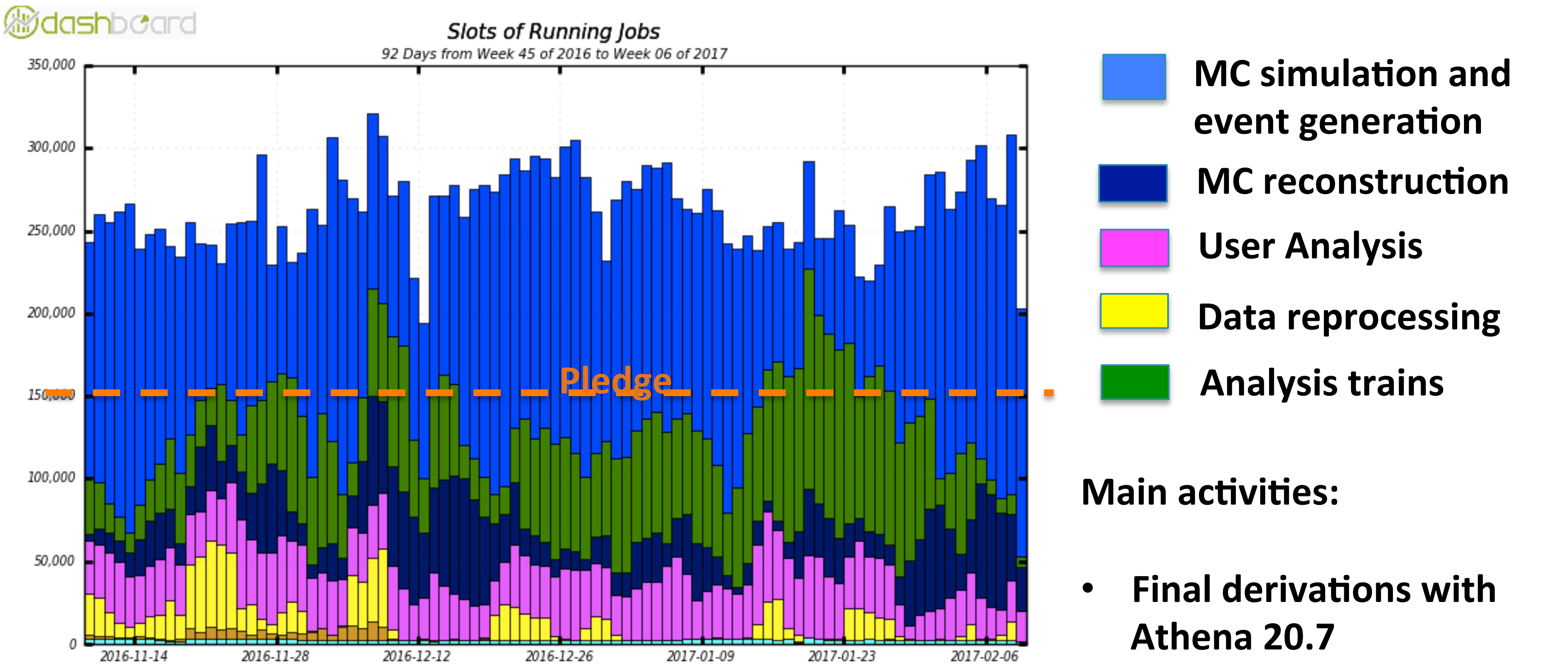
ATLAS Software and Computing Report

Simone Campana
Torre Wenaus

Introduction

- We continue our work in improving computing performance
- Mitigation strategies have been considered in case of computing resources shortage in 2017 and 2018
- All documented in the ATLAS internal note ATL-SOFT-INT-2017-001. Distributed to the LHCC Computing Referees and to the CRSG chair
- In this presentation we will touch quite a bit of this (the lower hanging fruits), not all. Feel free to ask

Processing activity – last 3 months



Main activities:

- **Final derivations with Athena 20.7**
- **Tails of MC15 simulation**

In 2016 we used 50% more CPU than our pledge, flat over the year. Our first "mitigation strategy" is to make sure we use everything we have

Plans for 2017

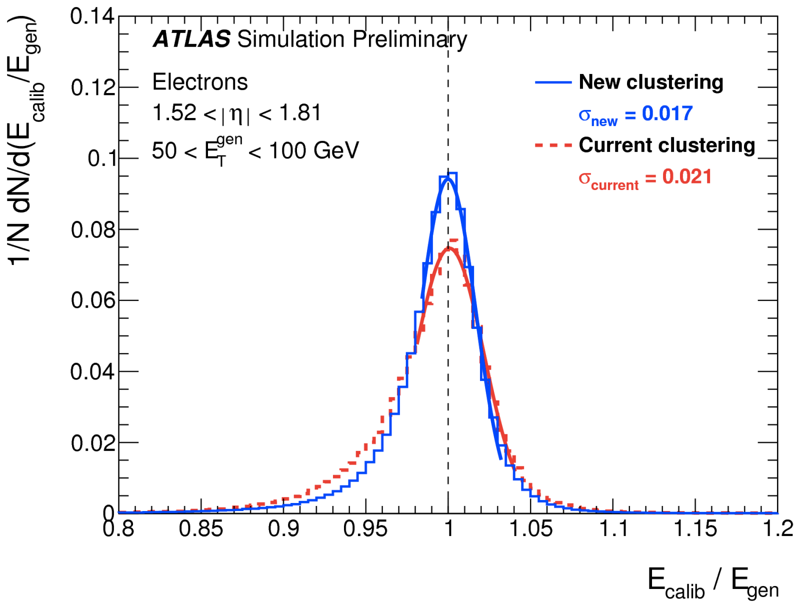
Start the new MC simulation campaign MC16, will provide samples for the remaining Run-2 analyses

MC16 will be reconstructed with Athena21, the last major release for Run-2

We will also reprocess all 2015+2016 data with Athena 21

MC16 and Athena21 are 6 months late with respect of the original plan: we preferred a careful validation rather than redoing a massive production in case of problem

This delay allowed to complete the 2016 physics program with the available resources, but create a challenge now for 2017 and 2018.



Distributions of the calibrated energy, Ecalib, divided by the generated energy, Egen, for a Monte Carlo sample of electrons generated with 1.52<|η|<1.81 and 50<EgenT<100 GeV, without pileup. The dashed-line (red) histogram shows the performance when the electrons are reconstructed and calibrated using the current clustering algorithm (so-called sliding window, used in release 20.7 up to 2016 data taking) while the full-line (blue) histogram is based on the new clustering algorithm (so-called super-clusters, used in release 21 starting from 2017 data taking).

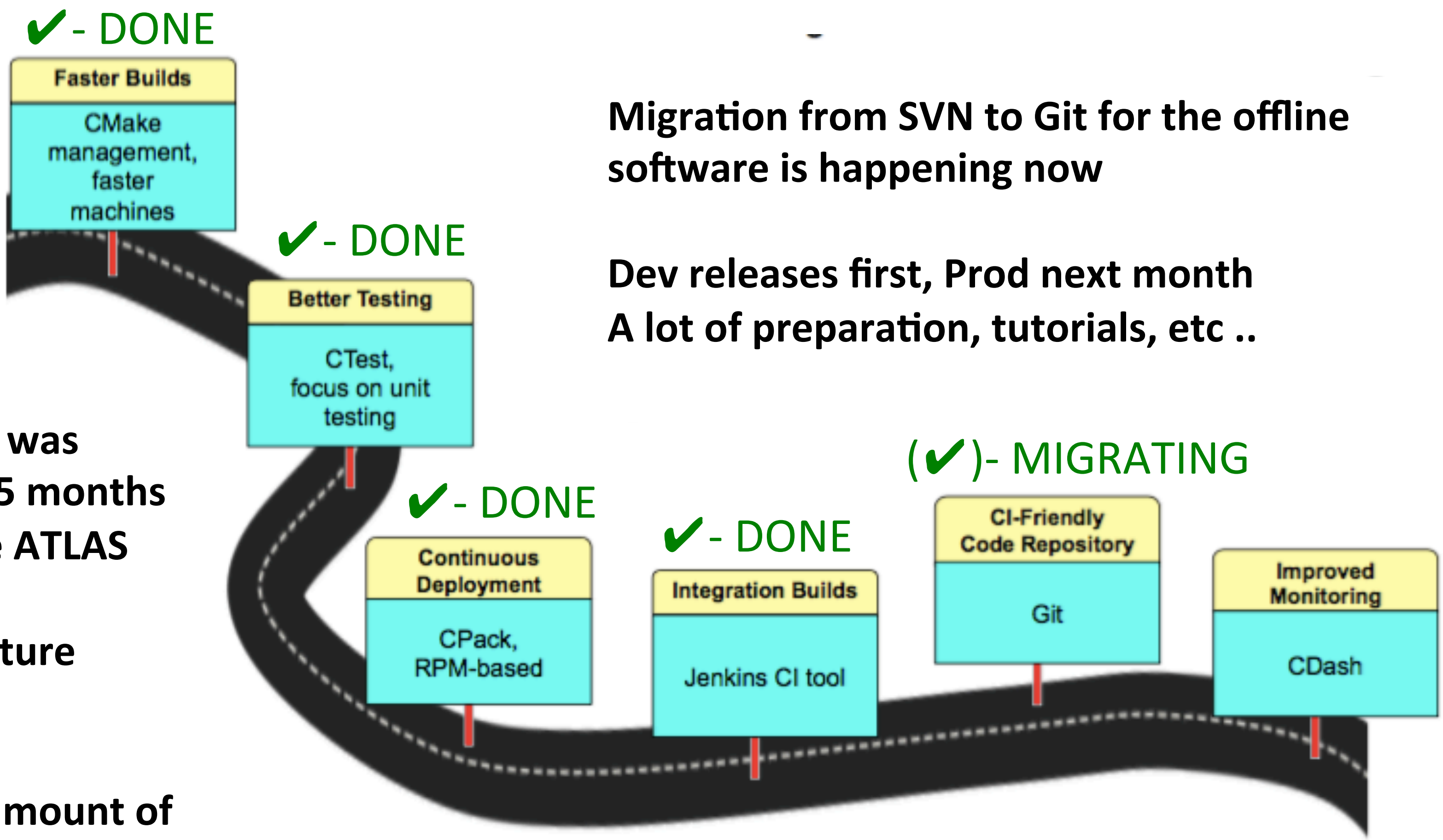
Software Infrastructure

Roadmap was defined 15 months ago in the ATLAS Software Infrastructure review

Massive amount of progress since then

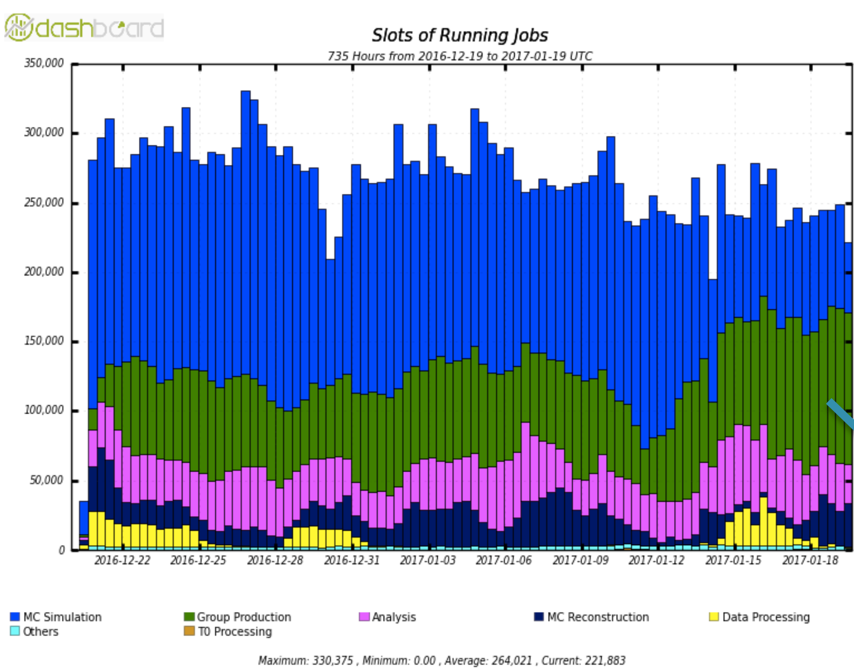
Migration from SVN to Git for the offline software is happening now

Dev releases first, Prod next month
A lot of preparation, tutorials, etc ..



Processing shares

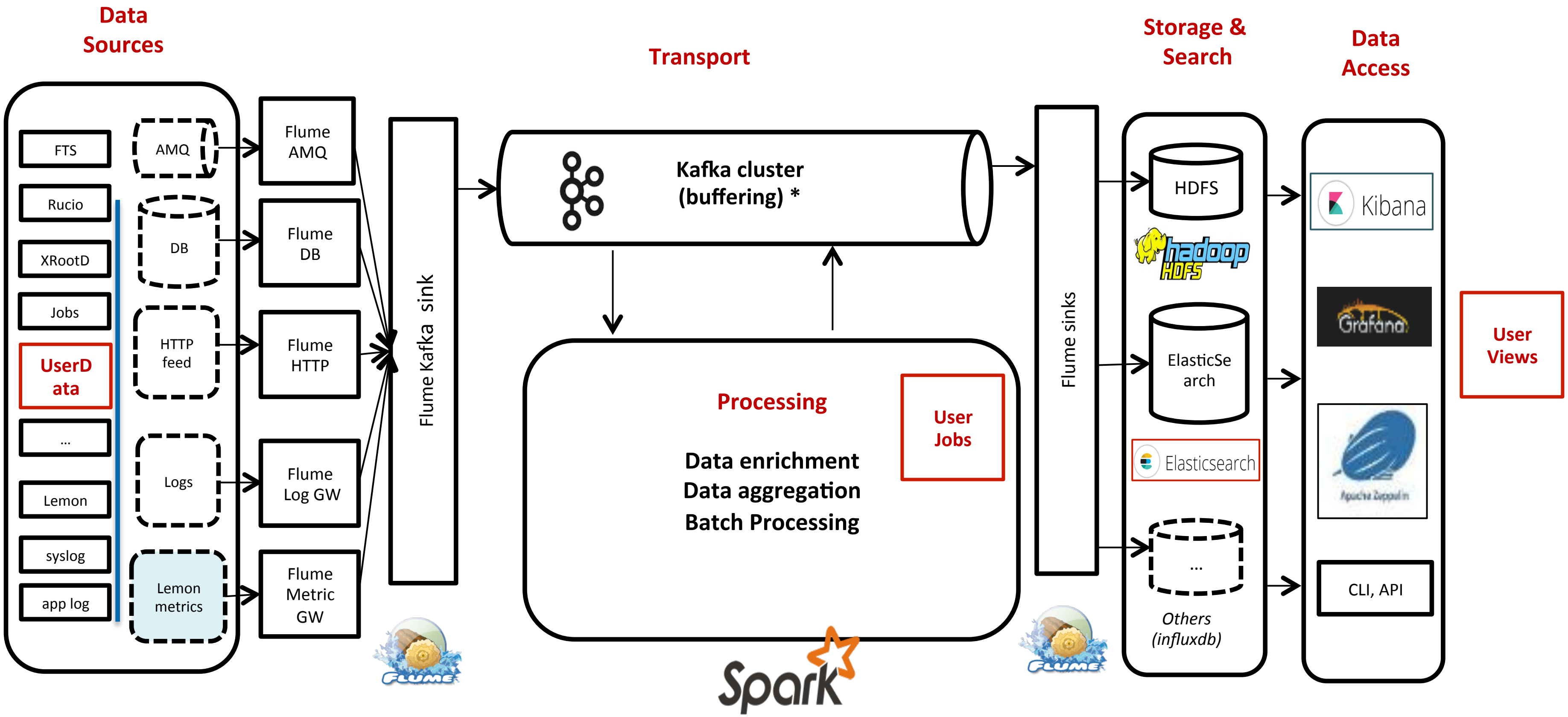
Because we use constantly 150% of our CPU resources, we need to organize the work properly. We implemented a refined mechanism of global shares in the Production System. We have also priorities inside the same share. It works extremely well.



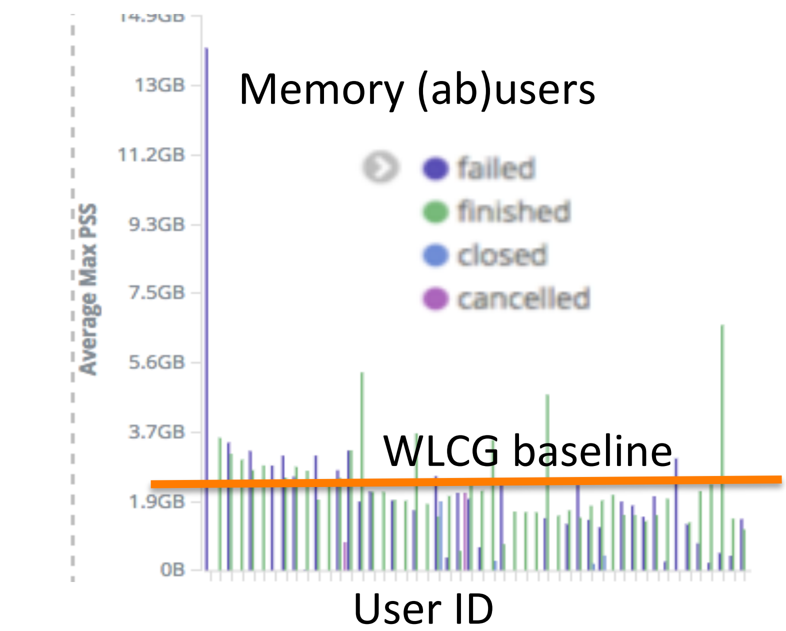
L1 Share	L2 Share	L3 Share	Actual HS06	Target HS06	Ratio	Queued
Analysis [20.0%]			327,847.73	593,416.39	55.25 %	1,752,697.06
Production [75.0%]			2,626,358.86	2,225,311.47	118.02 %	7,539,315.32
	MC root [37.1%]		2,224,649.30	1,101,639.34	201.94 %	4,518,355.33
		MC production [12.4%]	0.00	367,213.11	---	0.00
		MC 16 [12.4%]	56,042.77	367,213.11	15.26 %	117,007.22
		MC Default [12.4%]	2,168,606.53	367,213.11	590.56 %	4,401,348.10
	Derivations [14.9%]		54,414.27	440,655.74	12.35 %	34,620.84
		MC Derivations [4.5%]	52,125.20	132,196.72	39.43 %	30,269.79
		Data Derivations [10.4%]	2,289.07	308,459.02	0.74 %	4,351.05
	Reprocessing [7.4%]		348.17	220,327.87	0.16 %	10.37
		Reprocessing default [5.9%]	348.17	176,262.29	0.20 %	10.37
		Heavy Ion [1.5%]	0.00	44,065.57	---	0.00

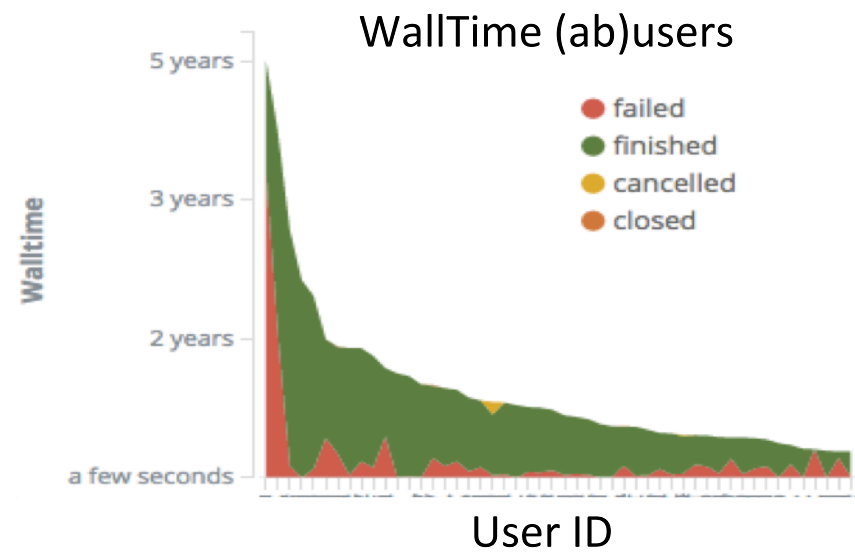
You ask for 100k cores for derivations
You get 100k cores for derivations

ATLAS invested considerable effort in adopting a Data Analytics infrastructure, relying on the building blocks of the CERN IT agile monitoring

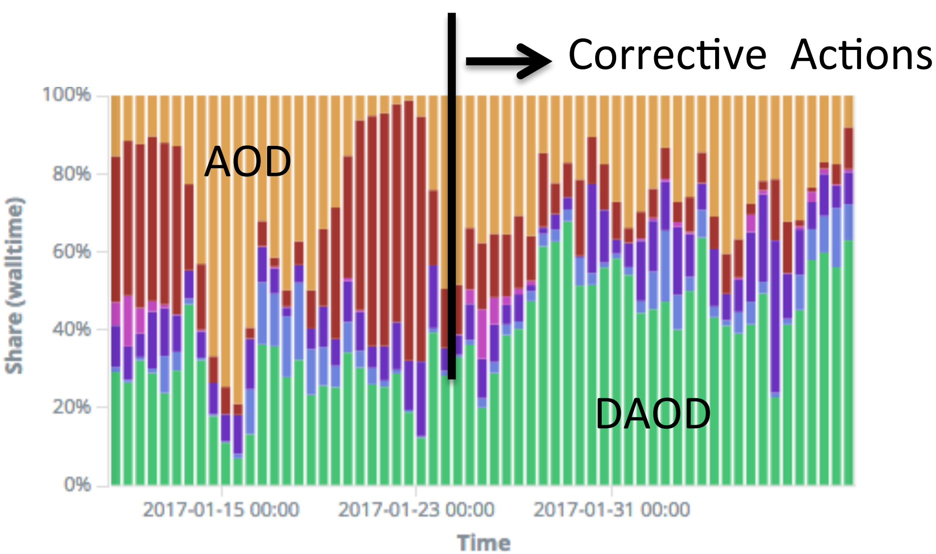


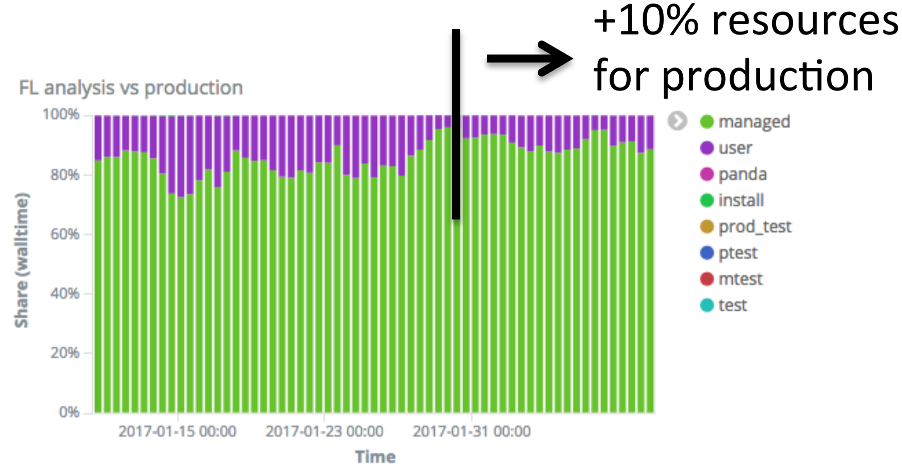
In return: monitoring properly distributed computing activities today saves considerable resources, improves the user experience and the quality of physics. The chaotic analysis example follows





Users should not run on AODs





AODs on TAPE only?

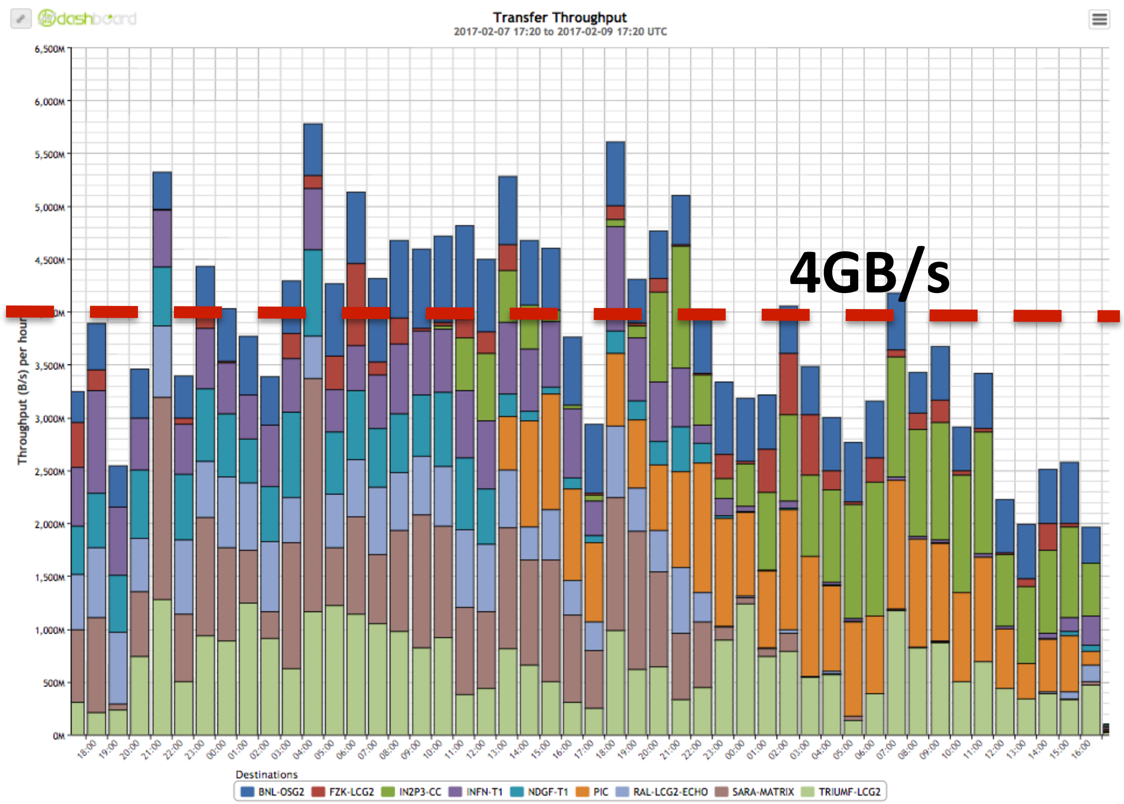
Would allow to save a lot of disk space, but implies running derivations from TAPE

Derivations is part of "analysis" so quick turnaround is needed

Tape staging tests: 4GB/s tape recall achieved , but no workload management system involved at this stage

This is not fast enough for a full derivation campaign, but is 2x faster than first attempt

We could foresee partial derivations (for dedicated samples) from tape once we are more confident.



A lot of work to do: not a 2017 target

AOD sizes

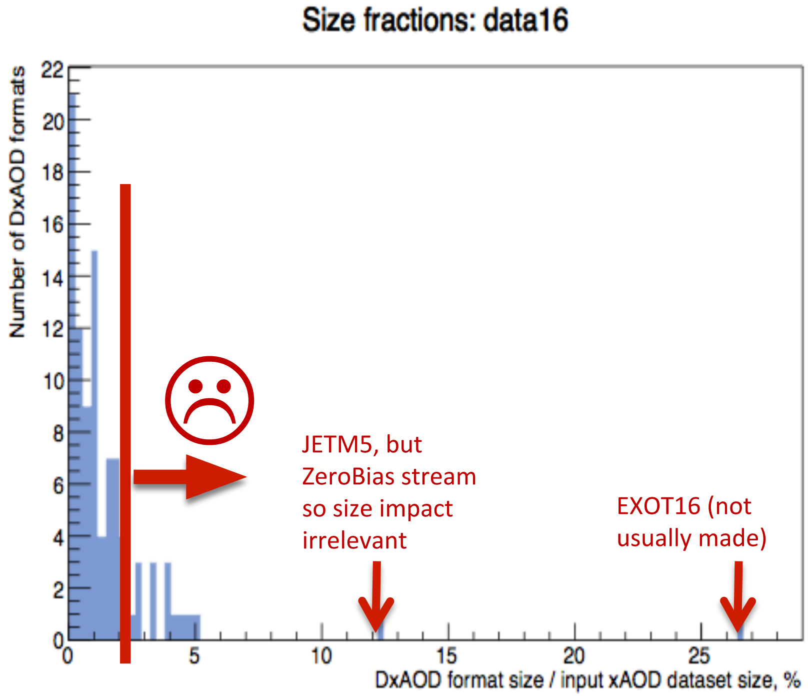
Excellent achievements of the Task Force looking into AOD size reduction
Work in progress, possible gains starting from 2018.

Size savings (KB/ev) w.r.t. baseline 21.0 AOD

	Data (1st pass)	Data (reprocessed)	Standard Model MC (ttbar)	Signal MC (ttbar)
Track pT 400->500 MeV and covariance matrix compression	26	26	27	27
Removal of negative E caloclusters & unused moments	4	4	12	12
Removal of unused PFlow moments	4	4	12	12
Removal of most jet containers (retain 3)	19	19	40	40
Removal of most flavour tagging (retain 1)	29	29	33	33
Removal of G4 truth	0	0	65	65
Use of AODSLIM	0	45	0	0
Use of AODSUPERSLIM	0	0	0	55
TOTAL SAVINGS	82	127	189	244
Current AOD size	420	420	582	582
% Saving	20%	30%	32%	42%
AOD size in Run-2 Computing Model	320	320	500	500
New AOD size in computing model	319	319	410	410

DAOD sizes

DAODs are very heavily scrutinized already. Changes in DAODs are more disruptive for users. We continue looking for reductions in DAOD sizes

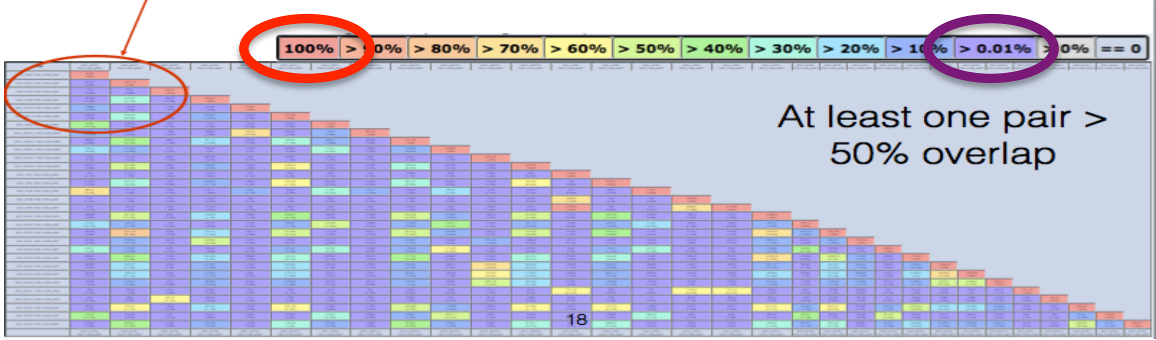


Event overlaps between DAOD forr

Dataset	DAOD_EXOT0. f708_m1606_p2667	DAOD_EXOT4. f708_m1606_p2667
DAOD_EXOT0. f708_m1606_p2667	502891 (100%)	
DAOD_EXOT4. f708_m1606_p2667	427343 (3.99%)	10647773 (100%)
DAOD_EXOT6. f708_m1606_p2667	2193 (0.24%)	72751 (0.66%)
DAOD_EXOT8. f708_m1606_p2667	281889 (2.12%)	5554767 (30.54%)

Red is bad
Diagonal is red by definition

Purple is Good

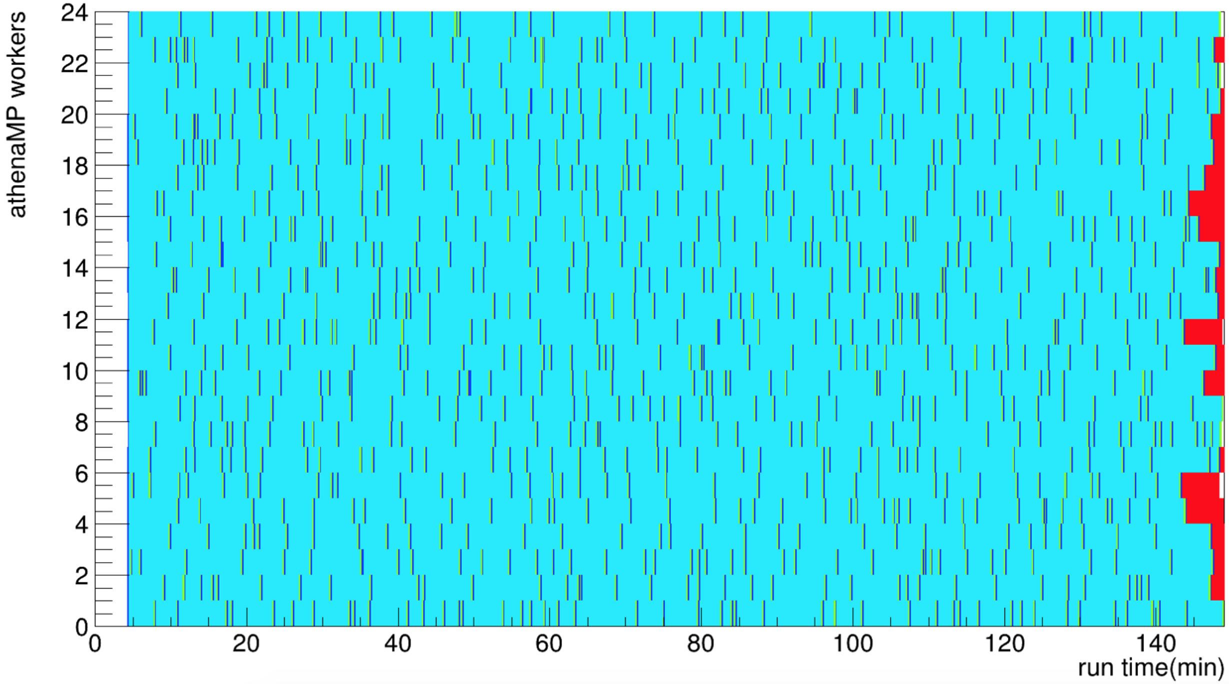


Few derivations above target in size, but no real major offender.

Little event overlap across derivations.

Event Service

The ATLAS Event Service allows to events processing with the smallest possible granularity (today): the single event. Obvious benefits for highly opportunistic resources (for example pre-emptible queues), but brings benefits everywhere (for example during queue draining time at Grid sites)



The EDISON HPC at NERSC:

The Event Service allows to leverage > 99% of the available cycles

The initial setup time was reduces to 5 minutes with a dedicated effort

Possible Mitigation Mechanisms

1) Reduce the HLT output rate to e.g. 750 Hz

Table 1: impact of raising the single lepton trigger thresholds for three example physics channels

Physics Analysis	Issues
Single top and Wt for e channel	≈ 40—50% efficiency loss
Higgs WW* → l ν l ν	≈ 10—20% efficiency loss eμ, μμ ≈ 40% efficiency loss for ee
Higgs WH → l ν bb for e channel	≈ 60% efficiency loss in low pT(W) < 90 GeV

Impact on physics too large. ATLAS discards this option.

Possible Mitigation Mechanisms

2) Parking Data until LS2

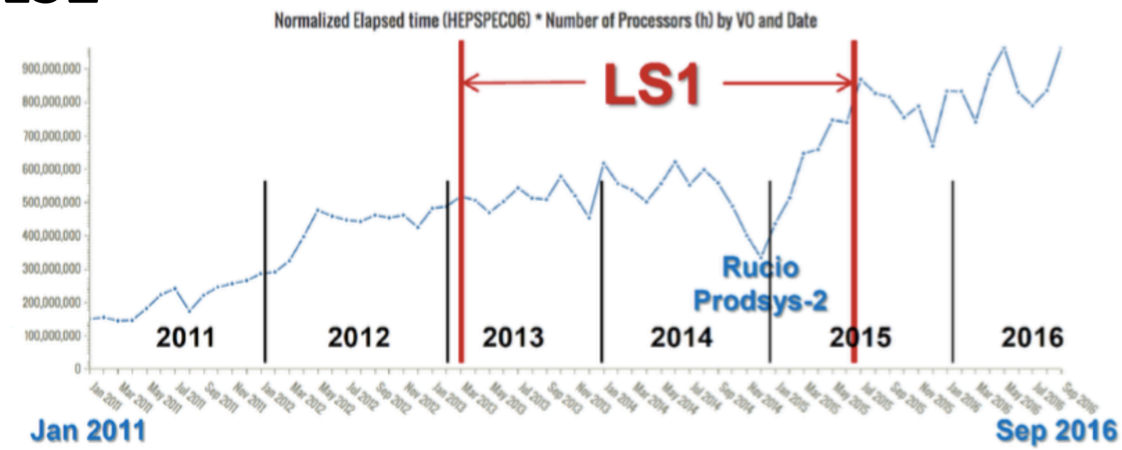


Figure 1: ATLAS grid CPU usage from January 2011 – September 2016. The dip during the holiday period in late 2014 / early 2015 was caused by the suspension of our grid production for the deployment of our new data management and production system Rucio and Prodsys-2.

- Implies considerable free resources in LS2 which won't be the case
- Implies delaying analyses, which has an impact on people's careers
- Not really a solution unless we can "park" also Monte Carlo simulation (the main CPU consumer)

Possible Mitigation Mechanisms

3) Reducing the amount of MC simulation

Implies reducing by O(40%) the amount of background samples.

The only feasible option among the three, but still with a considerable impact on the ATLAS physics program

Conclusions

- ATLAS Software and Computing is continuously optimized for performance. There are no low hanging fruits
- Several medium term improvements foreseen, will help in 2017 and 2018, catching up with delays in MC16
- Mitigation Strategies considered for 2017 and 2018. Some not appealing at all
- Funding agencies received very well the message at the end of 2016 and will contribute more than initially planned in 2017
- Because of our success in exploiting opportunistic CPU resources, storage remains our first priority