

Derivation Production

ATLAS Sites Jamboree,

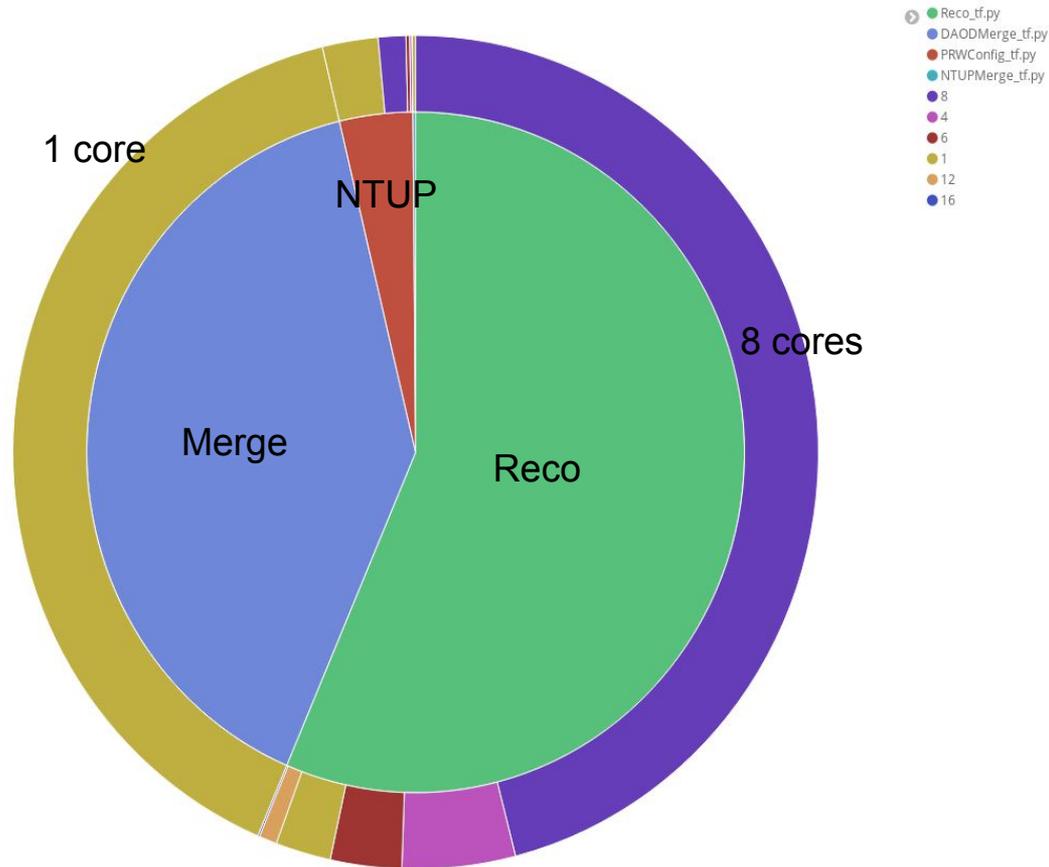
March 2018

Eirik Gramstad,

Monica Dobre, James Catmore

Derivation Production

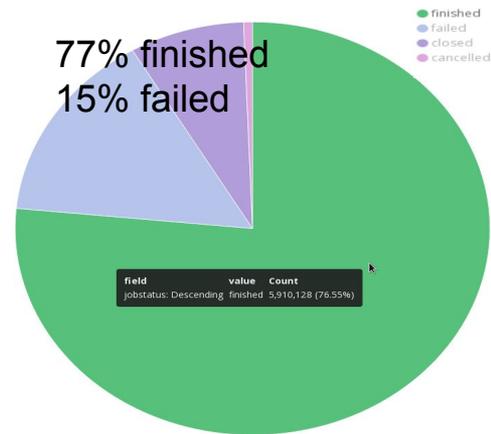
- ★ Derivation production covers:
 - **AODtoDAOD** (Reco_tf, AthenaMP on 8 cores)
 - **DAODMerge** (DAODMerge_tf, separate subsequent jobs on 1 core)
 - **NTUP_PILEUP** (PRWConfig_tf)
 - **NTUPMerge_tf** (merging of NTUP_PILEUP, done in subsequent jobs or in separate merge task)
- ★ Formats are configured into trains (share I/O burden, jet-building etc.) according to size fraction of input AOD



Derivation production in 2017

★ Open-ended production on data17

- Derivations made automatically for all reprocessed runs from Tier0 during data taking
- In total 105 formats (total of 627 runs processed):
- Communication with data prep. to avoid running over bad runs/wrong processing tag etc.
 - Tasks could be aborted quickly



1st batch: 69 formats

1st round - high priority: physics Main, all runs

- STDM4

1st round - normal priority: physics Main, all runs

- < 0.1%: HIGG2D5 FTAG3 TCAL1
- 0.1-0.15%: JETM10 EGAM4 STDM5 MUON2
- 0.16-0.2%: HIGG1D2 EXOT12 EGAM9
- 0.2-0.3%: SUSY2 SUSY12 JETM3 EGAM3 EXOT0 EXOT17
- 0.3-0.4%: JETM7 EXOT6 EGAM2 BPHY8
- 0.4%: BPHY7
- 0.4-0.5%: EXOT9 EGAM7
- 0.5-0.6%: SUSY16 EXOT20 STDM3
- 0.6-0.7%: SUSY18 JETM4 HIGG4D1 HIGG6D1
- 0.8-1%: EXOT18 SUSY3 EXOT5 STDM2
- 1-1.3%: JETM2 SUSY11 SUSY4 JETM6 EXOT19 EGAM1
- 1.3-1.6%: JETM1 HIGG2D4 EGAM8 SUSY1 SUSY5
- 1.6-1.8%: HIGG5D3 HIGG5D1
- 1.9-2.1%: SUSY10 EGAM5 JETM11
- 2.1-2.7%: TAUP1 EXOT13 EXOT2
- 2.9-3.4%: MUON1 SUSY6 JETM9 EXOT8 FTAG1
- 3.6-3.7%: HIGG2D1 HIGG5D2
- 4-5%: FTAG2 FTAG4
- 5-7%: HIGG6D2 MUON0 TAUP3
- HIGG4D4
- HIGG4D2
- HIGG8D1

2nd batch: 9 formats

Second round - high priority: physics Main, all runs+*

- EXOT3

Second round - normal priority: physics Main, all runs

- TOPQ1 TOPQ3
- TOPQ2
- TOPQ4
- TOPQ5
- EXOT10
- HIGG3D1
- SUSY9

3rd batch 8 formats

Third round - normal priority: physics Main, all runs

- SUSY8
- EXOT4
- EXOT15
- STDM9
- EXOT7 EXOT22
- HIGG1D1

Third round - normal priority: physics Main, data15 13TeV only

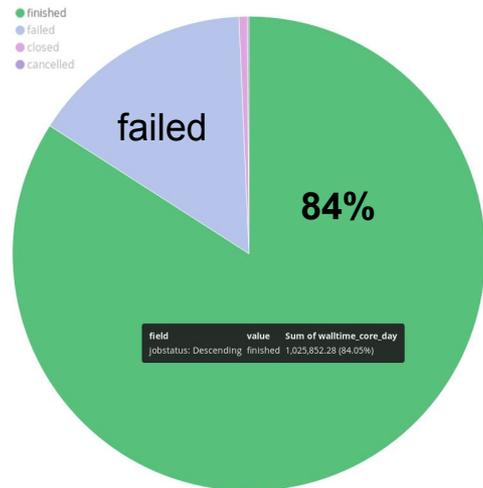
- HIGG1D2



Re-derivation campaign 2017

- submitted in 3 main batches with 3 different AthDerivation caches (21.2.8/9/10.0)
- ~90 formats on complete data15/16/17
 - Centrally submitted
- MC16a/c/d production
 - Submitted by group contacts, but requests were collected for some weeks before launching
- data completed in ~6 weeks (Nov. 17th - Dec. 31st) with an average of 95.000 slots

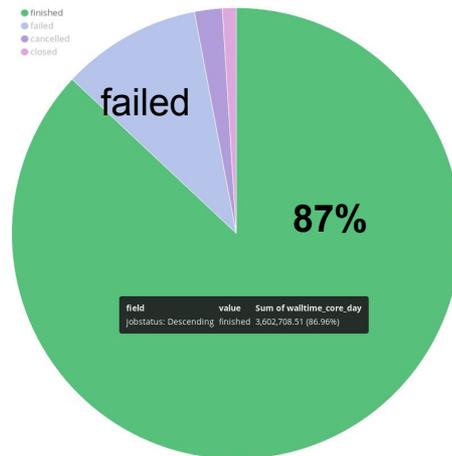
MC



Wall time vs. status

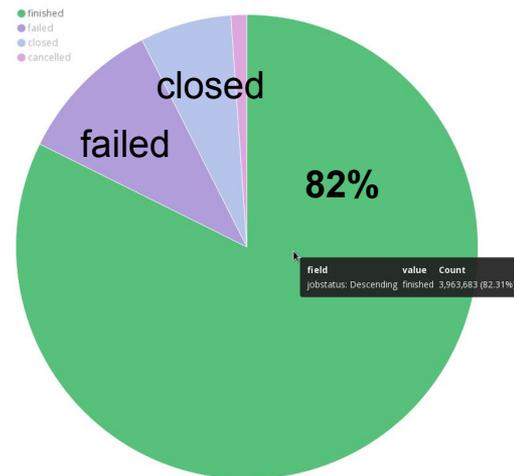
- ★ ~10% failed wall time (data)
- ★ Slightly higher for MC jobs

DATA



Fraction of jobs vs. status

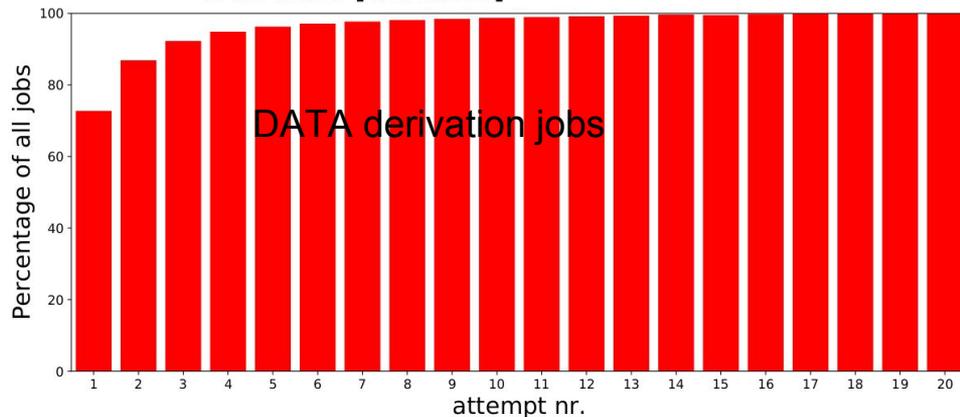
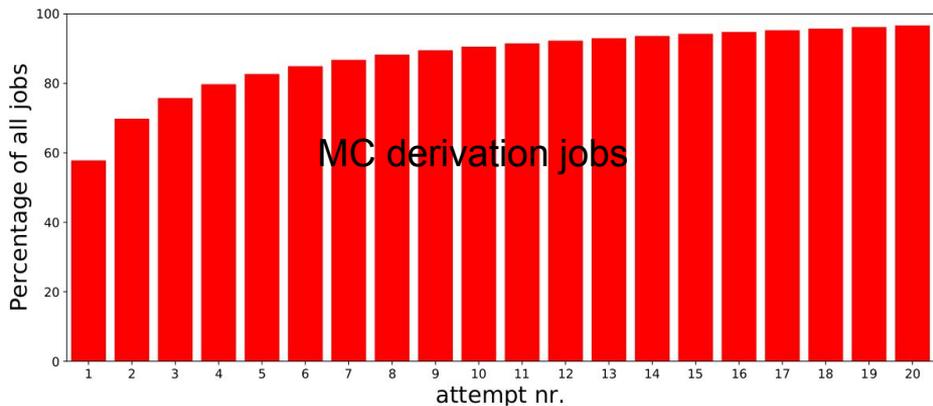
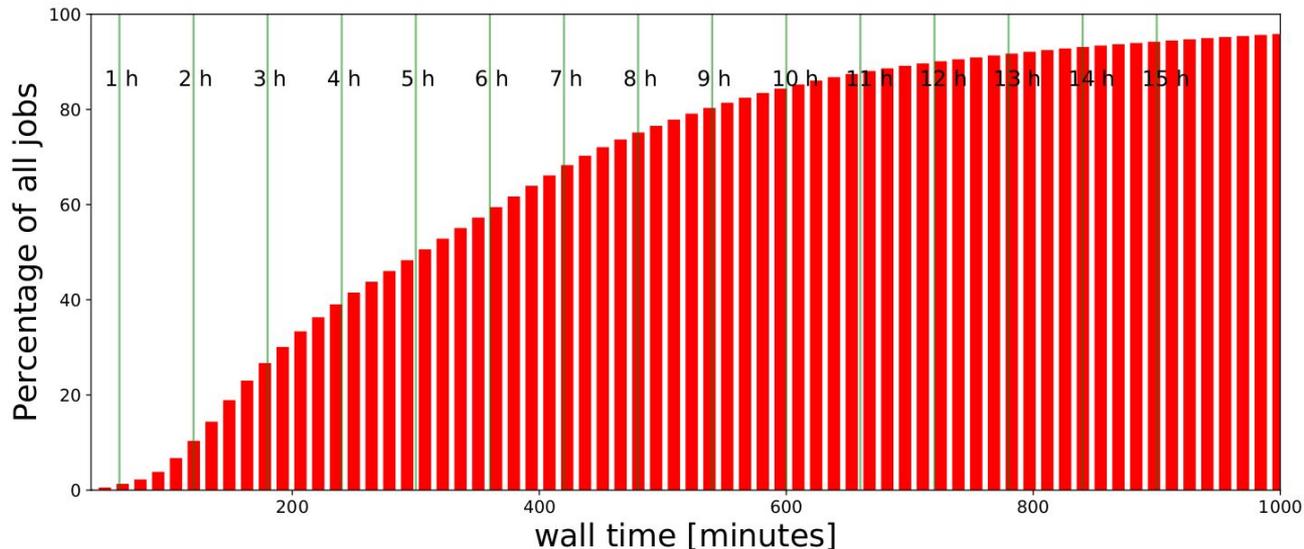
- ★ ~larger failure rate (32%) in MC compared to data (18%)
 - Mostly due to closed jobs



Job attempts/wall time

- ★ Plot to the right shows percentage of jobs finished after wall time (90% of jobs finished within 12 hours wall time)
- ★ Plots below show number of attempts for jobs to finish for MC and DATA (Reco)

DATA derivation jobs



Main issues in re-derivation campaign

★ A few software problems encountered

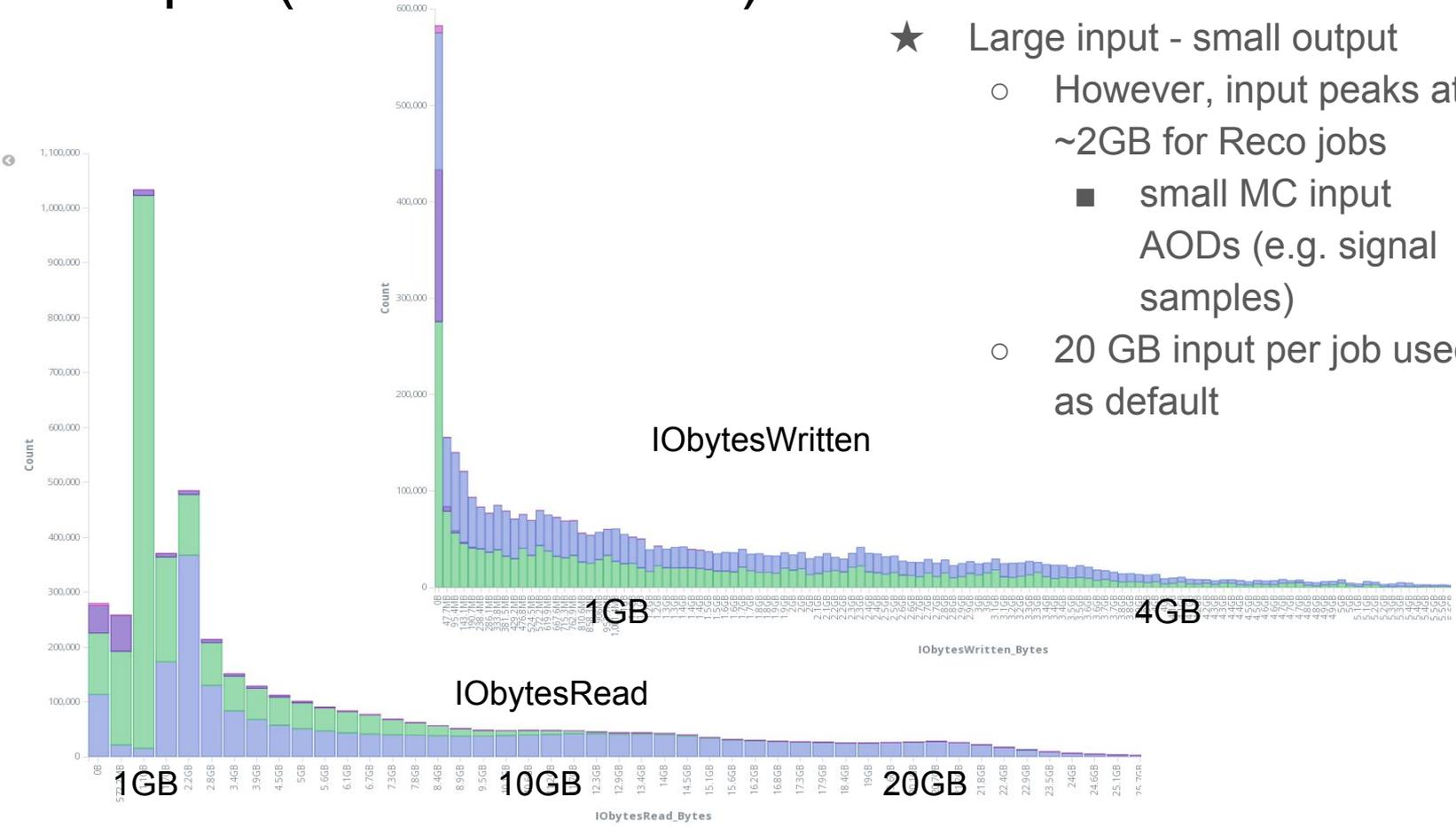
- Formats with too high memory consumption or too slow (were aborted)
- Problems merging empty files (fixed to a large extent by moving to SCORE)
- Some additional, rare, software problems (e.g. processing runs with non-standard conditions)
- Jobs with more than 200 input files crashing

★ Grid related issues

- Main recurring problem was related with STAGEIN of files (corrupted files at site)
 - Jobs were retried (at the same site) over and over again - needed to be reassigned and retried manually before succeeding
- Jobs run out of space on worker node (especially on MC)
 - Need to tune the input size per job to avoid too large output files, but difficult to know the exact size fraction for each format (depends on the MC AOD which is run over)
 - Easier with data (size fraction is fairly constant)

Input/output (MC derivations)

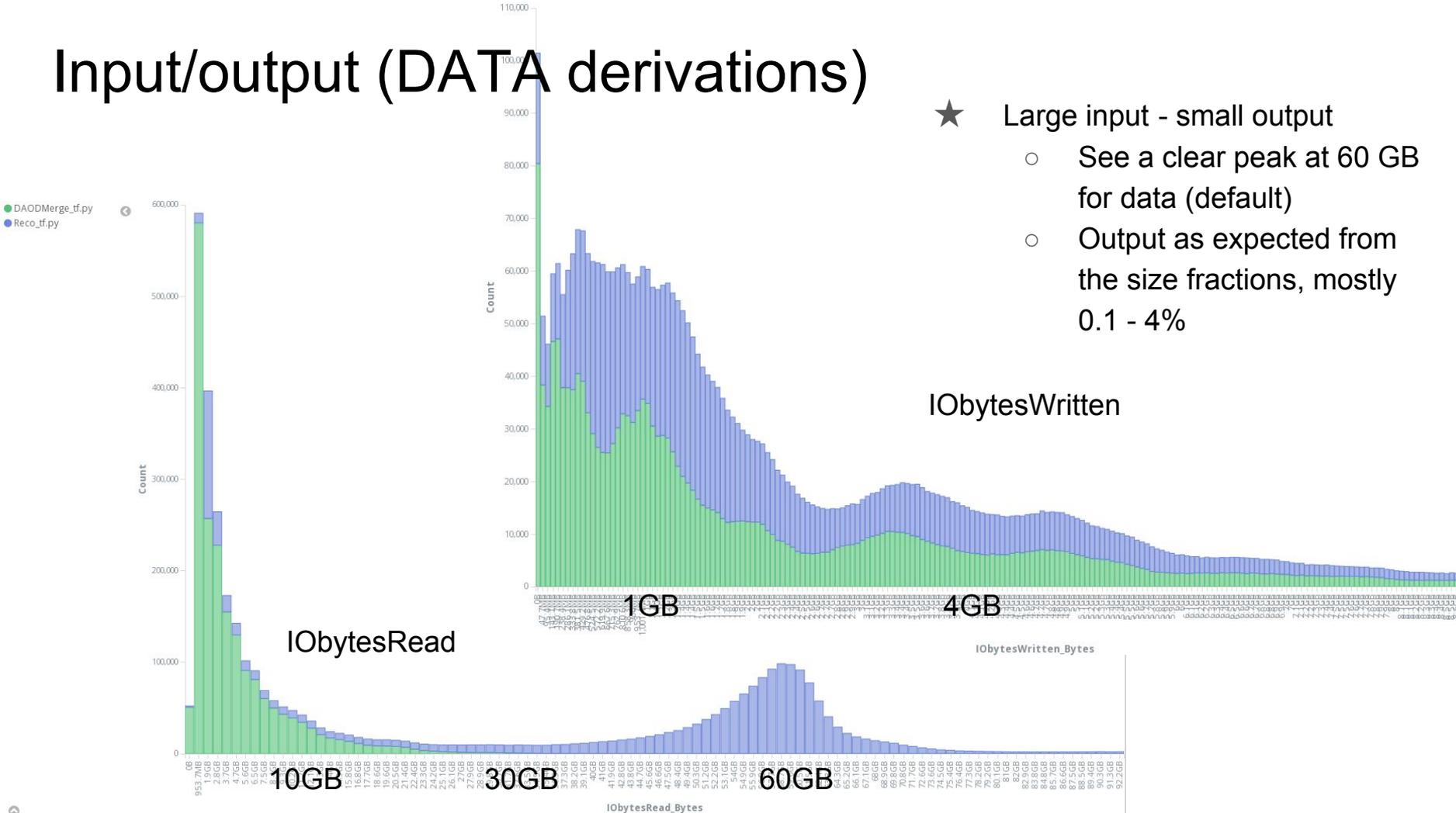
- Reco_tf.py
- DAODMerge_tf.py
- PRWConfig_tf.py
- NTUPMerge_tf.py



- ★ Large input - small output
 - However, input peaks at ~2GB for Reco jobs
 - small MC input AODs (e.g. signal samples)
 - 20 GB input per job used as default

Input/output (DATA derivations)

- ★ Large input - small output
 - See a clear peak at 60 GB for data (default)
 - Output as expected from the size fractions, mostly 0.1 - 4%



IObytesRead

IObytesWritten

10GB

30GB

IObytesRead_Bytes

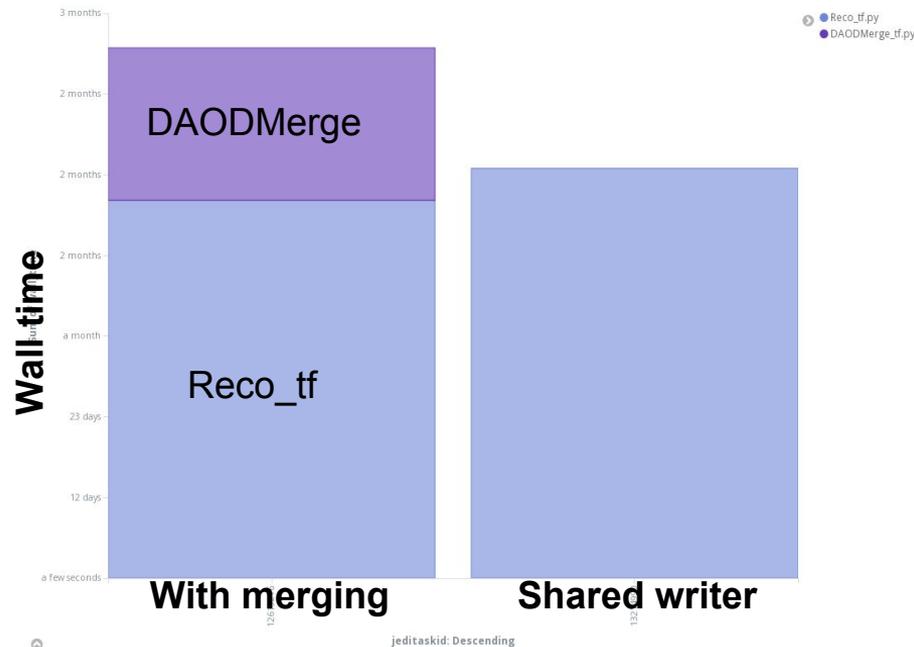
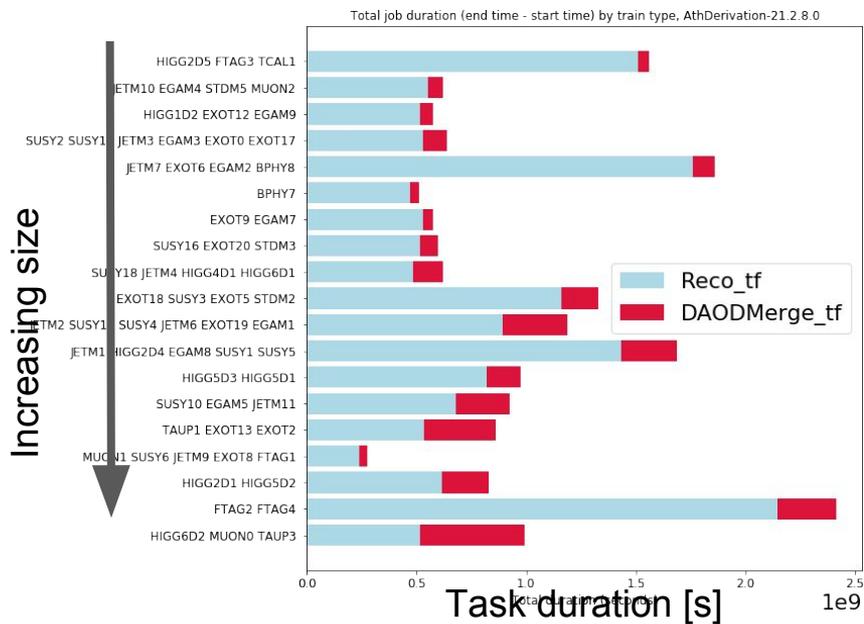
IObytesWritten_Bytes

60GB

Preparations for 2018 - open-ended production

★ Shared Writer

- Removes the merge step altogether
 - Will help both in terms of time, but also in failure rate (because merge is run as separate jobs in the derivation tasks)
 - Validation is almost done (soon to be put into production)



Preparations for 2018

- ★ Working on reducing the derivation sizes
- ★ Make sure we only run on derivations which are actually used
 - Each format needs a clear physics case and responsible person