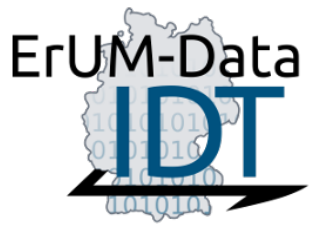
Analysis Workflows at ATLAS

... and current bottlenecks

Nikolai Hartmann

LMU Munich

April 17, 2020

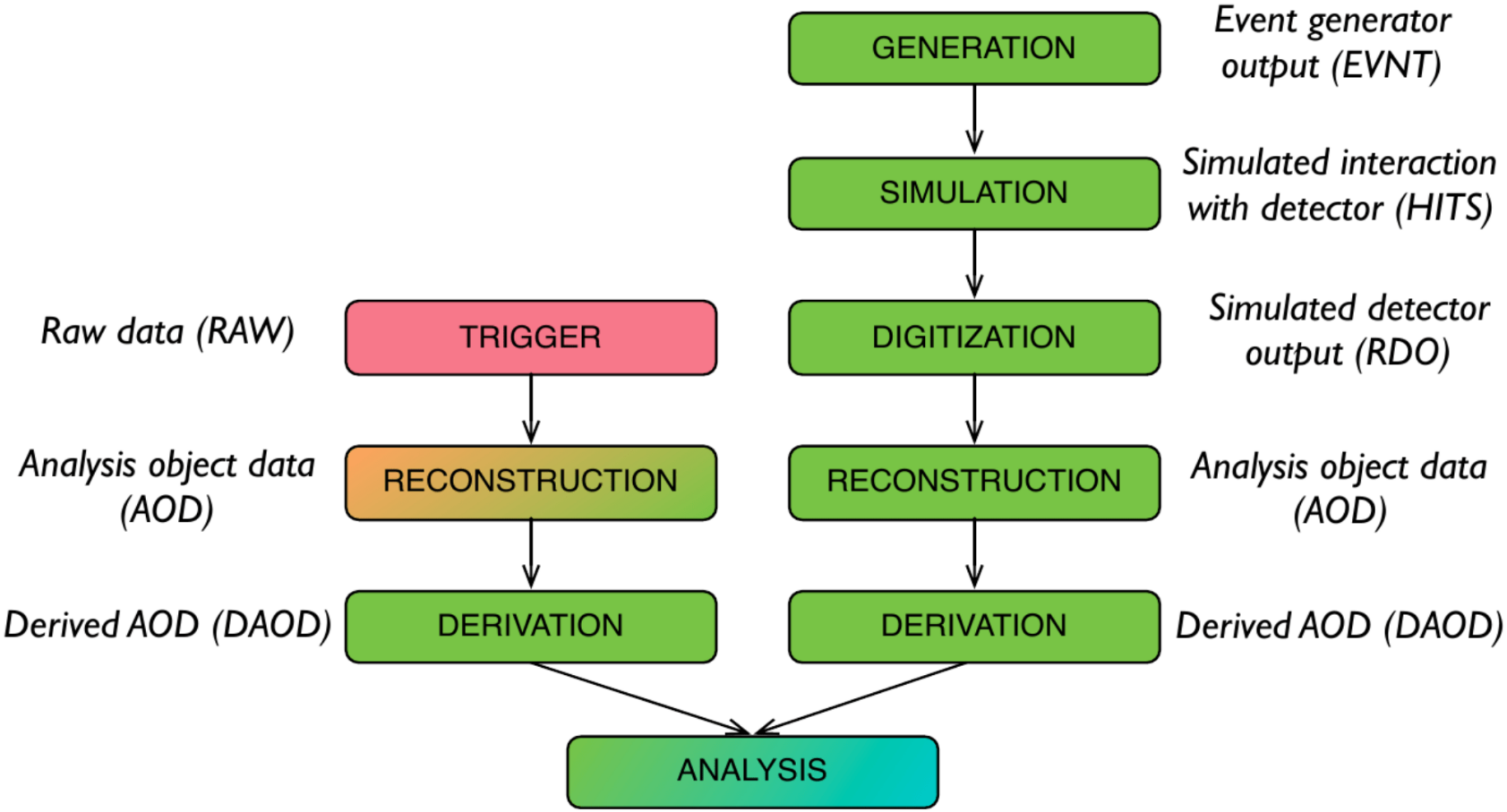


Outline

Overview based on personal (and people in our group) experience and on what was reported in a joint DOMA/HSF analysis meeting

- Current analysis workflows at ATLAS
- Workflow changes planned for the future
- Bottlenecks, turn-around times
- The role of new technologies/methods

Workflow at ATLAS



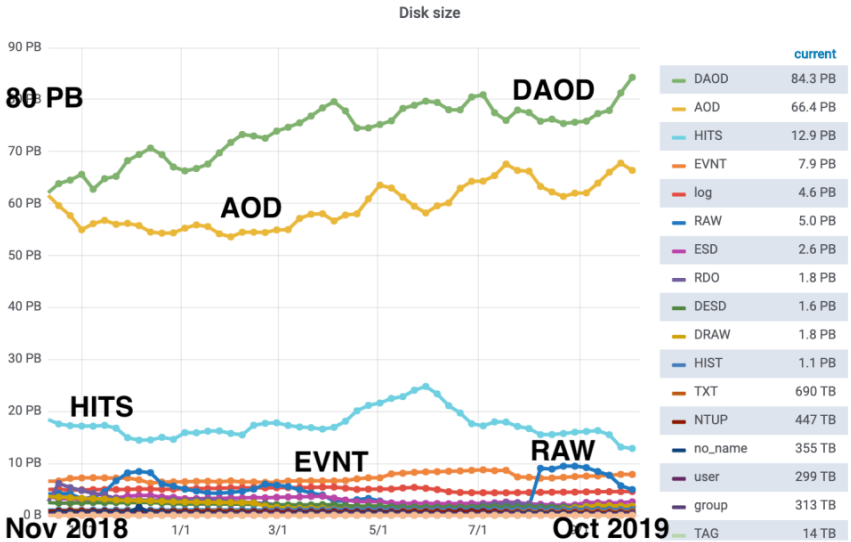
The analysis step

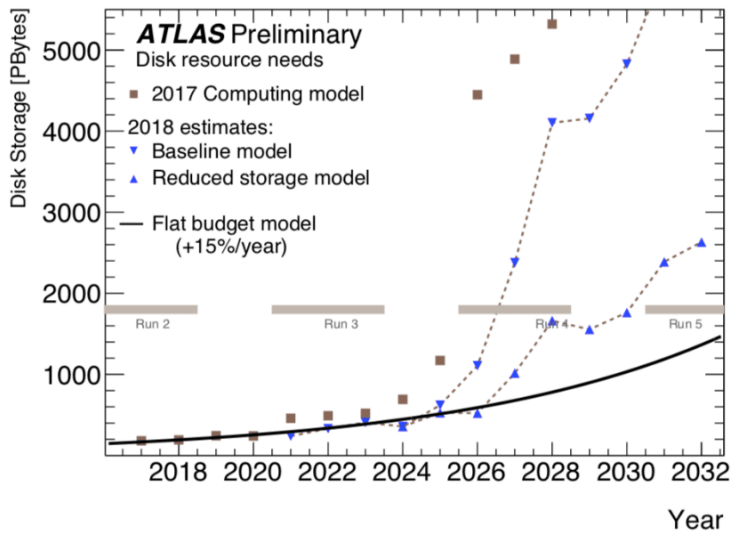
The last step typically consists of

- (optionally) producing yet another intermediate format
- (most of the time) producing a small ntuple from the intermediate format or directly from DAOD
- making plots/histograms and statistical analysis from the small ntuple or the intermediate format
- ntuples produced using custom code (that makes use of central tools from the performance groups)

The problem: too much disk space

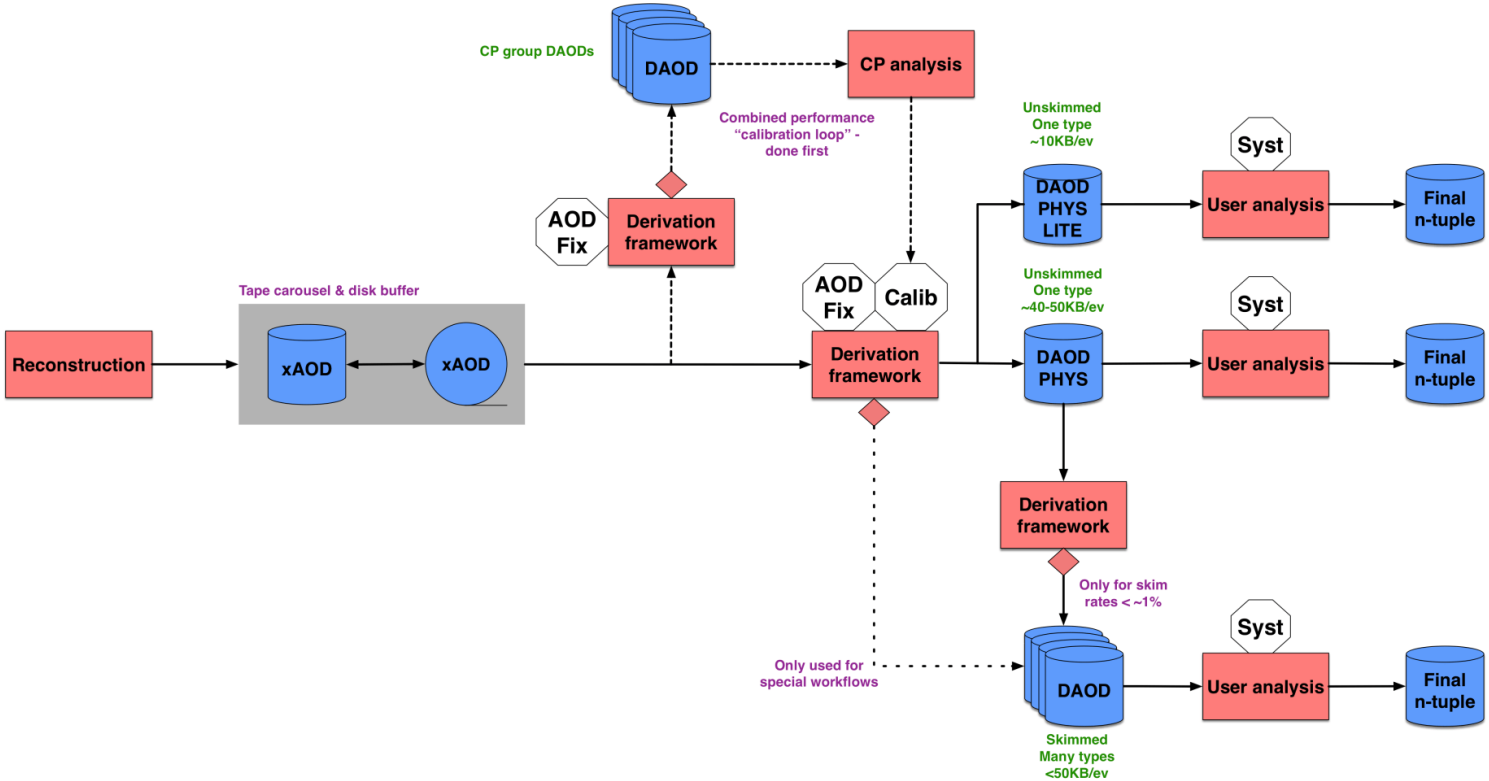
80 different DAOD formats take up too much space





ATL-SOFT-PROC-2020-002

Plans for the future



ATL-SOFT-PROC-2020-002

- replace DAODs by one small format (DAOD_PHYS, $\approx$ 50kB/evt)
- add another even smaller format with calibrations applied, simpler storage model potentially readable via columnar data analysis (DAOD_PHYSLITE, $\approx$ 10kB/evt)
- allows reduced storage and potential for better centralized workflows

Data sizes

Example: SUSY 1L analysis
(thanks to Eric Schanet for providing the numbers)

- Input DAOD files: total size $\approx$ 180 TiB, in 250k files, 6e9 events (background, signal, data)
- Ntuple output size: $\approx$ 100-150 GB either with or without systematics (systematics ntuples only written with minimal information needed for statistical analysis)
  $\rightarrow$ nominal ntuples sometimes skimmed further for quick plotting/studies

Turnaround times

Time for producing new ntuples from DAOD:

- Without systematics: few days
- With systematics: few weeks, worst case months (when also dealing with problems on the grid)

$\rightarrow$ rerunning over DAOD is a significant bottleneck!
$\rightarrow$ at least with systematics, processing is CPU limited ($\approx$ Hz speeds)
$\rightarrow$ needed when recommendations (calibration, systematics) change
$\rightarrow$ additional difficulty: all analysis run this at the same time (before important conferences)

In contrast: Analysis on final ntuple typically fast/negligible turnaround times (sometimes larger turnaround times for final statistical analysis)

See more examples from ATLAS and CMS from the DOMA-ACCESS/HSF-DAWG meeting

New technologies

What we might do more in the future:

- Interactive/In-memory analysis e.g with python tools in jupyter notebooks
  $\rightarrow$ alternative to ROOT TTree::Draw ntuple analysis
- Analysis on computing clusters via spark/dask, RDataFrame?
  $\rightarrow$ replace batch jobs?
  $\rightarrow$ needs to be easy to use, reliable (does anybody use PROOF?)
- Have services (like ServiceX) where events can be dynamically requested (potentially with some on-the-fly transformations)?
  $\rightarrow$ replace producing new ntuples for specific studies?
  $\rightarrow$ pre-calibrated datasets (ATLAS DAOD_PHYSLITE and CMS nanoAOD go in this direction) might offer wider possibilities

Summary and Discussion

- Current analysis workflow based on creating ROOT ntuples from central DAODs
  $\rightarrow$ some analysis introduce another step for more flexibility
  $\rightarrow$ in all cases rerunning on large DAODs is the major bottleneck!
- Ntuple analysis (on local machine) typically fast
  $\rightarrow$ if not, can often reduce ntuple size for specific studies
  $\rightarrow$ will stay an important workflow, maybe more interactive/in-memory analysis
- Concerning new technologies, we should ask the following questions:
  - Where can they provide new, useful possibilities for analysis?
  - Where can they help to circumvent the bottleneck of running a huge number of grid jobs?
- Having central datasets with CPU intense steps pre-done is crucial! (But keep in mind exotic signatures e.g. long-lived searches)