LHCC Referee Meeting

09/05/2017

ALICE Status Report

Predrag Buncic

CERN

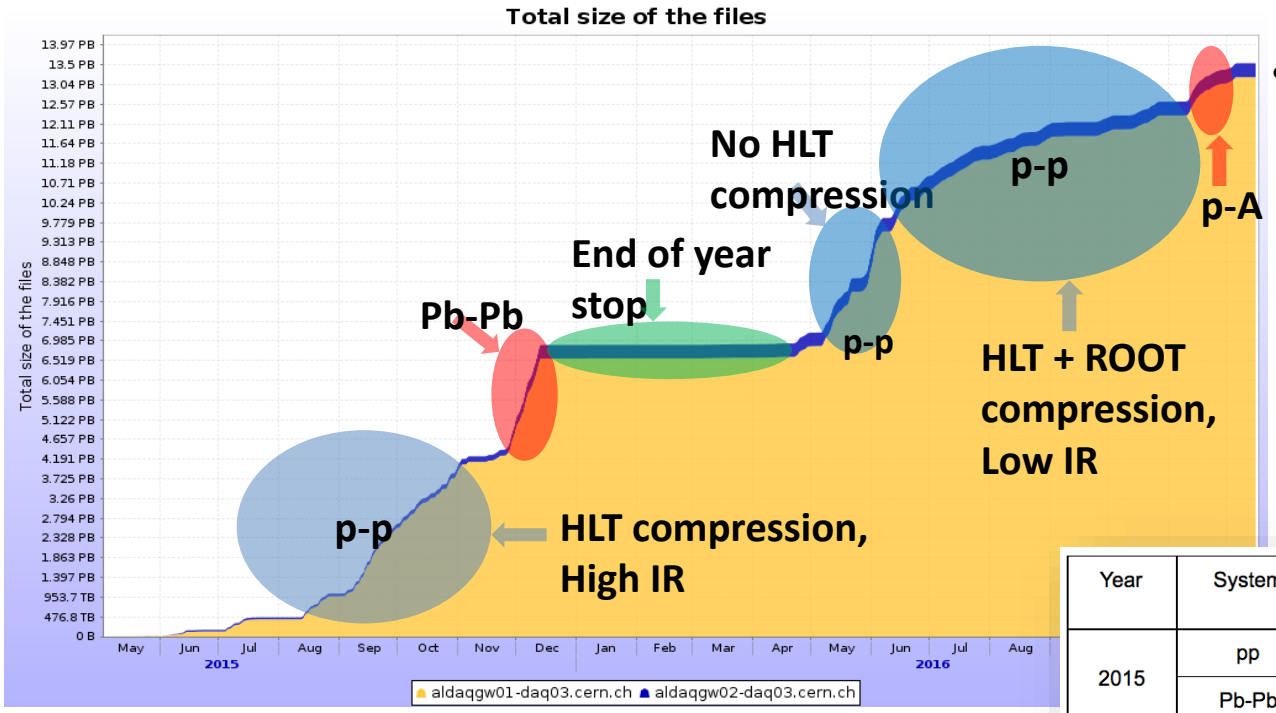
Run 2: Data taking objectives

- **For Pb-Pb collisions:**
  - Reach the target of 1 nb^{-1} integrated luminosity in Pb-Pb for rare triggers.
  - Increase the statistics of the unbiased data sample, including minimum bias and centrality triggered events.

- **For pp collisions:**
  - Collect a reference rare triggers sample with an integrated luminosity of 40 pb^{-1}, which is equivalent to the 1 nb^{-1} sample in Pb-Pb collisions.
  - Enlarge the statistics of the unbiased data sample, including minimum bias collisions at top energy.
  - Collect a reference sample of 10^9 events at the reference energy of 5.02 TeV

- **For p-Pb collisions:**
  - ✓ Enlarge the existing data sample, in particular the unbiased events sample at 5.02 TeV.

Run 2: Progress and data processing status



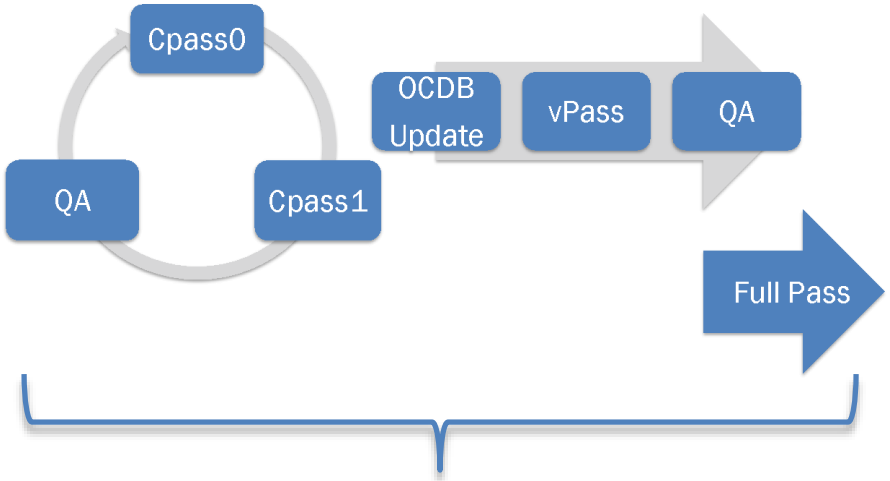
- Readout electronics upgrade (RCU2) with goal to increase the readout rate by a factor of 2
  - Originally planned for 2015, the final version installed from 2016

Year	System	Central barrel events	Average event size (MB)	Data volume (PB)
2015	pp	900 M	4.7	4.23
	Pb-Pb	210 M	12	2.52
2016	pp	1400 M	3.9	5.5
	p-Pb	880 M	1.7	1.5

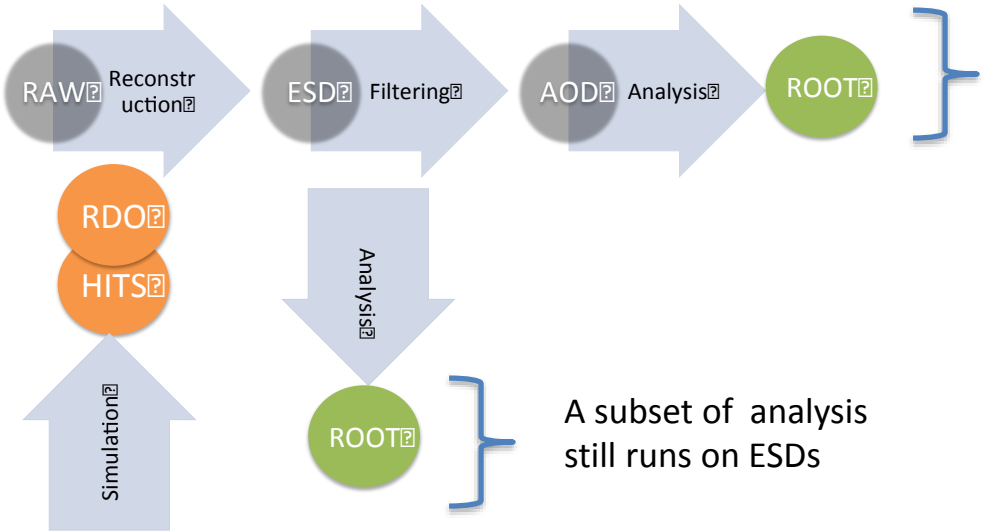
- Data taking follows strictly the approved physics programme
  - Better data quality by limiting the Interaction Rate
  - Continuously reducing RAW data volume by improving the HLT data compression
    - RAW data size reduction by factor ~4.3 to ~5.5 in 2016, expecting factor 6 in 2017

- Data processing – new high-precision calibration schema developed and certified
  - 2015 Pb-Pb and 2016 p-Pb data fully processed, p-p data processing at 90%
  - 2015 p-p data at 60%, processing will be completed by June 2017

Run 2: Computing model



- Three reconstruction passes are performed for every data set with increasing quality
- The first reconstruction pass is started after the end of data taking.
- The second and third reconstruction pass are performed after one month delay following the end of the previous pass.



Analysis on a growing set of AOD is run daily using Analysis Trains

- MC production is continuously spread over the year.
- One MC event per real pp event is produced and 0.175 MC events per real AA event.

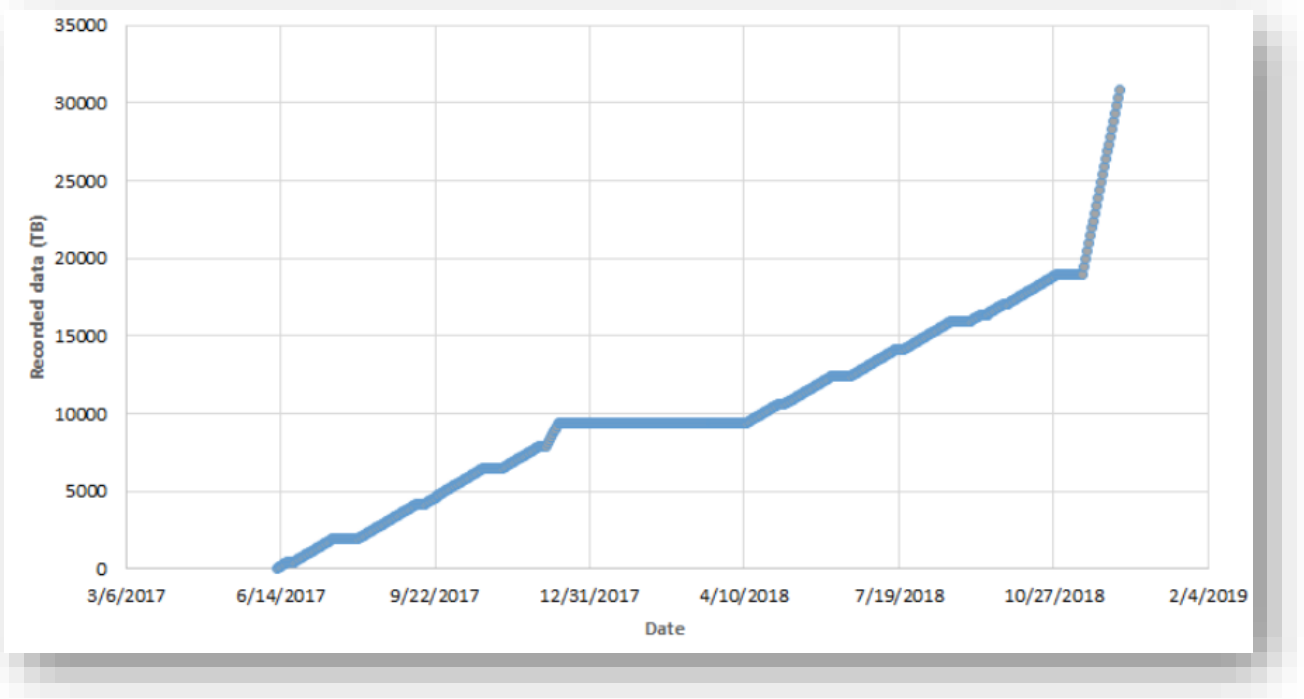
A subset of analysis still runs on ESDs

Changes in Computing Model Parameters

1. RCU2 not ready for 2015 pp/PbPb data taking and became fully operational in 2016
   - The most data intensive data taking with central PbPb collisions postponed to 2018
   - Requires CPU intensive simulation campaign
2. Larger than expected distortions in TPC observed at high IR
   - Fully corrected during 2016 at the cost of +5% CPU increase
3. Average observed pp event size larger than expected (3MB instead of 1.1MB)
   - Due to pileup that we cannot reject online
4. HLT compression improved
   - Compression 4.3 -> 5.5 -> 6
5. Improved data taking efficiency
   - ALICE efficiency 95%
   - LHC efficiency 40% -> 60%
6. Taking data at lower IR
   - Increased combined efficiency allows us to collect expected number of triggers while taking data at lower IR reducing the pileup, event size and processing time

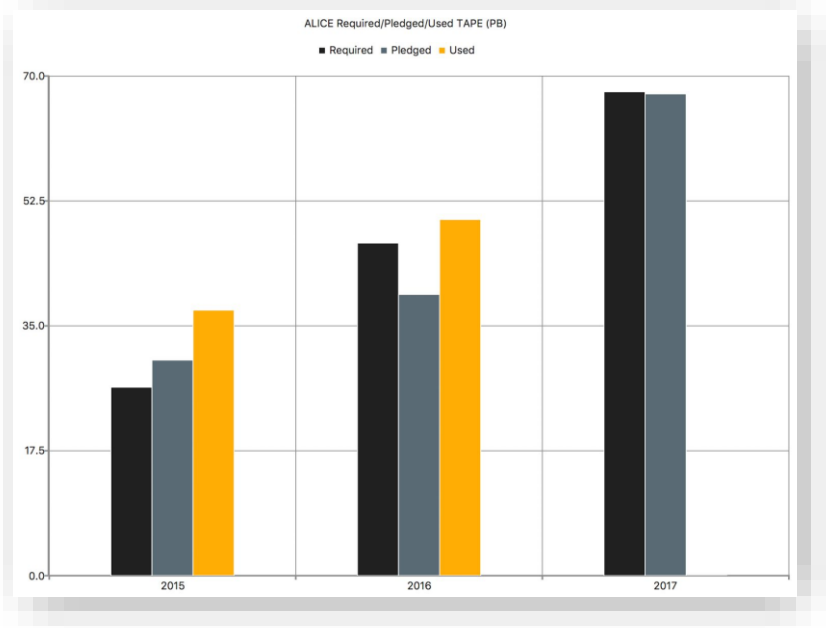
- **All these factors were taken into account in October 2016 when we updated our requirements for 2017-2019.**

Expectations for 2017 and 2018 data volume

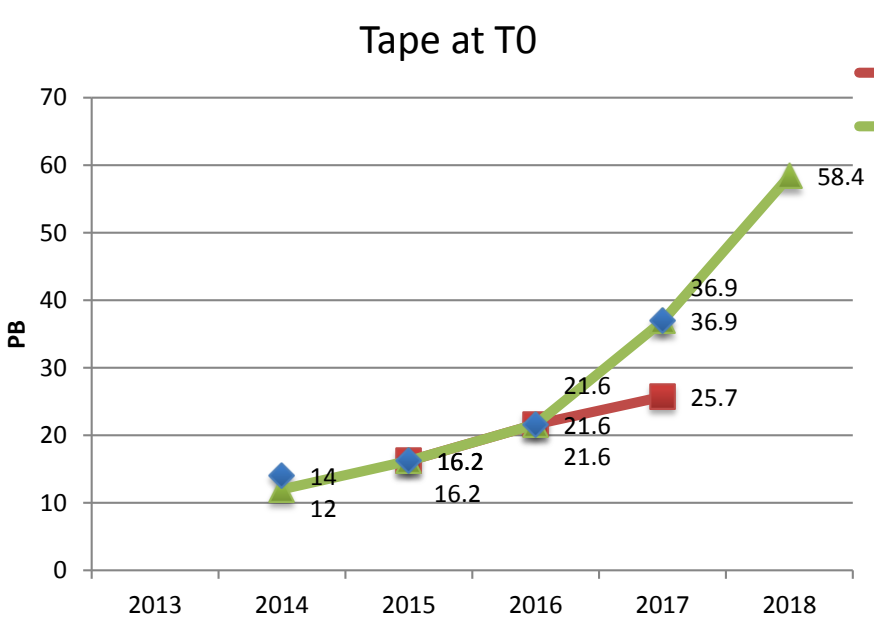


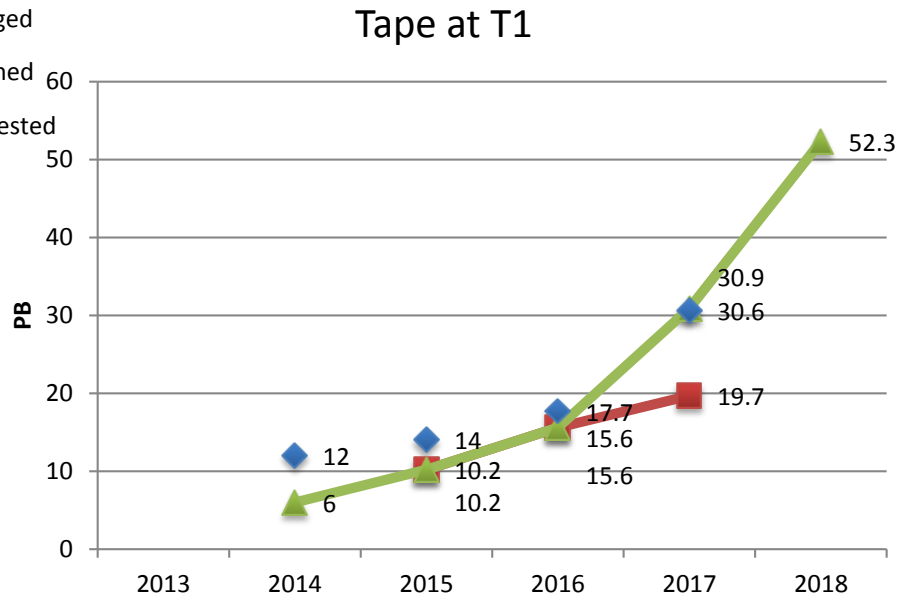
- During pp data taking mode will be set to limit the TPC readout rate to 400 Hz
  - The goal is to reach the statistics objective set for Run 2 in all trigger categories as well as at the reference energy of 5.02 TeV
  - The total amount of data recorded will be **17.5 PB**

- During the Pb-Pb run in 2018, assuming the HLT compression of a factor of 6 we anticipate a total readout rate of 10 GB/s
  - The total amount of data recorded will be **12 PB**

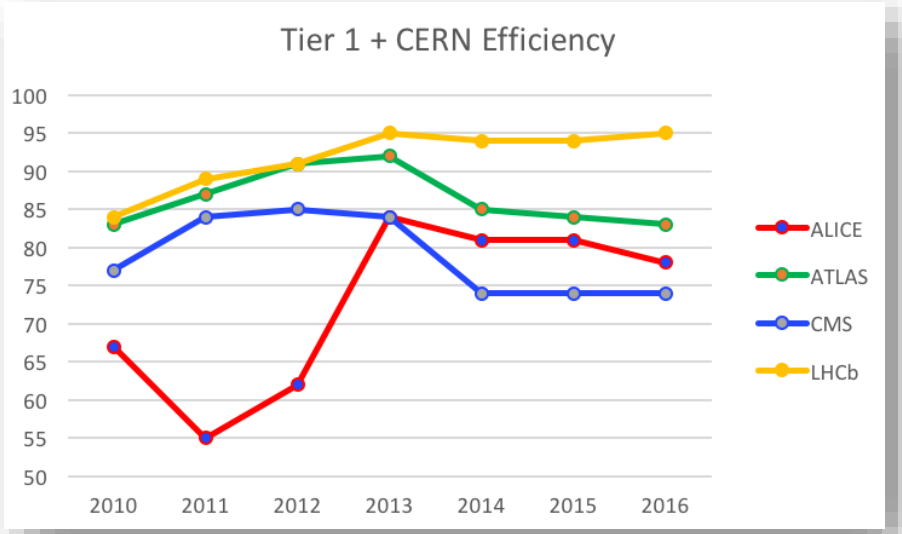


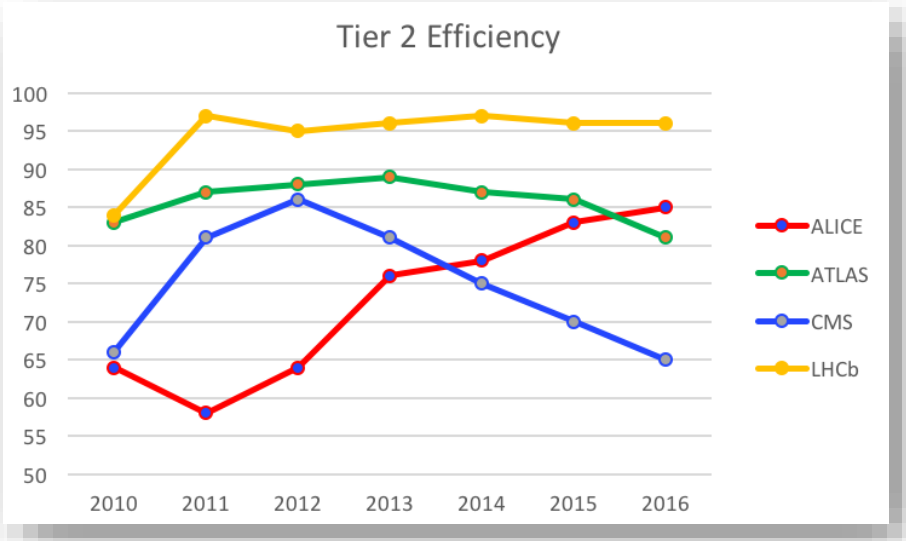
- Thanks to CERN and T1s, our tape request has been fully pledged in 2017

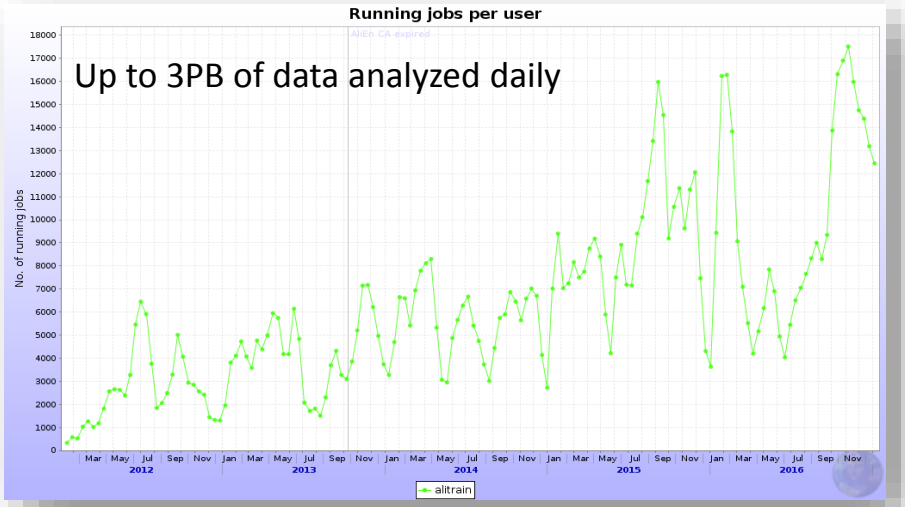


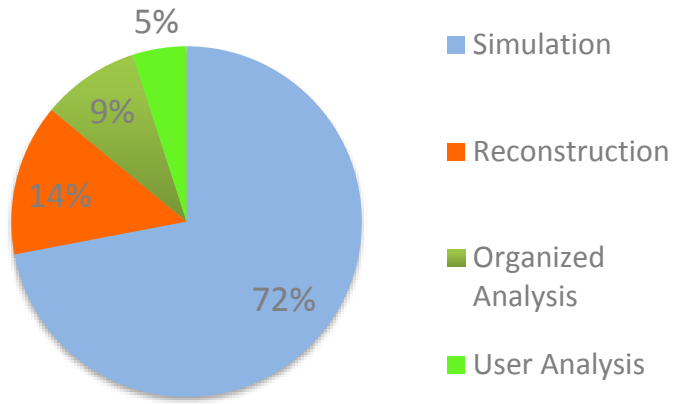


CPU Usage

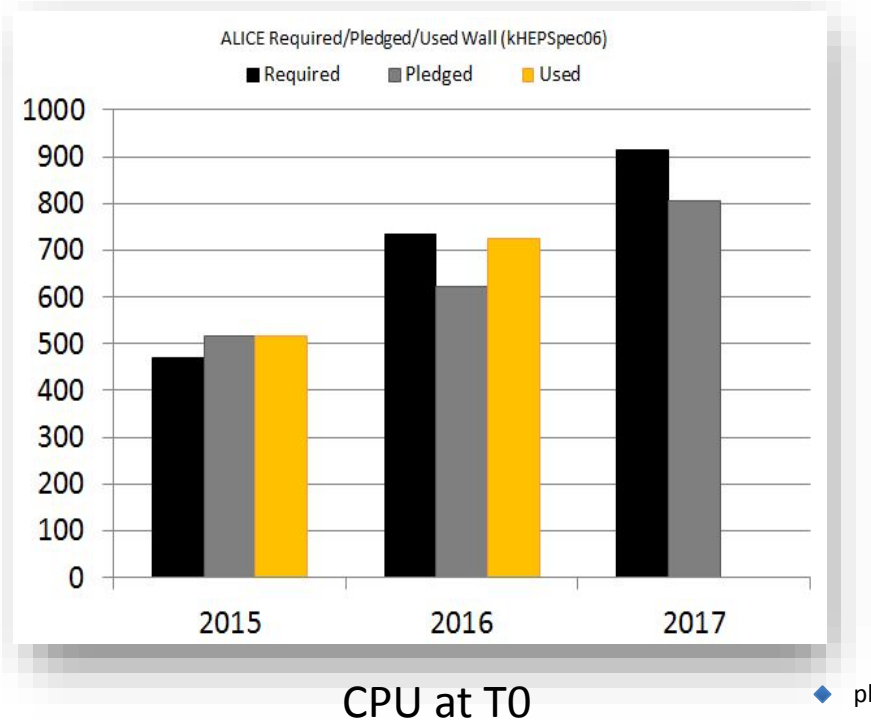




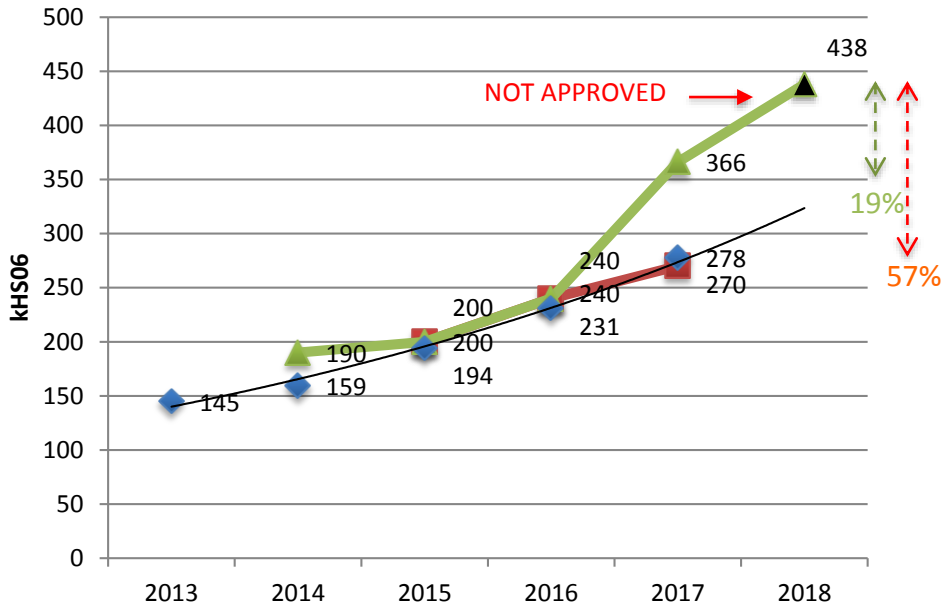




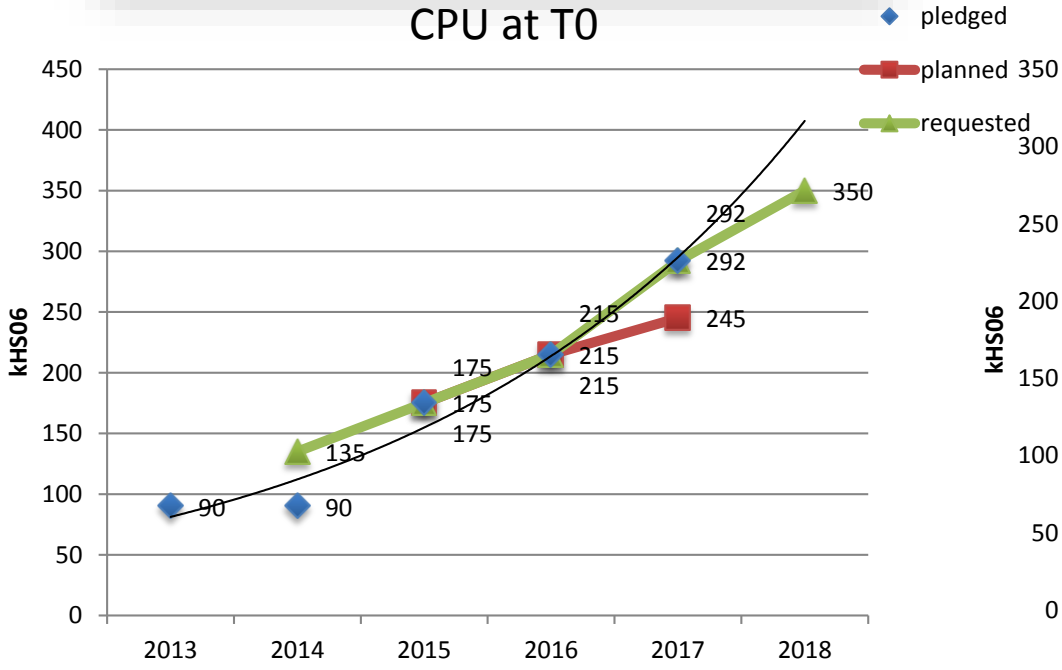
- Up to 130k concurrent jobs
- Available resources are effectively used in spite of increasing complexity and data volume



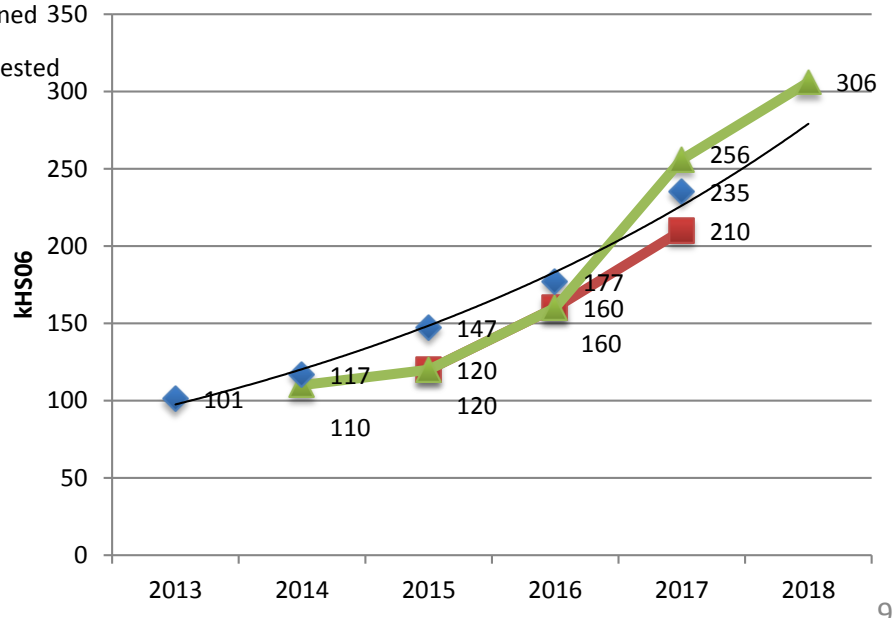
CPU at T2

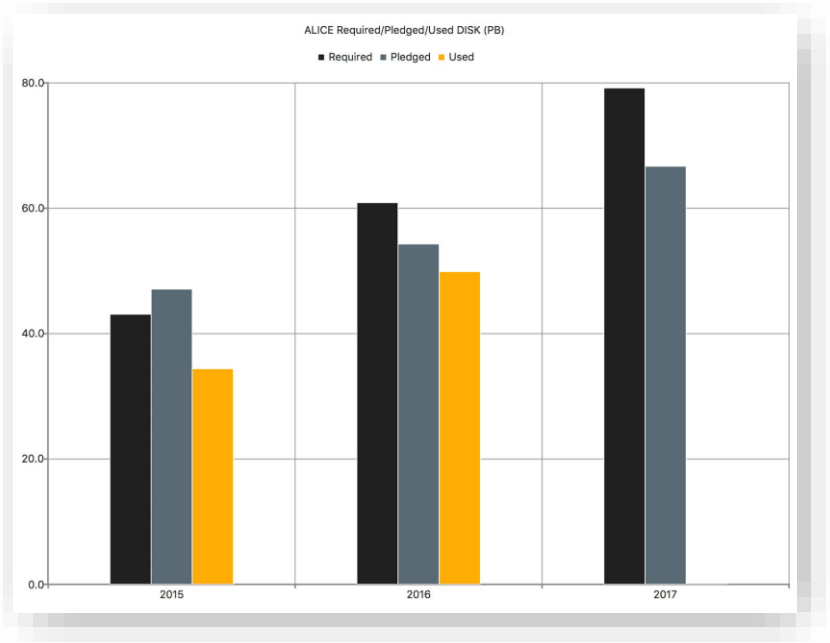


CPU at T0

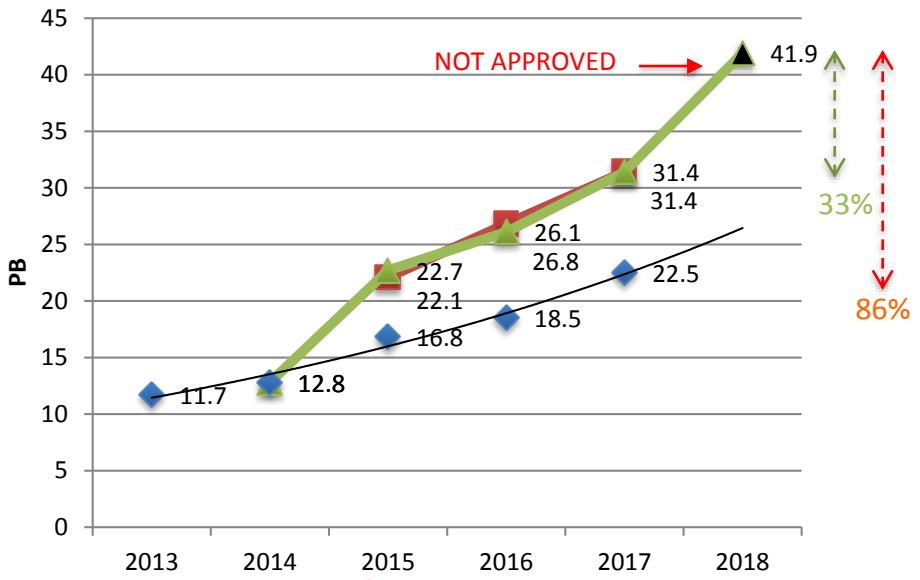


CPU at T1

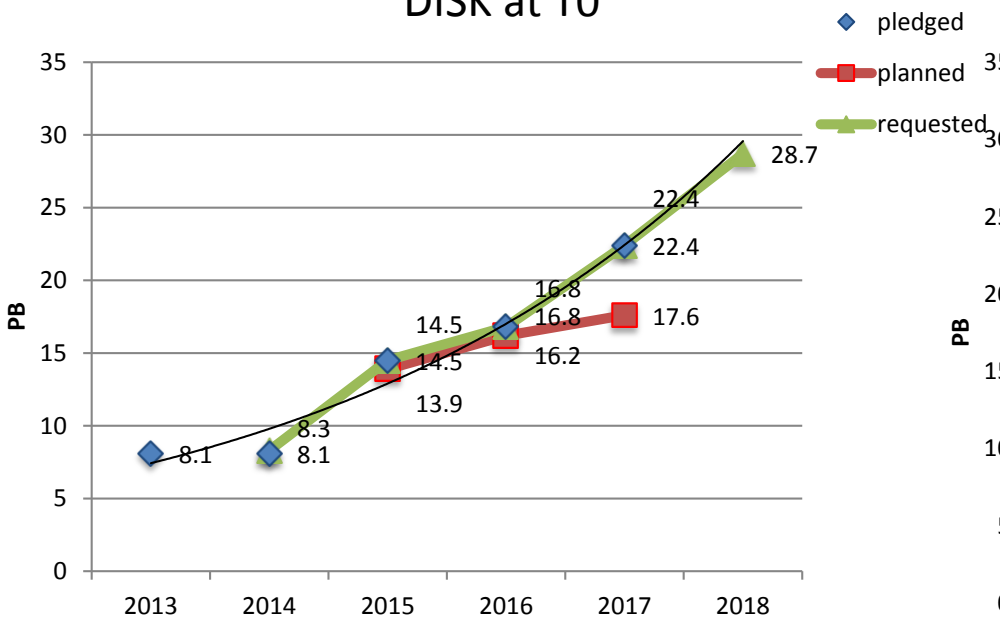




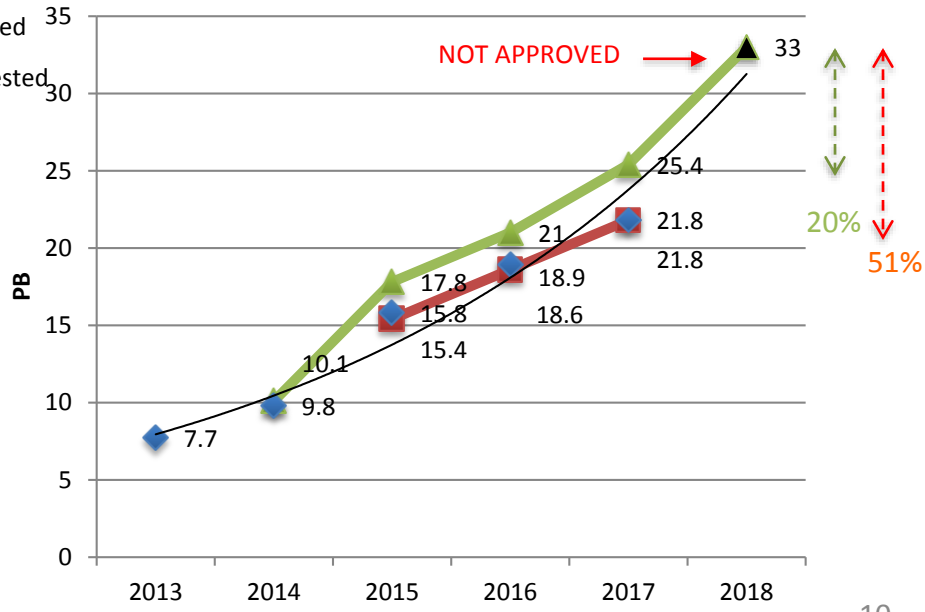
Disk at T2



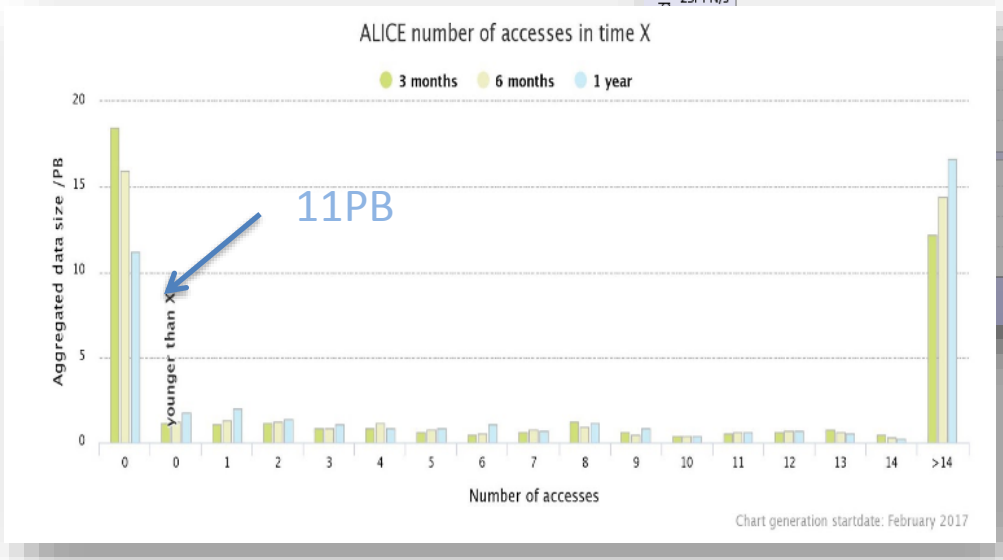
DISK at T0

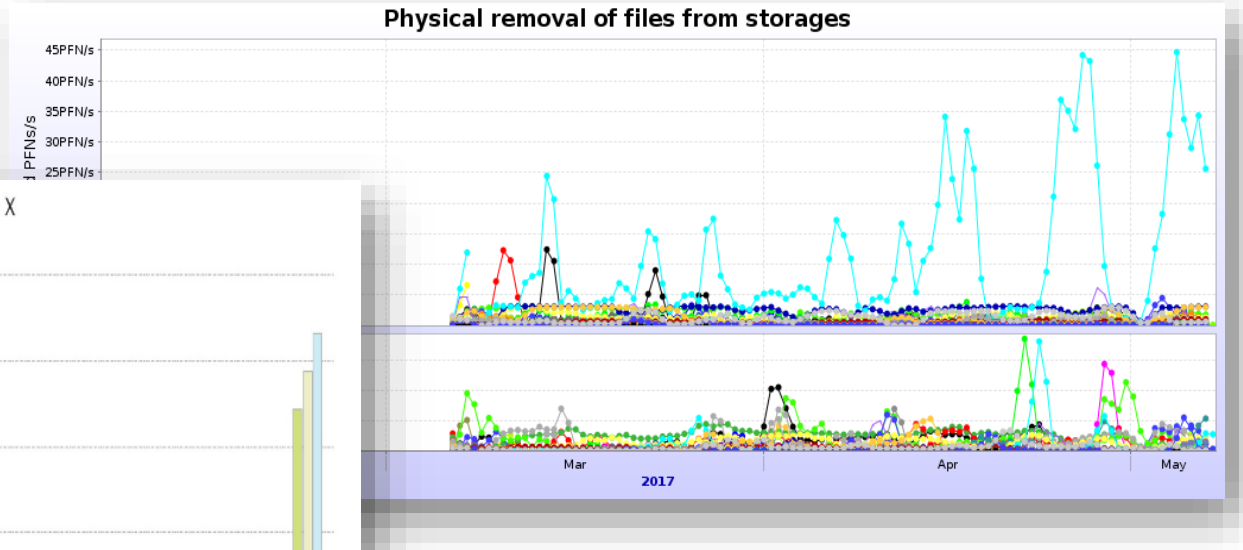


Disk at T1



Data popularity and cleanup

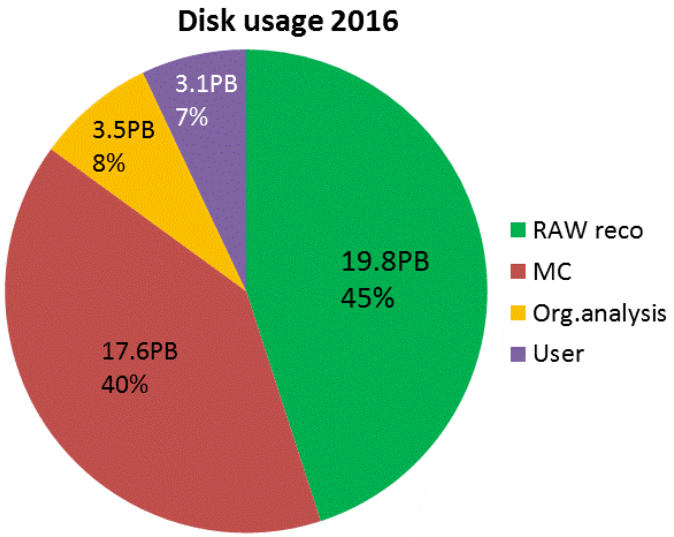


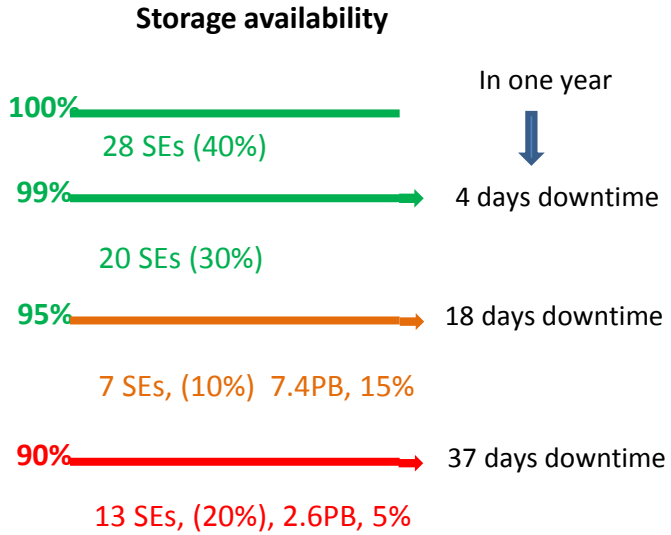


- 1.5PB already removed since February

- Relatively significant volume of data (11PB) that has not been accessed in past 12 months
  - If the window is enlarged to 2 years unused data portion drops down to 2.5 PB
  - We do not systematically back up ESD/AOD to tape, only RAW data
    - Regular cleanup procedure with cooperation of Physics Working Groups
  - Automated cleanup procedure would require
    - To systematically back up ESD/AOD to tape
    - An additional tape budget and larger disk buffer in front of tape system to accommodate ESD/AODs in addition to raw data

What occupies ALICE disk storage?

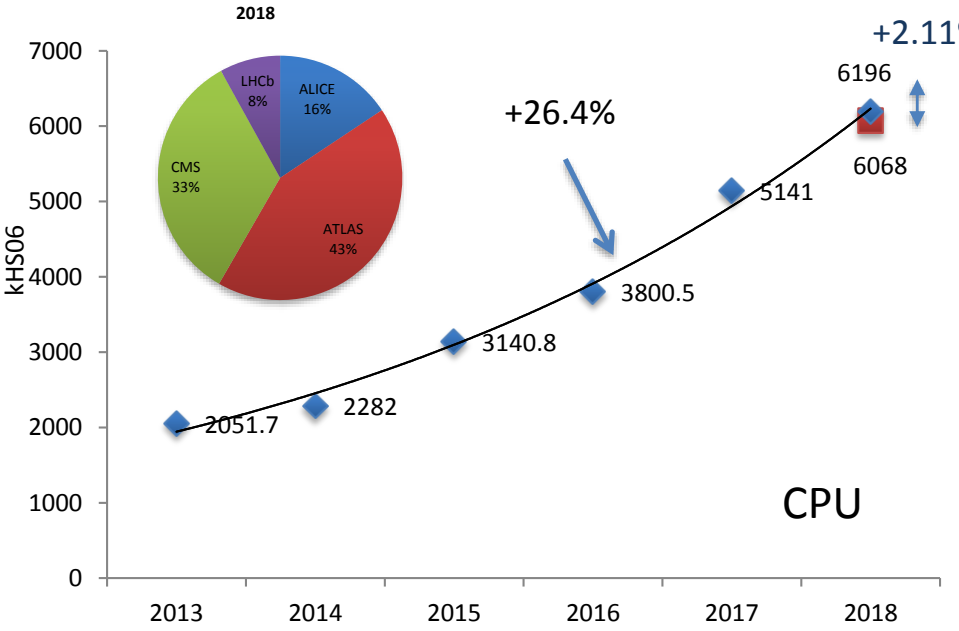


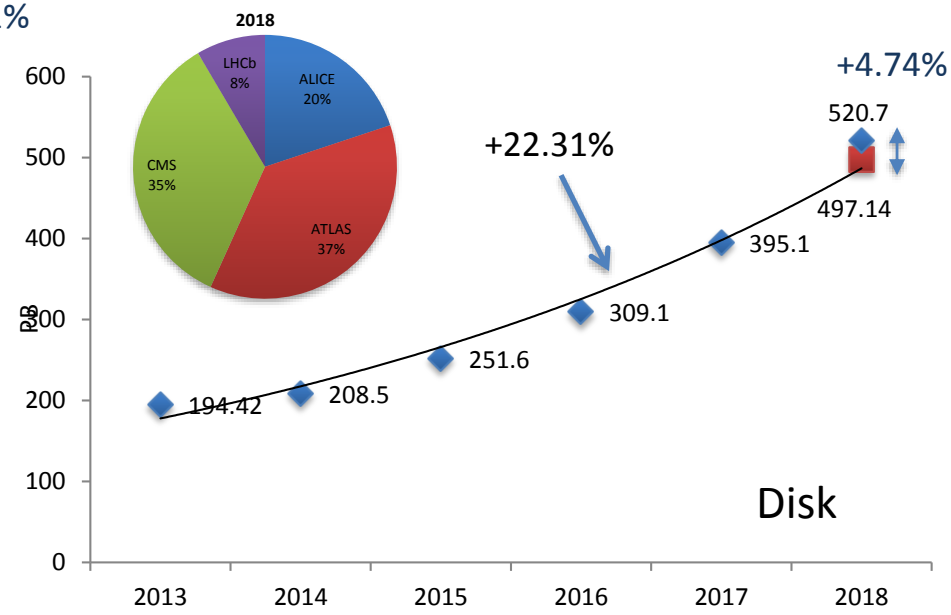


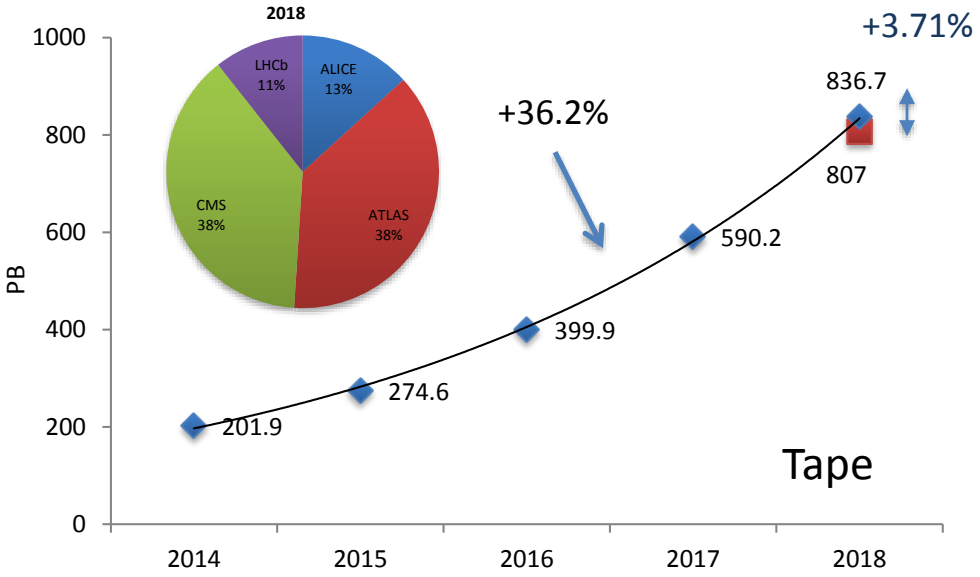
- Each raw event is stored only once on tape and temporary staged in staging disk buffer
- Old inactive MC productions are systematically removed
- All reconstruction output + associated MC
  - Recent ESDs are stored as single file w/o replicas due to lack of disk space
  - 2 replicas for AODs
- Already operating at the lowest limit with no room for more reduction of disk usage
  - ESD based analysis is at the pain threshold with up to 5% jobs failing in each round

Data occupying T0/T1/T2 Storage					
	Event Size [MB]	# of copies on disk		# of versions	# of copies on tape
		minimal	typical		
RAW	3 (pp) 11 (Pb-Pb)			1	2 (one at T0 + one at one of the T1s)
ESD	10 to 30% of RAW, depending on type of collision system and luminosity	1	~~2~~ 1	1-3	
AOD	10 to 15% of RAW, depending on type of collision system and luminosity	1	~~2~~ 2	1-4 per ESD version	
MC ESD	0.37 (pp) 2.7 (Pb-Pb)	1	~~2~~ 1	1	
MC AOD	30% of MC ESD	1	~~2~~ 2	2	

ALICE share of computing resources







- ALICE share of total resources remains balanced vs. the other experiments
- Historically, growth in all resource categories was faster than +20% year on year
- The impact of fully implemented ALICE 2018 request vs just +20% increase is relatively small

Projection 2017-2020 assuming 20% growth

		CPU			Disk			Tape		
	Tier	Requested[1)]	Pledged[2)]	missing	requested	pledged	missing	requested	pledged	missing
2017	0	292.00	292.00	0.00%	20.60	22.40	8.74%	34.5	36.9	6.96%
	1	256.00	235.48	-8.02%	24.40	21.81	-10.64%	26.60	30.60	15.04%
	2	366.00	323.05	-11.74%	31.20	26.39	-15.41%			
	Total	914.00	850.53	-6.94%	76.20	70.60	-7.35%	61.10	67.50	10.47%
2018	0	350.00	350.40	0.11%	27.00	26.88	-0.44%	55	55	
	1	306.00	282.58	-7.65%	32.00	26.17	-18.23%	41.00	41.00	
	2	438.00	387.66	-11.49%	41.00	31.67	-22.76%			
	Total	1094.00	1020.63	-6.71%	100.00	84.72	-15.28%	96.00	96.00	0.00%
2019	0	534.00	420.48	-21.26%	33.60	32.26	-4.00%	55	55	
	1	501.00	339.09	-32.32%	39.90	31.40	-21.31%	49.50	49.50	
	2	635.00	465.19	-26.74%	51.10	38.00	-25.63%			
	Total	1670.00	1224.76	-26.66%	124.60	101.66	-18.41%	104.50	104.50	0.00%
2020	0	534.00	504.58	-5.51%	33.60	38.71	15.20%	55	55	
	1	501.00	406.91	-18.78%	39.90	37.68	-5.57%	49.50	49.50	
	2	635.00	558.22	-12.09%	51.10	45.61	-10.75%			
	Total	1670.00	1469.71	-11.99%	124.60	121.99	-2.09%	104.50	104.50	0.00%

1) The resource request is based on April submission to C-RSG
2) In this table the resources provided by US, Japan and Austria are counted as fully pledged

- If we are forced to remain under 20% resource growth envelope and cannot find additional resource we won't be able to process 2018 pp and PbPb data during 2019
- The only solution would be to extend the processing time to 2019 + 2020
- This would put ALICE at disadvantage with respect to other LHC experiments in areas were they are competitive with us
- Manpower would remain tied to Run2 processing during time that we need to focus on preparations for Run3

Conclusions

- ALICE data taking
  - Strictly following the approved physics programme
  - Intensifies in second half of Run 2 as we profit from improved readout and LHC performance to catch up with our plans

- Pledged and delivered computing resources
  - Systematic under-pledges and under-delivery at T2s over several years
  - This artificially skews our 2018 requirements/pledges ratio prompting C-RSG to question our 2018 requirements

- 2018 data taking will be particularly data intensive
  - pp and PbPb runs with high multiplicity central triggers
  - If our request are not fulfilled, we will have to extend the processing time to 2019 + 2020

- Radical measures already taken
  - Reducing replicas, now at absolute minimum (no replicas for any file but AODs), aggressive disk cleanup

- We are in discussion with FAs and have positive initial feedback concerning the pledges but we need support for LHCC/C-RSG or else there will be no funding available for us in October