned for, the effort converged after about 6 months with a fully validated DPD and production launched
- The CPU and disk usage was thereby improved by a factor of about three
- In the meantime, a Study Group had been collecting information and had paved a way forward for the run 2 Analysis Model

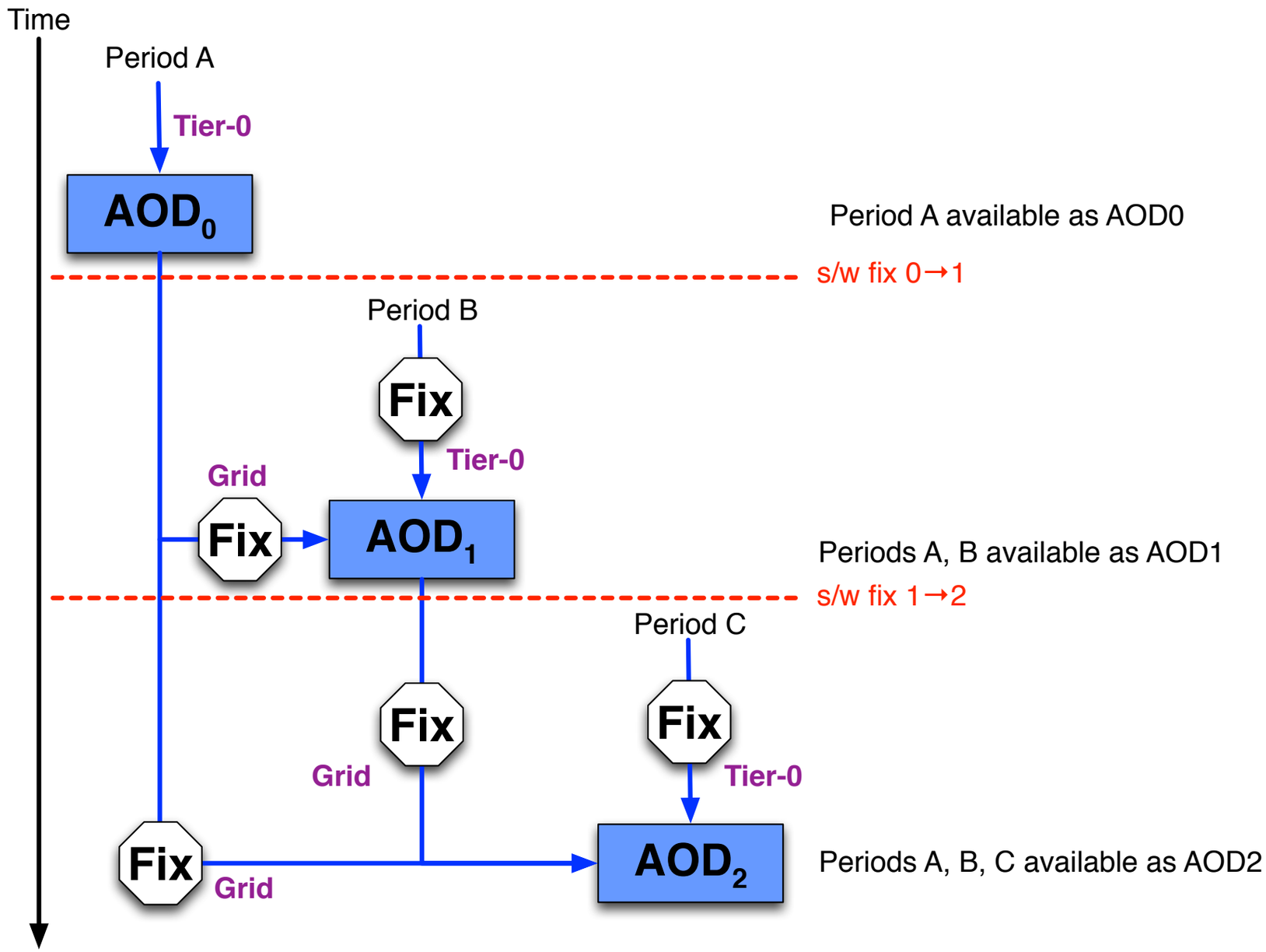
The run 1 analysis model vs time



- The AOD production software was frozen before running
- Software fixes built up which could be applied to AOD, e.g. in DPD production
- Over time, the amount of software fixes became non-negligible in terms of CPU
- These fixes were being repeated for every variant of DPD

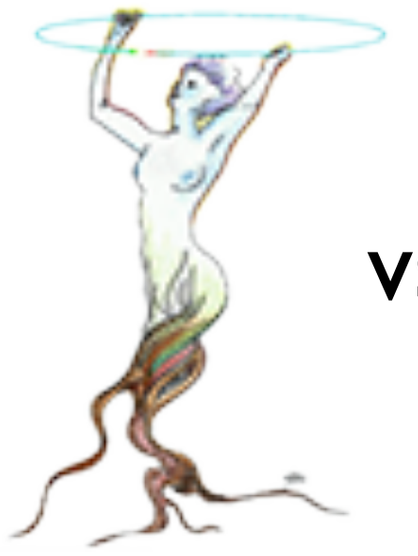
- One of the failures identified by the Study Group was freezing the AOD software...

AOD production in the run 2 analysis model



- Instead, allow updates to the AOD production software such that the AOD-"fixing" is done once

AOD format in the run 2 analysis model

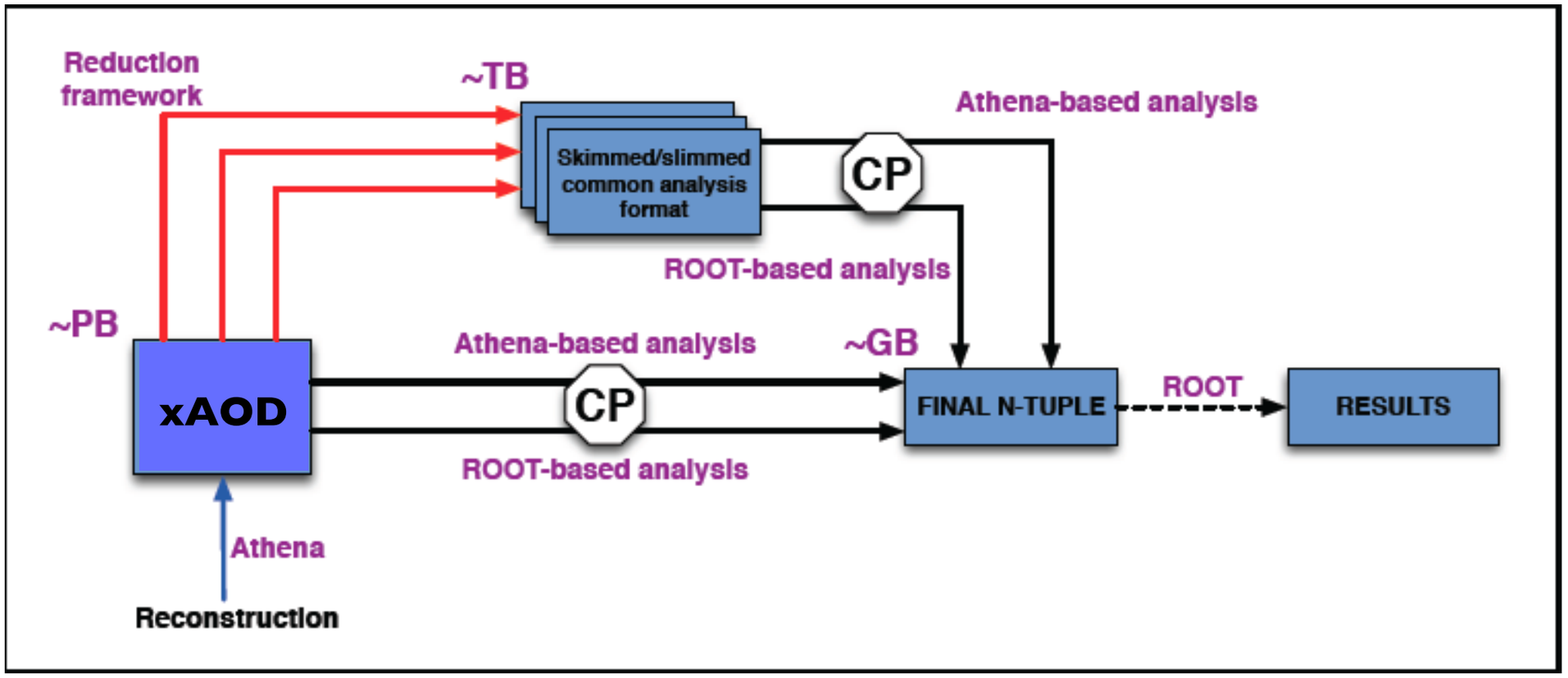


VS

- A matter of taste?

- Despite the many advantages of using athena, not least a powerful Event Data Model, the majority of ATLAS simply wanted to use Root to analyse their data
- The Study Group recommended a new AOD that was readable in both athena and Root

The run 2 analysis model - Trains



- The final piece in the run 2 analysis model is to employ the train model to produce the DPDs from the new AOD
- We will start with the Common DPD in the meantime

Summary

- The run 1 ATLAS analysis model was very successful, yielding around 270 publications at the time of writing
- Based on a philosophy of optimal flexibility for the groups, each of the groups developed their own analysis formats or DPDs independently, while sharing a common code base
- The computing resource usage eventually became significant, with both CPU and disk space causing delays and storage problems as total DPD production was similar to AOD production
- A Common DPD was created which improves the resource usage by a factor of three
- The run 2 analysis model will keep the best features of the run 1 model while optimising the computing resource usage
  - Reconstruction output will be directly ROOT-readable
  - The Train model will be used in DPD production
  - The AOD software will be allowed to be updated
- Thank you for your attention !

- Backup