

Disk capacities in 2017 and 2018

ALICE Offline week

12/11/2017

Current capacity

- Total pledged capacity for 2017: **66.9PB**
 - Delta from 2016: 12.6PB
 - Installed as of today: **61PB** (90%)
 - Largest deficit is at T2s
- Note that the disk is installed along the year, the figure above not yet final
- Used/free space: 46PB / **15PB(25%)**
 - Trying to keep free space balance across SEs

SE occupancy – top 15

SE Name	Size	Used	Free	Usage
6. CCIN2P3 - SE	2.629 PB	2.583 PB	47.67 TB	98.23%
45. RRC_KI_T1 - EOS	2.267 PB	2.212 PB	56 TB	97.59%
42. Prague - SE	1.591 PB	1.542 PB	50.69 TB	96.89%
28. KISTI_GSDC - SE2	1.446 PB	1.395 PB	52.87 TB	96.43%
53. Subatech - EOS	1.339 PB	1.288 PB	51.82 TB	96.22%
24. ITEP - SE	20.01 TB	19.23 TB	798.8 GB	96.1%
44. RRC-KI - SE	213.7 TB	204.2 TB	9.474 TB	95.57%
34. Legnaro - SE	1.128 PB	1.075 PB	54.53 TB	95.28%
10. CNAF - SE	3.405 PB	3.229 PB	180.9 TB	94.81%
15. GRIF_IPNO - SE	669.3 TB	620.3 TB	49.02 TB	92.68%
37. NIHAM - FILE	829.2 TB	766.2 TB	63.09 TB	92.39%
1. Bari - SE	1.201 PB	1.101 PB	101.6 TB	91.73%
13. FZK - SE	5.221 PB	4.691 PB	543.6 TB	89.83%
23. ISS - FILE	421.5 TB	378.4 TB	43.06 TB	89.78%
2. Birmingham - SE	418.4 TB	366.5 TB	51.87 TB	87.6%
14. Grenoble - SE	319.8 TB	279.9 TB	39.84 TB	87.54%
33. LBL - EOS	826 TB	719 TB	107 TB	87.05%
5. Catania - SE	1.1 PB	977.8 TB	148.6 TB	86.81%
39. ORNL - EOS	1.381 PB	1.188 PB	197.2 TB	86.05%

2017 capacity installed

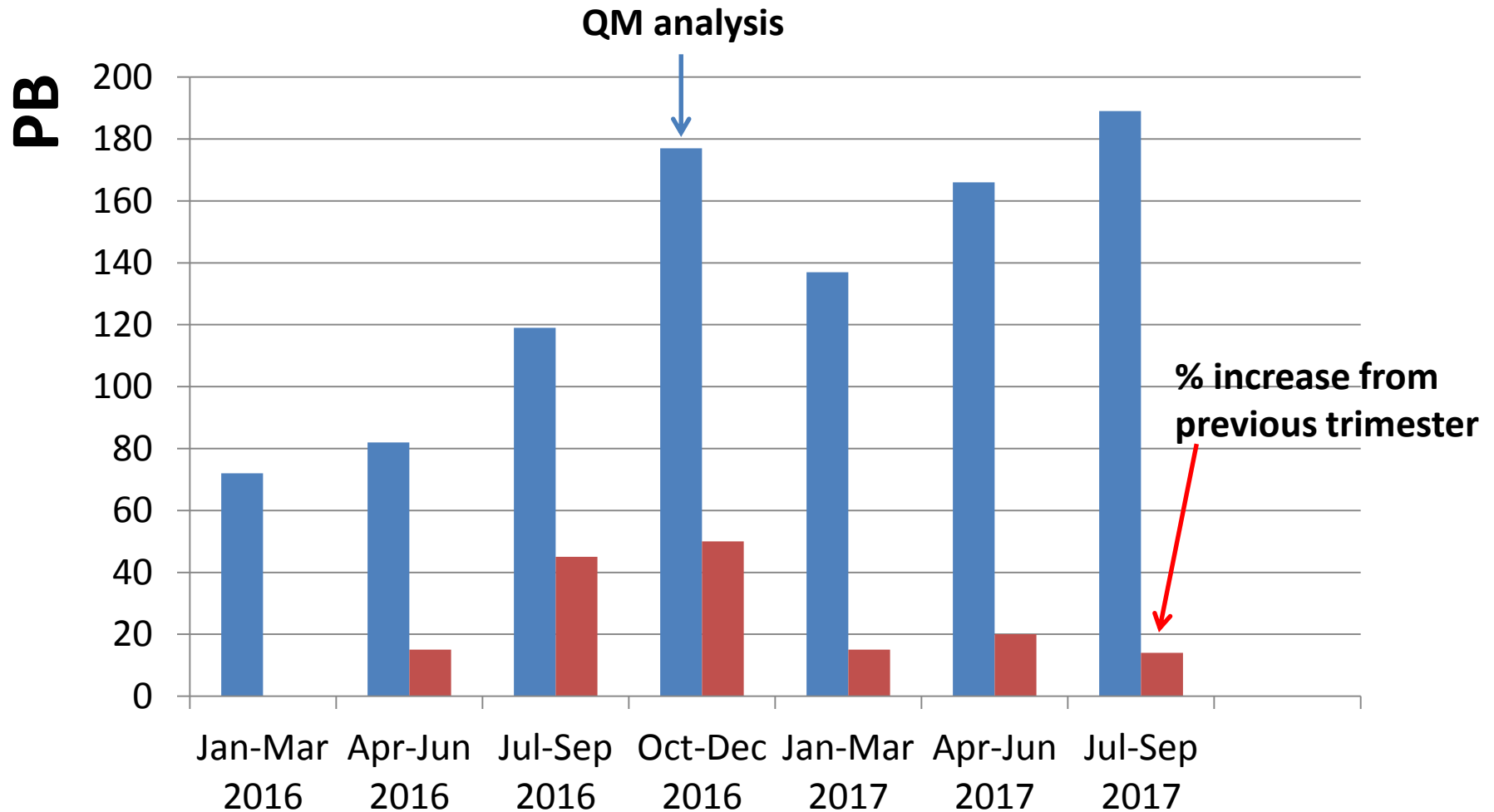
No increase expected

2017 capacity installed

What is being done

- Balancing the site storage is a challenging task, as the SEs become full fast
- Data removal is done on SEs with highest occupancy first (if possible)
- At some point, the SE is closed for writing, at the expense of network use – all data produced at a site with full SE goes to another site

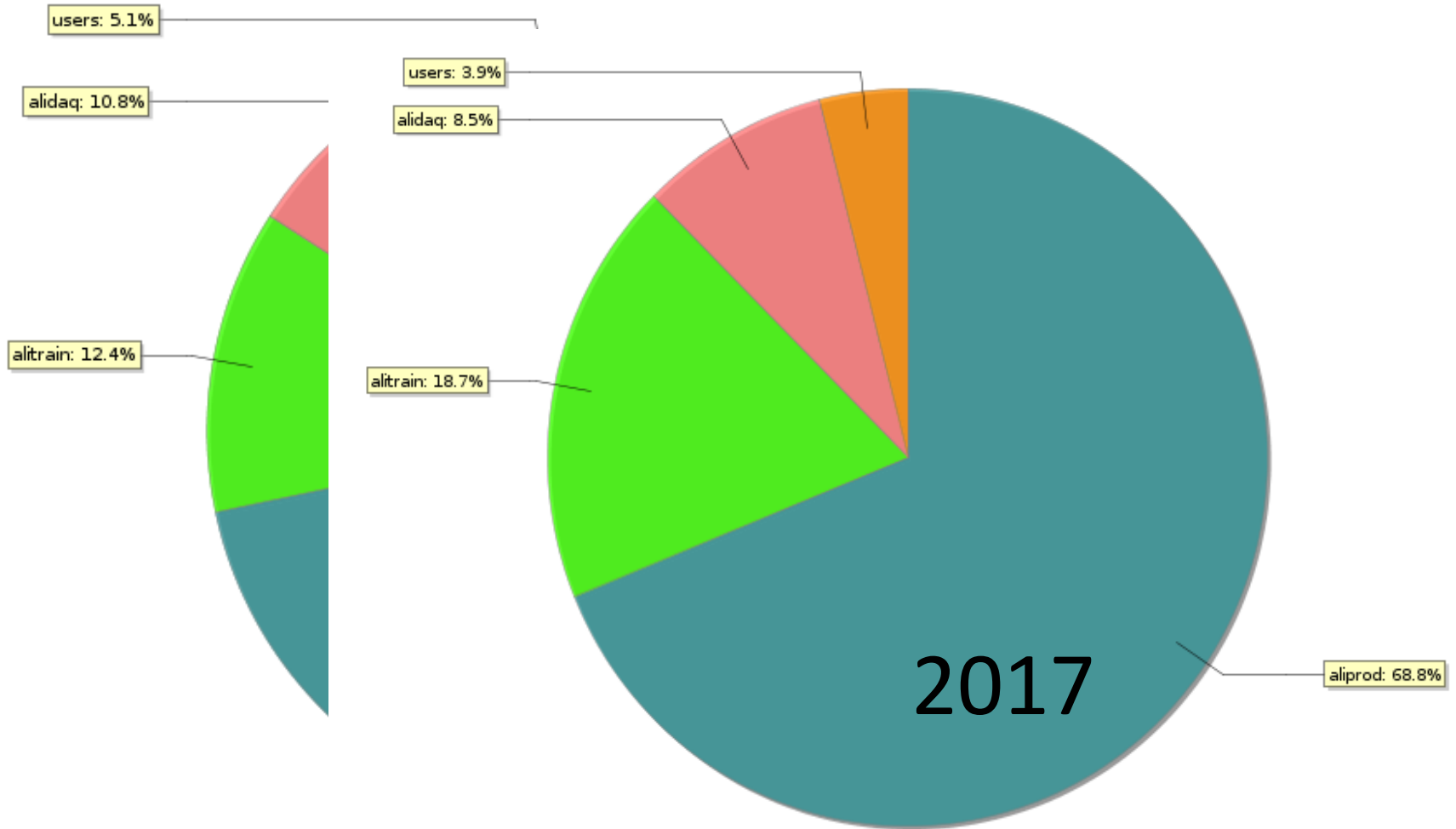
Evolution of disk use in analysis



Statistics

- We double the load on the storage every year
 - Capacity grows by 15% only
- The data is consumed by the organized analysis
 - Number of jobs doubled from 2016 (10k average) to 2017 (20k average)
 - 20% of total resource (103k average)
 - The amount of resources used for analysis is accelerating (was 12% in 2016)
- OK so far, but how does that project to Run 3?
- Very difficult to predict what will be needed even toward the end of Run2...

Analysis progression



Production statistics

- Occupied space from recent productions

Year	RAW [PB]	MC [PB]
2015	2.4	11
2016	2.4	2.6
2017	0.4	0.3
Total:	5	13.9

- 2017 – not all periods ready, MC mostly GP
- All of the above is Pass1 (short periods in 2015 have Pass2)
- Total occupied space from recent productions is **31%** from total disk

Cleanup procedures

- RAW and MC not accessed more than 1 year
 - Done, will be repeated
- User space quotas
 - Done, strict quotas applied
 - 25 users above quota and staying there
- Removal of intermediate files
 - Unmerged (per chunk) QA and AOD – a lot of files, cleanup started a week ago...

What is left

ALICE number of accesses in time X

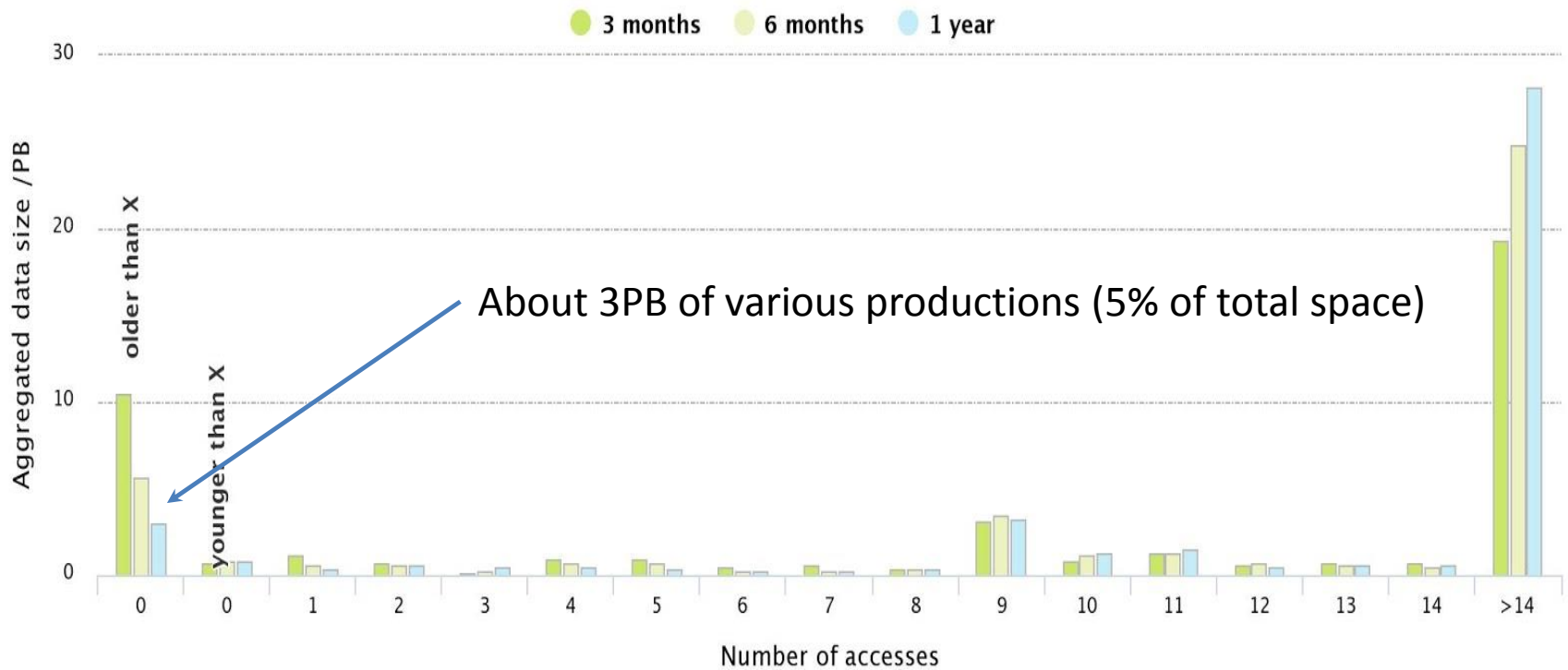


Chart generation startdate: July 2017

Further space gains

- Already at minimum with replicas
- Perhaps put some of the least accessed productions on tape (few PB gain)
 - Meaning those will be accessible **only** through planning process
 - Will affect RAW data reco in some form
- Scraping the bottom of the barrel

Projections for 2018/2019

- Largest period is Pb-Pb end of 2018
 - Projected to take 11PB of RAW data (to be compared to LHC15o 2PB in 2015)
 - ESDs/AODs are 1/3 of RAW
- If we assume the resulting ESDs/AODs will be of the same size (conservative) => projected **7.2 PB** disk (2 replicas) per RAW data production cycle
- LHC15o anchored MC: 72 productions, 8.6PB => projected **47.3 PB** disk for MC productions for 2018 Pb-Pb

What will be available/what to do

- Pledges for 2018 disk = **85.4 PB** (delta 2017 = **18.5 PB**)
 - Unclear how much will be the 2019 increase
- We should revise our MC productions needs for Pb-Pb MC
- We should be very careful with the production cycles
 - Good practice from this year is to remove immediately faulty periods
 - The LHC15/LHC16 Pass1 is effectively equivalent to Pass2
- Better is to repeat production from scratch (CPU cost only) than to keep previous production

What will be available/what to do

- Main consumer (reader) of data is organized analysis
 - Individual analysis steady, even slight decline
- The resource increase this year (20% use total) is above all historical projections
- Toward end of 2018/beginning of 2019 may become a serious bottleneck
 - Evaluation needed (similar to MC)